



Process Management Interface for Exascale (PMIx) Standard

Version 5.0

May 2023

This document describes the Process Management Interface for Exascale (PMIx) Standard, version 5.0.

Comments: Please provide comments on the PMIx Standard by filing issues on the document repository <https://github.com/pmix/pmix-standard/issues> or by sending them to the PMIx Community mailing list at <https://groups.google.com/forum/#!forum/pmix>. Comments should include the version of the PMIx standard you are commenting about, and the page, section, and line numbers that you are referencing. Please note that messages sent to the mailing list from an unsubscribed e-mail address will be ignored.

Copyright © 2018-2023 PMIx Administrative Steering Committee (ASC).

Permission to copy without fee all or part of this material is granted, provided the PMIx ASC copyright notice and the title of this document appear, and notice is given that copying is by permission of PMIx ASC.

This page intentionally left blank

Contents

1. Introduction	1
1.1. Background	1
1.2. PMIx Architecture Overview	1
1.3. Portability of Functionality	3
1.3.1. Attributes in PMIx	3
1.3.2. PMIx Roles	5
1.3.3. Application Binary Interface (ABI)	6
1.3.3.1. PMIx ABI Versioning	7
1.3.3.2. Linker versus Build ABI	8
2. PMIx Terms and Conventions	9
2.1. Notational Conventions	11
2.2. Semantics	13
2.3. Naming Conventions	13
2.4. Procedure Conventions	13
3. Data Structures and Types	15
3.1. Constants	16
3.1.1. PMIx Return Status Constants	17
3.1.1.1. User-Defined Error and Event Constants	19
3.2. Data Types	19
3.2.1. Key Structure	19
3.2.1.1. Key support macros	20
3.2.2. Namespace Structure	21
3.2.2.1. Namespace support macros	21
3.2.3. Rank Structure	22
3.2.3.1. Rank support macros	23
3.2.4. Process Structure	23
3.2.4.1. Process structure support macros	24

3.2.5.	Process State Structure	28
3.2.6.	Process Information Structure	29
3.2.6.1.	Process information structure support macros	29
3.2.7.	Job State Structure	31
3.2.8.	Value Structure	31
3.2.8.1.	Value structure support	32
3.2.9.	Info Structure	36
3.2.9.1.	Info structure support macros	36
3.2.9.2.	Info structure list macros	39
3.2.10.	Info Type Directives	41
3.2.10.1.	Info Directive support macros	42
3.2.11.	Environmental Variable Structure	44
3.2.11.1.	Environmental variable support macros	44
3.2.12.	Byte Object Type	46
3.2.12.1.	Byte object support macros	46
3.2.13.	Data Array Structure	48
3.2.13.1.	Data array support macros	48
3.2.14.	Argument Array Macros	49
3.2.15.	Set Environment Variable	52
3.3.	Generalized Data Types Used for Packing/Unpacking	53
3.4.	General Callback Functions	55
3.4.1.	Release Callback Function	55
3.4.2.	Op Callback Function	55
3.4.3.	Value Callback Function	56
3.4.4.	Info Callback Function	56
3.4.5.	Handler registration callback function	57
3.5.	PMIx Datatype Value String Representations	57
4.	Client Initialization and Finalization	61
4.1.	PMIx_Initialized	61
4.2.	PMIx_Get_version	62
4.3.	PMIx_Init	62
4.3.1.	Initialization events	65

4.3.2.	Initialization attributes	65
4.3.2.1.	Connection attributes	65
4.3.2.2.	Programming model attributes	65
4.4.	PMIx_Finalize	66
4.4.1.	Finalize attributes	67
4.5.	PMIx_Progress	67
5.	Data Access and Sharing	68
5.1.	Non-reserved keys	70
5.2.	Posting Key/Value Pairs	70
5.2.1.	PMIx_Put	71
5.2.1.1.	Scope of Put Data	72
5.2.2.	PMIx_Store_internal	72
5.2.3.	PMIx_Commit	73
5.3.	Retrieval rules for non-reserved keys	73
5.4.	PMIx_Get	74
5.4.1.	PMIx_Get_nb	77
5.4.2.	Retrieval attributes	80
6.	Reserved Keys	81
6.1.	Data realms	81
6.1.1.	Session realm attributes	82
6.1.2.	Job realm attributes	83
6.1.3.	Application realm attributes	85
6.1.4.	Process realm attributes	86
6.1.5.	Node realm keys	88
6.2.	Retrieval rules for reserved keys	89
6.2.1.	Accessing information: examples	90
6.2.1.1.	Session-level information	90
6.2.1.2.	Job-level information	91
6.2.1.3.	Application-level information	91
6.2.1.4.	Process-level information	92
6.2.1.5.	Node-level information	92

7. Query Operations	94
7.1. <code>PMIx_Query_info</code>	94
7.1.1. Query Structure	94
7.1.2. <code>PMIx_Query_info</code>	95
7.1.3. <code>PMIx_Query_info_nb</code>	100
7.1.4. Query keys	104
7.1.5. Query attributes	106
7.1.5.1. Query structure support macros	107
7.2. <code>PMIx_Resolve_peers</code>	108
7.2.1. <code>PMIx_Resolve_nodes</code>	109
7.3. Using Get vs Query	110
7.4. Accessing attribute support information	110
8. Synchronization	113
8.1. <code>PMIx_Fence</code>	113
8.2. <code>PMIx_Fence_nb</code>	115
8.2.1. Fence-related attributes	117
9. Publish/Lookup Operations	119
9.1. <code>PMIx_Publish</code>	119
9.2. <code>PMIx_Publish_nb</code>	120
9.3. Publish-specific constants	122
9.4. Publish-specific attributes	122
9.5. Publish-Lookup Datatypes	122
9.5.1. Range of Published Data	123
9.5.2. Data Persistence Structure	123
9.6. <code>PMIx_Lookup</code>	123
9.7. <code>PMIx_Lookup_nb</code>	125
9.7.1. Lookup Returned Data Structure	127
9.7.1.1. Lookup data structure support macros	127
9.7.2. Lookup Callback Function	130
9.8. Retrieval rules for published data	131
9.9. <code>PMIx_Unpublish</code>	131
9.10. <code>PMIx_Unpublish_nb</code>	132

10. Event Notification	135
10.1. Notification and Management	135
10.1.1. Events versus status constants	137
10.1.2. PMIx_Register_event_handler	137
10.1.3. Event registration constants	140
10.1.4. System events	140
10.1.5. Event handler registration and notification attributes	140
10.1.5.1. Fault tolerance event attributes	141
10.1.5.2. Hybrid programming event attributes	141
10.1.6. Notification Function	142
10.1.7. PMIx_Deregister_event_handler	143
10.1.8. PMIx_Notify_event	144
10.1.9. Notification Handler Completion Callback Function	147
10.1.9.1. Completion Callback Function Status Codes	148
11. Data Packing and Unpacking	149
11.1. Data Buffer Type	149
11.2. Support Macros	149
11.3. General Routines	151
11.3.1. PMIx_Data_pack	151
11.3.2. PMIx_Data_unpack	153
11.3.3. PMIx_Data_copy	154
11.3.4. PMIx_Data_print	155
11.3.5. PMIx_Data_copy_payload	156
11.3.6. PMIx_Data_load	156
11.3.7. PMIx_Data_unload	157
11.3.8. PMIx_Data_compress	158
11.3.9. PMIx_Data_decompress	158
11.3.10. PMIx_Data_embed	159
12. Process Management	161
12.1. Abort	161
12.1.1. PMIx_Abort	161

12.2.	Process Creation	162
12.2.1.	PMIx_Spawn	162
12.2.2.	PMIx_Spawn_nb	168
12.2.3.	Spawn-specific constants	173
12.2.4.	Spawn attributes	173
12.2.5.	Application Structure	176
	12.2.5.1. App structure support macros	177
	12.2.5.2. Spawn Callback Function	179
12.3.	Connecting and Disconnecting Processes	180
12.3.1.	PMIx_Connect	180
12.3.2.	PMIx_Connect_nb	182
12.3.3.	PMIx_Disconnect	184
12.3.4.	PMIx_Disconnect_nb	186
12.4.	Process Locality	187
12.4.1.	PMIx_Load_topology	187
12.4.2.	PMIx_Get_relative_locality	188
	12.4.2.1. Topology description	189
	12.4.2.2. Topology support macros	189
	12.4.2.3. Relative locality of two processes	191
	12.4.2.4. Locality keys	191
12.4.3.	PMIx_Parse_cpuset_string	191
12.4.4.	PMIx_Get_cpuset	192
	12.4.4.1. Binding envelope	192
12.4.5.	PMIx_Compute_distances	192
12.4.6.	PMIx_Compute_distances_nb	193
12.4.7.	Device Distance Callback Function	194
12.4.8.	Device type	195
12.4.9.	Device Distance Structure	196
12.4.10.	Device distance support macros	196
12.4.11.	Device distance attributes	198
13.	Job Management and Reporting	199
13.1.	Allocation Requests	199
13.1.1.	PMIx_Allocation_request	199

13.1.2.	PMIx_Allocation_request_nb	202
13.1.3.	Job Allocation attributes	204
13.1.4.	Job Allocation Directives	205
13.2.	Job Control	206
13.2.1.	PMIx_Job_control	206
13.2.2.	PMIx_Job_control_nb	208
13.2.3.	Job control constants	211
13.2.4.	Job control events	211
13.2.5.	Job control attributes	212
13.3.	Process and Job Monitoring	213
13.3.1.	PMIx_Process_monitor	213
13.3.2.	PMIx_Process_monitor_nb	215
13.3.3.	PMIx_Heartbeat	217
13.3.4.	Monitoring events	217
13.3.5.	Monitoring attributes	218
13.4.	Logging	218
13.4.1.	PMIx_Log	218
13.4.2.	PMIx_Log_nb	221
13.4.3.	Log attributes	224
14.	Process Sets and Groups	226
14.1.	Process Sets	226
14.1.1.	Process Set Constants	227
14.1.2.	Process Set Attributes	228
14.2.	Process Groups	228
14.2.1.	Relation to the host environment	228
14.2.2.	Construction procedure	229
14.2.3.	Destruct procedure	230
14.2.4.	Process Group Events	230
14.2.5.	Process Group Attributes	231
14.2.6.	PMIx_Group_construct	232
14.2.7.	PMIx_Group_construct_nb	235
14.2.8.	PMIx_Group_destruct	238
14.2.9.	PMIx_Group_destruct_nb	239

14.2.10. PMIx_Group_invite	240
14.2.11. PMIx_Group_invite_nb	243
14.2.12. PMIx_Group_join	245
14.2.13. PMIx_Group_join_nb	246
14.2.13.1. Group accept/decline directives	248
14.2.14. PMIx_Group_leave	248
14.2.15. PMIx_Group_leave_nb	249
15. Fabric Support Definitions	251
15.1. Fabric Support Events	254
15.2. Fabric Support Datatypes	254
15.2.1. Fabric Endpoint Structure	254
15.2.2. Fabric endpoint support macros	254
15.2.3. Fabric Coordinate Structure	256
15.2.4. Fabric coordinate support macros	256
15.2.5. Fabric Geometry Structure	257
15.2.6. Fabric geometry support macros	258
15.2.7. Fabric Coordinate Views	259
15.2.8. Fabric Link State	260
15.2.9. Fabric Operation Constants	260
15.2.10. Fabric registration structure	260
15.2.10.1. Static initializer for the fabric structure	263
15.2.10.2. Initialize the fabric structure	263
15.3. Fabric Support Attributes	263
15.4. Fabric Support Functions	266
15.4.1. PMIx_Fabric_register	267
15.4.2. PMIx_Fabric_register_nb	268
15.4.3. PMIx_Fabric_update	269
15.4.4. PMIx_Fabric_update_nb	269
15.4.5. PMIx_Fabric_deregister	270
15.4.6. PMIx_Fabric_deregister_nb	271

16. Security	272
16.1. Obtaining Credentials	272
16.1.1. PMIx_Get_credential	273
16.1.2. PMIx_Get_credential_nb	274
16.1.3. Credential Attributes	275
16.2. Validating Credentials	275
16.2.1. PMIx_Validate_credential	275
16.2.2. PMIx_Validate_credential_nb	277
17. Server-Specific Interfaces	279
17.1. Server Initialization and Finalization	279
17.1.1. PMIx_server_init	279
17.1.2. PMIx_server_finalize	282
17.1.3. Server Initialization Attributes	283
17.2. Server Support Functions	284
17.2.1. PMIx_generate_regex	284
17.2.2. PMIx_generate_ppn	285
17.2.3. PMIx_server_register_namespace	285
17.2.3.1. Namespace registration attributes	295
17.2.3.2. Assembling the registration information	296
17.2.4. PMIx_server_deregister_namespace	303
17.2.5. PMIx_server_register_resources	305
17.2.6. PMIx_server_deregister_resources	305
17.2.7. PMIx_server_register_client	306
17.2.8. PMIx_server_deregister_client	308
17.2.9. PMIx_server_setup_fork	308
17.2.10. PMIx_server_dmodex_request	309
17.2.10.1. Server Direct Modex Response Callback Function	310
17.2.11. PMIx_server_setup_application	311
17.2.11.1. Server Setup Application Callback Function	313
17.2.11.2. Server Setup Application Attributes	314
17.2.12. PMIx_Register_attributes	314
17.2.12.1. Attribute registration constants	316
17.2.12.2. Attribute registration structure	316

17.2.12.3.	Attribute registration structure descriptive attributes	317
17.2.12.4.	Attribute registration structure support macros	317
17.2.13.	PMIx_server_setup_local_support	319
17.2.14.	PMIx_server_IOF_deliver	320
17.2.15.	PMIx_server_collect_inventory	321
17.2.16.	PMIx_server_deliver_inventory	322
17.2.17.	PMIx_server_generate_locality_string	323
17.2.18.	PMIx_server_generate_cpuset_string	324
17.2.18.1.	Cpuset Structure	324
17.2.18.2.	Cpuset support macros	324
17.2.19.	PMIx_server_define_process_set	326
17.2.20.	PMIx_server_delete_process_set	326
17.3.	Server Function Pointers	327
17.3.1.	pmix_server_module_t Module	327
17.3.2.	pmix_server_client_connected_fn_t	329
17.3.3.	pmix_server_client_connected2_fn_t	329
17.3.4.	pmix_server_client_finalized_fn_t	331
17.3.5.	pmix_server_abort_fn_t	332
17.3.6.	pmix_server_fencenb_fn_t	333
17.3.6.1.	Modex Callback Function	336
17.3.7.	pmix_server_dmodex_req_fn_t	336
17.3.7.1.	Dmodex attributes	338
17.3.8.	pmix_server_publish_fn_t	338
17.3.9.	pmix_server_lookup_fn_t	340
17.3.10.	pmix_server_unpublish_fn_t	342
17.3.11.	pmix_server_spawn_fn_t	344
17.3.11.1.	Server spawn attributes	349
17.3.12.	pmix_server_connect_fn_t	349
17.3.13.	pmix_server_disconnect_fn_t	350
17.3.14.	pmix_server_register_events_fn_t	352
17.3.15.	pmix_server_deregister_events_fn_t	354
17.3.16.	pmix_server_notify_event_fn_t	356

17.3.17. <code>pmix_server_listener_fn_t</code>	357
17.3.17.1. PMIx Client Connection Callback Function	358
17.3.18. <code>pmix_server_query_fn_t</code>	358
17.3.19. <code>pmix_server_tool_connection_fn_t</code>	361
17.3.19.1. Tool connection attributes	363
17.3.19.2. PMIx Tool Connection Callback Function	363
17.3.20. <code>pmix_server_log_fn_t</code>	363
17.3.21. <code>pmix_server_alloc_fn_t</code>	365
17.3.22. <code>pmix_server_job_control_fn_t</code>	368
17.3.23. <code>pmix_server_monitor_fn_t</code>	370
17.3.24. <code>pmix_server_get_cred_fn_t</code>	373
17.3.24.1. Credential callback function	374
17.3.25. <code>pmix_server_validate_cred_fn_t</code>	375
17.3.26. Credential validation callback function	377
17.3.27. <code>pmix_server_iof_fn_t</code>	378
17.3.27.1. IOF delivery function	380
17.3.28. <code>pmix_server_stdin_fn_t</code>	381
17.3.29. <code>pmix_server_grp_fn_t</code>	382
17.3.29.1. Group Operation Constants	384
17.3.30. <code>pmix_server_fabric_fn_t</code>	384
18. Tools and Debuggers	387
18.1. Connection Mechanisms	387
18.1.1. Rendezvousing with a local server	390
18.1.2. Connecting to a remote server	391
18.1.3. Attaching to running jobs	391
18.1.4. Tool initialization attributes	392
18.1.5. Tool initialization environmental variables	392
18.1.6. Tool connection attributes	392
18.2. Launching Applications with Tools	393
18.2.1. Direct launch	394
18.2.2. Indirect launch	398
18.2.2.1. Initiator-based command line parsing	398
18.2.2.2. Intermediate Launcher (IL)-based command line parsing	402

18.2.3.	Tool spawn-related attributes	403
18.2.4.	Tool rendezvous-related events	403
18.3.	IO Forwarding	404
18.3.1.	Forwarding stdout/stderr	404
18.3.2.	Forwarding stdin	406
18.3.3.	IO Forwarding Channels	408
18.3.4.	IO Forwarding constants	408
18.3.5.	IO Forwarding attributes	408
18.4.	Debugger Support	410
18.4.1.	Co-Location of Debugger Daemons	412
18.4.2.	Co-Spawn of Debugger Daemons	413
18.4.3.	Debugger Agents	414
18.4.4.	Tracking the job lifecycle	414
18.4.4.1.	Job lifecycle events	415
18.4.4.2.	Job lifecycle attributes	416
18.4.5.	Debugger-related constants	416
18.4.6.	Debugger attributes	417
18.5.	Tool-Specific APIs	418
18.5.1.	PMIx_tool_init	418
18.5.2.	PMIx_tool_finalize	421
18.5.3.	PMIx_tool_disconnect	422
18.5.4.	PMIx_tool_attach_to_server	422
18.5.5.	PMIx_tool_get_servers	424
18.5.6.	PMIx_tool_set_server	424
18.5.7.	PMIx_IOF_pull	425
18.5.8.	PMIx_IOF_deregister	427
18.5.9.	PMIx_IOF_push	428
19.	Storage Support Definitions (<i>Provisional</i>)	431
19.1.	Storage support constants (<i>Provisional</i>)	431
19.2.	Storage support attributes (<i>Provisional</i>)	433

A. Python Bindings	435
A.1. Design Considerations	435
A.1.1. Error Codes vs Python Exceptions	435
A.1.2. Representation of Structured Data	435
A.2. Datatype Definitions	436
A.2.1. Example	442
A.3. Callback Function Definitions	442
A.3.1. IOF Delivery Function	442
A.3.2. Event Handler	443
A.3.3. Server Module Functions	444
A.3.3.1. Client Connected	444
A.3.3.2. Client Finalized	444
A.3.3.3. Client Aborted	444
A.3.3.4. Fence	445
A.3.3.5. Direct Modex	445
A.3.3.6. Publish	446
A.3.3.7. Lookup	446
A.3.3.8. Unpublish	447
A.3.3.9. Spawn	447
A.3.3.10. Connect	448
A.3.3.11. Disconnect	448
A.3.3.12. Register Events	449
A.3.3.13. Deregister Events	449
A.3.3.14. Notify Event	450
A.3.3.15. Query	450
A.3.3.16. Tool Connected	451
A.3.3.17. Log	451
A.3.3.18. Allocate Resources	452
A.3.3.19. Job Control	452
A.3.3.20. Monitor	453
A.3.3.21. Get Credential	453
A.3.3.22. Validate Credential	454
A.3.3.23. IO Forward	454

A.3.3.24. IO Push	455
A.3.3.25. Group Operations	455
A.3.3.26. Fabric Operations	456
A.4. PMIxCliient	457
A.4.1. Client.init	457
A.4.2. Client.initialized	457
A.4.3. Client.get_version	457
A.4.4. Client.finalize	458
A.4.5. Client.abort	458
A.4.6. Client.store_internal	458
A.4.7. Client.put	459
A.4.8. Client.commit	459
A.4.9. Client.fence	460
A.4.10. Client.get	460
A.4.11. Client.publish	461
A.4.12. Client.lookup	461
A.4.13. Client.unpublish	462
A.4.14. Client.spawn	462
A.4.15. Client.connect	463
A.4.16. Client.disconnect	463
A.4.17. Client.resolve_peers	464
A.4.18. Client.resolve_nodes	464
A.4.19. Client.query	465
A.4.20. Client.log	465
A.4.21. Client.allocation_request	466
A.4.22. Client.job_ctrl	466
A.4.23. Client.monitor	467
A.4.24. Client.get_credential	467
A.4.25. Client.validate_credential	468
A.4.26. Client.group_construct	468
A.4.27. Client.group_invite	469
A.4.28. Client.group_join	469
A.4.29. Client.group_leave	470

A.4.30.	Client.group_destruct	470
A.4.31.	Client.register_event_handler	471
A.4.32.	Client.deregister_event_handler	471
A.4.33.	Client.notify_event	472
A.4.34.	Client.fabric_register	472
A.4.35.	Client.fabric_update	473
A.4.36.	Client.fabric_deregister	473
A.4.37.	Client.load_topology	474
A.4.38.	Client.get_relative_locality	474
A.4.39.	Client.get_cpuset	475
A.4.40.	Client.parse_cpuset_string	475
A.4.41.	Client.compute_distances	475
A.4.42.	Client.error_string	476
A.4.43.	Client.proc_state_string	476
A.4.44.	Client.scope_string	477
A.4.45.	Client.persistence_string	477
A.4.46.	Client.data_range_string	478
A.4.47.	Client.info_directives_string	478
A.4.48.	Client.data_type_string	478
A.4.49.	Client.alloc_directive_string	479
A.4.50.	Client.iof_channel_string	479
A.4.51.	Client.job_state_string	480
A.4.52.	Client.get_attribute_string	480
A.4.53.	Client.get_attribute_name	480
A.4.54.	Client.link_state_string	481
A.4.55.	Client.device_type_string	481
A.4.56.	Client.progress	482
A.5.	PMIxServer	482
A.5.1.	Server.init	482
A.5.2.	Server.finalize	482
A.5.3.	Server.generate_regex	483
A.5.4.	Server.generate_ppn	483
A.5.5.	Server.generate_locality_string	484

A.5.6.	Server.generate_cpuset_string	484
A.5.7.	Server.register_nspace	484
A.5.8.	Server.deregister_nspace	485
A.5.9.	Server.register_resources	485
A.5.10.	Server.deregister_resources	486
A.5.11.	Server.register_client	486
A.5.12.	Server.deregister_client	486
A.5.13.	Server.setup_fork	487
A.5.14.	Server.dmodex_request	487
A.5.15.	Server.setup_application	488
A.5.16.	Server.register_attributes	488
A.5.17.	Server.setup_local_support	488
A.5.18.	Server.iof_deliver	489
A.5.19.	Server.collect_inventory	489
A.5.20.	Server.deliver_inventory	490
A.5.21.	Server.define_process_set	490
A.5.22.	Server.delete_process_set	491
A.5.23.	Server.register_resources	491
A.5.24.	Server.deregister_resources	492
A.6.	PMIxTool	492
A.6.1.	Tool.init	492
A.6.2.	Tool.finalize	492
A.6.3.	Tool.disconnect	493
A.6.4.	Tool.attach_to_server	493
A.6.5.	Tool.get_servers	494
A.6.6.	Tool.set_server	494
A.6.7.	Tool.iof_pull	494
A.6.8.	Tool.iof_deregister	495
A.6.9.	Tool.iof_push	495
A.7.	Example Usage	496
A.7.1.	Python Client	496
A.7.2.	Python Server	498

B. Use-Cases	502
B.1. Business Card Exchange for Process-to-Process Wire-up	502
B.1.1. Use Case Summary	502
B.1.2. Use Case Details	502
B.2. Debugging	506
B.2.1. Terminology	506
B.2.1.1. Tools vs Debuggers	506
B.2.1.2. Parallel Launching Methods	506
B.2.1.3. Process Synchronization	506
B.2.1.4. Process Acquisition	506
B.2.2. Use Case Details	507
B.2.2.1. Direct-Launch Debugger Tool	507
B.2.2.2. Indirect-Launch Debugger Tool	512
B.2.2.3. Attaching to a Running Job	516
B.2.2.4. Tool Interaction with RM	520
B.2.2.5. Environmental Parameter Directives for Applications and Launchers	521
B.3. Hybrid Applications	522
B.3.1. Use Case Summary	522
B.3.2. Use Case Details	522
B.3.2.1. Identifying Active Parallel Runtime Systems	522
B.3.2.2. Coordinating at Runtime	524
B.3.2.3. Coordinating at Runtime with Multiple Event Handlers	526
B.4. MPI Sessions	529
B.4.1. Use Case Summary	529
B.4.2. Use Case Details	530
B.5. Cross-Version Compatibility	532
B.5.1. Use Case Summary	532
B.5.2. Use Case Details	532
C. Revision History	535
C.1. Version 1.0: June 12, 2015	535
C.2. Version 2.0: Sept. 2018	536
C.2.1. Removed/Modified Application Programming Interfaces (APIs)	536
C.2.2. Deprecated constants	536

- C.2.3. Deprecated attributes 537
- C.3. Version 2.1: Dec. 2018 537
- C.4. Version 2.2: Jan 2019 538
- C.5. Version 3.0: Dec. 2018 538
 - C.5.1. Removed constants 538
 - C.5.2. Deprecated attributes 539
 - C.5.3. Removed attributes 539
- C.6. Version 3.1: Jan. 2019 539
- C.7. Version 3.2: Oct. 2020 540
 - C.7.1. Deprecated constants 540
 - C.7.2. Deprecated attributes 541
- C.8. Version 4.0: Dec. 2020 542
 - C.8.1. Added Constants 544
 - C.8.2. Added Attributes 547
 - C.8.3. Added Environmental Variables 558
 - C.8.4. Added Macros 559
 - C.8.5. Deprecated APIs 559
 - C.8.6. Deprecated constants 559
 - C.8.7. Removed constants 559
 - C.8.8. Deprecated attributes 560
 - C.8.9. Removed attributes 561
- C.9. Version 4.1: Oct. 2021 562
 - C.9.1. Removed constants 562
 - C.9.2. Added Functions (Provisional) 562
 - C.9.3. Added Data Structures (Provisional) 562
 - C.9.4. Added Macros (Provisional) 563
 - C.9.5. Added Constants (Provisional) 563
 - C.9.6. Added Attributes (Provisional) 563
- C.10. Version 5.0: May 2023 565
 - C.10.1. Added Functions (Provisional) 565
 - C.10.2. Added Macros (Provisional) 566
 - C.10.3. Added Constants (Provisional) 566
 - C.10.4. Added Attributes 566

C.10.5. Added Attributes (Provisional)	567
C.10.6. Deprecated constants	568
C.10.7. Deprecated attributes	568
C.10.8. Deprecated macros	569
C.10.9. Removed Constants	569
D. Acknowledgements	570
D.1. Version 5.0	570
D.2. Version 4.0	574
D.3. Version 3.0	575
D.4. Version 2.0	576
D.5. Version 1.0	577
Bibliography	579
Index	580
Index of APIs	582
Index of Support Macros	589
Index of Data Structures	593
Index of Constants	595
Index of Environmental Variables	604
Index of Attributes	605

CHAPTER 1

Introduction

1 Process Management Interface - Exascale (PMIx) is an application programming interface standard
2 that provides libraries and programming models with portable and well-defined access to commonly
3 needed services in distributed and parallel computing systems. A typical example of such a service
4 is the portable and scalable exchange of network addresses to establish communication channels
5 between the processes of a parallel application or service. As such, PMIx gives distributed system
6 software providers a better understanding of how programming models and libraries can interface
7 with and use system-level services. As a standard, PMIx provides APIs that allow for portable
8 access to these varied system software services and the functionalities they offer. Although these
9 services can be defined and implemented directly by the system software components providing
10 them, the community represented by the ASC feels that the development of a shared standard better
11 serves the community. As a result, PMIx enables programming languages and libraries to focus on
12 their core competencies without having to provide their own system-level services.

1.1 Background

14 The Process Management Interface (PMI) has been used for quite some time as a means of
15 exchanging wireup information needed for inter-process communication. Two versions (PMI-1 and
16 PMI-2 [2]) have been released as part of the MPICH effort, with PMI-2 demonstrating better
17 scaling properties than its PMI-1 predecessor.

18 PMI-1 and PMI-2 can be implemented using PMIx though PMIx is not a strict superset of either.
19 Since its introduction, PMIx has expanded on earlier PMI efforts by providing an extended version
20 of the PMI APIs which provide necessary functionality for launching and managing parallel
21 applications and tools at scale.

22 The increase in adoption has motivated the creation of this document to formally specify the
23 intended behavior of the PMIx APIs.

24 More information about the PMIx standard and affiliated projects can be found at the PMIx web
25 site: <https://pmix.org>

1.2 PMIx Architecture Overview

27 The presentation of the PMIx APIs within this document makes some basic assumptions about how
28 these APIs are used and implemented. These assumptions are generally made only to simplify the
29 presentation and explain PMIx with the expectation that most readers have similar concepts on how

1 computing systems are organized today. However, ultimately this document should only be
2 assumed to define a set of APIs.

3 A concept that is fundamental to PMIx is that a PMIx implementation might operate primarily as a
4 *messenger*, and not a *doer* — i.e., a PMIx implementation might rely heavily or fully on other
5 software components to provide functionality [1]. Since a PMIx implementation might only deliver
6 requests and responses to other software components, the API calls include ways to provide
7 arbitrary information to the backend components that actually implement the functionality. Also,
8 because PMIx implementations generally rely heavily on other system software, a PMIx
9 implementation might not be able to guarantee that a feature is available on all platforms the
10 implementation supports. These aspects are discussed in detail in the remainder of this chapter.

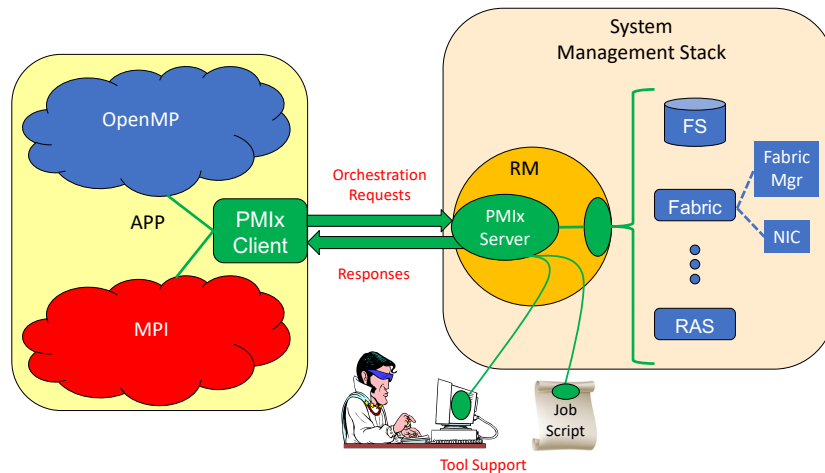


Figure 1.1.: PMIx-SMS Interactions

11 Fig. 1.1 shows a typical PMIx implementation in which the application is built against a PMIx
12 client library that contains the client-side APIs, attribute definitions, and communication support
13 for interacting with the local PMIx server. PMIx clients are processes which are started through the
14 PMIx infrastructure, either by the PMIx implementation directly or through a System Management
15 Software stack (SMS) component, and have registered as clients. A PMIx client is created in such a
16 way that the PMIx client library will have sufficient information available to authenticate with
17 the PMIx server. The PMIx server will have sufficient knowledge about the process which it
18 created, either directly or through other SMS, to authenticate the process and provide information
19 the process requests such as its identity and the identity of its peers.

20 As clients invoke PMIx APIs, it is possible that some client requests can be handled at the client
21 level. Other requests might require communication with the local PMIx server, which subsequently
22 might request services from the host SMS (represented here by a Resource Manager (RM)
23 daemon). The interaction between the PMIx server and SMS are achieved using callback functions
24 registered during server initialization. The host SMS can indicate its lack of support for any

1 operation by simply providing a *NULL* for the associated callback function, or can create a function
2 entry that returns *not supported* when called.

3 Recognizing the burden this places on SMS vendors, the PMIx community has included interfaces
4 by which the host SMS (containing the local PMIx service instance) can request support from local
5 SMS elements via the PMIx API. Once the SMS has transferred the request to an appropriate
6 location, a PMIx server interface can be used to pass the request between SMS subsystems. For
7 example, a request for network traffic statistics can utilize the PMIx networking abstractions to
8 retrieve the information from the Fabric Manager. This reduces the portability and interoperability
9 issues between the individual subsystems by transferring the burden of defining the interoperable
10 interfaces from the SMS subsystems to the PMIx community, which continues to work with those
11 providers to develop the necessary support.

12 Fig. 1.1 shows how tools can interact with the PMIx architecture. Tools, whether standalone or
13 embedded in job scripts, are an exception to the normal client registration process. A process can
14 register as a tool, provided the PMIx client library has adequate rendezvous information to connect
15 to the appropriate PMIx server (either hosted on the local machine or on a remote machine). This
16 allows processes which were not created by the PMIx infrastructure to request access to PMIx
17 functionality.

18 1.3 Portability of Functionality

19 It is difficult to define a portable API that will provide access to the many and varied features
20 underlying the operations for which PMIx provides access. For example, the options and features
21 provided to request the creation of new processes varied dramatically between different systems
22 existing at the time PMIx was introduced. Many RMs provide rich interfaces to specify the
23 resources assigned to processes. As a result, PMIx is faced with the challenge of attempting to meet
24 the seemingly conflicting goals of creating an API which allows access to these diverse features
25 while being portable across a wide range of existing software environments. In addition, the
26 functionalities required by different clients vary greatly. Producing a PMIx implementation which
27 can provide the needs of all possible clients on all of its target systems could be so burdensome as
28 to discourage PMIx implementations.

29 To help address this issue, the PMIx APIs are designed to allow resource managers and other
30 system management stack components to decide on support of a particular function and allow client
31 applications to query and adjust to the level of support available. PMIx clients should be written to
32 account for the possibility that a PMIx API might return an error code indicating that the call is not
33 supported. The PMIx community continues to look at ways to assist SMS implementers in their
34 decisions on what functionality to support by highlighting functions and attributes that are critical
35 to basic application execution (e.g., [PMIx_Get](#)) for certain classes of applications.

36 1.3.1 Attributes in PMIx

37 An area where differences between support on different systems can be challenging is regarding the
38 attributes that provide information to the client process and/or control the behavior of a PMIx API.

1 Most PMIx API calls can accept additional information or attributes specified in the form of
2 key/value pairs. These attributes provide information to the PMIx implementation that influence the
3 behavior of the API call. In addition to API calls being optional, support for the individual
4 attributes of an API call can vary between systems or implementations.

5 An application can adapt to the attribute support on a particular system in one of two ways. PMIx
6 provides an API to enable an application to query the attributes supported by a particular API (See
7 7.4). Through this API, the PMIx implementation can provide detailed information about the
8 attributes supported on a system for each API call queried. Alternatively, the application can mark
9 attributes as required using a flag within the `pmix_info_t` (See 3.2.9). If the required attribute is
10 not available on the system or the desired value for the attribute is not available, the call will return
11 the error code for *not supported*.

12 For example, the `PMIX_TIMEOUT` attribute can be used to specify the time (in seconds) before the
13 requested operation should time out. The intent of this attribute is to allow the client to avoid
14 “hanging” in a request that takes longer than the client wishes to wait, or may never return (e.g., a
15 `PMIx_Fence` that a blocked participant never enters).

16 The application can query the attribute support for `PMIx_Fence` and search whether
17 `PMIX_TIMEOUT` is listed as a supported attribute. The application can also set the required flag in
18 the `pmix_info_t` for that attribute when making the `PMIx_Fence` call. This will return an
19 error if this attribute is not supported. If the required flag is not set, the library and SMS host are
20 allowed to treat the attribute as optional, ignoring it if support is not available.

21 It is therefore critical that users and application implementers:

- 22 a) consider whether or not a given attribute is required, marking it accordingly; and
- 23 b) check the return status on all PMIx function calls to ensure support was present and that the
24 request was accepted. Note that for non-blocking APIs, a return of `PMIX_SUCCESS` only
25 indicates that the request had no obvious errors and is being processed – the eventual callback
26 will return the status of the requested operation itself.

27 PMIx clients (e.g., tools, parallel programming libraries) may find that they depend only on a small
28 subset of interfaces and attributes to work correctly. PMIx clients are strongly advised to define a
29 document itemizing the PMIx interfaces and associated attributes that are required for correct
30 operation, and are optional but recommended for full functionality. The PMIx standard cannot
31 define this list for all given PMIx clients, but such a list is valuable to RMs desiring to support these
32 clients.

33 A PMIx implementation may be able to support only a subset of the PMIx API and attributes on a
34 particular system due to either its own limitations or limitations of the SMS with which it
35 interfaces. A PMIx implementation may also provide additional attributes beyond those defined
36 herein in order to allow applications to access the full features of the underlying SMS. PMIx
37 implementations are strongly advised to document the PMIx interfaces and associated attributes
38 they support, with any annotations about behavior limitations. The PMIx standard cannot define
39 this support for implementations, but such documentation is valuable to PMIx clients desiring to
40 support a broad range of systems.

1 While a PMIx library implementer, or an SMS component server, may choose to support a
2 particular PMIx API, they are not required to support every attribute that might apply to it. This
3 would pose a significant barrier to entry for an implementer as there can be a broad range of
4 applicable attributes to a given API, at least some of which may rarely be used.

5 Note that an environment that does not include support for a particular attribute/API pair is not
6 “incomplete” or of lower quality than one that does include that support. Vendors must decide
7 where to invest their time based on the needs of their target markets, and it is perfectly reasonable
8 for them to perform cost/benefit decisions when considering what functions and attributes to
9 support.

10 Attributes in this document are organized according to their primary usage, either grouped with a
11 specific API or included in an appropriate functional chapter. Attributes in the PMIx Standard all
12 start with "**PMIX**" in their name, and many include a functional description as part of their name
13 (e.g., the use of "**PMIX_FABRIC_**" at the beginning of fabric-specific attributes). The PMIx
14 Standard also defines an attribute that can be used to indicate that an attribute variable has not yet
15 been set:

16 **PMIX_ATTR_UNDEF** "`pmix.undef`" (NULL)

17 A default attribute name signifying that the attribute field of a PMIx structure (e.g., a
18 `pmix_info_t`) has not yet been defined.

19 1.3.2 PMIx Roles

20 The role of a PMIx process in the PMIx universe is grouped into one of three categories based on
21 how it operates in the PMIx environment namely as a *client*, *server*, or *tool*. As a result, there are
22 three corresponding groupings of APIs each with their own initialization and finalization functions.
23 If a process initializes as either a *server* or a *tool* that process may also access all of the *client* APIs.

24 A process operating as a *client* is connected to the PMIx server instance within an RM when the
25 client calls the client PMIx initialization routine. The *client* is typically started directly or indirectly
26 (for example, by an intermediate script) by that RM. Additionally, a *client* may be started directly
27 by the user and then connect to an RM which is typically referred to as a *singleton* launch. A
28 process operating as a *server* is responsible for starting client processes and coordinating with other
29 server and tool processes in the same PMIx universe. Often processes operating as a *server* are part
30 of the Resource Manager (RM) infrastructure. A process operating as a *tool* is started
31 independently (e.g., via `fork/exec`) or by the RM and will connect to a PMIx *server* to interact with
32 the processes in the PMIx universe. An example of a *tool* process is a parallel debugger that will
33 connect to the server to assist with attaching to a set of client processes.

34 PMIx serves as a conduit between processes acting in these three different roles. As such, an API is
35 often described by how it interacts with processes operating in other roles in the PMIx universe.

Advice to PMIx library implementers

1 A PMIx implementation may support all or a subset of the API role groupings defined in the
2 standard. A common nomenclature is defined here to aid in identifying levels of conformance of an
3 implementation.

4 Note that it would not make sense for an implementation to exclude the *client* interfaces from their
5 implementation since they are also used by the *server* and *tool* roles. Therefore the *client* interfaces
6 represent the minimal set of required functionality for PMIx compliance.

7 A PMIx implementation that supports only the *client* APIs is said to be *client-role PMIx standard*
8 *compliant*. Similarly, a PMIx implementation that only supports the *client* and *tool* APIs is said to
9 be *client-role and tool-role PMIx standard compliant*. Finally, a PMIx implementation that only
10 supports the *client* and *server* APIs is said to be *client-role and server-role PMIx standard*
11 *compliant*.

12 A PMIx implementation that supports all three sets of the API role groupings is said to be
13 *client-role, server-role, and tool-role PMIx standard compliant*. These *client-role, server-role, and*
14 *tool-role PMIx standard compliant* implementations have the advantage of being able to support a
15 broad set of PMIx consumers in the different roles.

1.3.3 Application Binary Interface (ABI)

17 An Application Programming Interface (API) defines how data types, data structures, and functions
18 are represented in source code. An Application Binary Interface (ABI) defines how data types, data
19 structures, and functions are represented in machine code for a given system making it
20 platform-specific. An important aspect of an ABI is the size, layout, and alignment of data
21 structures. A stable ABI may allow a program compiled with one implementation of the PMIx
22 Standard to run with a different implementation of the PMIx Standard as long as both
23 implementations adhere to the same ABI.

24 The PMIx Standard strives to maintain a stable ABI to support applications and tools that rely on
25 more than one implementation of the PMIx Standard. To facilitate such interoperability the PMIx
26 ASC maintains a set of standardized headers that are versioned with the PMIx Standard that
27 applications and tools can reference ¹.

28 In recognition that there are circumstances where the ABI needs to be modified this section defines
29 some guidance for making such modifications. Additions to the PMIx interface can occur without
30 breaking ABI compatibility. Deprecating portions of the PMIx interface does not break ABI
31 compatibility but serves as a warning that the ABI may be impacted in the future. Removing
32 portions of the PMIx interface does break ABI compatibility. Modifications to the existing PMIx
33 interface do break ABI compatibility.

34 The PMIx ABI is comprised of the following:

¹PMIx Headers for ABI Compatibility <https://github.com/pmix/pmix-abi>

- Function signatures
- Constants and their values
- Structures and their membership (including field position)
- Attributes that are required to be implemented and their string representation

The ASC prohibits data structure and function signature changes that break the ABI (e.g., modifications to data structure member alignment and/or size). It is preferred that the authors create a new data structure or function and duplicate all impacted PMIx interfaces.

Rationale

Breaking the ABI of data structures and function signatures while using the same name prevents a PMIx implementation from supporting backward compatible and cross-version compatible PMIx libraries.

The value associated with attributes and constants are part of an API and not part of a *Linker ABI*. The PMIx Standard defines a *Build ABI* where the value associated with PMIx attributes and constants can be assumed based on the ABI version.

The PMIx Standard ABI is defined by a set of ASC maintained headers. The instantiation of the PMIx Standard ABI in binary form is platform-specific. The PMIx Standard and PMIx implementations must take care with data structures (e.g., `pmix_value_t`) to ensure that the offset and size of members within the structure along with the size of the structure itself remain stable for a given platform.

1.3.3.1 PMIx ABI Versioning

The PMIx Standard ABI is defined in two parts. The *PMIx Standard Stable ABI* represents the Stable PMIx Standard elements (see the PMIx Governance document). The *PMIx Standard Provisional ABI* represents the Provisional PMIx Standard elements (see the PMIx Governance document).

Both the Stable ABI and Provisional ABI are versioned with two increasing numbers:

- MAJOR incremented when the ABI changed in a backward-incompatible manner.
- MINOR incremented when functionality is added to the ABI in a backward-compatible manner.

The PMIx Standard ABI version numbers largely follow the Semantic Versioning 2.0.0 specification². However, a PMIx implementation may use a different version numbering technique for the objects associated with that implementation.

²Semantic Versioning <https://semver.org/spec/v2.0.0.html>

1 1.3.3.2 Linker versus Build ABI

2 The software development community often discusses ABI in two different ways, that of a *Linker*
3 *ABI* and of a *Build ABI*.

4 A *Linker ABI* defines a stable set of symbols (i.e., functions) against which a compiler will attempt
5 to link the binary. The Linker ABI does not specify constant values or macro definitions leaving
6 those to individual implementations to define. If Library A version **1** and Library A version **2** are
7 Linker ABI compatible then a program that is compiled against Library A version **1** can link
8 against Library A version **2**. Note that the reverse is not necessarily true as Library A version **2**
9 may define additional symbols not included in Library A version **1**. An indication of this
10 compatibility is useful when upgrading a library package on a system. Often libraries rely on
11 Semantic Versioning³ to signify breaks in the Linker ABI between versions of the library.

12 A *Build ABI* defines the full set of symbols, constants, and macros used by a compiler to generate
13 the resulting binary. If Library A and Library B are Build ABI compatible then a program
14 compiled against Library A will work when linked with Library B.

15 The PMIx Standard defines a Build ABI in the PMIx Standard ABI repository⁴. Any program built
16 against the headers defined in the PMIx Standard ABI version **X.Y** will work with any PMIx
17 implementation that is ABI compatible with version **X.Y**. Note that the PMIx implementation may
18 include additional items beyond the PMIX Standard ABI at version **X.Y** and still report being
19 PMIx Standard version **X.Y** compliant.

³Semantic Versioning <https://semver.org/>

⁴PMIx Headers for ABI Compatibility <https://github.com/pmix/pmix-abi>

CHAPTER 2

PMIx Terms and Conventions

1 In this chapter we describe some common terms and conventions used throughout this document.
2 The PMIx Standard has adopted the widespread use of key-value *attributes* to add flexibility to the
3 functionality expressed in the APIs. Accordingly, the ASC has chosen to require that the definition
4 of each standard API include the passing of an array of attributes. These provide a means of
5 customizing the behavior of the API as future needs emerge without having to alter or create new
6 variants of it. In addition, attributes provide a mechanism by which researchers can easily explore
7 new approaches to a given operation without having to modify the API itself.

8 In an effort to maintain long-term backward compatibility, PMIx does not include large numbers of
9 APIs that each focus on a narrow scope of functionality, but instead relies on the definition of fewer
10 generic APIs that include arrays of key-value attributes for “tuning” the function’s behavior. Thus,
11 modifications to the PMIx standard primarily consist of the definition of new attributes along with a
12 description of the APIs to which they relate and the expected behavior when used with those APIs.

13 The following terminology is used throughout this document:

- 14 • *session* refers to a set of resources assigned by the WorkLoad Manager (WLM) that has been
15 reserved for one or more users. A session is identified by a *session ID* that is unique within the
16 scope of the governing WLMs. Historically, High Performance Computing (HPC) sessions have
17 consisted of a static allocation of resources - i.e., a block of resources assigned to a user in
18 response to a specific request and managed as a unified collection. However, this is changing in
19 response to the growing use of dynamic programming models that require on-the-fly allocation
20 and release of system resources. Accordingly, the term *session* in this document refers to a
21 potentially dynamic entity, perhaps comprised of resources accumulated as a result of multiple
22 allocation requests that are managed as a single unit by the WLM.
- 23 • *job* refers to a set of one or more *applications* executed as a single invocation by the user within a
24 session with a unique identifier, the *job ID*, assigned by the RM or launcher. For example, the
25 command line “*mpiexec -n 1 app1 : -n 2 app2*” generates a single Multiple Program Multiple
26 Data (MPMD) job containing two applications. A user may execute multiple *jobs* within a given
27 session, either sequentially or concurrently.
- 28 • *namespace* refers to a character string value assigned by the RM to a *job*. All *applications*
29 executed as part of that *job* share the same *namespace*. The *namespace* assigned to each *job* must
30 be unique within the scope of the governing RM and often is implemented as a string
31 representation of the numerical *Job ID*. The *namespace* and *job* terms will be used
32 interchangeably throughout the document.

- 1 • *application* represents a set of identical, but not necessarily unique, execution contexts within a
2 *job*.
- 3 • *process* is assumed for ease of presentation to be an operating system process, also commonly
4 referred to as a *heavyweight* process. A process is often comprised of multiple *lightweight*
5 *threads*, commonly known as simply *threads*. However, it is not the intent of the PMIx Standard
6 to restrict the term process to a particular concept or implementation.
- 7 • *client* refers to a process that was registered with the PMIx server prior to being started, and
8 connects to that PMIx server via `PMIx_Init` using its assigned namespace and rank with the
9 information required to connect to that server being provided to the process at time of start of
10 execution.
- 11 • *tool* refers to a process that may or may not have been registered with the PMIx server prior to
12 being started and initializes using `PMIx_tool_init`.
- 13 • *clone* refers to a process that was directly started by a PMIx client (e.g., using *fork/exec*) and calls
14 `PMIx_Init`, thus connecting to its local PMIx server using the same namespace and rank as its
15 parent process.
- 16 • *rank* refers to the numerical location (starting from zero) of a process within the defined scope.
17 Thus, *job rank* is the rank of a process within its *job* and is synonymous with its unqualified
18 *rank*, while *application rank* is the rank of that process within its *application*.
- 19 • *peer* refers to another process within the same *job*.
- 20 • *workflow* refers to an orchestrated execution plan typically involving multiple *jobs* carried out
21 under the control of a *workflow manager*. An example workflow might first execute a
22 computational job to generate the flow of liquid through a complex cavity, followed by a
23 visualization job that takes the output of the first job as its input to produce an image output.
- 24 • *scheduler* refers to the component of the SMS responsible for scheduling of resource allocations.
25 This is also generally referred to as the *system workflow manager* - for the purposes of this
26 document, the *WLM* acronym will be used interchangeably to refer to the scheduler.
- 27 • *resource manager* is used in a generic sense to represent the subsystem that will host the PMIx
28 server library. This could be a vendor-supplied resource manager or a third-party agent such as a
29 programming model's runtime library.
- 30 • *host environment* is used interchangeably with *resource manager* to refer to the process hosting
31 the PMIx server library.
- 32 • *node* refers to a single operating system instance. Note that this may encompass one or more
33 physical objects.
- 34 • *package* refers to a single object that is either soldered or connected to a printed circuit board via
35 a mechanical socket. Packages may contain multiple chips that include (but are not limited to)
36 processing units, memory, and peripheral interfaces.

- 1 • *processing unit*, or *PU*, is the electronic circuitry within a computer that executes instructions.
2 Depending upon architecture and configuration settings, it may consist of a single hardware
3 thread or multiple hardware threads collectively organized as a *core*.
 - 4 • *fabric* is used in a generic sense to refer to the networks within the system regardless of speed or
5 protocol. Any use of the term *network* in the document should be considered interchangeable
6 with *fabric*.
 - 7 • *fabric device* (or *fabric devices*) refers to an operating system fabric interface, which may be
8 physical or virtual. Any use of the term Network Interface Card (NIC) in the document should be
9 considered interchangeable with *fabric device*.
 - 10 • *fabric plane* refers to a collection of fabric devices in a common logical or physical
11 configuration. Fabric planes are often implemented in HPC clusters as separate overlay or
12 physical networks controlled by a dedicated fabric manager.
 - 13 • *attribute* refers to a key-value pair comprised of a string key (represented by a `pmix_key_t`
14 structure) and an associated value containing a PMIx data type (e.g., boolean, integer, or a more
15 complex PMIx structure). Attributes are used both as directives when passed as qualifiers to
16 APIs (e.g., in a `pmix_info_t` array), and to identify the contents of information (e.g., to
17 specify that the contents of the corresponding `pmix_value_t` in a `pmix_info_t` represent
18 the `PMIX_UNIV_SIZE`).
 - 19 • *key* refers to the string component of a defined *attribute*. The PMIx Standard will often refer to
20 passing of a *key* to an API (e.g., to the `PMIx_Query_info` or `PMIx_Get` APIs) as a means of
21 identifying requested information. In this context, the *data type* specified in the *attribute's*
22 definition indicates the data type the caller should expect to receive in return. Note that not all
23 *attributes* can be used as *keys* as some have specific uses solely as API qualifiers.
 - 24 • *instant on* refers to a PMIx concept defined as: "All information required for setup and
25 communication (including the address vector of endpoints for every process) is available to each
26 process at start of execution"
- 27 The following sections provide an overview of the conventions used throughout the PMIx Standard
28 document.

29 2.1 Notational Conventions

30 Some sections of this document describe programming language specific examples or APIs. Text
31 that applies only to programs for which the base language is C is shown as follows:



1 C specific text...

2 `int foo = 42;`



3 Some text is for information only, and is not part of the normative specification. These take several
4 forms, described in their examples below:



5 Note: General text...



6 Throughout this document, the rationale for the design choices made in the interface specification is
7 set off in this section. Some readers may wish to skip these sections, while readers interested in
8 interface design may want to read them carefully.



9 Throughout this document, material aimed at users and that illustrates usage is set off in this
10 section. Some readers may wish to skip these sections, while readers interested in programming
11 with the PMIx API may want to read them carefully.



12 Throughout this document, material that is primarily commentary to PMIx library implementers is
13 set off in this section. Some readers may wish to skip these sections, while readers interested in
14 PMIx implementations may want to read them carefully.



15 Throughout this document, material that is primarily commentary aimed at host environments (e.g.,
16 RMs and RunTime Environments (RTEs)) providing support for the PMIx server library is set off in
17 this section. Some readers may wish to skip these sections, while readers interested in integrating
18 PMIx servers into their environment may want to read them carefully.



19 Attributes added in this version of the standard are shown in *magenta* to distinguish them from
20 those defined in prior versions, which are shown in *black*. Deprecated attributes are shown in *green*
21 and may be removed in a future version of the standard.

1 2.2 Semantics

2 The following terms will be taken to mean:

- 3 • *shall*, *must* and *will* indicate that the specified behavior is *required* of all conforming
4 implementations
- 5 • *should* and *may* indicate behaviors that a complete implementation would include, but are not
6 required of all conforming implementations

7 2.3 Naming Conventions

8 The PMIx standard has adopted the following conventions:

- 9 • PMIx constants and attributes are prefixed with "**PMIX_**".
- 10 • Structures and type definitions are prefixed with "**pmix_**".
- 11 • The string representation of attributes are prefixed with "**pmix**".
- 12 • Underscores are used to separate words in a function or variable name.
- 13 • Lowercase letters are used in PMIx client APIs except for the PMIx prefix (noted below) and the
14 first letter of the word following it. For example, **PMIx_Get_version**.
- 15 • PMIx server and tool APIs are all lower case letters following the prefix - e.g.,
16 **PMIx_server_register_nspace**.
- 17 • The **PMIX_** prefix is used to denote functions.
- 18 • The **pmix_** prefix is used to denote function pointer and type definitions.

19 Users shall not use the "**PMIX_**", "**PMIx_**", or "**pmix_**" prefixes for symbols in their code so as
20 to avoid symbol conflicts with PMIx implementations.

21 2.4 Procedure Conventions

22 While the current APIs are based on the C programming language, it is not the intent of the PMIx
23 Standard to preclude the use of other languages. Accordingly, the procedure specifications in the
24 PMIx Standard are written in a language-independent syntax with the arguments marked as IN,
25 OUT, or INOUT. The meanings of these are:

- 26 • IN: The call may use the input value but does not update the argument from the perspective of
27 the caller at any time during the calls execution,
- 28 • OUT: The call may update the argument but does not use its input value
- 29 • INOUT: The call may both use and update the argument.

1 Many PMIx interfaces, particularly nonblocking interfaces, use a **(void*)** callback data object
2 passed to the function that is then passed to the associated callback. On the client side, the callback
3 data object is an opaque, client-provided context that the client can pass to a non-blocking call.
4 When the nonblocking call completes, the callback data object is passed back to the client without
5 modification by the PMIx library, thus allowing the client to associate a context with that callback.
6 This is useful if there are many outstanding nonblocking calls.

7 A similar model is used for the server module functions (see [17.3.1](#)). In this case, the PMIx library
8 is making an upcall into its host via the PMIx server module callback function and passing a
9 specific callback function pointer and callback data object. The PMIx library expects the host to
10 call the cbfunc with the necessary arguments and pass back the original callback data object upon
11 completing the operation. This gives the server-side PMIx library the ability to associate a context
12 with the callback (since multiple operations may be outstanding). The host has no visibility into the
13 contents of the callback data object, nor is permitted to alter it in any way.

CHAPTER 3

Data Structures and Types

1 This chapter defines PMIx standard data structures (along with macros for convenient use), types,
2 and constants. These apply to all consumers of the PMIx interface. Where necessary for
3 clarification, the description of, for example, an attribute may be copied from this chapter into a
4 section where it is used.

5 A PMIx implementation may define additional attributes beyond those specified in this document.

▼ Advice to PMIx library implementers ▼

6 Structures, types, and macros in the PMIx Standard are defined in terms of the C-programming
7 language. Implementers wishing to support other languages should provide the equivalent
8 definitions in a language-appropriate manner.

9 If a PMIx implementation chooses to define additional attributes they should avoid using the
10 "**PMIX**" prefix in their name or starting the attribute string with a "**pmix**" prefix. This helps the
11 end user distinguish between what is defined by the PMIx standard and what is specific to that
12 PMIx implementation, and avoids potential conflicts with attributes defined by the Standard.

▼ Advice to users ▼

13 Use of increment/decrement operations on indices inside PMIx macros is discouraged due to
14 unpredictable behavior as the index may be cited more than once in the macro. The PMIx standard
15 only governs the existence and syntax of macros - it does not specify their implementation.

16 Users are also advised to use the macros and APIs for creating, loading, and releasing PMIx
17 structures to avoid potential issues with release of memory. For example, pointing a
18 `pmix_envar_t` element at a static string variable and then using `PMIX_ENVAR_DESTRUCT` to
19 clear it would generate an error as the static string had not been allocated.

1 3.1 Constants

2 PMIx defines a few values that are used throughout the standard to set the size of fixed arrays or as
3 a means of identifying values with special meaning. The community makes every attempt to
4 minimize the number of such definitions. The constants defined in this section may be used before
5 calling any PMIx library initialization routine. Additional constants associated with specific data
6 structures or types are defined in the section describing that data structure or type.

7 **PMIX_MAX_NSLEN 255** Maximum namespace string length as an integer.

▼ **Advice to PMIx library implementers** ▼

8 Namespace arrays in PMIx defined structures must reserve a space of size **PMIX_MAX_NSLEN+1**
9 to allow room for the **NULL** terminator

▲

10 **PMIX_MAX_KEYLEN 511** Maximum key string length as an integer.

▼ **Advice to PMIx library implementers** ▼

11 Key arrays in PMIx defined structures must reserve a space of size **PMIX_MAX_KEYLEN+1** to
12 allow room for the **NULL** terminator

▲

13 **PMIX_APP_WILDCARD UINT32_MAX** A value to indicate that the user wants the data for
14 the given key from every application that posted that key, or that the given value applies to all
15 applications within the given namespace.

1 3.1.1 PMIx Return Status Constants

2 The `pmix_status_t` type is an `int` compatible value for return status values. PMIx return
3 values other than `PMIX_SUCCESS` are required to always be negative. The return status value for a
4 successful operation is `PMIX_SUCCESS`, which must have an integer value of 0:

5 `PMIX_SUCCESS 0` Success.

▼ Advice to PMIx library implementers ▼

6 A PMIx implementation must define all of the return status constants defined in the PMIx standard,
7 even if the implementation will never return the specific value to the caller.

▲

▼ Advice to users ▼

8 Other than `PMIX_SUCCESS` (which is required to be zero), the integer value of any PMIx error
9 constant are negative and of greater magnitude (i.e. of larger absolute value)
10 than `PMIX_EXTERNAL_ERR_BASE`. For compatibility with future versions, users are advised to
11 always refer to constants by name, and not by their integer value which may be specific to an ABI
12 version.

▲

13 The presentation of each API in this document includes a list of return status constants which are
14 either specific to that API or are expected to be returned by the API in normal use.

15 In addition, the following are general constants covering a variety of possible reasons an
16 implementation of an API may return a constant other than one of the constants presented with the
17 API. Although implementations can define and return additional error constants, implementations
18 are encouraged to return one of the return constants listed with the API or in the list presented here
19 to encourage portability across implementations.

20 `PMIX_ERROR -1` General Error.

21 `PMIX_ERR_EXISTS -11` The requested operation would overwrite an existing value -
22 typically returned when an operation would overwrite an existing file or directory.

23 `PMIX_ERR_EXISTS_OUTSIDE_SCOPE -62` The requested key exists, but was posted in a
24 *scope* (see Section 5.2.1.1) that does not include the requester

25 `PMIX_ERR_INVALID_CRED -12` Invalid security credentials.

26 `PMIX_ERR_WOULD_BLOCK -15` Operation would block.

27 `PMIX_ERR_UNKNOWN_DATA_TYPE -16` The data type specified in an input to the PMIx
28 library is not recognized by the implementation.

29 `PMIX_ERR_TYPE_MISMATCH -18` The data type found in an object does not match the
30 expected data type as specified in the API call - e.g., a request to unpack a `PMIX_BOOL` value
31 from a buffer that does not contain a value of that type in the current unpack location.

32 `PMIX_ERR_UNPACK_INADEQUATE_SPACE -19` Inadequate space to unpack data - the
33 number of values in the buffer exceeds the specified number to unpack.

1 **PMIX_ERR_UNPACK_READ_PAST_END_OF_BUFFER -50** Unpacking past the end of the
2 provided buffer - the number of values in the buffer is less than the specified number to
3 unpack, or a request was made to unpack a buffer beyond the buffer's end.
4 **PMIX_ERR_UNPACK_FAILURE -20** The unpack operation failed for an unspecified reason.
5 **PMIX_ERR_PACK_FAILURE -21** The pack operation failed for an unspecified reason.
6 **PMIX_ERR_NO_PERMISSIONS -23** The user lacks permissions to execute the specified
7 operation.
8 **PMIX_ERR_TIMEOUT -24** Either a user-specified or system-internal timeout expired.
9 **PMIX_ERR_UNREACH -25** The specified target server or client process is not reachable -
10 i.e., a suitable connection either has not been or can not be made.
11 **PMIX_ERR_BAD_PARAM -27** One or more incorrect parameters (e.g., passing an attribute
12 with a value of the wrong type), or multiple parameters containing conflicting directives (e.g.,
13 multiple instances of the same attribute with different values, or different attributes specifying
14 conflicting behaviors), were passed to a PMIx API.
15 **PMIX_ERR_EMPTY -60** An array or list was given that has no members in it - i.e., the object
16 is empty.
17 **PMIX_ERR_RESOURCE_BUSY -28** Resource busy - typically seen when an attempt to
18 establish a connection to another process (e.g., a PMIx server) cannot be made due to a
19 communication failure.
20 **PMIX_ERR_OUT_OF_RESOURCE -29** Resource exhausted.
21 **PMIX_ERR_INIT -31** The requested operation requires that the PMIx library be initialized
22 prior to being called.
23 **PMIX_ERR_NOMEM -32** Out of memory.
24 **PMIX_ERR_NOT_FOUND -46** The requested information was not found.
25 **PMIX_ERR_NOT_SUPPORTED -47** The requested operation is not supported by either the
26 PMIx implementation or the host environment.
27 **PMIX_ERR_PARAM_VALUE_NOT_SUPPORTED -59** The requested operation is supported
28 by the PMIx implementation and (if applicable) the host environment. However, at least one
29 supplied parameter was given an unsupported value, and the operation cannot therefore be
30 executed as requested.
31 **PMIX_ERR_COMM_FAILURE -49** Communication failure - a message failed to be sent or
32 received, but the connection remains intact.
33 **PMIX_ERR_LOST_CONNECTION -61** Lost connection between server and client or tool.
34 **PMIX_ERR_INVALID_OPERATION -158** The requested operation is supported by the
35 implementation and host environment, but fails to meet a requirement (e.g., requesting to
36 *disconnect* from processes without first *connecting* to them, inclusion of conflicting
37 directives, or a request to perform an operation that conflicts with an ongoing one).
38 **PMIX_OPERATION_IN_PROGRESS -156** A requested operation is already in progress -
39 the duplicate request shall therefore be ignored.
40 **PMIX_OPERATION_SUCCEEDED -157** The requested operation was performed atomically
41 - no callback function will be executed.

1 **PMIX_ERR_PARTIAL_SUCCESS -52** The operation is considered successful but not all
2 elements of the operation were concluded (e.g., some members of a group construct operation
3 chose not to participate).

4 3.1.1.1 User-Defined Error and Event Constants

5 PMIx establishes a boundary for constants defined in the PMIx standard. Negative values larger
6 (i.e., more negative) than this (and any positive values greater than zero) are guaranteed not to
7 conflict with PMIx values.

8 **PMIX_EXTERNAL_ERR_BASE -3000** A starting point for user-level defined error and
9 event constants. Negative values that are more negative than the defined constant are
10 guaranteed not to conflict with PMIx values. Definitions should always be based on the
11 **PMIX_EXTERNAL_ERR_BASE** constant and not a specific value as the value of the constant
12 may change.

13 3.2 Data Types

14 This section defines various data types used by the PMIx APIs. The version of the standard in
15 which a particular data type was introduced is shown in the margin.

16 3.2.1 Key Structure

17 The **pmix_key_t** structure is a statically defined character array of length
18 **PMIX_MAX_KEYLEN+1**, thus supporting keys of maximum length **PMIX_MAX_KEYLEN** while
19 preserving space for a mandatory **NULL** terminator.

PMIx v2.0

```
▼ ─────────────────────────────────── C ─────────────────────────────────── ▼
```

```
20 typedef char pmix_key_t [PMIX_MAX_KEYLEN+1];
```

```
▲ ─────────────────────────────────── C ─────────────────────────────────── ▲
```

21 Characters in the key must be standard alphanumeric values supported by common utilities such as
22 *strcmp*.

▼ ─────────────────────────────────── Advice to users ─────────────────────────────────── ▼

23 References to keys in PMIx v1 were defined simply as an array of characters of size
24 **PMIX_MAX_KEYLEN+1**. The **pmix_key_t** type definition was introduced in version 2 of the
25 standard. The two definitions are code-compatible and thus do not represent a break in backward
26 compatibility.

27 Passing a **pmix_key_t** value to the standard *sizeof* utility can result in compiler warnings of
28 incorrect returned value. Users are advised to avoid using *sizeof(pmix_key_t)* and instead rely on
29 the **PMIX_MAX_KEYLEN** constant.

▲ ───────────────────────────────────

1 3.2.1.1 Key support macros

2 The following macros are provided for convenience when working with PMIx keys.

3 Check key macro

4 Compare the key in a `pmix_info_t` to a given value.

PMIx v3.0

▼ C ▼

5 **PMIX_CHECK_KEY**(a, b)

▲ C ▲

6 **IN** a

7 Pointer to the structure whose key is to be checked (pointer to `pmix_info_t`)

8 **IN** b

9 String value to be compared against (**char***)

10 Returns **true** if the key matches the given value

11 Check reserved key macro

12 Check if the given key is a PMIx *reserved* key as described in Chapter 6.

PMIx v4.0

▼ C ▼

13 **PMIX_CHECK_RESERVED_KEY**(a)

▲ C ▲

14 **IN** a

15 String value to be checked (**char***)

16 Returns **true** if the key is reserved by the Standard.

17 Load key macro

18 Load a key into a `pmix_info_t`.

PMIx v4.0

▼ C ▼

19 **PMIX_LOAD_KEY**(a, b)

▲ C ▲

20 **IN** a

21 Pointer to the structure whose key is to be loaded (pointer to `pmix_info_t`)

22 **IN** b

23 String value to be loaded (**char***)

24 No return value.

1 3.2.2 Namespace Structure

2 The `pmix_nspace_t` structure is a statically defined character array of length
3 `PMIX_MAX_NSLEN+1`, thus supporting namespaces of maximum length `PMIX_MAX_NSLEN`
4 while preserving space for a mandatory `NULL` terminator.

5

```
typedef char pmix_nspace_t[PMIX_MAX_NSLEN+1];
```

6 Characters in the namespace must be standard alphanumeric values supported by common utilities
7 such as `strcmp`.

Advice to users

8 References to namespace values in PMIx v1 were defined simply as an array of characters of size
9 `PMIX_MAX_NSLEN+1`. The `pmix_nspace_t` type definition was introduced in version 2 of the
10 standard. The two definitions are code-compatible and thus do not represent a break in backward
11 compatibility.

12 Passing a `pmix_nspace_t` value to the standard `sizeof` utility can result in compiler warnings of
13 incorrect returned value. Users are advised to avoid using `sizeof(pmix_nspace_t)` and instead rely
14 on the `PMIX_MAX_NSLEN` constant.

15 3.2.2.1 Namespace support macros

16 The following macros are provided for convenience when working with PMIx namespace
17 structures.

18 Check namespace macro

19 Compare the string in a `pmix_nspace_t` to a given value.

PMIx v3.0

20

```
PMIX_CHECK_NAMESPACE(a, b)
```

21 **IN** a

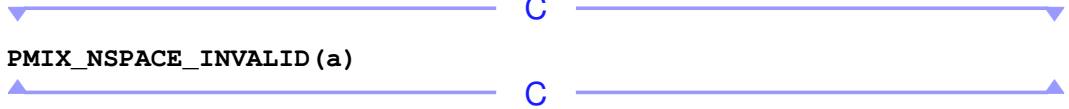

22 Pointer to the structure whose value is to be checked (pointer to `pmix_nspace_t`)

23 **IN** b

24 String value to be compared against (`char*`)

25 Returns `true` if the namespace matches the given value

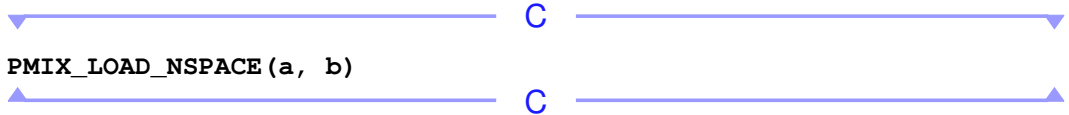

1 **Check invalid namespace macro**
2 Check if the provided `pmix_namespace_t` is invalid.

3 **PMIX_NAMESPACE_INVALID (a)**
4  C 

4 **IN a**
5 Pointer to the structure whose value is to be checked (pointer to `pmix_namespace_t`)
6 Returns `true` if the namespace is invalid (i.e., starts with a `NULL` resulting in a zero-length string
7 value)

8 **Load namespace macro**
9 Load a namespace into a `pmix_namespace_t`.

PMIx v4.0

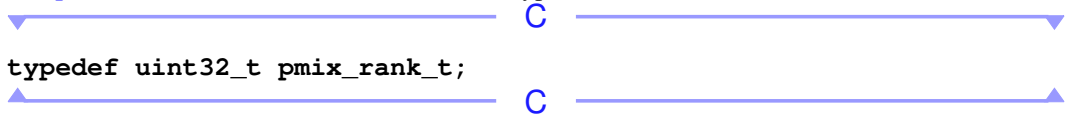

10 **PMIX_LOAD_NAMESPACE (a, b)**
11  C 

11 **IN a**
12 Pointer to the target structure (pointer to `pmix_namespace_t`)
13 **IN b**
14 String value to be loaded - if `NULL` is given, then the target structure will be initialized to
15 zero's (`char*`)
16 No return value.

17 3.2.3 Rank Structure

18 The `pmix_rank_t` structure is a `uint32_t` type for rank values.

PMIx v1.0

19 **typedef uint32_t pmix_rank_t;**
20  C 

21 The following constants can be used to set a variable of the type `pmix_rank_t`. All definitions
22 were introduced in version 1 of the standard unless otherwise marked. Valid rank values start at
23 zero.

24 **PMIX_RANK_UNDEF UINT32_MAX** A value to request job-level data where the information
25 itself is not associated with any specific rank, or when passing a `pmix_proc_t` identifier to
26 an operation that only references the namespace field of that structure.

27 **PMIX_RANK_WILDCARD UINT32_MAX-1** A value to indicate that the user wants the data
28 for the given key from every rank that posted that key.

29 **PMIX_RANK_LOCAL_NODE UINT32_MAX-2** Special rank value used to define groups of
ranks. This constant defines the group of all ranks on a local node.

1 **PMIX_RANK_LOCAL_PEERS** **UINT32_MAX-4** Special rank value used to define groups of
 2 ranks. This constant defines the group of all ranks on a local node within the same namespace
 3 as the current process.
 4 **PMIX_RANK_INVALID** **UINT32_MAX-3** An invalid rank value.
 5 **PMIX_RANK_VALID** **UINT32_MAX-50** Define an upper boundary for valid rank values.

6 3.2.3.1 Rank support macros

7 The following macros are provided for convenience when working with PMIx ranks.

8 **Check rank macro**

9 Check two ranks for equality, taking into account wildcard values

PMIx v4.0  C 

10 **PMIX_CHECK_RANK**(a, b)

 C 

11 **IN** a

12 Rank to be checked (**pmix_rank_t**)

13 **IN** b

14 Rank to be checked (**pmix_rank_t**)

15 Returns **true** if the ranks are equal, or at least one of the ranks is **PMIX_RANK_WILDCARD**

16 **Check rank is valid macro**

17 Check if the given rank is a valid value

Provisional  C 

v4.1

18 **PMIX_RANK_IS_VALID**(a)

 C 

19 **IN** a

20 Rank to be checked (**pmix_rank_t**)

21 Returns **true** if the given rank is valid (i.e., less than **PMIX_RANK_VALID**)

22 3.2.4 Process Structure

23 The **pmix_proc_t** structure is used to identify a single process in the PMIx universe. It contains
 24 a reference to the namespace and the **pmix_rank_t** within that namespace.

PMIx v1.0  C 

```
25 typedef struct pmix_proc {
26     pmix_namespace_t nspace;
27     pmix_rank_t rank;
28 } pmix_proc_t;
```

 C 

1 3.2.4.1 Process structure support macros

2 The following macros are provided to support the `pmix_proc_t` structure.

3 Static initializer for the proc structure

4 *(Provisional)*

5 Provide a static initializer for the `pmix_proc_t` fields.

PMIx v5.0

▼ C ————— ▼

6 **PMIX_PROC_STATIC_INIT**

▲ C ————— ▲

7 Initialize the proc structure

8 Initialize the `pmix_proc_t` fields.

PMIx v1.0

▼ C ————— ▼

9 **PMIX_PROC_CONSTRUCT (m)**

▲ C ————— ▲

10 **IN** m

11 Pointer to the structure to be initialized (pointer to `pmix_proc_t`)

12 Destruct the proc structure

13 Destruct the `pmix_proc_t` fields.

▼ C ————— ▼

14 **PMIX_PROC_DESTRUCT (m)**

▲ C ————— ▲

15 **IN** m

16 Pointer to the structure to be destructed (pointer to `pmix_proc_t`)

17 There is nothing to release here as the fields in `pmix_proc_t` are either a statically-declared array (the namespace) or a single value (the rank). However, the macro is provided for symmetry in the code and for future-proofing should some allocated field be included some day.

20 Create a proc array

21 Allocate and initialize an array of `pmix_proc_t` structures.

PMIx v1.0

▼ C ————— ▼

22 **PMIX_PROC_CREATE (m, n)**

▲ C ————— ▲

23 **INOUT** m

24 Address where the pointer to the array of `pmix_proc_t` structures shall be stored (handle)

25 **IN** n

26 Number of structures to be allocated (**size_t**)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Free a proc structure

Release a `pmix_proc_t` structure.



`PMIX_PROC_RELEASE` (m)



IN m

Pointer to a `pmix_proc_t` structure (handle)

Free a proc array

Release an array of `pmix_proc_t` structures.

PMIx v1.0



`PMIX_PROC_FREE` (m, n)



IN m

Pointer to the array of `pmix_proc_t` structures (handle)

IN n

Number of structures in the array (`size_t`)

Load a proc structure

Load values into a `pmix_proc_t`.

PMIx v2.0



`PMIX_PROC_LOAD` (m, n, r)



IN m

Pointer to the structure to be loaded (pointer to `pmix_proc_t`)

IN n

Namespace to be loaded (`pmix_namespace_t`)

IN r

Rank to be assigned (`pmix_rank_t`)

No return value. Deprecated in favor of `PMIX_LOAD_PROCID`

Compare identifiers

Compare two `pmix_proc_t` identifiers.

PMIx v3.0

1 **PMIX_CHECK_PROCID** (a, b) C

2 **IN a**
3 Pointer to a structure whose ID is to be compared (pointer to [pmix_proc_t](#))
4 **IN b**
5 Pointer to a structure whose ID is to be compared (pointer to [pmix_proc_t](#))

6 Returns **true** if the two structures contain matching namespaces and:

- 7 • the ranks are the same value
- 8 • one of the ranks is [PMIX_RANK_WILDCARD](#)

9 **Check if a process identifier is valid**

10 Check for invalid namespace or rank value

11 *Provisional*
v4.1

11 **PMIX_PROCID_INVALID** (a) C

12 **IN a**
13 Pointer to a structure whose ID is to be checked (pointer to [pmix_proc_t](#))

14 Returns **true** if the process identifier contains either an empty (i.e., invalid) *namespace* field or a *rank* field of
15 [PMIX_RANK_INVALID](#)

16 **Load a proclD structure**

17 Load values into a [pmix_proc_t](#).

17 *PMIx v4.0*

18 **PMIX_LOAD_PROCID** (m, n, r) C

19 **IN m**
20 Pointer to the structure to be loaded (pointer to [pmix_proc_t](#))
21 **IN n**
22 Namespace to be loaded ([pmix_namespace_t](#))
23 **IN r**
24 Rank to be assigned ([pmix_rank_t](#))

1 **Transfer a proclD structure**

2 Transfer contents of one `pmix_proc_t` value to another `pmix_proc_t`.



3 **PMIX_PROCID_XFER(d, s)**



4 **IN d**
5 Pointer to the target structure (pointer to `pmix_proc_t`)

6 **IN s**
7 Pointer to the source structure (pointer to `pmix_proc_t`)

8 **Construct a multi-cluster namespace**

9 Construct a multi-cluster identifier containing a cluster ID and a namespace.



10 **PMIX_MULTICLUSTER_NAMESPACE_CONSTRUCT(m, n, r)**



11 **IN m**
12 `pmix_namespace_t` structure that will contain the multi-cluster identifier (`pmix_namespace_t`)

13 **IN n**
14 Cluster identifier (`char*`)

15 **IN r**
16 Namespace to be loaded (`pmix_namespace_t`)

17 Combined length of the cluster identifier and namespace must be less than `PMIX_MAX_NSLEN-2`.

18 **Parse a multi-cluster namespace**

19 Parse a multi-cluster identifier into its cluster ID and namespace parts.



20 **PMIX_MULTICLUSTER_NAMESPACE_PARSE(m, n, r)**



21 **IN m**
22 `pmix_namespace_t` structure containing the multi-cluster identifier (pointer to `pmix_namespace_t`)

23 **IN n**
24 Location where the cluster ID is to be stored (`pmix_namespace_t`)

25 **IN r**
26 Location where the namespace is to be stored (`pmix_namespace_t`)

1 3.2.5 Process State Structure

2 The `pmix_proc_state_t` structure is a `uint8_t` type for process state values. The following constants
3 can be used to set a variable of the type `pmix_proc_state_t`.

Advice to users

4 The fine-grained nature of the following constants may exceed the ability of an RM to provide updated process
5 state values during the process lifetime. This is particularly true of states for short-lived processes.

6 `PMIX_PROC_STATE_UNDEF` 0 Undefined process state.
7 `PMIX_PROC_STATE_PREPPED` 1 Process is ready to be launched.
8 `PMIX_PROC_STATE_LAUNCH_UNDERWAY` 2 Process launch is underway.
9 `PMIX_PROC_STATE_RESTART` 3 Process is ready for restart.
10 `PMIX_PROC_STATE_TERMINATE` 4 Process is marked for termination.
11 `PMIX_PROC_STATE_RUNNING` 5 Process has been locally `fork`'ed by the RM.
12 `PMIX_PROC_STATE_CONNECTED` 6 Process has connected to PMIx server.
13 `PMIX_PROC_STATE_UNTERMINATED` 15 Define a "boundary" between the terminated states and
14 `PMIX_PROC_STATE_CONNECTED` so users can easily and quickly determine if a process is still
15 running or not. Any value less than this constant means that the process has not terminated.
16 `PMIX_PROC_STATE_TERMINATED` 20 Process has terminated and is no longer running.
17 `PMIX_PROC_STATE_ERROR` 50 Define a boundary so users can easily and quickly determine if a
18 process abnormally terminated. Any value above this constant means that the process has terminated
19 abnormally.
20 `PMIX_PROC_STATE_KILLED_BY_CMD` 51 Process was killed by a command.
21 `PMIX_PROC_STATE_ABORTED` 52 Process was aborted by a call to `PMIx_Abort`.
22 `PMIX_PROC_STATE_FAILED_TO_START` 53 Process failed to start.
23 `PMIX_PROC_STATE_ABORTED_BY_SIG` 54 Process aborted by a signal.
24 `PMIX_PROC_STATE_TERM_WO_SYNC` 55 Process exited without calling `PMIx_Finalize`.
25 `PMIX_PROC_STATE_COMM_FAILED` 56 Process communication has failed.
26 `PMIX_PROC_STATE_SENSOR_BOUND_EXCEEDED` 57 Process exceeded a specified sensor limit.
27 `PMIX_PROC_STATE_CALLED_ABORT` 58 Process called `PMIx_Abort`.
28 `PMIX_PROC_STATE_HEARTBEAT_FAILED` 59 Process failed to send heartbeat within specified time
29 limit.
30 `PMIX_PROC_STATE_MIGRATING` 60 Process failed and is waiting for resources before restarting.
31 `PMIX_PROC_STATE_CANNOT_RESTART` 61 Process failed and cannot be restarted.
32 `PMIX_PROC_STATE_TERM_NON_ZERO` 62 Process exited with a non-zero status.
33 `PMIX_PROC_STATE_FAILED_TO_LAUNCH` 63 Unable to launch process.

1 3.2.6 Process Information Structure

2 The `pmix_proc_info_t` structure defines a set of information about a specific process including its name,
3 location, and state.

```
4 typedef struct pmix_proc_info {  
5     /** Process structure */  
6     pmix_proc_t proc;  
7     /** Hostname where process resides */  
8     char *hostname;  
9     /** Name of the executable */  
10    char *executable_name;  
11    /** Process ID on the host */  
12    pid_t pid;  
13    /** Exit code of the process. Default: 0 */  
14    int exit_code;  
15    /** Current state of the process */  
16    pmix_proc_state_t state;  
17 } pmix_proc_info_t;
```

18 3.2.6.1 Process information structure support macros

19 The following macros are provided to support the `pmix_proc_info_t` structure.

20 Static initializer for the proc info structure

21 *(Provisional)*

22 Provide a static initializer for the `pmix_proc_info_t` fields.

PMIx v5.0

23 `PMIX_PROC_INFO_STATIC_INIT`

24 Initialize the process information structure

25 Initialize the `pmix_proc_info_t` fields.

PMIx v2.0

26 `PMIX_PROC_INFO_CONSTRUCT(m)`

27 **IN** m

28 Pointer to the structure to be initialized (pointer to `pmix_proc_info_t`)

1 **Destruct the process information structure**

2 Destruct the `pmix_proc_info_t` fields.

▼  

3 **PMIX_PROC_INFO_DESTRUCT** (m)

▲  

4 **IN** m

5 Pointer to the structure to be destructed (pointer to `pmix_proc_info_t`)

6 **Create a process information array**

7 Allocate and initialize a `pmix_proc_info_t` array.

PMIx v2.0

▼  

8 **PMIX_PROC_INFO_CREATE** (m, n)

▲  

9 **INOUT** m

10 Address where the pointer to the array of `pmix_proc_info_t` structures shall be stored (handle)

11 **IN** n

12 Number of structures to be allocated (`size_t`)

13 **Free a process information structure**

14 Release a `pmix_proc_info_t` structure.

PMIx v2.0

▼  

15 **PMIX_PROC_INFO_RELEASE** (m)

▲  

16 **IN** m

17 Pointer to a `pmix_proc_info_t` structure (handle)

18 **Free a process information array**

19 Release an array of `pmix_proc_info_t` structures.

PMIx v2.0

▼  

20 **PMIX_PROC_INFO_FREE** (m, n)

▲  

21 **IN** m

22 Pointer to the array of `pmix_proc_info_t` structures (handle)

23 **IN** n

24 Number of structures in the array (`size_t`)

1 3.2.7 Job State Structure

2 The `pmix_job_state_t` structure is a `uint8_t` type for job state values. The following constants can be
3 used to set a variable of the type `pmix_job_state_t`.

Advice to users

4 The fine-grained nature of the following constants may exceed the ability of an RM to provide updated job
5 state values during the job lifetime. This is particularly true for short-lived jobs.

6 `PMIX_JOB_STATE_UNDEF 0` Undefined job state.
7 `PMIX_JOB_STATE_AWAITING_ALLOC 1` Job is waiting for resources to be allocated to it.
8 `PMIX_JOB_STATE_LAUNCH_UNDERWAY 2` Job launch is underway.
9 `PMIX_JOB_STATE_RUNNING 3` All processes in the job have been spawned and are executing.
10 `PMIX_JOB_STATE_SUSPENDED 4` All processes in the job have been suspended.
11 `PMIX_JOB_STATE_CONNECTED 5` All processes in the job have connected to their PMIx server.
12 `PMIX_JOB_STATE_UNTERMINATED 15` Define a “boundary” between the terminated states and
13 `PMIX_JOB_STATE_TERMINATED` so users can easily and quickly determine if a job is still running
14 or not. Any value less than this constant means that the job has not terminated.
15 `PMIX_JOB_STATE_TERMINATED 20` All processes in the job have terminated and are no longer
16 running - typically will be accompanied by the job exit status in response to a query.
17 `PMIX_JOB_STATE_TERMINATED_WITH_ERROR 50` Define a boundary so users can easily and
18 quickly determine if a job abnormally terminated - typically will be accompanied by a job-related error
19 code in response to a query Any value above this constant means that the job terminated abnormally.

20 3.2.8 Value Structure

21 The `pmix_value_t` structure is used to represent the value passed to `PMIx_Put` and retrieved by
22 `PMIx_Get`, as well as many of the other PMIx functions.

23 A collection of values may be specified under a single key by passing a `pmix_value_t` containing an array
24 of type `pmix_data_array_t`, with each array element containing its own object. All members shown
25 below were introduced in version 1 of the standard unless otherwise marked.

PMIx v1.0

```

1  typedef struct pmix_value {
2      pmix_data_type_t type;
3      union {
4          bool flag;
5          uint8_t byte;
6          char *string;
7          size_t size;
8          pid_t pid;
9          int integer;
10         int8_t int8;
11         int16_t int16;
12         int32_t int32;
13         int64_t int64;
14         unsigned int uint;
15         uint8_t uint8;
16         uint16_t uint16;
17         uint32_t uint32;
18         uint64_t uint64;
19         float fval;
20         double dval;
21         struct timeval tv;
22         time_t time; // version 2.0
23         pmix_status_t status; // version 2.0
24         pmix_rank_t rank; // version 2.0
25         pmix_proc_t *proc; // version 2.0
26         pmix_byte_object_t bo;
27         pmix_persistence_t persist; // version 2.0
28         pmix_scope_t scope; // version 2.0
29         pmix_data_range_t range; // version 2.0
30         pmix_proc_state_t state; // version 2.0
31         pmix_proc_info_t *pinfo; // version 2.0
32         pmix_data_array_t *darray; // version 2.0
33         void *ptr; // version 2.0
34         pmix_alloc_directive_t adir; // version 2.0
35     } data;
36 } pmix_value_t;

```

37 3.2.8.1 Value structure support

38 The following macros and APIs are provided to support the `pmix_value_t` structure.

39 Static initializer for the value structure

40 *(Provisional)*

41 Provide a static initializer for the `pmix_value_t` fields.

PMIx v5.0

1			C	
		PMIX_VALUE_STATIC_INIT		
2		Initialize the value structure		
3		Initialize the <code>pmix_value_t</code> fields.		
	<i>PMIx v1.0</i>		C	
4		PMIX_VALUE_CONSTRUCT (m)		
5		IN m		
6		Pointer to the structure to be initialized (pointer to <code>pmix_value_t</code>)		
7		Destruct the value structure		
8		Destruct the <code>pmix_value_t</code> fields.		
	<i>PMIx v1.0</i>		C	
9		PMIX_VALUE_DESTRUCT (m)		
10		IN m		
11		Pointer to the structure to be destructed (pointer to <code>pmix_value_t</code>)		
12		Create a value array		
13		Allocate and initialize an array of <code>pmix_value_t</code> structures.		
	<i>PMIx v1.0</i>		C	
14		PMIX_VALUE_CREATE (m, n)		
15		INOUT m		
16		Address where the pointer to the array of <code>pmix_value_t</code> structures shall be stored (handle)		
17		IN n		
18		Number of structures to be allocated (<code>size_t</code>)		
19		Free a value structure		
20		Release a <code>pmix_value_t</code> structure.		
	<i>PMIx v4.0</i>		C	
21		PMIX_VALUE_RELEASE (m)		
22		IN m		
23		Pointer to a <code>pmix_value_t</code> structure (handle)		

1 **Free a value array**
2 Release an array of `pmix_value_t` structures.

▼ C 

3 `PMIX_VALUE_FREE(m, n)`

▲ C 

4 **IN** `m`
5 Pointer to the array of `pmix_value_t` structures (handle)
6 **IN** `n`
7 Number of structures in the array (`size_t`)

8 **Load a value structure**

9 **Summary**

10 Load data into a `pmix_value_t` structure.

11 **Format**

PMIx v5.0

▼ C 

```
12 pmix_status_t  
13 PMIx_Value_load(pmix_value_t *val,  
14                 const void *data,  
15                 pmix_data_type_t type);
```

▲ C 

16 **IN** `val`
17 The `pmix_value_t` into which the data is to be loaded (pointer to `pmix_value_t`)
18 **IN** `data`
19 Pointer to the data value to be loaded (handle)
20 **IN** `type`
21 Type of the provided data value (`pmix_data_type_t`)

22 Returns `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant.

23 **Description**

24 Copy the provided data into the `pmix_value_t`. Any data stored in the source value can be modified or
25 free'd without affecting the copied data once the function has completed.

26 **Unload a value structure**

27 **Summary**

28 Unload data from a `pmix_value_t` structure.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

Format

C

```
pmix_status_t  
PMIx_Value_unload(pmix_value_t *val,  
                  void **data,  
                  size_t *sz);
```

C

IN *val*

The `pmix_value_t` from which the data is to be unloaded (pointer to `pmix_value_t`)

INOUT *data*

Pointer to the location where the data value is to be returned (handle)

INOUT *sz*

Pointer to return the size of the unloaded value (handle)

Returns `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant.

Description

Return a copy of the data in the `pmix_value_t`. The source value can be modified or free'd without affecting the copied data once the function has completed.

Advice to users

Memory will be allocated and the pointer to that data will be in the `data` argument - the source `pmix_value_t` will not be altered. The user is responsible for releasing the returned data.

Transfer data between value structures

Summary

Transfer the data value between two `pmix_value_t` structures.

Format

PMIx v5.0

C

```
pmix_status_t  
PMIx_Value_xfer(pmix_value_t *dest,  
                const pmix_value_t *src);
```

C

IN *dest*

Pointer to the `pmix_value_t` destination (handle)

IN *src*

Pointer to the `pmix_value_t` source (handle)

Returns `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant.

Description

Copy the data in the source `pmix_value_t` into the destination `pmix_value_t`. The source value can be modified or free'd without affecting the copied data once the function has completed.

1 Retrieve a numerical value from a value struct

2 Retrieve a numerical value from a `pmix_value_t` structure.

▼ `C` ▼

3 `PMIX_VALUE_GET_NUMBER(s, m, n, t)`

▲ `C` ▲

4 **OUT** `s`

5 Status code for the request (`pmix_status_t`)

6 **IN** `m`

7 Pointer to the `pmix_value_t` structure (handle)

8 **OUT** `n`

9 Variable to be set to the value (match expected type)

10 **IN** `t`

11 Type of number expected in `m` (`pmix_data_type_t`)

12 Sets the provided variable equal to the numerical value contained in the given `pmix_value_t`, returning
13 success if the data type of the value matches the expected type and `PMIX_ERR_BAD_PARAM` if it doesn't

14 3.2.9 Info Structure

15 The `pmix_info_t` structure defines a key/value pair with associated directive. All fields were defined in
16 version 1.0 unless otherwise marked.

PMIx v1.0 ▼ `C` ▼

```
17 typedef struct pmix_info_t {  
18     pmix_key_t key;  
19     pmix_info_directives_t flags; // version 2.0  
20     pmix_value_t value;  
21 } pmix_info_t;
```

▲ `C` ▲

22 3.2.9.1 Info structure support macros

23 The following macros are provided to support the `pmix_info_t` structure.

24 Static initializer for the info structure

25 *(Provisional)*

26 Provide a static initializer for the `pmix_info_t` fields.

PMIx v5.0 ▼ `C` ▼

27 `PMIX_INFO_STATIC_INIT`

▲ `C` ▲

1 **Initialize the info structure**

2 Initialize the `pmix_info_t` fields.



3 **PMIX_INFO_CONSTRUCT (m)**



4 **IN m**

5 Pointer to the structure to be initialized (pointer to `pmix_info_t`)

6 **Destruct the info structure**

7 Destruct the `pmix_info_t` fields.

PMIx v1.0



8 **PMIX_INFO_DESTRUCT (m)**



9 **IN m**

10 Pointer to the structure to be destructed (pointer to `pmix_info_t`)

11 **Create an info array**

12 Allocate and initialize an array of info structures.

PMIx v1.0



13 **PMIX_INFO_CREATE (m, n)**



14 **INOUT m**

15 Address where the pointer to the array of `pmix_info_t` structures shall be stored (handle)

16 **IN n**

17 Number of structures to be allocated (`size_t`)

18 **Free an info array**

19 Release an array of `pmix_info_t` structures.

PMIx v1.0



20 **PMIX_INFO_FREE (m, n)**



21 **IN m**

22 Pointer to the array of `pmix_info_t` structures (handle)

23 **IN n**

24 Number of structures in the array (`size_t`)

25 **Load key and value data into a info struct**

26 **Summary**

27 Load data into a `pmix_info_t` structure.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

Format

C

```
pmix_status_t  
PMIx_Info_load(pmix_info_t *info,  
               const char* key,  
               const void *data,  
               pmix_data_type_t type);
```

C

- IN info**
The [pmix_info_t](#) into which the data is to be loaded (handle)
- IN key**
Pointer to the key to be loaded (handle)
- IN data**
Pointer to the data value to be loaded (handle)
- IN type**
Type of the provided data value ([pmix_data_type_t](#))

Returns [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant.

Description

Copy the provided data into the [pmix_info_t](#). Any data stored in the source parameters can be modified or free'd without affecting the copied data once the function has completed.

Copy data between info structures

Summary

Copy all data between two [pmix_info_t](#) structures.

Format

C

```
pmix_status_t  
PMIx_Info_xfer(pmix_info_t *dest,  
               pmix_info_t *src);
```

C

- IN dest**
The [pmix_info_t](#) into which the data is to be copied (handle)
- IN src**
The [pmix_info_t](#) from which the data is to be copied (handle)

Returns [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant.

Description

Copy the data in the source [pmix_info_t](#) into the destination. Any data stored in the source structure can be modified or free'd without affecting the copied data once the function has completed.

1 **Test a boolean info struct**

2 A special macro for checking if a boolean `pmix_info_t` is `true`.

```
▼ C
```

3 **PMIX_INFO_TRUE (m)**

```
▲ C
```

4 **IN** `m`

5 Pointer to a `pmix_info_t` structure (handle)

6 A `pmix_info_t` structure is considered to be of type `PMIX_BOOL` and value `true` if:

- 7 • the structure reports a type of `PMIX_UNDEF`, or
- 8 • the structure reports a type of `PMIX_BOOL` and the data flag is `true`

9 **3.2.9.2 Info structure list macros**

10 Constructing an array of `pmix_info_t` is a fairly common operation. The following macros are provided to
11 simplify this construction.

12 **Start a list of `pmix_info_t` structures**

13 **Summary**

14 Initialize a list of `pmix_info_t` structures. The actual list is opaque to the caller and is
15 implementation-dependent.

16 **Format**

PMIx v5.0

```
▼ C
```

17 `void*`

18 `PMIx_Info_list_start (void);`

```
▲ C
```

19 **Description**

20 Note that the returned pointer will be initialized to an opaque structure whose elements are
21 implementation-dependent. The caller must not modify or dereference the object.

22 **Add a `pmix_info_t` structure to a list**

23 **Summary**

24 Add a `pmix_info_t` structure containing the specified value to the provided list.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

Format

C

```
pmix_status_t
PMIx_Info_list_add(void *ptr,
                   const char *key,
                   const void *value,
                   pmix_data_type_t type);
```

C

IN ptr

A void* pointer initialized via [PMIx_Info_list_start](#) (handle)

IN key

String key to be loaded - must be less than or equal to [PMIX_MAX_KEYLEN](#) in length (handle)

IN value

Pointer to the data value to be loaded (handle)

IN type

Type of the provided data value ([pmix_data_type_t](#))

Returns [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant.

Description

Copy the provided key and data into a [pmix_info_t](#) on the list. The key and any data stored in the source value can be modified or free'd without affecting the copied data once the function has completed.

Transfer a [pmix_info_t](#) structure to a list

Summary

Transfer the information in a [pmix_info_t](#) structure to a structure on the provided list.

Format

C

```
pmix_status_t
PMIx_Info_list_xfer(void *ptr,
                   const pmix_info_t *src);
```

C

IN ptr

A void* pointer initialized via [PMIx_Info_list_start](#) (handle)

IN src

Pointer to the source [pmix_info_t](#) (pointer to [pmix_info_t](#))

Returns [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant.

Description

All data (including key, value, and directives) will be copied into a destination [pmix_info_t](#) on the list. The source [pmix_info_t](#) may be free'd without affecting the copied data once the function has completed.

Convert a [pmix_info_t](#) list to an array

Summary

Transfer the information in the provided [pmix_info_t](#) list to a [pmix_data_array_t](#) array

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Format

C

```
pmix_status_t
PMIx_Info_list_convert(void *ptr,
                      pmix_data_array_t *par);
```

C

- IN ptr**
A `void*` pointer initialized via `PMIx_Info_list_start` (handle)
- IN par**
Pointer to an instantiated `pmix_data_array_t` structure where the `pmix_info_t` array is to be stored (pointer to `pmix_data_array_t`)

Returns `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant.

Description

Information collected in the provided list of `pmix_info_t` will be transferred to a `pmix_data_array_t` containing `pmix_info_t` structures.

Release a `pmix_info_t` list

Summary

Release the provided `pmix_info_t` list

Format

C

```
void
PMIx_Info_list_release(void *ptr);
```

C

- IN ptr**
A `void*` pointer initialized via `PMIx_Info_list_start` (handle)

Description

Information contained in the `pmix_info_t` on the list shall be released in addition to whatever backing storage the implementation may have allocated to support construction of the list.

PMIx v5.0

3.2.10 Info Type Directives

PMIx v2.0

The `pmix_info_directives_t` structure is a `uint32_t` type that defines the behavior of command directives via `pmix_info_t` arrays. By default, the values in the `pmix_info_t` array passed to a PMIx are *optional*.

Advice to users

A PMIx implementation or PMIx-enabled RM may ignore any `pmix_info_t` value passed to a PMIx API that it does not support or does not recognize if it is not explicitly marked as `PMIX_INFO_REQD`. This is because the values specified default to optional, meaning they can be ignored in such circumstances. This may lead to unexpected behavior when porting between environments or PMIx implementations if the user is relying on the behavior specified by the `pmix_info_t` value. Users relying on the behavior defined by the `pmix_info_t` are advised to set the `PMIX_INFO_REQD` flag using the `PMIX_INFO_REQUIRED` macro.

Advice to PMIx library implementers

The top 16-bits of the `pmix_info_directives_t` are reserved for internal use by PMIx library implementers - the PMIx standard will *not* specify their intent, leaving them for customized use by implementers. Implementers are advised to use the provided `PMIX_INFO_IS_REQUIRED` macro for testing this flag, and must return `PMIX_ERR_NOT_SUPPORTED` as soon as possible to the caller if the required behavior is not supported.

The following constants were introduced in version 2.0 (unless otherwise marked) and can be used to set a variable of the type `pmix_info_directives_t`.

PMIX_INFO_REQD `0x00000001` The behavior defined in the `pmix_info_t` array is required, and not optional. This is a bit-mask value.

PMIX_INFO_REQD_PROCESSED `0x00000004` Mark that this required attribute has been processed. A required attribute can be handled at any level - the PMIx client library might take care of it, or it may be resolved by the PMIx server library, or it may pass up to the host environment for handling. If a level does not recognize or support the required attribute, it is required to pass it upwards to give the next level an opportunity to process it. Thus, the host environment (or the server library if the host does not support the given operation) must know if a lower level has handled the requirement so it can return a `PMIX_ERR_NOT_SUPPORTED` error status if the host itself cannot meet the request. Upon processing the request, the level must therefore mark the attribute with this directive to alert any subsequent levels that the requirement has been met.

PMIX_INFO_ARRAY_END `0x00000002` Mark that this `pmix_info_t` struct is at the end of an array created by the `PMIX_INFO_CREATE` macro. This is a bit-mask value.

PMIX_INFO_DIR_RESERVED `0xffff0000` A bit-mask identifying the bits reserved for internal use by implementers - these currently are set as `0xffff0000`.

Advice to PMIx server hosts

Host environments are advised to use the provided `PMIX_INFO_IS_REQUIRED` macro for testing this flag and must return `PMIX_ERR_NOT_SUPPORTED` as soon as possible to the caller if the required behavior is not supported.

3.2.10.1 Info Directive support macros

The following macros are provided to support the setting and testing of `pmix_info_t` directives.

Mark an info structure as required

Set the `PMIX_INFO_REQD` flag in a `pmix_info_t` structure.

PMIx v2.0

```
PMIX_INFO_REQUIRED(info);
```

IN `info`

Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

This macro simplifies the setting of the `PMIX_INFO_REQD` flag in `pmix_info_t` structures.

1 **Mark an info structure as optional**

2 Unsets the `PMIX_INFO_REQD` flag in a `pmix_info_t` structure.

PMIx v2.0

```
▼ _____ C _____ ▼
```

3 `PMIX_INFO_OPTIONAL(info);`

```
▲ _____ C _____ ▲
```

4 **IN** `info`

5 Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

6 This macro simplifies marking a `pmix_info_t` structure as *optional*.

7 **Test an info structure for *required* directive**

8 Test the `PMIX_INFO_REQD` flag in a `pmix_info_t` structure, returning `true` if the flag is set.

PMIx v2.0

```
▼ _____ C _____ ▼
```

9 `PMIX_INFO_IS_REQUIRED(info);`

```
▲ _____ C _____ ▲
```

10 **IN** `info`

11 Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

12 This macro simplifies the testing of the required flag in `pmix_info_t` structures.

13 **Test an info structure for *optional* directive**

14 Test a `pmix_info_t` structure, returning `true` if the structure is *optional*.

PMIx v2.0

```
▼ _____ C _____ ▼
```

15 `PMIX_INFO_IS_OPTIONAL(info);`

```
▲ _____ C _____ ▲
```

16 **IN** `info`

17 Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

18 Test the `PMIX_INFO_REQD` flag in a `pmix_info_t` structure, returning `true` if the flag is *not* set.

19 **Mark a required attribute as processed**

20 Mark that a required `pmix_info_t` structure has been processed.

PMIx v4.0

```
▼ _____ C _____ ▼
```

21 `PMIX_INFO_PROCESSED(info);`

```
▲ _____ C _____ ▲
```

22 **IN** `info`

23 Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

24 Set the `PMIX_INFO_REQD_PROCESSED` flag in a `pmix_info_t` structure indicating that is has been
25 processed.

1 **Test if a required attribute has been processed**

2 Test that a required `pmix_info_t` structure has been processed.

▼ `PMIX_INFO_WAS_PROCESSED` C `(info)` ▼

3 `PMIX_INFO_WAS_PROCESSED`(info);

▲ `PMIX_INFO_WAS_PROCESSED` C `(info)` ▲

4 **IN** info

5 Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

6 Test the `PMIX_INFO_REQD_PROCESSED` flag in a `pmix_info_t` structure.

7 **Test an info structure for *end of array* directive**

8 Test a `pmix_info_t` structure, returning `true` if the structure is at the end of an array created by the

9 `PMIX_INFO_CREATE` macro.

PMIx v2.2

▼ `PMIX_INFO_IS_END` C `(info)` ▼

10 `PMIX_INFO_IS_END`(info);

▲ `PMIX_INFO_IS_END` C `(info)` ▲

11 **IN** info

12 Pointer to the `pmix_info_t` (pointer to `pmix_info_t`)

13 This macro simplifies the testing of the end-of-array flag in `pmix_info_t` structures.

14 **3.2.11 Environmental Variable Structure**

PMIx v3.0

15 Define a structure for specifying environment variable modifications. Standard environment variables (e.g.,

16 `PATH`, `LD_LIBRARY_PATH`, and `LD_PRELOAD`) take multiple arguments separated by delimiters.

17 Unfortunately, the delimiters depend upon the variable itself - some use semi-colons, some colons, etc. Thus,

18 the operation requires not only the name of the variable to be modified and the value to be inserted, but also

19 the separator to be used when composing the aggregate value.

▼ `pmix_env_var_t` C `(info)` ▼

```
20 typedef struct {  
21     char *envar;  
22     char *value;  
23     char separator;  
24 } pmix_env_var_t;
```

▲ `pmix_env_var_t` C `(info)` ▲

25 **3.2.11.1 Environmental variable support macros**

26 The following macros are provided to support the `pmix_env_var_t` structure.

1 **Static initializer for the envvar structure**

2 *(Provisional)*

3 Provide a static initializer for the `pmix_envvar_t` fields.



4 `PMIX_ENVVAR_STATIC_INIT`



5 **Initialize the envvar structure**

6 Initialize the `pmix_envvar_t` fields.

PMIx v3.0



7 `PMIX_ENVVAR_CONSTRUCT (m)`



8 **IN** `m`

9 Pointer to the structure to be initialized (pointer to `pmix_envvar_t`)

10 **Destruct the envvar structure**

11 Clear the `pmix_envvar_t` fields.

PMIx v3.0



12 `PMIX_ENVVAR_DESTRUCT (m)`



13 **IN** `m`

14 Pointer to the structure to be destructed (pointer to `pmix_envvar_t`)

15 **Create an envvar array**

16 Allocate and initialize an array of `pmix_envvar_t` structures.

PMIx v3.0



17 `PMIX_ENVVAR_CREATE (m, n)`



18 **INOUT** `m`

19 Address where the pointer to the array of `pmix_envvar_t` structures shall be stored (handle)

20 **IN** `n`

21 Number of structures to be allocated (`size_t`)

22 **Free an envvar array**

23 Release an array of `pmix_envvar_t` structures.

PMIx v3.0



24 `PMIX_ENVVAR_FREE (m, n)`



25 **IN** `m`

26 Pointer to the array of `pmix_envvar_t` structures (handle)

27 **IN** `n`

28 Number of structures in the array (`size_t`)

1 Load an envar structure

2 Load values into a `pmix_envvar_t`.

PMIx v2.0

3 `PMIX_ENVAR_LOAD(m, e, v, s)`

4 **IN** `m`

Pointer to the structure to be loaded (pointer to `pmix_envvar_t`)

5 **IN** `e`

Environmental variable name (`char*`)

6 **IN** `v`

Value of variable (`char*`)

7 **IN** `v`

Separator character (`char`)

12 3.2.12 Byte Object Type

13 The `pmix_byte_object_t` structure describes a raw byte sequence.

PMIx v1.0

```
14 typedef struct pmix_byte_object {  
15     char *bytes;  
16     size_t size;  
17 } pmix_byte_object_t;
```

18 3.2.12.1 Byte object support macros

19 The following macros support the `pmix_byte_object_t` structure.

20 Static initializer for the byte object structure

21 *(Provisional)*

22 Provide a static initializer for the `pmix_byte_object_t` fields.

PMIx v5.0

23 `PMIX_BYTE_OBJECT_STATIC_INIT`

24 Initialize the byte object structure

25 Initialize the `pmix_byte_object_t` fields.

PMIx v2.0

26 `PMIX_BYTE_OBJECT_CONSTRUCT(m)`

27 **IN** `m`

Pointer to the structure to be initialized (pointer to `pmix_byte_object_t`)

1 **Destruct the byte object structure**

2 Clear the `pmix_byte_object_t` fields.



3 `PMIX_BYTE_OBJECT_DESTRUCT(m)`



4 **IN** `m`

5 Pointer to the structure to be destructed (pointer to `pmix_byte_object_t`)

6 **Create a byte object structure**

7 Allocate and initialize an array of `pmix_byte_object_t` structures.

PMIx v2.0



8 `PMIX_BYTE_OBJECT_CREATE(m, n)`



9 **INOUT** `m`

10 Address where the pointer to the array of `pmix_byte_object_t` structures shall be stored (handle)

11 **IN** `n`

12 Number of structures to be allocated (`size_t`)

13 **Free a byte object array**

14 Release an array of `pmix_byte_object_t` structures.

PMIx v2.0



15 `PMIX_BYTE_OBJECT_FREE(m, n)`



16 **IN** `m`

17 Pointer to the array of `pmix_byte_object_t` structures (handle)

18 **IN** `n`

19 Number of structures in the array (`size_t`)

20 **Load a byte object structure**

21 Load values into a `pmix_byte_object_t`.

PMIx v2.0



22 `PMIX_BYTE_OBJECT_LOAD(b, d, s)`



23 **IN** `b`

24 Pointer to the structure to be loaded (pointer to `pmix_byte_object_t`)

25 **IN** `d`

26 Pointer to the data to be loaded (`char*`)

27 **IN** `s`

28 Number of bytes in the data array (`size_t`)

1 3.2.13 Data Array Structure

2 The `pmix_data_array_t` structure defines an array data structure.

```
3 typedef struct pmix_data_array {  
4     pmix_data_type_t type;  
5     size_t size;  
6     void *array;  
7 } pmix_data_array_t;
```

8 3.2.13.1 Data array support macros

9 The following macros support the `pmix_data_array_t` structure.

10 Static initializer for the data array structure

11 *(Provisional)*

12 Provide a static initializer for the `pmix_data_array_t` fields.

PMIx v5.0

```
13 PMIX_DATA_ARRAY_STATIC_INIT
```

14 Initialize a data array structure

15 Initialize the `pmix_data_array_t` fields, allocating memory for the array of the indicated type.

PMIx v2.2

```
16 PMIX_DATA_ARRAY_CONSTRUCT(m, n, t)
```

17 **IN** m

18 Pointer to the structure to be initialized (pointer to `pmix_data_array_t`)

19 **IN** n

20 Number of elements in the array (`size_t`)

21 **IN** t

22 PMIx data type of the array elements (`pmix_data_type_t`)

23 Destruct a data array structure

24 Destruct the `pmix_data_array_t`, releasing the memory in the array.

PMIx v2.2

```
25 PMIX_DATA_ARRAY_DESTRUCT(m)
```

26 **IN** m

27 Pointer to the structure to be destructed (pointer to `pmix_data_array_t`)

1 Create a data array structure

2 Allocate memory for the `pmix_data_array_t` object itself, and then allocate memory for the array of the
3 indicated type.

▼ C ▼

4 `PMIX_DATA_ARRAY_CREATE(m, n, t)`

▲ C ▲

5 **INOUT** `m`

6 Variable to be set to the address of the structure (pointer to `pmix_data_array_t`)

7 **IN** `n`

8 Number of elements in the array (`size_t`)

9 **IN** `t`

10 PMIx data type of the array elements (`pmix_data_type_t`)

11 Free a data array structure

12 Release the memory in the array, and then release the `pmix_data_array_t` object itself.

PMIx v2.2

▼ C ▼

13 `PMIX_DATA_ARRAY_FREE(m)`

▲ C ▲

14 **IN** `m`

15 Pointer to the structure to be released (pointer to `pmix_data_array_t`)

16 3.2.14 Argument Array Macros

17 The following macros support the construction and release of `NULL`-terminated `argv` arrays of strings.

18 Argument array extension

19 Append a string to a `NULL`-terminated, `argv`-style array of strings.

▼ C ▼

20 `PMIX_ARGV_APPEND(r, a, b);`

▲ C ▲

21 **OUT** `r`

22 Status code indicating success or failure of the operation (`pmix_status_t`)

23 **INOUT** `a`

24 Argument list (pointer to `NULL`-terminated array of strings)

25 **IN** `b`

26 Argument to append to the list (string)

27 This function helps the caller build the `argv` portion of `pmix_app_t` structure, arrays of keys for querying,
28 or other places where `argv`-style string arrays are required.

▼ Advice to users ▼

29 The provided argument is copied into the destination array - thus, the source string can be free'd without
30 affecting the array once the macro has completed.

▲

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

Argument array prepend

Prepend a string to a NULL-terminated, argv-style array of strings.

```

▼ _____ C _____ ▼
PMIX_ARGV_PREPEND(r, a, b);
▲ _____ C _____ ▲

```

OUT r

Status code indicating success or failure of the operation ([pmix_status_t](#))

INOUT a

Argument list (pointer to NULL-terminated array of strings)

IN b

Argument to append to the list (string)

This function helps the caller build the **argv** portion of [pmix_app_t](#) structure, arrays of keys for querying, or other places where argv-style string arrays are required.

Advice to users

The provided argument is copied into the destination array - thus, the source string can be free'd without affecting the array once the macro has completed.

Argument array extension - unique

Append a string to a NULL-terminated, argv-style array of strings, but only if the provided argument doesn't already exist somewhere in the array.

```

▼ _____ C _____ ▼
PMIX_ARGV_APPEND_UNIQUE(r, a, b);
▲ _____ C _____ ▲

```

OUT r

Status code indicating success or failure of the operation ([pmix_status_t](#))

INOUT a

Argument list (pointer to NULL-terminated array of strings)

IN b

Argument to append to the list (string)

This function helps the caller build the **argv** portion of [pmix_app_t](#) structure, arrays of keys for querying, or other places where argv-style string arrays are required.

Advice to users

The provided argument is copied into the destination array - thus, the source string can be free'd without affecting the array once the macro has completed.

1 **Argument array release**
2 Free an argv-style array and all of the strings that it contains.

▼  C 

3 **PMIX_ARGV_FREE(a);**

▲  C 

4 **IN a**
5 Argument list (pointer to NULL-terminated array of strings)

6 This function releases the array and all of the strings it contains.

7 **Argument array split**
8 Split a string into a NULL-terminated argv array.

▼  C 

9 **PMIX_ARGV_SPLIT(a, b, c);**

▲  C 

10 **OUT a**
11 Resulting argv-style array (**char****)

12 **IN b**
13 String to be split (**char***)

14 **IN c**
15 Delimiter character (**char**)

16 Split an input string into a NULL-terminated argv array. Do not include empty strings in the resulting array.

▼  **Advice to users** 

17 All strings are inserted into the argv array by value; the newly-allocated array makes no references to the
18 src_string argument (i.e., it can be freed after calling this function without invalidating the output argv array)

▲ 

19 **Argument array join**
20 Join all the elements of an argv array into a single newly-allocated string.

▼  C 

21 **PMIX_ARGV_JOIN(a, b, c);**

▲  C 

22 **OUT a**
23 Resulting string (**char***)

24 **IN b**
25 Argv-style array to be joined (**char****)

26 **IN c**
27 Delimiter character (**char**)

28 Join all the elements of an argv array into a single newly-allocated string.

1 **Argument array count**
2 Return the length of a NULL-terminated argv array.

▼  

3 **PMIX_ARGV_COUNT**(*r*, *a*);

▲  

4 **OUT** *r*
5 Number of strings in the array (integer)

6 **IN** *a*
7 Argv-style array (**char****)

8 Count the number of elements in an argv array

9 **Argument array copy**
10 Copy an argv array, including copying all of its strings.

▼  

11 **PMIX_ARGV_COPY**(*a*, *b*);

▲  

12 **OUT** *a*
13 New argv-style array (**char****)

14 **IN** *b*
15 Argv-style array (**char****)

16 Copy an argv array, including copying all of its strings.

17 3.2.15 Set Environment Variable

18 **Summary**
19 Set an environment variable in a NULL-terminated, env-style array.

▼  

20 **PMIX_SETENV**(*r*, *name*, *value*, *env*);

▲  

21 **OUT** *r*
22 Status code indicating success or failure of the operation (**pmix_status_t**)

23 **IN** *name*
24 Argument name (string)

25 **IN** *value*
26 Argument value (string)

27 **INOUT** *env*
28 Environment array to update (pointer to array of strings)

Description

Similar to `setenv` from the C API, this allows the caller to set an environment variable in the specified `env` array, which could then be passed to the `pmix_app_t` structure or any other destination.

Advice to users

The provided name and value are copied into the destination environment array - thus, the source strings can be free'd without affecting the array once the macro has completed.

3.3 Generalized Data Types Used for Packing/Unpacking

The `pmix_data_type_t` structure is a `uint16_t` type for identifying the data type for packing/unpacking purposes. New data type values introduced in this version of the Standard are shown in **magenta**.

Advice to PMIx library implementers

The following constants can be used to set a variable of the type `pmix_data_type_t`. Data types in the PMIx Standard are defined in terms of the C-programming language. Implementers wishing to support other languages should provide the equivalent definitions in a language-appropriate manner. Additionally, a PMIx implementation may choose to add additional types.

<code>PMIX_UNDEF</code>	0	Undefined.
<code>PMIX_BOOL</code>	1	Boolean (converted to/from native <code>true/false</code>) (<code>bool</code>).
<code>PMIX_BYTE</code>	2	A byte of data (<code>uint8_t</code>).
<code>PMIX_STRING</code>	3	<code>NULL</code> terminated string (<code>char*</code>).
<code>PMIX_SIZE</code>	4	Size <code>size_t</code> .
<code>PMIX_PID</code>	5	Operating Process IDentifier (PID) (<code>pid_t</code>).
<code>PMIX_INT</code>	6	Integer (<code>int</code>).
<code>PMIX_INT8</code>	7	8-byte integer (<code>int8_t</code>).
<code>PMIX_INT16</code>	8	16-byte integer (<code>int16_t</code>).
<code>PMIX_INT32</code>	9	32-byte integer (<code>int32_t</code>).
<code>PMIX_INT64</code>	10	64-byte integer (<code>int64_t</code>).
<code>PMIX_UINT</code>	11	Unsigned integer (<code>unsigned int</code>).
<code>PMIX_UINT8</code>	12	Unsigned 8-byte integer (<code>uint8_t</code>).
<code>PMIX_UINT16</code>	13	Unsigned 16-byte integer (<code>uint16_t</code>).
<code>PMIX_UINT32</code>	14	Unsigned 32-byte integer (<code>uint32_t</code>).
<code>PMIX_UINT64</code>	15	Unsigned 64-byte integer (<code>uint64_t</code>).
<code>PMIX_FLOAT</code>	16	Float (<code>float</code>).
<code>PMIX_DOUBLE</code>	17	Double (<code>double</code>).
<code>PMIX_TIMEVAL</code>	18	Time value (<code>struct timeval</code>).
<code>PMIX_TIME</code>	19	Time (<code>time_t</code>).
<code>PMIX_STATUS</code>	20	Status code <code>pmix_status_t</code> .
<code>PMIX_VALUE</code>	21	Value (<code>pmix_value_t</code>).
<code>PMIX_PROC</code>	22	Process (<code>pmix_proc_t</code>).

1 **PMIX_APP** 23 Application context.
2 **PMIX_INFO** 24 Info object.
3 **PMIX_PDATA** 25 Pointer to data.
4 **PMIX_BYTE_OBJECT** 27 Byte object ([pmix_byte_object_t](#)).
5 **PMIX_KVAL** 28 Key/value pair.
6 **PMIX_PERSIST** 30 Persistence ([pmix_persistence_t](#)).
7 **PMIX_POINTER** 31 Pointer to an object (**void***).
8 **PMIX_SCOPE** 32 Scope ([pmix_scope_t](#)).
9 **PMIX_DATA_RANGE** 33 Range for data ([pmix_data_range_t](#)).
10 **PMIX_COMMAND** 34 PMIx command code (used internally).
11 **PMIX_INFO_DIRECTIVES** 35 Directives flag for [pmix_info_t](#) ([pmix_info_directives_t](#)).
12 **PMIX_DATA_TYPE** 36 Data type code ([pmix_data_type_t](#)).
13 **PMIX_PROC_STATE** 37 Process state ([pmix_proc_state_t](#)).
14 **PMIX_PROC_INFO** 38 Process information ([pmix_proc_info_t](#)).
15 **PMIX_DATA_ARRAY** 39 Data array ([pmix_data_array_t](#)).
16 **PMIX_PROC_RANK** 40 Process rank ([pmix_rank_t](#)).
17 **PMIX_PROC_NAMESPACE** 60 Process namespace ([pmix_namespace_t](#)). %
18 **PMIX_QUERY** 41 Query structure ([pmix_query_t](#)).
19 **PMIX_COMPRESSED_STRING** 42 String compressed with zlib (**char***).
20 **PMIX_COMPRESSED_BYTE_OBJECT** *(Provisional)* 59 Byte object whose bytes have been
21 compressed with zlib ([pmix_byte_object_t](#)).
22 **PMIX_ALLOC_DIRECTIVE** 43 Allocation directive ([pmix_alloc_directive_t](#)).
23 **PMIX_IOF_CHANNEL** 45 Input/output forwarding channel ([pmix_iof_channel_t](#)).
24 **PMIX_ENVAR** 46 Environmental variable structure ([pmix_envar_t](#)).
25 **PMIX_COORD** 47 Structure containing fabric coordinates ([pmix_coord_t](#)).
26 **PMIX_REGATTR** 48 Structure supporting attribute registrations ([pmix_regattr_t](#)).
27 **PMIX_REGEX** 49 Regular expressions - can be a valid NULL-terminated string or an arbitrary array of
28 bytes.
29 **PMIX_JOB_STATE** 50 Job state ([pmix_job_state_t](#)).
30 **PMIX_LINK_STATE** 51 Link state ([pmix_link_state_t](#)).
31 **PMIX_PROC_CPASET** 52 Structure containing the binding bitmap of a process ([pmix_cpuset_t](#)).
32 **PMIX_GEOMETRY** 53 Geometry structure containing the fabric coordinates of a specified
33 device. ([pmix_geometry_t](#)).
34 **PMIX_DEVICE_DIST** 54 Structure containing the minimum and maximum relative distance from the
35 caller to a given fabric device. ([pmix_device_distance_t](#)).
36 **PMIX_ENDPOINT** 55 Structure containing an assigned endpoint for a given fabric device.
37 ([pmix_endpoint_t](#)).
38 **PMIX_TOPO** 56 Structure containing the topology for a given node. ([pmix_topology_t](#)).
39 **PMIX_DEVTYPE** 57 Bitmask containing the types of devices being referenced.
40 ([pmix_device_type_t](#)).
41 **PMIX_LOCTYPE** 58 Bitmask describing the relative location of another process.
42 ([pmix_locality_t](#)).
43 **PMIX_STOR_MEDIUM** 66 Bitmask specifying different types of storage mediums.
44 ([pmix_storage_medium_t](#)).

1 **PMIX_STOR_ACCESS** 67 Bitmask specifying different levels of storage accessibility (i.e., from where
2 a storage system may be accessed). ([pmix_storage_accessibility_t](#)).
3 **PMIX_STOR_PERSIST** 68 Bitmask specifying different levels of persistence for a particular storage
4 system. ([pmix_storage_persistence_t](#)).
5 **PMIX_STOR_ACCESS_TYPE** 69 Bitmask specifying different storage system access types.
6 ([pmix_storage_access_type_t](#)).
7 **PMIX_DATA_TYPE_MAX** 500 A starting point for implementer-specific data types. Values above this
8 are guaranteed not to conflict with PMIx values. Definitions should always be based on the
9 [PMIX_DATA_TYPE_MAX](#) constant and not a specific value as the value of the constant may change.

10 3.4 General Callback Functions

11 PMIx provides blocking and nonblocking versions of most APIs. In the nonblocking versions, a callback is
12 activated upon completion of the the operation. This section describes many of those callbacks.

13 3.4.1 Release Callback Function

14 Summary

15 The [pmix_release_cbfunc_t](#) is used by the [pmix_modex_cbfunc_t](#) and
16 [pmix_info_cbfunc_t](#) operations to indicate that the callback data may be reclaimed/freed by the caller.

17 Format

PMIx v1.0

```
18 typedef void (*pmix_release_cbfunc_t)
19         (void *cbdata);
```

20 INOUT cbdata

21 Callback data passed to original API call (memory reference)

22 Description

23 Since the data is “owned” by the host server, provide a callback function to notify the host server that we are
24 done with the data so it can be released.

25 3.4.2 Op Callback Function

26 Summary

27 The [pmix_op_cbfunc_t](#) is used by operations that simply return a status.

PMIx v1.0

```
28 typedef void (*pmix_op_cbfunc_t)
29         (pmix_status_t status, void *cbdata);
```

30 IN status

31 Status associated with the operation (handle)

32 IN cbdata

33 Callback data passed to original API call (memory reference)

1
2
3
4

Description

Used by a wide range of PMIx API's including [PMIx_Fence_nb](#), [pmix_server_client_connected2_fn_t](#), [PMIx_server_register_nspace](#). This callback function is used to return a status to an often nonblocking operation.

3.4.3 Value Callback Function

Summary

The [pmix_value_cbfunc_t](#) is used by [PMIx_Get_nb](#) to return data.

PMIx v1.0

```
typedef void (*pmix_value_cbfunc_t)
    (pmix_status_t status,
     pmix_value_t *kv, void *cbdata);
```

- IN status**
Status associated with the operation (handle)
- IN kv**
Key/value pair representing the data ([pmix_value_t](#))
- IN cbdata**
Callback data passed to original API call (memory reference)

Description

A callback function for calls to [PMIx_Get_nb](#). The *status* indicates if the requested data was found or not. A pointer to the [pmix_value_t](#) structure containing the found data is returned. The pointer will be **NULL** if the requested data was not found.

3.4.4 Info Callback Function

Summary

The [pmix_info_cbfunc_t](#) is a general information callback used by various APIs.

PMIx v2.0

```
typedef void (*pmix_info_cbfunc_t)
    (pmix_status_t status,
     pmix_info_t info[], size_t ninfo,
     void *cbdata,
     pmix_release_cbfunc_t release_fn,
     void *release_cbdata);
```

- IN status**
Status associated with the operation ([pmix_status_t](#))
- IN info**
Array of [pmix_info_t](#) returned by the operation (pointer)
- IN ninfo**
Number of elements in the *info* array ([size_t](#))

- 1 **IN** **cbdata**
- 2 Callback data passed to original API call (memory reference)
- 3 **IN** **release_fn**
- 4 Function to be called when done with the *info* data (function pointer)
- 5 **IN** **release_cbdata**
- 6 Callback data to be passed to *release_fn* (memory reference)

7 **Description**

8 The *status* indicates if requested data was found or not. An array of `pmix_info_t` will contain the key/value
9 pairs.

10 **3.4.5 Handler registration callback function**

11 **Summary**

12 Callback function for calls to register handlers, e.g., event notification and IOF requests.

13 **Format**

PMIx v3.0

```

14    typedef void (*pmix_hdlr_reg_cbfunc_t)
15                (pmix_status_t status,
16                size_t refid,
17                void *cbdata);

```

- 18 **IN** **status**
- 19 **PMIX_SUCCESS** or an appropriate error constant (`pmix_status_t`)
- 20 **IN** **refid**
- 21 reference identifier assigned to the handler by PMIx, used to deregister the handler (`size_t`)
- 22 **IN** **cbdata**
- 23 object provided to the registration call (pointer)

24 **Description**

25 Callback function for calls to register handlers, e.g., event notification and IOF requests.

26 **3.5 PMIx Datatype Value String Representations**

27 Provide a string representation for several types of values. Note that the provided string is statically defined
28 and must NOT be `free`'d.

29 **Summary**

30 String representation of a `pmix_status_t`.

PMIx v1.0

```

31    const char*
32    PMIx_Error_string(pmix_status_t status);

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Summary

String representation of a [pmix_proc_state_t](#).

▼  

const char*

PMIx_Proc_state_string(pmix_proc_state_t state);

▲  

Summary

String representation of a [pmix_scope_t](#).

PMIx v2.0

▼  

const char*

PMIx_Scope_string(pmix_scope_t scope);

▲  

Summary

String representation of a [pmix_persistence_t](#).

PMIx v2.0

▼  

const char*

PMIx_Persistence_string(pmix_persistence_t persist);

▲  

Summary

String representation of a [pmix_data_range_t](#).

PMIx v2.0

▼  

const char*

PMIx_Data_range_string(pmix_data_range_t range);

▲  

Summary

String representation of a [pmix_info_directives_t](#).

PMIx v2.0

▼  

const char*

PMIx_Info_directives_string(pmix_info_directives_t directives);

▲  

Summary

String representation of a [pmix_data_type_t](#).

PMIx v2.0

▼  

const char*

PMIx_Data_type_string(pmix_data_type_t type);

▲  

1 **Summary**
2 String representation of a `pmix_alloc_directive_t`.

```
▼ _____ C _____ ▼  
3 const char*  
4 PMIx_Alloc_directive_string(pmix_alloc_directive_t directive);  
▲ _____ C _____ ▲
```

5 **Summary**
6 String representation of a `pmix_iof_channel_t`.

PMIx v3.0

```
▼ _____ C _____ ▼  
7 const char*  
8 PMIx_IOF_channel_string(pmix_iof_channel_t channel);  
▲ _____ C _____ ▲
```

9 **Summary**
10 String representation of a `pmix_job_state_t`.

PMIx v4.0

```
▼ _____ C _____ ▼  
11 const char*  
12 PMIx_Job_state_string(pmix_job_state_t state);  
▲ _____ C _____ ▲
```

13 **Summary**
14 String representation of a PMIx attribute.

PMIx v4.0

```
▼ _____ C _____ ▼  
15 const char*  
16 PMIx_Get_attribute_string(char *attributename);  
▲ _____ C _____ ▲
```

17 **Summary**
18 Return the PMIx attribute name corresponding to the given attribute string.

PMIx v4.0

```
▼ _____ C _____ ▼  
19 const char*  
20 PMIx_Get_attribute_name(char *attributestring);  
▲ _____ C _____ ▲
```

21 **Summary**
22 String representation of a `pmix_link_state_t`.

PMIx v4.0

```
▼ _____ C _____ ▼  
23 const char*  
24 PMIx_Link_state_string(pmix_link_state_t state);  
▲ _____ C _____ ▲
```


1
2

Summary

String representation of a `pmix_device_type_t`.



C

3
4

```
const char*
```

```
PMIx_Device_type_string(pmix_device_type_t type);
```



C

CHAPTER 4

Client Initialization and Finalization

1 The PMIx library is required to be initialized and finalized around the usage of most PMIx functions or
2 macros. The APIs that may be used outside of the initialized and finalized region are noted. All other APIs
3 must be used inside this region.

4 There are three sets of initialization and finalization functions depending upon the role of the process in the
5 PMIx Standard - those associated with the PMIx *client* are defined in this chapter. Similar functions
6 corresponding to the roles of *server* and *tool* are defined in Chapters 17 and 18, respectively.

7 Note that a process can only call *one* of the initialization/finalization functional pairs from the set of three -
8 e.g., a process that calls the client initialization function cannot also call the tool or server initialization
9 functions, and must call the corresponding client finalization function. Regardless of the role assumed by the
10 process, all processes have access to the client APIs. Thus, the *server* and *tool* roles can be considered
11 supersets of the PMIx client.

4.1 PMIx_Initialized

Summary

13 Determine if the PMIx library has been initialized. This function may be used outside of the initialized and
14 finalized region, and is usable by servers and tools in addition to clients.
15

Format

16 *PMIx v1.0*

17 `int PMIx_Initialized(void)`

18 A value of **1** (true) will be returned if the PMIx library has been initialized, and **0** (false) otherwise.

Rationale

19 The return value is an integer for historical reasons as that was the signature of prior PMI libraries.

Description

20 Check to see if the PMIx library has been initialized using any of the init functions: `PMIx_Init`,
21 `PMIx_server_init`, or `PMIx_tool_init`.
22

1 4.2 PMIx_Get_version

2 Summary

3 Get the PMIx version information. This function may be used outside of the initialized and finalized region,
4 and is usable by servers and tools in addition to clients.

5 Format

6 `const char* PMIx_Get_version(void)`

7 Description

8 Get the PMIx version string. Note that the provided string is statically defined and must *not* be free'd.

9 4.3 PMIx_Init

10 Summary

11 Initialize the PMIx client library

12 Format

PMIx v1.2

13 `pmix_status_t`
14 `PMIx_Init(pmix_proc_t *proc,`
15 `pmix_info_t info[], size_t ninfo)`

16 INOUT `proc`

17 `proc` structure (handle)

18 IN `info`

19 Array of `pmix_info_t` structures (array of handles)

20 IN `ninfo`

21 Number of elements in the `info` array (`size_t`)

22 Returns `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant.

23 **Optional Attributes**

The following attributes are optional for implementers of PMIx libraries:

24 **`PMIX_USOCK_DISABLE`** "`pmix.usock.disable`" (`bool`)

25 Disable legacy UNIX socket (usock) support. If the library supports Unix socket connections, this
26 attribute may be supported for disabling it.

27 **`PMIX_SOCKET_MODE`** "`pmix.sockmode`" (`uint32_t`)

28 POSIX `mode_t` (9 bits valid). If the library supports socket connections, this attribute may be
29 supported for setting the socket mode.

30 **`PMIX_SINGLE_LISTENER`** "`pmix.sing.listnr`" (`bool`)

1 Use only one rendezvous socket, letting priorities and/or environment parameters select the active
2 transport. If the library supports multiple methods for clients to connect to servers, this attribute may
3 be supported for disabling all but one of them.

4 **PMIX_TCP_REPORT_URI** "pmix.tcp.repuri" (char*)
5 If provided, directs that the TCP Uniform Resource Identifier (URI) be reported and indicates the
6 desired method of reporting: '-' for stdout, '+' for stderr, or filename. If the library supports TCP
7 socket connections, this attribute may be supported for reporting the URI.

8 **PMIX_TCP_IF_INCLUDE** "pmix.tcp.ifinclude" (char*)
9 Comma-delimited list of devices and/or Classless Inter-Domain Routing (CIDR) notation to include
10 when establishing the TCP connection. If the library supports TCP socket connections, this attribute
11 may be supported for specifying the interfaces to be used.

12 **PMIX_TCP_IF_EXCLUDE** "pmix.tcp.ifexclude" (char*)
13 Comma-delimited list of devices and/or CIDR notation to exclude when establishing the TCP
14 connection. If the library supports TCP socket connections, this attribute may be supported for
15 specifying the interfaces that are *not* to be used.

16 **PMIX_TCP_IPV4_PORT** "pmix.tcp.ipv4" (int)
17 The IPv4 port to be used.. If the library supports IPV4 connections, this attribute may be supported
18 for specifying the port to be used.

19 **PMIX_TCP_IPV6_PORT** "pmix.tcp.ipv6" (int)
20 The IPv6 port to be used. If the library supports IPV6 connections, this attribute may be supported
21 for specifying the port to be used.

22 **PMIX_TCP_DISABLE_IPV4** "pmix.tcp.disipv4" (bool)
23 Set to **true** to disable IPv4 family of addresses. If the library supports IPV4 connections, this
24 attribute may be supported for disabling it.

25 **PMIX_TCP_DISABLE_IPV6** "pmix.tcp.disipv6" (bool)
26 Set to **true** to disable IPv6 family of addresses. If the library supports IPV6 connections, this
27 attribute may be supported for disabling it.

28 **PMIX_EXTERNAL_PROGRESS** "pmix.evext" (bool)
29 The host shall progress the PMIx library via calls to **PMIx_Progress**

30 **PMIX_EVENT_BASE** "pmix.evbase" (void*)
31 Pointer to an **event_base** to use in place of the internal progress thread. All PMIx library events are
32 to be assigned to the provided event base. The event base *must* be compatible with the event library
33 used by the PMIx implementation - e.g., either both the host and PMIx library must use libevent, or
34 both must use libev. Cross-matches are unlikely to work and should be avoided - it is the responsibility
35 of the host to ensure that the PMIx implementation supports (and was built with) the appropriate event
36 library.

37 If provided, the following attributes are used by the event notification system for inter-library coordination:

38 **PMIX_PROGRAMMING_MODEL** "pmix.pgm.model" (char*)
39 Programming model being initialized (e.g., "MPI" or "OpenMP").

40 **PMIX_MODEL_LIBRARY_NAME** "pmix.mdl.name" (char*)

1 Programming model implementation ID (e.g., "OpenMPI" or "MPICH").

2 **PMIX_MODEL_LIBRARY_VERSION** "pmix.mdl.vrs" (char*)

3 Programming model version string (e.g., "2.1.1").

4 **PMIX_THREADING_MODEL** "pmix.threads" (char*)

5 Threading model used (e.g., "pthreads").

6 **PMIX_MODEL_NUM_THREADS** "pmix.mdl.nthrds" (uint64_t)

7 Number of active threads being used by the model.

8 **PMIX_MODEL_NUM_CPUS** "pmix.mdl.ncpu" (uint64_t)

9 Number of cpus being used by the model.

10 **PMIX_MODEL_CPU_TYPE** "pmix.mdl.cputype" (char*)

11 Granularity - "hwthread", "core", etc.

12 **PMIX_MODEL_AFFINITY_POLICY** "pmix.mdl.tap" (char*)

13 Thread affinity policy - e.g.: "master" (thread co-located with master thread), "close" (thread located
14 on cpu close to master thread), "spread" (threads load-balanced across available cpus).



15 Description

16 Initialize the PMIx client, returning the process identifier assigned to this client's application in the provided
17 **pmix_proc_t** struct. Passing a value of **NULL** for this parameter is allowed if the user wishes solely to
18 initialize the PMIx system and does not require return of the identifier at that time.

19 When called, the PMIx client shall check for the required connection information of the local PMIx server and
20 establish the connection. If the information is not found, or the server connection fails, then an appropriate
21 error constant shall be returned.

22 If successful, the function shall return **PMIX_SUCCESS** and fill the *proc* structure (if provided) with the
23 server-assigned namespace and rank of the process within the application. In addition, all startup information
24 provided by the resource manager shall be made available to the client process via subsequent calls to
25 **PMIx_Get**.

26 The PMIx client library shall be reference counted, and so multiple calls to **PMIx_Init** are allowed by the
27 standard. Thus, one way for an application process to obtain its namespace and rank is to simply call
28 **PMIx_Init** with a non-NULL *proc* parameter. Note that each call to **PMIx_Init** must be balanced with a
29 call to **PMIx_Finalize** to maintain the reference count.

30 Each call to **PMIx_Init** may contain an array of **pmix_info_t** structures passing directives to the PMIx
31 client library as per the above attributes.

32 Multiple calls to **PMIx_Init** shall not include conflicting directives. The **PMIx_Init** function will return
33 an error when directives that conflict with prior directives are encountered.

1 4.3.1 Initialization events

2 The following events are typically associated with calls to `PMIx_Init`:

3 `PMIX_MODEL_DECLARED` -147 Model declared.
4 `PMIX_MODEL_RESOURCES` -151 Resource usage by a programming model has changed.
5 `PMIX_OPENMP_PARALLEL_ENTERED` -152 An OpenMP parallel code region has been entered.
6 `PMIX_OPENMP_PARALLEL_EXITED` -153 An OpenMP parallel code region has completed.

7 4.3.2 Initialization attributes

8 The following attributes influence the behavior of `PMIx_Init`.

9 4.3.2.1 Connection attributes

10 These attributes are used to describe a TCP socket for rendezvous with the local RM by passing them into the
11 relevant initialization API - thus, they are not typically accessed via the `PMIx_Get` API.

12 `PMIX_TCP_REPORT_URI` "pmix.tcp.repuri" (char*)
13 If provided, directs that the TCP URI be reported and indicates the desired method of reporting: '-'
14 for stdout, '+' for stderr, or filename.
15 `PMIX_TCP_URI` "pmix.tcp.uri" (char*)
16 The URI of the PMIx server to connect to, or a file name containing it in the form of `file:<name`
17 `of file containing it>`.
18 `PMIX_TCP_IF_INCLUDE` "pmix.tcp.ifinclude" (char*)
19 Comma-delimited list of devices and/or CIDR notation to include when establishing the TCP
20 connection.
21 `PMIX_TCP_IF_EXCLUDE` "pmix.tcp.ifexclude" (char*)
22 Comma-delimited list of devices and/or CIDR notation to exclude when establishing the TCP
23 connection.
24 `PMIX_TCP_IPV4_PORT` "pmix.tcp.ipv4" (int)
25 The IPv4 port to be used..
26 `PMIX_TCP_IPV6_PORT` "pmix.tcp.ipv6" (int)
27 The IPv6 port to be used.
28 `PMIX_TCP_DISABLE_IPV4` "pmix.tcp.disipv4" (bool)
29 Set to `true` to disable IPv4 family of addresses.
30 `PMIX_TCP_DISABLE_IPV6` "pmix.tcp.disipv6" (bool)
31 Set to `true` to disable IPv6 family of addresses.

32 4.3.2.2 Programming model attributes

33 These attributes are associated with programming models.

34 `PMIX_PROGRAMMING_MODEL` "pmix.pgm.model" (char*)
35 Programming model being initialized (e.g., "MPI" or "OpenMP").
36 `PMIX_MODEL_LIBRARY_NAME` "pmix.mdl.name" (char*)
37 Programming model implementation ID (e.g., "OpenMPI" or "MPICH").
38 `PMIX_MODEL_LIBRARY_VERSION` "pmix.mld.vrs" (char*)
39 Programming model version string (e.g., "2.1.1").
40 `PMIX_THREADING_MODEL` "pmix.threads" (char*)

1 Threading model used (e.g., “pthreads”).

2 **PMIX_MODEL_NUM_THREADS** "pmix.mdl.nthrds" (uint64_t)

3 Number of active threads being used by the model.

4 **PMIX_MODEL_NUM_CPUS** "pmix.mdl.ncpu" (uint64_t)

5 Number of cpus being used by the model.

6 **PMIX_MODEL_CPU_TYPE** "pmix.mdl.cputype" (char*)

7 Granularity - “hwthread”, “core”, etc.

8 **PMIX_MODEL_PHASE_NAME** "pmix.mdl.phase" (char*)

9 User-assigned name for a phase in the application execution (e.g., “cfd reduction”).

10 **PMIX_MODEL_PHASE_TYPE** "pmix.mdl.ptype" (char*)

11 Type of phase being executed (e.g., “matrix multiply”).

12 **PMIX_MODEL_AFFINITY_POLICY** "pmix.mdl.tap" (char*)

13 Thread affinity policy - e.g.: "master" (thread co-located with master thread), "close" (thread located

14 on cpu close to master thread), "spread" (threads load-balanced across available cpus).

15 4.4 PMIx_Finalize

16 Summary

17 Finalize the PMIx client library.

18 Format

PMIx v1.0

19 **pmix_status_t**

20 **PMIx_Finalize**(const pmix_info_t info[], size_t ninfo)

21 **IN info**

22 Array of **pmix_info_t** structures (array of handles)

23 **IN ninfo**

24 Number of elements in the *info* array (**size_t**)

25 Returns **PMIX_SUCCESS** or a negative value indicating the error.

Optional Attributes

26 The following attributes are optional for implementers of PMIx libraries:

27 **PMIX_EMBED_BARRIER** "pmix.embed.barrier" (bool)

28 Execute a blocking fence operation before executing the specified operation. **PMIx_Finalize** does

29 not include an internal barrier operation by default. This attribute directs **PMIx_Finalize** to

30 execute a barrier as part of the finalize operation.

31 Description

32 Decrement the PMIx client library reference count. When the reference count reaches zero, the library will

33 finalize the PMIx client, closing the connection with the local PMIx server and releasing all internally

34 allocated memory.

1 **4.4.1 Finalize attributes**

2 The following attribute influences the behavior of `PMIx_Finalize`.

3 `PMIX_EMBED_BARRIER` "pmix.embed.barrier" (bool)

4 Execute a blocking fence operation before executing the specified operation. `PMIx_Finalize` does
5 not include an internal barrier operation by default. This attribute directs `PMIx_Finalize` to
6 execute a barrier as part of the finalize operation.

7 **4.5 PMIx_Progress**

8 **Summary**

9 Progress the PMIx library.

10 *PMIx v4.0* **Format**



11 `void`

12 `PMIx_Progress(void)`



13 **Description**

14 Progress the PMIx library. Note that special care must be taken to avoid deadlocking in PMIx callback
15 functions and APIs.

CHAPTER 5

Data Access and Sharing

1 In PMIx key-value pairs are the primary way that information is shared between processes in the PMIx
2 universe. A key-value pair consists of a string key and a value data structure which includes the type of the
3 value's data. Each key is unique within the context or domain in which it exists. The context or domain of a
4 key in PMIx is called a data realm. This chapter describes how key-value pairs are made available and
5 accessed by clients.

6 A key-value pair has a scope defined by its data realm. PMIx defines data realms for sessions, jobs,
7 applications, nodes, processes and a special realm for publish/lookup information. These are explained more
8 fully in chapters 6 and 9.

9 The source or provider of key-value pairs can be the host environment or a PMIx client. The host environment
10 can provide key-values pairs in any realm except for the publish/lookup realm. PMIx clients can only provide
11 key-values pairs within their own process realm and the publish/lookup realm. The value of a key is not
12 necessarily static and can be modified by the original provider of the key.

13 PMIx classifies keys as either reserved or non-reserved. Only keys provided by a host environment or PMIx
14 implementation can begin with "**pmix**" and are called reserved keys. Host environments and PMIx
15 implementations must only provide keys that begin with "**pmix**" to avoid conflicts with keys that are provided
16 by clients. Reserved keys are documented in Chapter 6. Each standardized reserved key has an associated
17 macro that is documented along with its description. Likewise clients can only provide keys that do not begin
18 with "**pmix**" to avoid conflicting with keys provided by the host environment or PMIx implementation. The
19 string representation and the macro are equivalent methods to refer to a key. Throughout this document, the
20 macro is used rather than its string value to refer to reserved keys. Clients cannot create reserved keys, nor can
21 they modify the value of reserved keys. The way reserved keys are accessed by clients is different than for
22 non-reserved keys as explained in Chapter 6.

23 Host environments are not required to provide all keys that are documented in the standard. However, when
24 providing keys that are documented, they must adhere to the behavior defined by the standard for reserved
25 keys. An implementation cannot provide a reserved key that behaves differently than what is in the standard.
26 Implementations are encouraged to standardize any non-standardized, reserved keys that they provide to avoid
27 conflicting with other implementations or efforts to standardize the same key.

28 Host environments are permitted to provide non-reserved keys and non-standardized reserved keys. However
29 there are important caveats to providing either of these categories of keys that are explained in Section 5.1 and
30 Chapter 6. Host environments should avoid providing reserved keys which are not standardized. If the
31 reserved key is later standardized to behave in a way that is different from the behavior provided by the host
32 environment, the host environment will not be compliant with the PMIx standard.

33 Reserved keys are set by the host environment and are made available at client initialization. Each key is
34 associated with a data realm and the retrieval API call will indicate the realm being queried. Reserved keys are
35 used to access information about a client's execution context. After initialization, the client can access
36 reserved keys using the **PMIx_Get** API. Common information that a client may access includes, but is not

1 limited to, the name of the host it is executing on, the number of peer processes in its namespace, the number
2 of peer processes in its application, and its job rank.

3 Non-reserved keys are provided by clients. A call to **PMIx_Put** will result in a key-value being associated
4 with the calling client within its process data realm. There is no mechanism to retract a key-value that has been
5 made available by a client through **PMIx_Put**. The key can be assigned a new value, but cannot be removed
6 entirely. Although key values can originate from different sources, they are always retrieved using the
7 **PMIx_Get** and **PMIx_Get_nb** API.

8 The publish/lookup data realm is accessed through a separate set of APIs. **PMIx_Put** cannot add or modify
9 key-values within the publish/lookup realm and **PMIx_Lookup** cannot access key-values outside the
10 publish/lookup realm. This data realm is described in detail in chapter 9. Although **PMIx_Publish** and
11 **PMIx_Lookup** are analogous to **PMIx_Put** and **PMIx_Get** in that both pairs of APIs provide functionality
12 for exposing and retrieving key-values, the semantics vary significantly. For example, **PMIx_Lookup**
13 includes the ability to wait until a key is defined before returning. **PMIx_Publish** can restrict the scope of
14 who can access data to its own namespace, the host environment, the session it is executing in, all processes or
15 even custom scopes, while **PMIx_Put** can restrict the scope of who can access data to the node it is executing
16 on, remote nodes, all nodes or only itself. The publish/lookup data realm is useful for advertising information
17 that is not necessarily specific to one process to other processes in the PMIx universe. The process accessing
18 this information does *not* need to know the identity of the process that provided the data.

19 PMIx does not provide a mechanism to asynchronously notify a process about the availability of key-value
20 information once it is made available by another process. However, the nonblocking accessor interfaces (e.g.,
21 **PMIx_Get_nb**, **PMIx_Lookup_nb**) may provide a degree of asynchronous notification on information
22 availability.

23 Process related key-value exchanges allow a PMIx process to share information specific to itself, and access
24 information specific to one or more processes in the PMIx universe. These interactions occur within the
25 process data realm. The 'put/commit/get' exchange pattern is often used to exchange process related
26 information. Optionally, a 'put/commit/fence/get' exchange pattern adds the 'fence' synchronization (and
27 possible collective exchange) for applications that desire it. Commonly, these exchange patterns are used in a
28 *business card exchange* (a.k.a. *modex exchange*) where one PMIx client shares its connectivity information,
29 then other PMIx clients access that information to establish a connection with that client. In some
30 environments that support "instant-on" all connectivity information for PMIx clients is stored in the job-level
31 information at process creation time and is accessible to the clients without the need to perform any additional
32 key-value exchange.

33 Keys can exist in multiple data realms, possibly with different values. PMIx clients can access available
34 information associated with any of the data realms. For example, a client can access the number of nodes
35 (**PMIX_NUM_NODES**) used by a session, job, or application. Rather than having three different attributes, a
36 single attribute is used but with the data realm context of the query specified as additional attributes. Examples
37 of these access patterns are presented in Section 6.2.1.

38 PMIx clients can share key-value pairs associated with themselves by using the **PMIx_Put** function. The
39 **PMIx_Put** function automatically associates the key-value pair with the calling process, thus making it
40 specific to that process. A client may call **PMIx_Put** as many times as necessary and the data is not available
41 to other processes until explicitly committed. A client must call **PMIx_Commit** to make accessible all
42 key-value pairs previously put by this process to all other processes in the PMIx universe. This put and commit
43 pattern provides implementors the opportunity to make individual **PMIx_Put** calls efficient local operations,
44 and then make the whole set of key-value pairs accessible in a single step.

1 PMIx clients can access the key-value pairs associated with any process data realm in the PMIx universe
2 (including the calling process) by passing the specific process name of the target process to the `PMIx_Get`
3 and `PMIx_Get_nb` functions. The PMIx server local to the calling process will retrieve that key-value pair
4 from the PMIx server associated with the target process. Clients can also access session, job, application,
5 node, and namespace level information by using the `PMIx_Get` and `PMIx_Get_nb` functions as shown in
6 Section 6.2.1. The completion semantics for `PMIx_Get` and `PMIx_Get_nb` differ depending on the type of
7 key and its availability. See Sections 5.3 and 6.2. For example, if a non-reserved key is not available, the
8 `PMIx_Get` or `PMIx_Get_nb` call will not complete, by default, until that key-value pair becomes available.

9 5.1 Non-reserved keys

10 *Non-reserved keys* are keys whose string representation begin with a prefix other than "`pmix`". Such keys are
11 defined by a client when information needs to be exchanged between processes, for example, where connection
12 information is required and the host environment does not support the *instant on* option or where the host
13 environment does not provide a required piece of data. Other than the prefix, there are no restrictions on the
14 use or content of non-reserved keys.

15 PMIx provides support for two methods of exchanging non-reserved keys:

- 16 • Global, collective exchange of the information prior to retrieval. This is accomplished by executing a
17 barrier operation that includes collection and exchange of the data provided by each process such that each
18 process has access to the full set of data from all participants once the operation has completed. PMIx
19 provides the `PMIx_Fence` function (or its non-blocking equivalent) for this purpose, accompanied by the
20 `PMIX_COLLECT_DATA` qualifier.
- 21 • Direct, on-demand retrieval of the information. No barrier or global exchange is conducted in this case.
22 Instead, information is retrieved from the host where that process is executing upon request - i.e., a call to
23 `PMIx_Get` results in a data exchange with the PMIx server on the remote host. Various caching strategies
24 may be employed by the host environment and/or PMIx implementation to reduce the number of retrievals.
25 Note that this method requires that the host environment both know the location of the posting process and
26 support direct information retrieval.

27 Both of the above methods are based on retrieval from a specific process - i.e., the *proc* argument to
28 `PMIx_Get` must include both the namespace and the rank of the process that posted the information.
29 However, in some cases, non-reserved keys are provided on a globally unique basis and the retrieving process
30 has no knowledge of the identity of the process posting the key. This is typically found in legacy applications
31 (where the originating process identifier is often embedded in the key itself) and in unstructured applications
32 that lack rank-related behavior. In these cases, the key remains associated with the namespace of the process
33 that posted it, but is retrieved by use of the `PMIX_RANK_UNDEF` rank. In addition, the keys must be globally
34 exchanged prior to retrieval as there is no way for the host to otherwise locate the source for the information.

35 Note that the retrieval rules for non-reserved keys (detailed in Section 5.3) differ significantly from those used
36 for reserved keys.

37 5.2 Posting Key/Value Pairs

38 PMIx clients can post non-reserved key-value pairs associated with themselves by using `PMIx_Put`.
39 Alternatively, PMIx clients can cache both reserved and non-reserved key-value pairs accessible only by the
40 caller via the `PMIx_Store_internal` API.

1 5.2.1 PMIx_Put

2 Summary

3 Stage a key/value pair in preparation for being made accessible to processes.

4 *PMIx v1.0* Format

C

```
5 pmix_status_t  
6 PMIx_Put(pmix_scope_t scope,  
7         const pmix_key_t key,  
8         pmix_value_t *val);
```

C

9 **IN scope**
10 Distribution scope of the provided value (handle)
11 **IN key**
12 key ([pmix_key_t](#))
13 **IN value**
14 Reference to a [pmix_value_t](#) structure (handle)

15 Returns PMIX_SUCCESS or one of the following error codes when the condition described occurs:

- 16 • [PMIX_ERR_BAD_PARAM](#) indicating a reserved key is provided in the *key* argument.

17 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
18 error code or an implementation defined error code as described in Section 3.1.1.

19 Description

20 Post a key-value pair for distribution. Depending upon the PMIx implementation, the posted value may be
21 locally cached in the client’s PMIx library until [PMIx_Commit](#) is called.

22 The provided *scope* determines the ability of other processes to access the posted data, as defined in
23 Section 5.2.1.1 on page 72. Specific implementations may support different scope values, but all
24 implementations must support at least [PMIX_GLOBAL](#).

25 The [pmix_value_t](#) structure supports both string and binary values. PMIx implementations are required to
26 support heterogeneous environments by properly converting binary values between host architectures, and will
27 copy the provided *value* into internal memory prior to returning from [PMIx_Put](#).

Advice to users

28 Note that keys starting with a string of “[pmix](#)” must not be used in calls to [PMIx_Put](#). Thus, applications
29 should never use a defined “PMIX” attribute as the key in a call to [PMIx_Put](#).

1 5.2.1.1 Scope of Put Data

2 The `pmix_scope_t` structure is a `uint8_t` type that defines the availability of data passed to `PMIx_Put`.
3 The following constants can be used to set a variable of the type `pmix_scope_t`. All definitions were
4 introduced in version 1 of the standard unless otherwise marked.

5 Specific implementations may support different scope values, but all implementations must support at least
6 `PMIX_GLOBAL`. If a specified scope value is not supported, then the `PMIx_Put` call must return
7 `PMIX_ERR_NOT_SUPPORTED`.

8 `PMIX_SCOPE_UNDEF` 0 Undefined scope.

9 `PMIX_LOCAL` 1 The data is intended only for other application processes on the same node. Data
10 marked in this way will not be included in data packages sent to remote requesters - i.e., it is only
11 available to processes on the local node.

12 `PMIX_REMOTE` 2 The data is intended solely for applications processes on remote nodes. Data marked
13 in this way will not be shared with other processes on the same node - i.e., it is only available to
14 processes on remote nodes.

15 *PMIx v2.0* `PMIX_GLOBAL` 3 The data is to be shared with all other requesting processes, regardless of location.

16 `PMIX_INTERNAL` 4 The data is intended solely for this process and is not shared with other processes.

17 5.2.2 `PMIx_Store_internal`

18 Summary

19 Store some data locally for retrieval by other areas of the process.

20 *PMIx v1.0* Format

```
21 pmix_status_t  
22 PMIx_Store_internal(const pmix_proc_t *proc,  
23                    const pmix_key_t key,  
24                    pmix_value_t *val);
```

25 **IN** `proc`
26 process reference (handle)

27 **IN** `key`
28 key to retrieve (string)

29 **IN** `val`
30 Value to store (handle)

31 Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

- 32 • `PMIX_ERR_BAD_PARAM` indicating a reserved key is provided in the `key` argument.

33 If none of the above return codes are appropriate, then an implementation must return either a general `PMIx`
34 error code or an implementation defined error code as described in Section 3.1.1.

35 Description

36 Store some data locally for retrieval by other areas of the process. This is data that has only internal scope - it
37 will never be posted externally. Typically used to cache data obtained by means outside of `PMIx` so that it can
38 be accessed by various areas of the process.

1 5.2.3 PMIx_Commit

2 Summary

3 Make available to other processes all key-value pairs previously staged via [PMIx_Put](#).

4 *PMIx v1.0* Format

5 `pmix_status_t PMIx_Commit(void);`

6 Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

7 Description

8 PMIx implementations may choose to locally cache non-reserved keys prior to submitting them for
9 distribution. [PMIx_Commit](#) initiates the operation of making available previously staged key-value pairs to
10 other processes. Depending on the implementation, this may involve transmitting the entire collection of data
11 posted by the process to a server. [PMIx_Commit](#) is an asynchronous operation that will immediately return
12 to the caller while the data is staged in the background.

Advice to users

13 Users are advised to always include the call to [PMIx_Commit](#) in case the local implementation requires it.
14 Note that posted data will not necessarily be circulated during [PMIx_Commit](#). Availability of the data by
15 other processes upon completion of [PMIx_Commit](#) therefore still relies upon the exchange mechanisms
16 described at the beginning of this chapter.

17 5.3 Retrieval rules for non-reserved keys

18 Since non-reserved keys cannot, by definition, have been provided by the host environment, their retrieval
19 follows significantly different rules than those defined for reserved keys (as detailed in Section 6). [PMIx_Get](#)
20 for a non-reserved key will obey the following precedence search:

- 21 1. If the [PMIX_GET_REFRESH_CACHE](#) attribute is given, then the request is first forwarded to the local
22 PMIx server which will then update the client's cache. Note that this may not, depending upon
23 implementation details, result in any action.
- 24 2. Check the local PMIx client cache for the requested key - if not found and either the [PMIX_OPTIONAL](#) or
25 [PMIX_GET_REFRESH_CACHE](#) attribute was given, the search will stop at this point and return the
26 [PMIX_ERR_NOT_FOUND](#) status.
- 27 3. Request the information from the local PMIx server. The server will check its cache for the specified key. If
28 the value still isn't found and the [PMIX_IMMEDIATE](#) attribute was given, then the library shall return the
29 [PMIX_ERR_NOT_FOUND](#) error constant to the requester. Otherwise, the PMIx server library will take one
30 of the following actions:

- If the target process has a rank of `PMIX_RANK_UNDEF`, then this indicates that the key being requested is globally unique and *not* associated with a specific process. In this case, the server shall hold the request until either the data appears at the server or, if given, the `PMIX_TIMEOUT` is reached. In the latter case, the server will return the `PMIX_ERR_TIMEOUT` status. Note that the server may, depending on PMIx implementation, never respond if the caller failed to specify a `PMIX_TIMEOUT` and the requested key fails to arrive at the server.
- If the target process is *local* (i.e., attached to the same PMIx server), then the server will hold the request until either the target process provides the data or, if given, the `PMIX_TIMEOUT` is reached. In the latter case, the server will return the `PMIX_ERR_TIMEOUT` status. Note that data which is posted via `PMIx_Put` but not staged with `PMIx_Commit` may, depending upon implementation, never appear at the server.
- If the target process is *remote* (i.e., not attached to the same PMIx server), the server will either:
 - If the host has provided the `pmix_server_dmodex_req_fn_t` module function interface, then the server shall pass the request to its host for servicing. The host is responsible for determining the location of the target process and passing the request to the PMIx server at that location.

When the remote data request is received, the target PMIx server will check its cache for the specified key. If the key is not present, the request shall be held until either the target process provides the data or, if given, the `PMIX_TIMEOUT` is reached. In the latter case, the server will return the `PMIX_ERR_TIMEOUT` status. The host shall convey the result back to the originating PMIx server, which will reply to the requesting client with the result of the request when the host provides it.

Note that the target server may, depending on PMIx implementation, never respond if the caller failed to specify a `PMIX_TIMEOUT` and the target process fails to post the requested key.
 - if the host does not support the `pmix_server_dmodex_req_fn_t` interface, then the server will immediately respond to the client with the `PMIX_ERR_NOT_FOUND` status

Advice to PMIx library implementers

While there is no requirement that all PMIx implementations follow the client-server paradigm used in the above description, implementers are required to provide behaviors consistent with the described search pattern.

Advice to users

Users are advised to always specify the `PMIX_TIMEOUT` value when retrieving non-reserved keys to avoid potential deadlocks should the specified key not become available.

5.4 PMIx_Get

Summary

Retrieve a key/value pair from the client's namespace.

Format

C

```
pmix_status_t
PMIx_Get(const pmix_proc_t *proc, const pmix_key_t key,
         const pmix_info_t info[], size_t ninfo,
         pmix_value_t **val);
```

C

IN **proc**
Process identifier - a **NULL** value may be used in place of the caller's ID (handle)

IN **key**
Key to retrieve ([pmix_key_t](#))

IN **info**
Array of info structures (array of handles)

IN **ninfo**
Number of elements in the *info* array (integer)

OUT **val**
value (handle)

A successful return indicates that the requested data has been returned in the manner requested (e.g., in a provided static memory location).

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_ERR_BAD_PARAM** A bad parameter was passed to the function call - e.g., the request included the **PMIX_GET_STATIC_VALUES** directive, but the provided storage location was **NULL**
- **PMIX_ERR_EXISTS_OUTSIDE_SCOPE** The requested key exists, but was posted in a *scope* (see Section 5.2.1.1) that does not include the requester.
- **PMIX_ERR_NOT_FOUND** The requested data was not available.

If none of the above return codes are appropriate, then an implementation must return either a general **PMIX** error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

The following attributes are required to be supported by all **PMIX** libraries:

PMIX_OPTIONAL "pmix.optional" (**bool**)

Look only in the client's local data store for the requested value - do not request data from the **PMIX** server if not found.

PMIX_IMMEDIATE "pmix.immediate" (**bool**)

Specified operation should immediately return an error from the **PMIX** server if the requested data cannot be found - do not request it from the host **RM**.

PMIX_DATA_SCOPE "pmix.scope" (**pmix_scope_t**)

Scope of the data to be searched in a **PMIx_Get** call.

PMIX_SESSION_INFO "pmix.ssn.info" (**bool**)

Return information regarding the session realm of the target process.

1 **PMIX_JOB_INFO** "pmix.job.info" (bool)
2 Return information regarding the job realm corresponding to the namespace in the target process'
3 identifier.

4 **PMIX_APP_INFO** "pmix.app.info" (bool)
5 Return information regarding the application realm to which the target process belongs - the namespace
6 of the target process serves to identify the job containing the target application. If information about an
7 application other than the one containing the target process is desired, then the attribute array must
8 contain a **PMIX_APPNUM** attribute identifying the desired target application. This is useful in cases
9 where there are multiple applications and the mapping of processes to applications is unclear.

10 **PMIX_NODE_INFO** "pmix.node.info" (bool)
11 Return information from the node realm regarding the node upon which the specified process is
12 executing. If information about a node other than the one containing the specified process is desired,
13 then the attribute array must also contain either the **PMIX_NODEID** or **PMIX_HOSTNAME** attribute
14 identifying the desired target. This is useful for requesting information about a specific node even if the
15 identity of processes running on that node are not known.

16 **PMIX_GET_STATIC_VALUES** "pmix.get.static" (bool)
17 Request that the data be returned in the provided storage location. The caller is responsible for
18 destructing the **pmix_value_t** using the **PMIX_VALUE_DESTRUCT** macro when done.

19 **PMIX_GET_POINTER_VALUES** "pmix.get.pntrs" (bool)
20 Request that any pointers in the returned value point directly to values in the key-value store. The user
21 *must not* release any returned data pointers.

22 **PMIX_GET_REFRESH_CACHE** "pmix.get.refresh" (bool)
23 When retrieving data for a remote process, refresh the existing local data cache for the process in case
24 new values have been put and committed by the process since the last refresh. Local process
25 information is assumed to be automatically updated upon posting by the process. A **NULL** key will
26 cause all values associated with the process to be refreshed - otherwise, only the indicated key will be
27 updated. A process rank of **PMIX_RANK_WILDCARD** can be used to update job-related information in
28 dynamic environments. The user is responsible for subsequently updating refreshed values they may
29 have cached in their own local memory.

▲-----▲
▼-----▼ **Optional Attributes** -----▼

30 The following attributes are optional for host environments:

31 **PMIX_TIMEOUT** "pmix.timeout" (int)
32 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
33 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
34 (client, server, and host) simultaneously timing the operation.

▲-----▲

Description

Retrieve information for the specified *key* associated with the process identified in the given `pmix_proc_t`. See Chapters 6 and 5.1 for details on rules governing retrieval of information. Information will be returned according to provided directives:

- In the absence of any directive, the returned `pmix_value_t` shall be an allocated memory object. The caller is responsible for releasing the object when done.
- If `PMIX_GET_POINTER_VALUES` is given, then the function shall return a pointer to a `pmix_value_t` in the PMIx library's memory that contains the requested information.
- If `PMIX_GET_STATIC_VALUES` is given, then the function shall return the information in the provided `pmix_value_t` pointer. In this case, the caller must provide storage for the structure and pass the pointer to that storage in the *val* parameter. If the implementation cannot return a static value, then the call to `PMIx_Get` must return the `PMIX_ERR_NOT_SUPPORTED` status.

Retrieve information for the specified *key* associated with the process identified in the given `pmix_proc_t`. See Chapters 6 and 5.1 for details on rules governing retrieval of information. Information will be returned according to provided directives:

- In the absence of any directive, the returned `pmix_value_t` shall be an allocated memory object. The caller is responsible for releasing the object when done.
- If `PMIX_GET_POINTER_VALUES` is given, then the function shall return a pointer to a `pmix_value_t` in the PMIx library's memory that contains the requested information.
- If `PMIX_GET_STATIC_VALUES` is given, then the function shall return the information in the provided `pmix_value_t` pointer. In this case, the caller must provide storage for the structure and pass the pointer to that storage in the *val* parameter.

This is a blocking operation - the caller will block until the retrieval rules of Section 6.2 or 5.3 are met.

The *info* array is used to pass user directives regarding the get operation.

5.4.1 PMIx_Get_nb

Summary

Nonblocking `PMIx_Get` operation.

Format

PMIx v1.0

```
pmix_status_t
PMIx_Get_nb(const pmix_proc_t *proc, const char key[],
            const pmix_info_t info[], size_t ninfo,
            pmix_value_cbfunc_t cbfunc, void *cbdata);
```

IN `proc`
Process identifier - a `NULL` value may be used in place of the caller's ID (handle)

IN `key`
Key to retrieve (string)

1 **IN info**
 2 Array of info structures (array of handles)
 3 **IN ninfo**
 4 Number of elements in the *info* array (integer)
 5 **IN cbfunc**
 6 Callback function (function reference)
 7 **IN cbdata**
 8 Data to be passed to the callback function (memory reference)

9 A successful return indicates that the request is being processed and the result will be returned in the provided
 10 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
 11 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

12 If executed, the status returned in the provided callback function will be one of the following constants:

- 13 • **PMIX_SUCCESS** The requested data has been returned.
- 14 • **PMIX_ERR_EXISTS_OUTSIDE_SCOPE** The requested key exists, but was posted in a *scope* (see Section
 15 5.2.1.1) that does not include the requester.
- 16 • **PMIX_ERR_NOT_FOUND** The requested data was not available.
- 17 • a non-zero PMIx error constant indicating a reason for the request's failure.

▼ ----- Required Attributes ----- ▼

18 The following attributes are required to be supported by all PMIx libraries:

19 **PMIX_OPTIONAL** "pmix.optional" (bool)
 20 Look only in the client's local data store for the requested value - do not request data from the PMIx
 21 server if not found.

22 **PMIX_IMMEDIATE** "pmix.immediate" (bool)
 23 Specified operation should immediately return an error from the PMIx server if the requested data
 24 cannot be found - do not request it from the host RM.

25 **PMIX_DATA_SCOPE** "pmix.scope" (pmix_scope_t)
 26 Scope of the data to be searched in a **PMIx_Get** call.

27 **PMIX_SESSION_INFO** "pmix.ssn.info" (bool)
 28 Return information regarding the session realm of the target process.

29 **PMIX_JOB_INFO** "pmix.job.info" (bool)
 30 Return information regarding the job realm corresponding to the namespace in the target process'
 31 identifier.

32 **PMIX_APP_INFO** "pmix.app.info" (bool)
 33 Return information regarding the application realm to which the target process belongs - the namespace
 34 of the target process serves to identify the job containing the target application. If information about an
 35 application other than the one containing the target process is desired, then the attribute array must
 36 contain a **PMIX_APPNUM** attribute identifying the desired target application. This is useful in cases
 37 where there are multiple applications and the mapping of processes to applications is unclear.

38 **PMIX_NODE_INFO** "pmix.node.info" (bool)

1 Return information from the node realm regarding the node upon which the specified process is
2 executing. If information about a node other than the one containing the specified process is desired,
3 then the attribute array must also contain either the `PMIX_NODEID` or `PMIX_HOSTNAME` attribute
4 identifying the desired target. This is useful for requesting information about a specific node even if the
5 identity of processes running on that node are not known.

6 `PMIX_GET_POINTER_VALUES` "`pmix.get.pntrs`" (bool)

7 Request that any pointers in the returned value point directly to values in the key-value store. The user
8 *must not* release any returned data pointers.

9 `PMIX_GET_REFRESH_CACHE` "`pmix.get.refresh`" (bool)

10 When retrieving data for a remote process, refresh the existing local data cache for the process in case
11 new values have been put and committed by the process since the last refresh. Local process
12 information is assumed to be automatically updated upon posting by the process. A `NULL` key will
13 cause all values associated with the process to be refreshed - otherwise, only the indicated key will
14 be updated. A process rank of `PMIX_RANK_WILDCARD` can be used to update job-related information in
15 dynamic environments. The user is responsible for subsequently updating refreshed values they may
16 have cached in their own local memory.

17
18 The following attributes are required for host environments that support this operation:

19 `PMIX_WAIT` "`pmix.wait`" (int)

20 Caller requests that the PMIx server wait until at least the specified number of values are found (a value
21 of zero indicates *all* and is the default).

▲-----▲
▼-----▼ **Optional Attributes** -----▼

22 The following attributes are optional for host environments that support this operation:

23 `PMIX_TIMEOUT` "`pmix.timeout`" (int)

24 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
25 `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers
26 (client, server, and host) simultaneously timing the operation.

▲-----▲
Description

27 The callback function will be executed once the retrieval rules of Chapters 6 or 5.3 are met. See `PMIx_Get`
28 for a full description. Note that the non-blocking form of this function cannot support the
29 `PMIX_GET_STATIC_VALUES` attribute as the user cannot pass in the required pointer to storage for the
30 result.
31

1 5.4.2 Retrieval attributes

2 The following attributes are defined for use by retrieval APIs:

3 **PMIX_OPTIONAL** "pmix.optional" (bool)

4 Look only in the client's local data store for the requested value - do not request data from the PMIx
5 server if not found.

6 **PMIX_IMMEDIATE** "pmix.immediate" (bool)

7 Specified operation should immediately return an error from the PMIx server if the requested data
8 cannot be found - do not request it from the host RM.

9 **PMIX_GET_POINTER_VALUES** "pmix.get.pntrs" (bool)

10 Request that any pointers in the returned value point directly to values in the key-value store. The user
11 *must not* release any returned data pointers.

12 **PMIX_GET_STATIC_VALUES** "pmix.get.static" (bool)

13 Request that the data be returned in the provided storage location. The caller is responsible for
14 destructing the `pmix_value_t` using the `PMIX_VALUE_DESTRUCT` macro when done.

15 **PMIX_GET_REFRESH_CACHE** "pmix.get.refresh" (bool)

16 When retrieving data for a remote process, refresh the existing local data cache for the process in case
17 new values have been put and committed by the process since the last refresh. Local process
18 information is assumed to be automatically updated upon posting by the process. A `NULL` key will
19 cause all values associated with the process to be refreshed - otherwise, only the indicated key will be
20 updated. A process rank of `PMIX_RANK_WILDCARD` can be used to update job-related information in
21 dynamic environments. The user is responsible for subsequently updating refreshed values they may
22 have cached in their own local memory.

23 **PMIX_DATA_SCOPE** "pmix.scope" (pmix_scope_t)

24 Scope of the data to be searched in a `PMIx_Get` call.

25 **PMIX_TIMEOUT** "pmix.timeout" (int)

26 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
27 `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers
28 (client, server, and host) simultaneously timing the operation.

29 **PMIX_WAIT** "pmix.wait" (int)

30 Caller requests that the PMIx server wait until at least the specified number of values are found (a value
31 of zero indicates *all* and is the default).

CHAPTER 6

Reserved Keys

1 *Reserved* keys are keys whose string representation begin with a prefix of "**pmix**". By definition, reserved
2 keys are provided by the host environment and the PMIx server, and are required to be available at client start
3 of execution. PMIx clients and tools are therefore prohibited from posting reserved keys.

4 Host environments may opt to define non-standardized reserved keys. All reserved keys, whether standardized
5 or non-standardized, follow the same retrieval rules. Users are advised to check both the local PMIx
6 implementation and host environment documentation for a list of any non-standardized reserved keys they
7 must avoid, and to learn of any non-standard keys that may require special handling.

8 6.1 Data realms

9 PMIx information spans a wide range of sources. In some cases, there are multiple overlapping sources for the
10 same type of data - e.g., the session, job, and application can each provide information on the number of nodes
11 involved in their respective area. In order to resolve the ambiguity, a *data realm* is used to identify the scope to
12 which the referenced data applies. Thus, a reference to an attribute that isn't specific to a realm (e.g., the
13 **PMIX_NUM_NODES** attribute) must be accompanied by a corresponding attribute identifying the realm to
14 which the request pertains if it differs from the default.

15 PMIx defines five *data realms* to resolve the ambiguities, as captured in the following attributes used in
16 **PMIx_Get** for retrieving information from each of the realms:

17 **PMIX_SESSION_INFO** "**pmix.ssn.info**" (**bool**)

18 Return information regarding the session realm of the target process.

19 **PMIX_JOB_INFO** "**pmix.job.info**" (**bool**)

20 Return information regarding the job realm corresponding to the namespace in the target process'
21 identifier.

22 **PMIX_APP_INFO** "**pmix.app.info**" (**bool**)

23 Return information regarding the application realm to which the target process belongs - the namespace
24 of the target process serves to identify the job containing the target application. If information about an
25 application other than the one containing the target process is desired, then the attribute array must
26 contain a **PMIX_APPNUM** attribute identifying the desired target application. This is useful in cases
27 where there are multiple applications and the mapping of processes to applications is unclear.

28 **PMIX_PROC_INFO** "**pmix.proc.info**" (**bool**)

29 Return information regarding the target process. This attribute is technically not required as the
30 **PMIx_Get** API specifically identifies the target process in its parameters. However, it is included here
31 for completeness.

32 **PMIX_NODE_INFO** "**pmix.node.info**" (**bool**)

1 Return information from the node realm regarding the node upon which the specified process is
2 executing. If information about a node other than the one containing the specified process is desired,
3 then the attribute array must also contain either the **PMIX_NODEID** or **PMIX_HOSTNAME** attribute
4 identifying the desired target. This is useful for requesting information about a specific node even if the
5 identity of processes running on that node are not known.

Advice to users

6 If information about a session other than the one containing the requesting process is desired, then the attribute
7 array must contain a **PMIX_SESSION_ID** attribute identifying the desired target session. This is required as
8 many environments only guarantee unique namespaces within a session, and not across sessions.

9 Determining the target within a realm varies between realms and is explained in detail in the realm
10 descriptions below. Note that several attributes can be either queried as a key or set as an attribute to specify
11 the target within a realm. The attributes **PMIX_SESSION_ID**, **PMIX_NAMESPACE** and **PMIX_APPNUM** can be
12 used in both ways.

13 6.1.1 Session realm attributes

14 If information about a session other than the one containing the requesting process is desired, then the *info*
15 array passed to **PMIx_Get** must contain a **PMIX_SESSION_ID** attribute identifying the desired target
16 session. This is required as many environments only guarantee unique namespaces within a session, and not
17 across sessions.

18 Note that the *proc* argument of **PMIx_Get** is ignored when referencing session-related information.

19 The following keys, by default, request session-level information. They will return information about the
20 caller's session unless a **PMIX_SESSION_ID** attribute is specified in the *info* array passed to **PMIx_Get**:

21 **PMIX_CLUSTER_ID** "pmix.clid" (char*)

22 A string name for the cluster this allocation is on.

23 **PMIX_UNIV_SIZE** "pmix.univ.size" (uint32_t)

24 Maximum number of process that can be simultaneously executing in a session. Note that this attribute
25 is equivalent to the **PMIX_MAX_PROCS** attribute for the *session* realm - it is included in the PMIx
26 Standard for historical reasons.

27 **PMIX_TMPDIR** "pmix.tmpdir" (char*)

28 Full path to the top-level temporary directory assigned to the session.

29 **PMIX_TDIR_RM CLEAN** "pmix.tdir.rmclean" (bool)

30 The Resource Manager will remove any directories or files it creates in **PMIX_TMPDIR**.

31 **PMIX_HOSTNAME_KEEP_FQDN** "pmix.fqdn" (bool)

32 Fully Qualified Domain Names (FQDNs) are being retained by the PMIx library.

33 **PMIX_RM_NAME** "pmix.rm.name" (char*)

34 String name of the RM.

35 **PMIX_RM_VERSION** "pmix.rm.version" (char*)

36 RM version string.

1 The following session-related keys default to the realms described in their descriptions but can be retrieved
2 from the session realm by setting the [PMIX_SESSION_INFO](#) attribute in the *info* array passed to [PMIx_Get](#):

3 **PMIX_ALLOCATED_NODELIST** "pmix.alist" (char*)

4 Comma-delimited list or regular expression of all nodes in the specified realm regardless of whether or
5 not they currently host processes. Defaults to the *job* realm.

6 **PMIX_NUM_ALLOCATED_NODES** "pmix.num.anodes" (uint32_t)

7 Number of nodes in the specified realm regardless of whether or not they currently host processes.
8 Defaults to the *job* realm.

9 **PMIX_MAX_PROCS** "pmix.max.size" (uint32_t)

10 Maximum number of processes that can be simultaneously executed in the specified realm. Typically,
11 this is a constraint imposed by a scheduler or by user settings in a hostfile or other resource description.
12 Defaults to the *job* realm.

13 **PMIX_NODE_LIST** "pmix.nlist" (char*)

14 Comma-delimited list of nodes currently hosting processes in the specified realm. Defaults to the *job*
15 realm.

16 **PMIX_NUM_SLOTS** "pmix.num.slots" (uint32_t)

17 Maximum number of processes that can simultaneously be executing in the specified realm. Note that
18 this attribute is the equivalent to [PMIX_MAX_PROCS](#) - it is included in the PMIx Standard for
19 historical reasons. Defaults to the *job* realm.

20 **PMIX_NUM_NODES** "pmix.num.nodes" (uint32_t)

21 Number of nodes currently hosting processes in the specified realm. Defaults to the *job* realm.

22 **PMIX_NODE_MAP** "pmix.nmap" (char*)

23 Regular expression of nodes currently hosting processes in the specified realm - see [17.2.3.2](#) for an
24 explanation of its generation. Defaults to the *job* realm.

25 **PMIX_NODE_MAP_RAW** "pmix.nmap.raw" (char*)

26 Comma-delimited list of nodes containing procs within the specified realm. Defaults to the *job* realm.

27 **PMIX_PROC_MAP** "pmix.pmap" (char*)

28 Regular expression describing processes on each node in the specified realm - see [17.2.3.2](#) for an
29 explanation of its generation. Defaults to the *job* realm.

30 **PMIX_PROC_MAP_RAW** "pmix.pmap.raw" (char*)

31 Semi-colon delimited list of strings, each string containing a comma-delimited list of ranks on the
32 corresponding node within the specified realm. Defaults to the *job* realm.

33 **PMIX_ANL_MAP** "pmix.anlmap" (char*)

34 Process map equivalent to [PMIX_PROC_MAP](#) expressed in Argonne National Laboratory's
35 PMI-1/PMI-2 notation. Defaults to the *job* realm.

36 6.1.2 Job realm attributes

37 Job-related information can be retrieved by requesting a key which defaults to the job realm or by including the
38 [PMIX_JOB_INFO](#) attribute in the *info* array passed to [PMIx_Get](#). For job-related keys the target job is
39 specified by setting the namespace of the target job in the *proc* argument and specifying a rank of
40 [PMIX_RANK_WILDCARD](#) in the *proc* argument passed to [PMIx_Get](#).

41 If information is requested about a namespace in a session other than the one containing the requesting process,
42 then the *info* array must contain a [PMIX_SESSION_ID](#) attribute identifying the desired target session. This is
43 required as many environments only guarantee unique namespaces within a session, and not across sessions.

1 The following keys, by default, request job-level information: They will return information about the job
2 indicated in *proc*:

3 **PMIX_JOBID** "pmix.jobid" (char*)

4 Job identifier assigned by the scheduler to the specified job - may be identical to the namespace, but is
5 often a numerical value expressed as a string (e.g., "12345.3").

6 **PMIX_NPROC_OFFSET** "pmix.offset" (pmix_rank_t)

7 Starting global rank of the specified job. The returned value is the same as the value of
8 **PMIX_GLOBAL_RANK** of rank 0 of the specified job.

9 **PMIX_MAX_PROCS** "pmix.max.size" (uint32_t)

10 Maximum number of processes that can be simultaneously executed in the specified realm. Typically,
11 this is a constraint imposed by a scheduler or by user settings in a hostfile or other resource description.
12 Defaults to the *job* realm. In this context, this is the maximum number of processes that can be
13 simultaneously executed in the specified job, which may be a subset of the number allocated to the
14 overall session.

15 **PMIX_NUM_SLOTS** "pmix.num.slots" (uint32_t)

16 Maximum number of processes that can simultaneously be executing in the specified realm. Note that
17 this attribute is the equivalent to **PMIX_MAX_PROCS** - it is included in the PMIx Standard for
18 historical reasons. Defaults to the *job* realm. In this context, this is the maximum number of process
19 that can be simultaneously executing within the specified job, which may be a subset of the number
20 allocated to the overall session. Jobs may reserve a subset of their assigned maximum processes for
21 dynamic operations such as **PMIx_Spawn**.

22 **PMIX_NUM_NODES** "pmix.num.nodes" (uint32_t)

23 Number of nodes currently hosting processes in the specified realm. Defaults to the *job* realm. In this
24 context, this is the number of nodes currently hosting processes in the specified job, which may be a
25 subset of the nodes allocated to the overall session. Jobs may reserve a subset of their assigned nodes
26 for dynamic operations such as **PMIx_Spawn** - i.e., not all nodes may have executing processes from
27 this job at a given point in time.

28 **PMIX_NODE_MAP** "pmix.nmap" (char*)

29 Regular expression of nodes currently hosting processes in the specified realm - see 17.2.3.2 for an
30 explanation of its generation. Defaults to the *job* realm. In this context, this is the regular expression
31 of nodes currently hosting processes in the specified job.

32 **PMIX_NODE_LIST** "pmix.nlist" (char*)

33 Comma-delimited list of nodes currently hosting processes in the specified realm. Defaults to the *job*
34 realm. In this context, this is the comma-delimited list of nodes currently hosting processes in the
35 specified job.

36 **PMIX_PROC_MAP** "pmix.pmap" (char*)

37 Regular expression describing processes on each node in the specified realm - see 17.2.3.2 for an
38 explanation of its generation. Defaults to the *job* realm. In this context, this is the regular expression
39 describing processes on each node in the specified job.

40 **PMIX_ANL_MAP** "pmix.anlmap" (char*)

41 Process map equivalent to **PMIX_PROC_MAP** expressed in Argonne National Laboratory's
42 PMI-1/PMI-2 notation. Defaults to the *job* realm. In this context, this is the process mapping in
43 Argonne National Laboratory's PMI-1/PMI-2 notation of the processes in the specified job.

1 **PMIX_CMD_LINE** "pmix.cmd.line" (char*)
2 Command line used to execute the specified job (e.g., "mpirun -n 2 -map-by foo ./myapp : -n 4
3 ./myapp2"). If the job was created by a call to **PMIx_Spawn**, the string is an in-order concatenation of
4 the values of **PMIX_APP_ARGV** for each application in the job using the character ':' as a separator.

5 **PMIX_NSDIR** "pmix.nmdir" (char*)
6 Full path to the temporary directory assigned to the specified job, under **PMIX_TMPDIR**.

7 **PMIX_JOB_SIZE** "pmix.job.size" (uint32_t)
8 Total number of processes in the specified job across all contained applications. Note that this value
9 can be different from **PMIX_MAX_PROCS**. For example, users may choose to subdivide an allocation
10 (running several jobs in parallel within it), and dynamic programming models may support adding and
11 removing processes from a running *job* on-the-fly. In the latter case, PMIx events may be used to notify
12 processes within the job that the job size has changed.

13 **PMIX_JOB_NUM_APPS** "pmix.job.napps" (uint32_t)
14 Number of applications in the specified job.

15 **PMIX_LOCAL_PEERS** "pmix.lpeers" (char*)
16 Comma-delimited list of ranks that are executing on the local node within the specified namespace –
17 shortcut for **PMIx_Resolve_peers** for the local node.

18 **PMIX_LOCALLDR** "pmix.llldr" (pmix_rank_t)
19 Lowest rank within the specified job on the node (defaults to current node in absence of
20 **PMIX_HOSTNAME** or **PMIX_NODEID** qualifier).

21 **PMIX_LOCAL_CPUSSETS** "pmix.lcpus" (pmix_data_array_t)
22 A **pmix_data_array_t** array of string representations of the Processing Unit (PU) binding bitmaps
23 applied to each local *peer* on the caller's node upon launch. Each string shall begin with the name of
24 the library that generated it (e.g., "hwloc") followed by a colon and the bitmap string itself. The array
25 shall be in the same order as the processes returned by **PMIX_LOCAL_PEERS** for that namespace.

26 **PMIX_LOCAL_SIZE** "pmix.local.size" (uint32_t)
27 Number of processes in the specified job or application on the caller's node. Defaults to job unless the
28 **PMIX_APP_INFO** and the **PMIX_APPNUM** qualifiers are given.

29 6.1.3 Application realm attributes

30 Application-related information can be retrieved by requesting a key which defaults to the application realm or
31 by including the **PMIX_APP_INFO** attribute in the *info* array passed to **PMIx_Get**. If the **PMIX_APPNUM**
32 qualifier is given, then the query shall return the corresponding value for the given application within the
33 namespace specified in the *proc* argument of the query (a **NULL** value for the *proc* argument equates to the
34 namespace of the caller). If the **PMIX_APPNUM** qualifier is not included, then the retrieval shall default to the
35 application containing the process specified by *proc*. If the rank specified in *proc* is
36 **PMIX_RANK_WILDCARD**, then the application number shall default to that of the calling process if the
37 namespace is its own job, or a value of zero if the namespace is that of a different job.

38 The following keys, by default, request application-level information. They will return information about the
39 application indicated in *proc*:

40 **PMIX_APPLDR** "pmix.aldr" (pmix_rank_t)
41 Lowest rank in the specified application.

42 **PMIX_APP_SIZE** "pmix.app.size" (uint32_t)
43 Number of processes in the specified application, regardless of their execution state - i.e., this number
44 may include processes that either failed to start or have already terminated.

1 **PMIX_APP_ARGV** "pmix.app.argv" (char*)
2 Consolidated argv passed to the spawn command for the given application (e.g., `./myapp arg1 arg2`
3 `arg3`).
4 **PMIX_APP_MAP_TYPE** "pmix.apmap.type" (char*)
5 Type of mapping used to layout the application (e.g., `cyclic`).
6 **PMIX_APP_MAP_REGEX** "pmix.apmap.regex" (char*)
7 Regular expression describing the result of the process mapping.

8 The following application-related keys default to the realms described in their descriptions but can be retrieved
9 from the application realm by setting the **PMIX_APP_INFO** attribute in the *info* array passed to **PMIx_Get**:

10 **PMIX_NUM_NODES** "pmix.num.nodes" (uint32_t)
11 Number of nodes currently hosting processes in the specified realm. Defaults to the *job* realm. In this
12 context, this is the number of nodes currently hosting processes in the specified application, which may
13 be a subset of the nodes allocated to the overall session.

14 **PMIX_MAX_PROCS** "pmix.max.size" (uint32_t)
15 Maximum number of processes that can be simultaneously executed in the specified realm. Typically,
16 this is a constraint imposed by a scheduler or by user settings in a hostfile or other resource description.
17 Defaults to the *job* realm. In this context, this is the maximum number of processes that can be
18 executed in the specified application, which may be a subset of the number allocated to the overall
19 session and job.

20 **PMIX_NUM_SLOTS** "pmix.num.slots" (uint32_t)
21 Maximum number of processes that can simultaneously be executing in the specified realm. Note that
22 this attribute is the equivalent to **PMIX_MAX_PROCS** - it is included in the PMI Standard for
23 historical reasons. Defaults to the *job* realm. In this context, this is the number of slots assigned to the
24 specified application, which may be a subset of the slots allocated to the overall session and job.

25 **PMIX_NODE_MAP** "pmix.nmap" (char*)
26 Regular expression of nodes currently hosting processes in the specified realm - see 17.2.3.2 for an
27 explanation of its generation. Defaults to the *job* realm. In this context, this is the regular expression
28 of nodes currently hosting processes in the specified application.

29 **PMIX_NODE_LIST** "pmix.nlist" (char*)
30 Comma-delimited list of nodes currently hosting processes in the specified realm. Defaults to the *job*
31 realm. In this context, this is the comma-delimited list of nodes currently hosting processes in the
32 specified application.

33 **PMIX_PROC_MAP** "pmix.pmap" (char*)
34 Regular expression describing processes on each node in the specified realm - see 17.2.3.2 for an
35 explanation of its generation. Defaults to the *job* realm. In this context, this is the regular expression
36 describing processes on each node in the specified application.

37 6.1.4 Process realm attributes

38 Process-related information can be retrieved by requesting a key which defaults to the process realm or by
39 including the **PMIX_PROC_INFO** attribute in the *info* array passed to **PMIx_Get**. The target process is
40 specified by the namespace and rank of the *proc* argument to **PMIx_Get**. For process-related keys (other than

1 **PMIX_PROCID** and **PMIX_NAMESPACE**) the target process is specified by setting the namespace and rank of the
2 target process in the *proc* argument passed to **PMIx_Get**. If information is requested about a process in a
3 session other than the one containing the requesting process, then an attribute identifying the target session
4 must be provided. This is required as many environments only guarantee unique namespaces within a session,
5 and not across sessions.

6 The following keys, by default, request process-level information: They will return information about the
7 process indicated in *proc*:

8 **PMIX_APPNUM** "pmix.appnum" (uint32_t)

9 The application number within the job in which the specified process is a member.

10 **PMIX_RANK** "pmix.rank" (pmix_rank_t)

11 Process rank within the job, starting from zero.

12 **PMIX_NAMESPACE** "pmix.namespace" (char*)

13 Namespace of the job - may be a numerical value expressed as a string, but is often an alphanumeric
14 string carrying information solely of use to the system. Required to be unique within the scope of the
15 host environment. One cannot retrieve the namespace of an arbitrary process since that would require
16 already knowing the namespace of that process. However, a process' own namespace can be retrieved
17 by passing a NULL value of *proc* to **PMIx_Get**.

18 **PMIX_SESSION_ID** "pmix.session.id" (uint32_t)

19 Session identifier assigned by the scheduler.

20 **PMIX_GLOBAL_RANK** "pmix.grank" (pmix_rank_t)

21 Rank of the specified process spanning across all jobs in this session, starting with zero. Note that no
22 ordering of the jobs is implied when computing this value. As jobs can start and end at random times,
23 this is defined as a continually growing number - i.e., it is not dynamically adjusted as individual jobs
24 and processes are started or terminated.

25 **PMIX_APP_RANK** "pmix.apprank" (pmix_rank_t)

26 Rank of the specified process within its application.

27 **PMIX_PARENT_ID** "pmix.parent" (pmix_proc_t)

28 Process identifier of the parent process of the specified process - typically used to identify the
29 application process that caused the job containing the specified process to be spawned (e.g., the process
30 that called **PMIx_Spawn**). This attribute is only provided for a process if it was created by a call to
31 **PMIx_Spawn** or **PMIx_Spawn_nb**.

32 **PMIX_EXIT_CODE** "pmix.exit.code" (int)

33 Exit code returned when the specified process terminated.

34 **PMIX_PROCID** "pmix.procid" (pmix_proc_t)

35 The caller's process identifier. The value returned is identical to what **PMIx_Init** or
36 **PMIx_tool_init** provides. The process identifier in the **PMIx_Get** call is ignored when
37 requesting this key.

38 **PMIX_LOCAL_RANK** "pmix.lrank" (uint16_t)

39 Rank of the specified process on its node - refers to the numerical location (starting from zero) of the
40 process on its node when counting only those processes from the same job that share the node, ordered
41 by their overall rank within that job.

42 **PMIX_NODE_RANK** "pmix.nrank" (uint16_t)

43 Rank of the specified process on its node spanning all jobs- refers to the numerical location (starting
44 from zero) of the process on its node when counting all processes (regardless of job) that share the
45 node, ordered by their overall rank within the job. The value represents a snapshot in time when the

1 specified process was started on its node and is not dynamically adjusted as processes from other jobs
2 are started or terminated on the node.

3 **PMIX_PACKAGE_RANK** "pmix.pkgrank" (uint16_t)

4 Rank of the specified process on the *package* where this process resides - refers to the numerical
5 location (starting from zero) of the process on its package when counting only those processes from the
6 same job that share the package, ordered by their overall rank within that job. Note that processes that
7 are not bound to PUs within a single specific package cannot have a package rank.

8 **PMIX_PROC_PID** "pmix.ppid" (pid_t)

9 Operating system PID of specified process.

10 **PMIX_PROCDIR** "pmix.pdir" (char*)

11 Full path to the subdirectory under **PMIX_NSDIR** assigned to the specified process.

12 **PMIX_CPUSSET** "pmix.cpuset" (char*)

13 A string representation of the PU binding bitmap applied to the process upon launch. The string shall
14 begin with the name of the library that generated it (e.g., "hwloc") followed by a colon and the bitmap
15 string itself.

16 **PMIX_CPUSSET_BITMAP** "pmix.bitmap" (pmix_cpuset_t*)

17 Bitmap applied to the process upon launch.

18 **PMIX_CREDENTIAL** "pmix.cred" (char*)

19 Security credential assigned to the process.

20 **PMIX_SPAWNED** "pmix.spawned" (bool)

21 **true** if this process resulted from a call to **PMIx_Spawn**. Lack of inclusion (i.e., a return status of
22 **PMIX_ERR_NOT_FOUND**) corresponds to a value of **false** for this attribute.

23 **PMIX_REINCARNATION** "pmix.reinc" (uint32_t)

24 Number of times this process has been re-instantiated - i.e, a value of zero indicates that the process
25 has never been restarted.

26 In addition, process-level information includes functional attributes directly associated with a process - for
27 example, the process-related fabric attributes included in Section 15.3 or the distance attributes of Section
28 12.4.11.

29 6.1.5 Node realm keys

30 Node-related information can be retrieved by requesting a key which defaults to the node realm or by including
31 the **PMIX_NODE_INFO** attribute in the *info* array passed to **PMIx_Get**. The target node defaults to the local
32 node unless a different node is specified in the *info* array using either the **PMIX_HOSTNAME** or
33 **PMIX_NODEID**. Some node related keys are an exception to this rule and are listed separately at the end of
34 this section. These special keys can only target the local node and also require that a namespace be specified
35 using the *proc* argument to **PMIx_Get**.

36 The following keys, by default, request node-level information. They will return information about either the
37 local node or the node specified by **PMIX_HOSTNAME** or **PMIX_NODEID**:

38 **PMIX_HOSTNAME** "pmix.hname" (char*)

39 Name of the host, as returned by the **gethostname** utility or its equivalent.

40 **PMIX_HOSTNAME_ALIASES** "pmix.alias" (char*)

41 Comma-delimited list of names by which the target node is known.

42 **PMIX_NODEID** "pmix.nodeid" (uint32_t)

1 Node identifier expressed as the node's index (beginning at zero) in an array of nodes within the active
2 session. The value must be unique and directly correlate to the `PMIX_HOSTNAME` of the node - i.e.,
3 users can interchangeably reference the same location using either the `PMIX_HOSTNAME` or
4 corresponding `PMIX_NODEID`.

5 `PMIX_NODE_SIZE` "pmix.node.size" (uint32_t)

6 Number of processes across all jobs that are executing upon the node.

7 `PMIX_AVAIL_PHYS_MEMORY` "pmix.pmem" (uint64_t)

8 Total available physical memory on a node.

9 `PMIX_LOCAL_PROCS` "pmix.lprocs" (pmix_proc_t array)

10 Array of `pmix_proc_t` of all processes executing on the local node – shortcut for
11 `PMIx_Resolve_peers` for the local node and a `NULL` namespace argument. The process identifier
12 is ignored for this attribute.

13 `PMIX_NODE_OVERSUBSCRIBED` "pmix.ndosub" (bool) (Provisional)

14 True if the number of processes from this job on this node exceeds the number of slots allocated to it

15 In addition, node-level information includes functional attributes directly associated with a node - for example,
16 the node-related fabric attributes included in Section 15.3.

17 6.2 Retrieval rules for reserved keys

18 The retrieval rules for reserved keys are relatively simple as the keys, if provided by an implementation, are
19 required, by definition, to be available when the client begins execution. Accordingly, `PMIx_Get` for a
20 reserved key first checks the local PMIx Client cache (per the data realm rules of the prior section) for the
21 target key. If the information is not found, then the `PMIX_ERR_NOT_FOUND` error constant is returned unless
22 the target process belongs to a different namespace from that of the requester.

23 In the case where the target and requester's namespaces differ, then the request is forwarded to the local PMIx
24 server. Upon receiving the request, the server shall check its data storage for the specified namespace. If it
25 already knows about this namespace, then it shall attempt to lookup the specified key, returning the value if it
26 is found or the `PMIX_ERR_NOT_FOUND` error constant.

27 If the server does not have a copy of the information for the specified namespace, then the server shall take one
28 of the following actions:

- 29 1. If the request included the `PMIX_IMMEDIATE` attribute, then the server will respond to the client with the
30 `PMIX_ERR_NOT_FOUND` status.
- 31 2. If the host has provided the Direct Business Card Exchange (DBCX) module function interface
32 (`pmix_server_dmodex_req_fn_t`), then the server shall pass the request to its host for servicing.
33 The host is responsible for identifying a source of information on the specified namespace and retrieving it.
34 The host is required to retrieve *all* of the information regarding the target namespace and return it to the
35 requesting server in anticipation of follow-on requests. If the host cannot retrieve the namespace
36 information, then it must respond with the `PMIX_ERR_NOT_FOUND` error constant unless the
37 `PMIX_TIMEOUT` is given and reached (in which case, the host must respond with the
38 `PMIX_ERR_TIMEOUT` constant).

39 Once the the PMIx server receives the namespace information, the server shall search it (again adhering to
40 the prior data realm rules) for the requested key, returning the value if it is found or the
41 `PMIX_ERR_NOT_FOUND` error constant.

3. If the host does not support the DBCX interface, then the server will respond to the client with the `PMIX_ERR_NOT_FOUND` status

6.2.1 Accessing information: examples

This section provides examples illustrating methods for accessing information from the various realms. The intent of the examples is not to provide comprehensive coding guidance, but rather to further illustrate the use of `PMIx_Get` for obtaining information on a *session*, *job*, *application*, *process*, and *node*.

6.2.1.1 Session-level information

The `PMIx_Get` API does not include an argument for specifying the *session* associated with the information being requested. Thus, requests for keys that are not specifically for session-level information must be accompanied by the `PMIX_SESSION_INFO` qualifier.

Example requests are shown below:

```
pmix_info_t info;
pmix_value_t *value;
pmix_status_t rc;
pmix_proc_t myproc, wildcard;

/* initialize the client library */
PMIx_Init(&myproc, NULL, 0);

/* get the #slots in our session */
PMIX_PROC_LOAD(&wildcard, myproc.nspace, PMIX_RANK_WILDCARD);
rc = PMIx_Get(&wildcard, PMIX_UNIV_SIZE, NULL, 0, &value);

/* get the #nodes in our session */
PMIx_Info_load(&info, PMIX_SESSION_INFO, NULL, PMIX_BOOL);
rc = PMIx_Get(&wildcard, PMIX_NUM_NODES, &info, 1, &value);
```

Information regarding a different session can be requested by adding the `PMIX_SESSION_ID` attribute identifying the target session. In this case, the *proc* argument to `PMIx_Get` will be ignored:

C

```

1  pmix_info_t info[2];
2  pmix_value_t *value;
3  pmix_status_t rc;
4  pmix_proc_t myproc;
5  uint32_t sid;
6
7  /* initialize the client library */
8  PMIx_Init(&myproc, NULL, 0);
9
10 /* get the #nodes in a different session */
11 sid = 12345;
12 PMIx_Info_load(&info[0], PMIX_SESSION_INFO, NULL, PMIX_BOOL);
13 PMIx_Info_load(&info[1], PMIX_SESSION_ID, &sid, PMIX_UINT32);
14 rc = PMIx_Get(NULL, PMIX_NUM_NODES, info, 2, &value);

```

C

6.2.1.2 Job-level information

Information regarding a job can be obtained by the methods detailed in Section 6.1.2. Example requests are shown below:

C

```

18 pmix_info_t info;
19 pmix_value_t *value;
20 pmix_status_t rc;
21 pmix_proc_t myproc, wildcard;
22
23 /* initialize the client library */
24 PMIx_Init(&myproc, NULL, 0);
25
26 /* get the #apps in our job */
27 PMIX_PROC_LOAD(&wildcard, myproc.nspace, PMIX_RANK_WILDCARD);
28 rc = PMIx_Get(&wildcard, PMIX_JOB_NUM_APPS, NULL, 0, &value);
29
30 /* get the #nodes in our job */
31 PMIx_Info_load(&info, PMIX_JOB_INFO, NULL, PMIX_BOOL);
32 rc = PMIx_Get(&wildcard, PMIX_NUM_NODES, &info, 1, &value);

```

C

6.2.1.3 Application-level information

Information regarding an application can be obtained by the methods described in Section 6.1.3. Example requests are shown below:


```

1  pmix_info_t info;
2  pmix_value_t *value;
3  pmix_status_t rc;
4  pmix_proc_t myproc, otherproc;
5  uint32_t appsize, appnum;
6
7  /* initialize the client library */
8  PMIx_Init(&myproc, NULL, 0);
9
10 /* get the #processes in our application */
11 rc = PMIx_Get(&myproc, PMIX_APP_SIZE, NULL, 0, &value);
12 appsize = value->data.uint32;
13
14 /* get the #nodes in an application containing "otherproc".
15  * For this use-case, assume that we are in the first application
16  * and we want the #nodes in the second application - use the
17  * rank of the first process in that application, remembering
18  * that ranks start at zero */
19 PMIX_PROC_LOAD(&otherproc, myproc.namespace, appsize);
20
21 /* Since "otherproc" refers to a process in the second application,
22  * we can simply mark that we want the info for this key from the
23  * application realm */
24 PMIx_Info_load(&info, PMIX_APP_INFO, NULL, PMIX_BOOL);
25 rc = PMIx_Get(&otherproc, PMIX_NUM_NODES, &info, 1, &value);
26
27 /* alternatively, we can directly ask for the #nodes in
28  * the second application in our job, again remembering that
29  * application numbers start with zero. Since we are asking
30  * for application realm information about a specific appnum
31  * within our own namespace, the process identifier can be NULL */
32 appnum = 1;
33 PMIx_Info_load(&appinfo[0], PMIX_APP_INFO, NULL, PMIX_BOOL);
34 PMIx_Info_load(&appinfo[1], PMIX_APPNUM, &appnum, PMIX_UINT32);
35 rc = PMIx_Get(NULL, PMIX_NUM_NODES, appinfo, 2, &value);

```

36 6.2.1.4 Process-level information

37 Process-level information is accessed by providing the namespace and rank of the target process. In the
38 absence of any directive as to the level of information being requested, the PMIx library will always return the
39 process-level value. See Section 6.1.4 for details.

40 6.2.1.5 Node-level information

41 Information regarding a node within the system can be obtained by the methods described in Section 6.1.5.
42 Example requests are shown below:

```
1 pmix_info_t info[2];
2 pmix_value_t *value;
3 pmix_status_t rc;
4 pmix_proc_t myproc, otherproc;
5 uint32_t nodeid;
6
7 /* initialize the client library */
8 PMIx_Init(&myproc, NULL, 0);
9
10 /* get the #procs on our node */
11 rc = PMIx_Get(&myproc, PMIX_NODE_SIZE, NULL, 0, &value);
12
13 /* get the #slots on another node */
14 PMIx_Info_load(&info[0], PMIX_NODE_INFO, NULL, PMIX_BOOL);
15 PMIx_Info_load(&info[1], PMIX_HOSTNAME, "remotehost", PMIX_STRING);
16 rc = PMIx_Get(NULL, PMIX_MAX_PROCS, info, 2, &value);
17
18 /* get the total #procs on the remote node - note that we don't
19  * actually need to include the "PMIX_NODE_INFO" attribute here,
20  * but (a) it does no harm and (b) it allowed us to simply reuse
21  * the prior info array
22 rc = PMIx_Get(NULL, PMIX_NODE_SIZE, info, 2, &value);
```

CHAPTER 7

Query Operations

1 This chapter presents mechanisms for generalized queries that access information about the host environment
2 and the system in general. The chapter presents the concept of a query followed by a detailed explanation of
3 the query APIs provided. The chapter compares the use of these APIs with [PMIx_Get](#). The chapter
4 concludes with detailed information about how to use the query interface to access information about what
5 PMIx APIs an implementation supports as well as what attributes each supported API supports.

6 7.1 PMIx_Query_info

7 As the level of interaction between applications and the host SMS grows, so too does the need for the
8 application to query the SMS regarding its capabilities and state information. PMIx provides a generalized
9 query interface for this purpose, along with a set of standardized attribute keys to support a range of requests.
10 This includes requests to determine the status of scheduling queues and active allocations, the scope of API
11 and attribute support offered by the SMS, namespaces of active jobs, location and information about a job's
12 processes, and information regarding available resources.

13 An example use-case for the [PMIx_Query_info_nb](#) API is to ensure clean job completion. Time-shared
14 systems frequently impose maximum run times when assigning jobs to resource allocations. To shut down
15 gracefully (e.g., to write a checkpoint before termination) it is necessary for an application to periodically
16 query the resource manager for the time remaining in its allocation. This is especially true on systems for
17 which allocation times may be shortened or lengthened from the original time limit. Many resource managers
18 provide APIs to dynamically obtain this information, but each API is specific to the resource manager. PMIx
19 supports this use-case by defining an attribute key ([PMIX_TIME_REMAINING](#)) that can be used with the
20 [PMIx_Query_info_nb](#) interface to obtain the number of seconds remaining in the current job allocation.

21 PMIx sometimes provides multiple methods by which an application can obtain information or services. For
22 this example, note that one could alternatively use the [PMIx_Register_event_handler](#) API to register
23 for an event indicating incipient job termination, and then use the [PMIx_Job_control_nb](#) API to request
24 that the host SMS generate an event a specified amount of time prior to reaching the maximum run time.

25 7.1.1 Query Structure

26 A PMIx query structure is composed of one or more keys and a list of qualifiers which provide additional
27 information to describe the query. Keys which use the same qualifiers can be placed in the same query for
28 compactness, though it is permissible to put each key in its own query.

29 The [pmix_query_t](#) structure is used by the [PMIx_Query_info](#) APIs to describe a single query
30 operation.

PMIx v2.0

```

1 typedef struct pmix_query {
2     char **keys;
3     pmix_info_t *qualifiers;
4     size_t nqual;
5 } pmix_query_t;

```

6 where:

- 7 • *keys* is a **NULL**-terminated argv-style array of strings
- 8 • *qualifiers* is an array of `pmix_info_t` describing constraints on the query
- 9 • *nqual* is the number of elements in the *qualifiers* array

10 The following APIs support query of various session and environment values.

11 7.1.2 PMIx_Query_info

12 Summary

13 Query information about the system in general.

14 Format

PMIx v4.0

```

15 pmix_status_t
16 PMIx_Query_info(pmix_query_t queries[], size_t nqueries,
17                pmix_info_t *info[], size_t *ninfo);

```

18 IN queries

19 Array of query structures (array of handles)

20 IN nqueries

21 Number of elements in the *queries* array (integer)

22 INOUT info

23 Address where a pointer to an array of `pmix_info_t` containing the results of the query can be
24 returned (memory reference)

25 INOUT ninfo

26 Address where the number of elements in *info* can be returned (handle)

27 A successful return indicates that all data was found and has been returned.

28 Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

- 29 • `PMIX_ERR_NOT_FOUND` None of the requested data was available.
- 30 • `PMIX_ERR_PARTIAL_SUCCESS` Some of the requested data was found. The *info* array shall contain an
31 element for each query key that returned a value.

1 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
2 error code or an implementation defined error code as described in Section 3.1.1.

3 If a value other than **PMIX_SUCCESS** or **PMIX_ERR_PARTIAL_SUCCESS** is returned, the *info* array shall
4 be **NULL** and *ninfo* zero.

▼----- Required Attributes -----▼

5 A call to this API can specify multiple queries. Each query is composed of a list of keys and a list of attributes
6 which can influence that query. PMIx libraries and host environments that support this API are required to
7 support the following attributes which are specified on a per-query basis:

8 **PMIX_QUERY_REFRESH_CACHE** "pmix.qry.rfsh" (bool)

9 Retrieve updated information from server. NO QUALIFIERS.

10 **PMIX_SESSION_INFO** "pmix.ssn.info" (bool)

11 Return information regarding the session realm of the target process.

12 **PMIX_JOB_INFO** "pmix.job.info" (bool)

13 Return information regarding the job realm corresponding to the namespace in the target process'
14 identifier.

15 **PMIX_APP_INFO** "pmix.app.info" (bool)

16 Return information regarding the application realm to which the target process belongs - the namespace
17 of the target process serves to identify the job containing the target application. If information about an
18 application other than the one containing the target process is desired, then the attribute array must
19 contain a **PMIX_APPNUM** attribute identifying the desired target application. This is useful in cases
20 where there are multiple applications and the mapping of processes to applications is unclear.

21 **PMIX_NODE_INFO** "pmix.node.info" (bool)

22 Return information from the node realm regarding the node upon which the specified process is
23 executing. If information about a node other than the one containing the specified process is desired,
24 then the attribute array must also contain either the **PMIX_NODEID** or **PMIX_HOSTNAME** attribute
25 identifying the desired target. This is useful for requesting information about a specific node even if the
26 identity of processes running on that node are not known.

27 **PMIX_PROC_INFO** "pmix.proc.info" (bool)

28 Return information regarding the target process. This attribute is technically not required as the
29 **PMIx_Get** API specifically identifies the target process in its parameters. However, it is included here
30 for completeness.

31 **PMIX_PROCID** "pmix.procid" (pmix_proc_t)

32 The caller's process identifier. The value returned is identical to what **PMIx_Init** or
33 **PMIx_tool_init** provides. The process identifier in the **PMIx_Get** call is ignored when
34 requesting this key. In this context, specifies the process ID whose information is being requested -
35 e.g., a query asking for the **pmix_proc_info_t** of a specified process. Only required when the
36 request is for information on a specific process.

37 **PMIX_NAMESPACE** "pmix.namespace" (char*)

1 Namespace of the job - may be a numerical value expressed as a string, but is often an alphanumeric
2 string carrying information solely of use to the system. Required to be unique within the scope of the
3 host environment. One cannot retrieve the namespace of an arbitrary process since that would require
4 already knowing the namespace of that process. However, a process' own namespace can be retrieved
5 by passing a NULL value of *proc* to **PMIx_Get**. Specifies the namespace of the process whose
6 information is being requested. Must be accompanied by the **PMIX_RANK** attribute. Only required
7 when the request is for information on a specific process.

8 **PMIX_RANK** "pmix.rank" (pmix_rank_t)

9 Process rank within the job, starting from zero. Specifies the rank of the process whose information is
10 being requested. Must be accompanied by the **PMIX_NAMESPACE** attribute. Only required when the
11 request is for information on a specific process.

12 **PMIX_QUERY_ATTRIBUTE_SUPPORT** "pmix.qry.attrs" (bool)

13 Query list of supported attributes for specified APIs. REQUIRED QUALIFIERS: one or more of
14 **PMIX_CLIENT_FUNCTIONS**, **PMIX_SERVER_FUNCTIONS**, **PMIX_TOOL_FUNCTIONS**, and
15 **PMIX_HOST_FUNCTIONS**.

16 **PMIX_CLIENT_ATTRIBUTES** "pmix.client.attrs" (bool)

17 Request attributes supported by the PMIx client library.

18 **PMIX_SERVER_ATTRIBUTES** "pmix.srvr.attrs" (bool)

19 Request attributes supported by the PMIx server library.

20 **PMIX_HOST_ATTRIBUTES** "pmix.host.attrs" (bool)

21 Request attributes supported by the host environment.

22 **PMIX_TOOL_ATTRIBUTES** "pmix.setup.env" (bool)

23 Request attributes supported by the PMIx tool library functions.

24 Note that inclusion of both the **PMIX_PROCID** directive and either the **PMIX_NAMESPACE** or the **PMIX_RANK**
25 attribute will return a **PMIX_ERR_BAD_PARAM** result, and that the inclusion of a process identifier must
26 apply to all keys in that **pmix_query_t**. Queries for information on multiple specific processes therefore
27 requires submitting multiple **pmix_query_t** structures, each referencing one process. Directives which are
28 not applicable to a key are ignored.



29 An implementation is not required to support any particular keys. If a key is unsupported, the implementation
30 should handle that key in the same way that it is required to handle a key which it cannot find. The following
31 keys may be specified in a query:

32 **PMIX_QUERY_ATTRIBUTE_SUPPORT** "pmix.qry.attrs" (bool)

33 Query list of supported attributes for specified APIs. REQUIRED QUALIFIERS: one or more of
34 **PMIX_CLIENT_FUNCTIONS**, **PMIX_SERVER_FUNCTIONS**, **PMIX_TOOL_FUNCTIONS**, and
35 **PMIX_HOST_FUNCTIONS**.

36 **PMIX_QUERY_NAMESPACES** "pmix.qry.ns" (char*)

37 Request a comma-delimited list of active namespaces. NO QUALIFIERS.

38 **PMIX_QUERY_JOB_STATUS** "pmix.qry.jst" (pmix_status_t)

1 Status of a specified, currently executing job. REQUIRED QUALIFIER: **PMIX_NAMESPACE** indicating
2 the namespace whose status is being queried.

3 **PMIX_QUERY_QUEUE_LIST** "pmix.qry.qlst" (char*)
4 Request a comma-delimited list of scheduler queues. NO QUALIFIERS.

5 **PMIX_QUERY_QUEUE_STATUS** "pmix.qry.qst" (char*)
6 Returns status of a specified scheduler queue, expressed as a string. OPTIONAL QUALIFIERS:
7 **PMIX_ALLOC_QUEUE** naming specific queue whose status is being requested.

8 **PMIX_QUERY_PROC_TABLE** "pmix.qry.ptable" (char*)
9 Returns a (**pmix_data_array_t**) array of **pmix_proc_info_t**, one entry for each process in
10 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: **PMIX_NAMESPACE**
11 indicating the namespace whose process table is being queried.

12 **PMIX_QUERY_LOCAL_PROC_TABLE** "pmix.qry.lptable" (char*)
13 Returns a (**pmix_data_array_t**) array of **pmix_proc_info_t**, one entry for each process in
14 the specified namespace executing on the same node as the requester, ordered by process job rank.
15 REQUIRED QUALIFIER: **PMIX_NAMESPACE** indicating the namespace whose local process table is
16 being queried. OPTIONAL QUALIFIER: **PMIX_HOSTNAME** indicating the host whose local process
17 table is being queried. By default, the query assumes that the host upon which the request was made is
18 to be used.

19 **PMIX_QUERY_SPAWN_SUPPORT** "pmix.qry.spawn" (bool)
20 Return a comma-delimited list of supported spawn attributes. NO QUALIFIERS.

21 **PMIX_QUERY_DEBUG_SUPPORT** "pmix.qry.debug" (bool)
22 Return a comma-delimited list of supported debug attributes. NO QUALIFIERS.

23 **PMIX_QUERY_MEMORY_USAGE** "pmix.qry.mem" (bool)
24 Return information on memory usage for the processes indicated in the qualifiers. OPTIONAL
25 QUALIFIERS: **PMIX_NAMESPACE** and **PMIX_RANK**, or **PMIX_PROCID** of specific process(es) whose
26 memory usage is being requested.

27 **PMIX_QUERY_REPORT_AVG** "pmix.qry.avg" (bool)
28 Report only average values for sampled information. NO QUALIFIERS.

29 **PMIX_QUERY_REPORT_MINMAX** "pmix.qry.minmax" (bool)
30 Report minimum and maximum values. NO QUALIFIERS.

31 **PMIX_QUERY_ALLOC_STATUS** "pmix.query.alloc" (char*)
32 String identifier of the allocation whose status is being requested. NO QUALIFIERS.

33 **PMIX_TIME_REMAINING** "pmix.time.remaining" (char*)
34 Query number of seconds (**uint32_t**) remaining in allocation for the specified namespace.
35 OPTIONAL QUALIFIERS: **PMIX_NAMESPACE** of the namespace whose info is being requested (defaults
36 to allocation containing the caller).

37 **PMIX_SERVER_URI** "pmix.srvr.uri" (char*)
38 URI of the PMIx server to be contacted. Requests the URI of the specified PMIx server's PMIx
39 connection. Defaults to requesting the information for the local PMIx server.

1 **PMIX_CLIENT_AVG_MEMORY** "pmix.cl.mem.avg" (float)
 2 Average Megabytes of memory used by client processes on node. OPTIONAL QUALIFIERS:
 3 **PMIX_HOSTNAME** or **PMIX_NODEID** (defaults to caller's node).

4 **PMIX_DAEMON_MEMORY** "pmix.dmn.mem" (float)
 5 Megabytes of memory currently used by the RM daemon on the node. OPTIONAL QUALIFIERS:
 6 **PMIX_HOSTNAME** or **PMIX_NODEID** (defaults to caller's node).

7 **PMIX_QUERY_AUTHORIZATIONS** "pmix.qry.auths" (bool)
 8 Return operations the PMIx tool is authorized to perform. NO QUALIFIERS.

9 **PMIX_PROC_PID** "pmix.ppid" (pid_t)
 10 Operating system PID of specified process.

11 **PMIX_PROC_STATE_STATUS** "pmix.proc.state" (pmix_proc_state_t)
 12 State of the specified process as of the last report - may not be the actual current state based on update
 13 rate.

14 **Description**

15 Query information about the system in general. This can include a list of active namespaces, fabric topology,
 16 etc. Also can be used to query node-specific info such as the list of peers executing on a given node. The host
 17 environment is responsible for exercising appropriate access control on the information.

18 The returned *status* indicates if requested data was found or not. The returned *info* array will contain a
 19 **PMIX_QUERY_RESULTS** element for each query of the *queries* array. If qualifiers were included in the
 20 query, then the first element of each results array shall contain the **PMIX_QUERY_QUALIFIERS** key with a
 21 **pmix_data_array_t** containing the qualifiers. The remaining **pmix_info_t** shall contain the results of
 22 the query, one entry for each key that was found. Note that duplicate keys in the *queries* array shall result in
 23 duplicate responses within the constraints of the accompanying qualifiers. The caller is responsible for
 24 releasing the returned array.

25 The **PMIx_Query_info** operation is expected to be called between initialization (e.g., **PMIx_Init**) and
 26 finalization (e.g., **PMIx_Finalize**) with the exception of calling the blocking query API exclusively with the
 27 following set of query attributes:

- 28 • **PMIX_QUERY_STABLE_ABI_VERSION**
 29 • **PMIX_QUERY_PROVISIONAL_ABI_VERSION**

30 If this operation is called exclusively with one or more of those attributes then the result is computed locally to
 31 the PMIx implementation and returns the associated value.

▼----- Rationale -----▼

32 Applications that wish to query the ABI level supported by the PMIx implementation wish to do so before
 33 calling any initialization routine. This allows such applications to query the ABI level and take action based on
 34 that information which might include calling the initialization routine in a different manner and/or setting up
 35 the environment before the PMIx implementation is initialized.



Advice to PMIx library implementers

It is recommended that information returned from `PMIx_Query_info` be locally cached so that retrieval by subsequent calls to `PMIx_Get`, `PMIx_Query_info`, or `PMIx_Query_info_nb` can succeed with minimal overhead. The local cache shall be checked prior to querying the PMIx server and/or the host environment. Queries that include the `PMIX_QUERY_REFRESH_CACHE` attribute shall bypass the local cache and retrieve a new value for the query, refreshing the values in the cache upon return.

7.1.3 `PMIx_Query_info_nb`

Summary

Query information about the system in general.

Format

PMIx v2.0

C

```
pmix_status_t
PMIx_Query_info_nb(pmix_query_t queries[], size_t nqueries,
                  pmix_info_cbfunc_t cbfunc, void *cbdata);
```

C

IN `queries`

Array of query structures (array of handles)

IN `nqueries`

Number of elements in the *queries* array (integer)

IN `cbfunc`

Callback function `pmix_info_cbfunc_t` (function reference)

IN `cbdata`

Data to be passed to the callback function (memory reference)

A successful return indicates that the request has been accepted for processing. The provided callback function will only be executed upon successful return of the operation. Note that the library must not invoke the callback function prior to returning from the API.

Returns `PMIX_SUCCESS` or a negative value indicating the error.

If executed, the status returned in the provided callback function will be one of the following constants:

- `PMIX_SUCCESS` All data was found and has been returned.
- `PMIX_ERR_NOT_FOUND` None of the requested data was available. The *info* array will be `NULL` and *ninfo* zero.
- `PMIX_ERR_PARTIAL_SUCCESS` Some of the requested data was found. The *info* array shall contain an element for each query key that returned a value.
- `PMIX_ERR_NOT_SUPPORTED` The host RM does not support this function. The *info* array will be `NULL` and *ninfo* zero.
- a non-zero PMIx error constant indicating a reason for the request's failure. The *info* array will be `NULL` and *ninfo* zero.

Required Attributes

PMIx libraries and host environments that support this API are required to support the following attributes:

PMIX_QUERY_REFRESH_CACHE "pmix.qry.rfsh" (bool)

Retrieve updated information from server. NO QUALIFIERS.

PMIX_SESSION_INFO "pmix.ssn.info" (bool)

Return information regarding the session realm of the target process.

PMIX_JOB_INFO "pmix.job.info" (bool)

Return information regarding the job realm corresponding to the namespace in the target process' identifier.

PMIX_APP_INFO "pmix.app.info" (bool)

Return information regarding the application realm to which the target process belongs - the namespace of the target process serves to identify the job containing the target application. If information about an application other than the one containing the target process is desired, then the attribute array must contain a **PMIX_APPNUM** attribute identifying the desired target application. This is useful in cases where there are multiple applications and the mapping of processes to applications is unclear.

PMIX_NODE_INFO "pmix.node.info" (bool)

Return information from the node realm regarding the node upon which the specified process is executing. If information about a node other than the one containing the specified process is desired, then the attribute array must also contain either the **PMIX_NODEID** or **PMIX_HOSTNAME** attribute identifying the desired target. This is useful for requesting information about a specific node even if the identity of processes running on that node are not known.

PMIX_PROC_INFO "pmix.proc.info" (bool)

Return information regarding the target process. This attribute is technically not required as the **PMIx_Get** API specifically identifies the target process in its parameters. However, it is included here for completeness.

PMIX_PROCID "pmix.procid" (pmix_proc_t)

The caller's process identifier. The value returned is identical to what **PMIx_Init** or **PMIx_tool_init** provides. The process identifier in the **PMIx_Get** call is ignored when requesting this key. In this context, specifies the process ID whose information is being requested - e.g., a query asking for the **pmix_proc_info_t** of a specified process. Only required when the request is for information on a specific process.

PMIX_NAMESPACE "pmix.namespace" (char*)

Namespace of the job - may be a numerical value expressed as a string, but is often an alphanumeric string carrying information solely of use to the system. Required to be unique within the scope of the host environment. One cannot retrieve the namespace of an arbitrary process since that would require already knowing the namespace of that process. However, a process' own namespace can be retrieved by passing a NULL value of *proc* to **PMIx_Get**. Specifies the namespace of the process whose information is being requested. Must be accompanied by the **PMIX_RANK** attribute. Only required when the request is for information on a specific process.

PMIX_RANK "pmix.rank" (pmix_rank_t)

1 Process rank within the job, starting from zero. Specifies the rank of the process whose information is
2 being requested. Must be accompanied by the **PMIX_NAMESPACE** attribute. Only required when the
3 request is for information on a specific process.

4 **PMIX_CLIENT_ATTRIBUTES** "pmix.client.attrs" (bool)
5 Request attributes supported by the PMIx client library.

6 **PMIX_SERVER_ATTRIBUTES** "pmix.srvr.attrs" (bool)
7 Request attributes supported by the PMIx server library.

8 **PMIX_HOST_ATTRIBUTES** "pmix.host.attrs" (bool)
9 Request attributes supported by the host environment.

10 **PMIX_TOOL_ATTRIBUTES** "pmix.setup.env" (bool)
11 Request attributes supported by the PMIx tool library functions.

12 Note that inclusion of both the **PMIX_PROCID** directive and either the **PMIX_NAMESPACE** or the **PMIX_RANK**
13 attribute will return a **PMIX_ERR_BAD_PARAM** result, and that the inclusion of a process identifier must
14 apply to all keys in that **pmix_query_t**. Queries for information on multiple specific processes therefore
15 requires submitting multiple **pmix_query_t** structures, each referencing one process. Directives which are
16 not applicable to a key are ignored.



17 An implementation is not required to support any particular keys. If a key is unsupported, the implementation
18 should handle that key in the same way that it is required to handle a key which it cannot find. The following
19 keys may be specified in a query:

20 **PMIX_QUERY_ATTRIBUTE_SUPPORT** "pmix.qry.attrs" (bool)
21 Query list of supported attributes for specified APIs. REQUIRED QUALIFIERS: one or more of
22 **PMIX_CLIENT_FUNCTIONS**, **PMIX_SERVER_FUNCTIONS**, **PMIX_TOOL_FUNCTIONS**, and
23 **PMIX_HOST_FUNCTIONS**.

24 **PMIX_QUERY_NAMESPACES** "pmix.qry.ns" (char*)
25 Request a comma-delimited list of active namespaces. NO QUALIFIERS.

26 **PMIX_QUERY_JOB_STATUS** "pmix.qry.jst" (pmix_status_t)
27 Status of a specified, currently executing job. REQUIRED QUALIFIER: **PMIX_NAMESPACE** indicating
28 the namespace whose status is being queried.

29 **PMIX_QUERY_QUEUE_LIST** "pmix.qry qlst" (char*)
30 Request a comma-delimited list of scheduler queues. NO QUALIFIERS.

31 **PMIX_QUERY_QUEUE_STATUS** "pmix.qry.qst" (char*)
32 Returns status of a specified scheduler queue, expressed as a string. OPTIONAL QUALIFIERS:
33 **PMIX_ALLOC_QUEUE** naming specific queue whose status is being requested.

34 **PMIX_QUERY_PROC_TABLE** "pmix.qry.ptable" (char*)
35 Returns a (**pmix_data_array_t**) array of **pmix_proc_info_t**, one entry for each process in
36 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: **PMIX_NAMESPACE**
37 indicating the namespace whose process table is being queried.

38 **PMIX_QUERY_LOCAL_PROC_TABLE** "pmix.qry.lptable" (char*)

1 Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in
2 the specified namespace executing on the same node as the requester, ordered by process job rank.
3 REQUIRED QUALIFIER: `PMIX_NAMESPACE` indicating the namespace whose local process table is
4 being queried. OPTIONAL QUALIFIER: `PMIX_HOSTNAME` indicating the host whose local process
5 table is being queried. By default, the query assumes that the host upon which the request was made is
6 to be used.

7 **PMIX_QUERY_SPAWN_SUPPORT** "`pmix.qry.spawn`" (`bool`)
8 Return a comma-delimited list of supported spawn attributes. NO QUALIFIERS.

9 **PMIX_QUERY_DEBUG_SUPPORT** "`pmix.qry.debug`" (`bool`)
10 Return a comma-delimited list of supported debug attributes. NO QUALIFIERS.

11 **PMIX_QUERY_MEMORY_USAGE** "`pmix.qry.mem`" (`bool`)
12 Return information on memory usage for the processes indicated in the qualifiers. OPTIONAL
13 QUALIFIERS: `PMIX_NAMESPACE` and `PMIX_RANK`, or `PMIX_PROCID` of specific process(es) whose
14 memory usage is being requested.

15 **PMIX_QUERY_REPORT_AVG** "`pmix.qry.avg`" (`bool`)
16 Report only average values for sampled information. NO QUALIFIERS.

17 **PMIX_QUERY_REPORT_MINMAX** "`pmix.qry.minmax`" (`bool`)
18 Report minimum and maximum values. NO QUALIFIERS.

19 **PMIX_QUERY_ALLOC_STATUS** "`pmix.query.alloc`" (`char*`)
20 String identifier of the allocation whose status is being requested. NO QUALIFIERS.

21 **PMIX_TIME_REMAINING** "`pmix.time.remaining`" (`char*`)
22 Query number of seconds (`uint32_t`) remaining in allocation for the specified namespace.
23 OPTIONAL QUALIFIERS: `PMIX_NAMESPACE` of the namespace whose info is being requested (defaults
24 to allocation containing the caller).

25 **PMIX_SERVER_URI** "`pmix.srvr.uri`" (`char*`)
26 URI of the PMIx server to be contacted. Requests the URI of the specified PMIx server's PMIx
27 connection. Defaults to requesting the information for the local PMIx server.

28 **PMIX_CLIENT_AVG_MEMORY** "`pmix.cl.mem.avg`" (`float`)
29 Average Megabytes of memory used by client processes on node. OPTIONAL QUALIFIERS:
30 `PMIX_HOSTNAME` or `PMIX_NODEID` (defaults to caller's node).

31 **PMIX_DAEMON_MEMORY** "`pmix.dmn.mem`" (`float`)
32 Megabytes of memory currently used by the RM daemon on the node. OPTIONAL QUALIFIERS:
33 `PMIX_HOSTNAME` or `PMIX_NODEID` (defaults to caller's node).

34 **PMIX_QUERY_AUTHORIZATIONS** "`pmix.qry.auths`" (`bool`)
35 Return operations the PMIx tool is authorized to perform. NO QUALIFIERS.

36 **PMIX_PROC_PID** "`pmix.ppid`" (`pid_t`)
37 Operating system PID of specified process.

38 **PMIX_PROC_STATE_STATUS** "`pmix.proc.state`" (`pmix_proc_state_t`)

1 State of the specified process as of the last report - may not be the actual current state based on update
2 rate.

3 Description

4 Non-blocking form of the [PMIx_Query_info](#) API.

5 7.1.4 Query keys

6 The following keys may be queried using the [PMIx_Query_info](#) and [PMIx_Query_info_nb](#) APIs:

7 **PMIX_QUERY_SUPPORTED_KEYS** "pmix.qry.keys" (char*)

8 Returns comma-delimited list of keys supported by the query function. NO QUALIFIERS.

9 **PMIX_QUERY_SUPPORTED_QUALIFIERS** "pmix.qry.qual" (char*)

10 Return comma-delimited list of qualifiers supported by a query on the provided key, instead of actually
11 performing the query on the key. NO QUALIFIERS.

12 **PMIX_QUERY_NAMESPACES** "pmix.qry.ns" (char*)

13 Request a comma-delimited list of active namespaces. NO QUALIFIERS.

14 **PMIX_QUERY_NAMESPACE_INFO** "pmix.qry.nsinfo" (pmix_data_array_t*)

15 Return an array of active namespace information - each element will itself contain an array including
16 the namespace plus the command line of the application executing within it. OPTIONAL

17 QUALIFIERS: [PMIX_NAMESPACE](#) of specific namespace whose info is being requested.

18 **PMIX_QUERY_JOB_STATUS** "pmix.qry.jst" (pmix_status_t)

19 Status of a specified, currently executing job. REQUIRED QUALIFIER: [PMIX_NAMESPACE](#) indicating
20 the namespace whose status is being queried.

21 **PMIX_QUERY_QUEUE_LIST** "pmix.qry.qlst" (char*)

22 Request a comma-delimited list of scheduler queues. NO QUALIFIERS.

23 **PMIX_QUERY_QUEUE_STATUS** "pmix.qry.qst" (char*)

24 Returns status of a specified scheduler queue, expressed as a string. OPTIONAL QUALIFIERS:

25 [PMIX_ALLOC_QUEUE](#) naming specific queue whose status is being requested.

26 **PMIX_QUERY_PROC_TABLE** "pmix.qry.ptable" (char*)

27 Returns a ([pmix_data_array_t](#)) array of [pmix_proc_info_t](#), one entry for each process in
28 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: [PMIX_NAMESPACE](#)
29 indicating the namespace whose process table is being queried.

30 **PMIX_QUERY_LOCAL_PROC_TABLE** "pmix.qry.lptable" (char*)

31 Returns a ([pmix_data_array_t](#)) array of [pmix_proc_info_t](#), one entry for each process in
32 the specified namespace executing on the same node as the requester, ordered by process job rank.

33 REQUIRED QUALIFIER: [PMIX_NAMESPACE](#) indicating the namespace whose local process table is

34 being queried. OPTIONAL QUALIFIER: [PMIX_HOSTNAME](#) indicating the host whose local process
35 table is being queried. By default, the query assumes that the host upon which the request was made is

36 to be used.

37 **PMIX_QUERY_AUTHORIZATIONS** "pmix.qry.auths" (bool)

38 Return operations the PMIx tool is authorized to perform. NO QUALIFIERS.

39 **PMIX_QUERY_SPAWN_SUPPORT** "pmix.qry.spawn" (bool)

40 Return a comma-delimited list of supported spawn attributes. NO QUALIFIERS.

41 **PMIX_QUERY_DEBUG_SUPPORT** "pmix.qry.debug" (bool)

42 Return a comma-delimited list of supported debug attributes. NO QUALIFIERS.

1 **PMIX_QUERY_MEMORY_USAGE** "pmix.qry.mem" (bool)
2 Return information on memory usage for the processes indicated in the qualifiers. OPTIONAL
3 QUALIFIERS: **PMIX_NAMESPACE** and **PMIX_RANK**, or **PMIX_PROCID** of specific process(es) whose
4 memory usage is being requested.

5 **PMIX_TIME_REMAINING** "pmix.time.remaining" (char*)
6 Query number of seconds (**uint32_t**) remaining in allocation for the specified namespace.
7 OPTIONAL QUALIFIERS: **PMIX_NAMESPACE** of the namespace whose info is being requested (defaults
8 to allocation containing the caller).

9 **PMIX_QUERY_ATTRIBUTE_SUPPORT** "pmix.qry.attrs" (bool)
10 Query list of supported attributes for specified APIs. REQUIRED QUALIFIERS: one or more of
11 **PMIX_CLIENT_FUNCTIONS**, **PMIX_SERVER_FUNCTIONS**, **PMIX_TOOL_FUNCTIONS**, and
12 **PMIX_HOST_FUNCTIONS**.

13 **PMIX_QUERY_NUM_PSETS** "pmix.qry.psetnum" (**size_t**)
14 Return the number of process sets defined in the specified range (defaults to
15 **PMIX_RANGE_SESSION**).

16 **PMIX_QUERY_PSET_NAMES** "pmix.qry.psets" (**pmix_data_array_t***)
17 Return a **pmix_data_array_t** containing an array of strings of the process set names defined in
18 the specified range (defaults to **PMIX_RANGE_SESSION**).

19 **PMIX_QUERY_PSET_MEMBERSHIP** "pmix.qry.pmems" (**pmix_data_array_t***)
20 Return an array of **pmix_proc_t** containing the members of the specified process set.

21 **PMIX_QUERY_AVAIL_SERVERS** "pmix.qry.asrvrs" (**pmix_data_array_t***)
22 Return an array of **pmix_info_t**, each element itself containing a **PMIX_SERVER_INFO_ARRAY**
23 entry holding all available data for a server on this node to which the caller might be able to connect.

24 **PMIX_QUERY_STABLE_ABI_VERSION** "pmix.qry.stabiver" (char *)
25 Query the PMIx Standard Stable ABI version(s) supported by the PMIx library. The version returned
26 will be of the form "**MAJOR.MINOR**". If multiple versions are supported then a comma-separated list
27 of version numbers will be returned. See Section 1.3.3.1 for versioning advice. NO QUALIFIERS.

28 **PMIX_QUERY_PROVISIONAL_ABI_VERSION** "pmix.qry.prabiver" (char *)
29 Query the PMIx Standard Provisional ABI version(s) supported by the PMIx library. The version
30 returned will be of the form "**MAJOR.MINOR**". If multiple versions are supported then a
31 comma-separated list of version numbers will be returned. See Section 1.3.3.1 for versioning advice.
32 NO QUALIFIERS.

33 These keys are used to query memory available and used in the system.

34 **PMIX_AVAIL_PHYS_MEMORY** "pmix.pmem" (**uint64_t**)
35 Total available physical memory on a node. OPTIONAL QUALIFIERS: **PMIX_HOSTNAME** or
36 **PMIX_NODEID** (defaults to caller's node).

37 **PMIX_DAEMON_MEMORY** "pmix.dmn.mem" (float)
38 Megabytes of memory currently used by the RM daemon on the node. OPTIONAL QUALIFIERS:
39 **PMIX_HOSTNAME** or **PMIX_NODEID** (defaults to caller's node).

40 **PMIX_CLIENT_AVG_MEMORY** "pmix.cl.mem.avg" (float)
41 Average Megabytes of memory used by client processes on node. OPTIONAL QUALIFIERS:
42 **PMIX_HOSTNAME** or **PMIX_NODEID** (defaults to caller's node).

1 7.1.5 Query attributes

2 Attributes used to direct behavior of the [PMIx_Query_info](#) and [PMIx_Query_info_nb](#) APIs:

- 3 **PMIX_QUERY_RESULTS** "pmix.qry.res" ([pmix_data_array_t](#))
4 Contains an array of query results for a given [pmix_query_t](#) passed to the [PMIx_Query_info](#)
5 APIs. If qualifiers were included in the query, then the first element of the array shall be the
6 [PMIX_QUERY_QUALIFIERS](#) attribute containing those qualifiers. Each of the remaining elements
7 of the array is a [pmix_info_t](#) containing the query key and the corresponding value returned by the
8 query. This attribute is solely for reporting purposes and cannot be used in [PMIx_Get](#) or other query
9 operations.
- 10 **PMIX_QUERY_QUALIFIERS** "pmix.qryquals" ([pmix_data_array_t](#))
11 Contains an array of qualifiers that were included in the query that produced the provided results. This
12 attribute is solely for reporting purposes and cannot be used in [PMIx_Get](#) or other query operations.
- 13 **PMIX_QUERY_REFRESH_CACHE** "pmix.qry.rfsh" ([bool](#))
14 Retrieve updated information from server. NO QUALIFIERS.
- 15 **PMIX_QUERY_LOCAL_ONLY** "pmix.qry.local" ([bool](#))
16 Constrain the query to local information only. NO QUALIFIERS.
- 17 **PMIX_QUERY_REPORT_AVG** "pmix.qry.avg" ([bool](#))
18 Report only average values for sampled information. NO QUALIFIERS.
- 19 **PMIX_QUERY_REPORT_MINMAX** "pmix.qry.minmax" ([bool](#))
20 Report minimum and maximum values. NO QUALIFIERS.
- 21 **PMIX_QUERY_ALLOC_STATUS** "pmix.query.alloc" ([char*](#))
22 String identifier of the allocation whose status is being requested. NO QUALIFIERS.
- 23 **PMIX_SERVER_INFO_ARRAY** "pmix.srv.arr" ([pmix_data_array_t](#))
24 Array of [pmix_info_t](#) about a given server, starting with its [PMIX_NAMESPACE](#) and including at least
25 one of the rendezvous-required pieces of information.

26 The following attributes are used as qualifiers in queries regarding attribute support within the PMIx
27 implementation and/or the host environment: **PMIX_CLIENT_FUNCTIONS**

- 28 "pmix.client.fns" ([bool](#))
29 Request a list of functions supported by the PMIx client library.
- 30 **PMIX_CLIENT_ATTRIBUTES** "pmix.client.attrs" ([bool](#))
31 Request attributes supported by the PMIx client library.
- 32 **PMIX_SERVER_FUNCTIONS** "pmix.srvr.fns" ([bool](#))
33 Request a list of functions supported by the PMIx server library.
- 34 **PMIX_SERVER_ATTRIBUTES** "pmix.srvr.attrs" ([bool](#))
35 Request attributes supported by the PMIx server library.
- 36 **PMIX_HOST_FUNCTIONS** "pmix.srvr.fns" ([bool](#))
37 Request a list of functions supported by the host environment.
- 38 **PMIX_HOST_ATTRIBUTES** "pmix.host.attrs" ([bool](#))
39 Request attributes supported by the host environment.
- 40 **PMIX_TOOL_FUNCTIONS** "pmix.tool.fns" ([bool](#))
41 Request a list of functions supported by the PMIx tool library.
- 42 **PMIX_TOOL_ATTRIBUTES** "pmix.setup.env" ([bool](#))
43 Request attributes supported by the PMIx tool library functions.

1 7.1.5.1 Query structure support macros

2 The following macros are provided to support the `pmix_query_t` structure.

3 Static initializer for the query structure

4 *(Provisional)*

5 Provide a static initializer for the `pmix_query_t` fields.

6 *PMIx v5.0*  

`PMIX_QUERY_STATIC_INIT`

7 Initialize the query structure

8 Initialize the `pmix_query_t` fields

9 *PMIx v2.0*  

`PMIX_QUERY_CONSTRUCT (m)`

IN `m`

Pointer to the structure to be initialized (pointer to `pmix_query_t`)

12 Destruct the query structure

13 Destruct the `pmix_query_t` fields

14 *PMIx v2.0*  

`PMIX_QUERY_DESTRUCT (m)`

IN `m`

Pointer to the structure to be destructed (pointer to `pmix_query_t`)

17 Create a query array

18 Allocate and initialize an array of `pmix_query_t` structures

19 *PMIx v2.0*  

`PMIX_QUERY_CREATE (m, n)`

INOUT `m`

Address where the pointer to the array of `pmix_query_t` structures shall be stored (handle)

IN `n`

Number of structures to be allocated (`size_t`)

1 **Free a query structure**
2 Release a `pmix_query_t` structure



3 `PMIX_QUERY_RELEASE` (m)



4 **IN** m
5 Pointer to a `pmix_query_t` structure (handle)

6 **Free a query array**
7 Release an array of `pmix_query_t` structures

PMIx v2.0



8 `PMIX_QUERY_FREE` (m, n)



9 **IN** m
10 Pointer to the array of `pmix_query_t` structures (handle)

11 **IN** n
12 Number of structures in the array (`size_t`)

13 **Create the info array of query qualifiers**
14 Create an array of `pmix_info_t` structures for passing query qualifiers, updating the `nqual` field of the
15 `pmix_query_t` structure.

PMIx v2.2



16 `PMIX_QUERY_QUALIFIERS_CREATE` (m, n)



17 **IN** m
18 Pointer to the `pmix_query_t` structure (handle)

19 **IN** n
20 Number of qualifiers to be allocated (`size_t`)

21 7.2 `PMIx_Resolve_peers`

22 There are a number of common queries for which PMIx provides convenience routines. These APIs provide
23 simplified access to commonly requested queries. Due to their simplified interface, these APIs cannot be
24 customized through the use of attributes. If a more specialized version of these queries are required, similar
25 functionality can often be accessed through the `PMIx_Query_info` or `PMIx_Query_info_nb` APIs.

26 **Summary**

27 Obtain the array of processes within the specified namespace that are executing on a given node.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Format

C

```
pmix_status_t  
PMIx_Resolve_peers(const char *nodename,  
const pmix_namespace_t nspace,  
pmix_proc_t **procs, size_t *nprocs);
```

C

IN *nodename*

Name of the node to query - **NULL** can be used to denote the current local node (string)

IN *namespace*

namespace (string)

OUT *procs*

Array of process structures (array of handles)

OUT *nprocs*

Number of elements in the *procs* array (integer)

Returns **PMIX_SUCCESS** or a negative value indicating the error.

Description

Given a *nodename*, return the array of processes within the specified *namespace* that are executing on that node. If the *namespace* is **NULL**, then all processes on the node will be returned. If the specified node does not currently host any processes, then the returned array will be **NULL**, and *nprocs* will be zero. The caller is responsible for releasing the *procs* array when done with it. The **PMIX_PROC_FREE** macro is provided for this purpose.

7.2.1 PMIx_Resolve_nodes

Summary

Return a list of nodes hosting processes within the given namespace.

Format

PMIx v1.0

C

```
pmix_status_t  
PMIx_Resolve_nodes(const char *namespace, char **nodelist);
```

C

IN *namespace*

Namespace (string)

OUT *nodelist*

Comma-delimited list of nodenames (string)

Returns **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant.

Description

Given a *namespace*, return the list of nodes hosting processes within that namespace. The returned string will contain a comma-delimited list of nodenames. The caller is responsible for releasing the string when done with it.

1 7.3 Using Get vs Query

2 Both `PMIx_Get` and `PMIx_Query_info` can be used to retrieve information about the system. In general,
3 the *get* operation should be used to retrieve:

- 4 • information provided by the host environment at time of job start. This includes information on the number
5 of processes in the job, their location, and possibly their communication endpoints.
- 6 • information posted by processes via the `PMIx_Put` function.

7 This information is largely considered to be *static*, although this will not necessarily be true for environments
8 supporting dynamic programming models or fault tolerance. Note that the `PMIx_Get` function only accesses
9 information about execution environments - i.e., its scope is limited to values pertaining to a specific *session*,
10 *job*, *application*, *process*, or *node*. It cannot be used to obtain information about areas such as the status of
11 queues in the WLM.

12 In contrast, the *query* option should be used to access:

- 13 • system-level information (such as the available WLM queues) that would generally not be included in
14 job-level information provided at job start.
- 15 • dynamic information such as application and queue status, and resource utilization statistics. Note that the
16 `PMIX_QUERY_REFRESH_CACHE` attribute must be provided on each query to ensure current data is
17 returned.
- 18 • information created post job start, such as process tables.
- 19 • information requiring more complex search criteria than supported by the simpler `PMIx_Get` API.
- 20 • queries focused on retrieving multi-attribute blocks of data with a single request, thus bypassing the
21 single-key limitation of the `PMIx_Get` API.

22 In theory, all information can be accessed via `PMIx_Query_info` as the local cache is typically the same
23 datastore searched by `PMIx_Get`. However, in practice, the overhead associated with the *query* operation may
24 (depending upon implementation) be higher than the simpler *get* operation due to the need to construct and
25 process the more complex `pmix_query_t` structure. Thus, requests for a single key value are likely to be
26 accomplished faster with `PMIx_Get` versus the *query* operation.

27 7.4 Accessing attribute support information

28 Information as to which attributes are supported by either the PMIx implementation or its host environment
29 can be obtained via the `PMIx_Query_info` APIs. The `PMIX_QUERY_ATTRIBUTE_SUPPORT` attribute
30 must be listed as the first entry in the *keys* field of the `pmix_query_t` structure, followed by the name of the
31 function whose attribute support is being requested - support for multiple functions can be requested
32 simultaneously by simply adding the function names to the array of *keys*. Function names *must* be given as
33 user-level API names - e.g., “`PMIx_Get`”, “`PMIx_server_setup_application`”, or
34 “`PMIx_tool_attach_to_server`”.

35 The desired levels of attribute support are provided as qualifiers. Multiple levels can be requested
36 simultaneously by simply adding elements to the *qualifiers* array. Each qualifier should contain the desired
37 level attribute with the boolean value set to indicate whether or not that level is to be included in the returned
38 information. Failure to provide any levels is equivalent to a request for all levels. Supported levels include:

- 1 • **PMIX_CLIENT_FUNCTIONS** "pmix.client.fns" (bool)
2 Request a list of functions supported by the PMIx client library.
- 3 • **PMIX_CLIENT_ATTRIBUTES** "pmix.client.attrs" (bool)
4 Request attributes supported by the PMIx client library.
- 5 • **PMIX_SERVER_FUNCTIONS** "pmix.srvr.fns" (bool)
6 Request a list of functions supported by the PMIx server library.
- 7 • **PMIX_SERVER_ATTRIBUTES** "pmix.srvr.attrs" (bool)
8 Request attributes supported by the PMIx server library.
- 9 • **PMIX_HOST_FUNCTIONS** "pmix.srvr.fns" (bool)
10 Request a list of functions supported by the host environment.
- 11 • **PMIX_HOST_ATTRIBUTES** "pmix.host.attrs" (bool)
12 Request attributes supported by the host environment.
- 13 • **PMIX_TOOL_FUNCTIONS** "pmix.tool.fns" (bool)
14 Request a list of functions supported by the PMIx tool library.
- 15 • **PMIX_TOOL_ATTRIBUTES** "pmix.setup.env" (bool)
16 Request attributes supported by the PMIx tool library functions.

17 Unlike other queries, queries for attribute support can result in the number of returned `pmix_info_t`
18 structures being different from the number of queries. Each element in the returned array will correspond to a
19 pair of specified attribute level and function in the query, where the *key* is the function and the *value* contains a
20 `pmix_data_array_t` of `pmix_info_t`. Each element of the array is marked by a *key* indicating the
21 requested attribute *level* with a *value* composed of a `pmix_data_array_t` of `pmix_regattr_t`, each
22 describing a supported attribute for that function, as illustrated in Fig. 7.1 below where the requestor asked for
23 supported attributes of `PMIx_Get` at the *client* and *server* levels, plus attributes of
24 `PMIx_Allocation_request` at all levels.

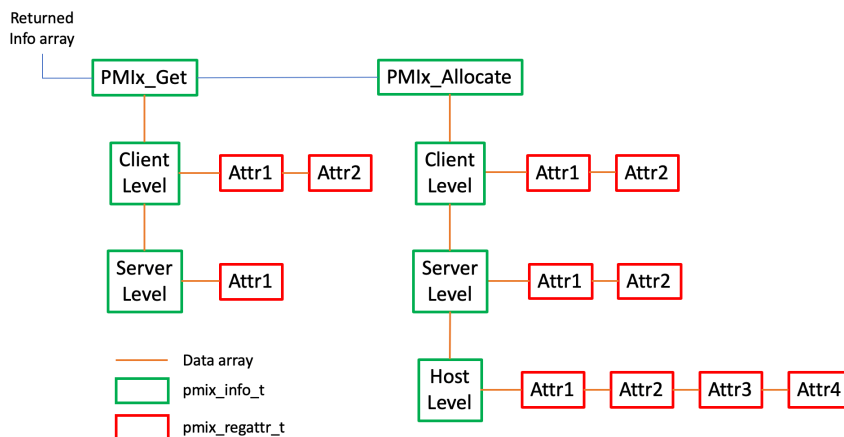


Figure 7.1.: Returned information hierarchy for attribute support request

1 The array of returned structures, and their child arrays, are subject to the return rules for the
2 **PMIx_Query_info_nb** API. For example, a request for supported attributes of the **PMIx_Get** function
3 that includes the *host* level will return values for the *client* and *server* levels, plus an array element with a *key*
4 of **PMIX_HOST_ATTRIBUTES** and a value type of **PMIX_UNDEF** indicating that no attributes are supported
5 at that level.

CHAPTER 8

Synchronization

1 Applications may need to synchronize their operations at various points in their execution. Depending on a
2 variety of factors (e.g., the programming model and where the synchronization point lies), the application may
3 choose to execute the operation using PMIx to access the communication capabilities of the host environment's
4 infrastructure. This is particularly useful in situations where communication libraries are not yet initialized by
5 the application. Synchronization operations also offer an opportunity for processes to exchange data at a known
6 point in their execution. For example, communication libraries within different processes can synchronize to
7 exchange information on communication endpoints for subsequent wireup of messaging protocols.

8 PMIx clients can use the `PMIx_Fence` and `PMIx_Fence_nb` functions to synchronize a set of processes.
9 The fence operation can be useful after an application performs a number of `PMIx_Put` operations to
10 coordinate with other processes that the data is available for access. This avoids unsuccessful `PMIx_Get` calls
11 that might otherwise be invoked before the cooresponding `PMIx_Put` call is complete.

12 In its default form, the fence operation acts as a barrier between the processes and does not exchange data.
13 Clients can pass the `PMIX_COLLECT_DATA` attribute to request that the `PMIx_Fence` and
14 `PMIx_Fence_nb` functions exchange all committed data between all involved servers during the
15 synchronization operation. This will make local to each process the data put by other processes resulting in
16 faster resolution of `PMIx_Get` and `PMIx_Get_nb` function calls at the cost of a synchronous data exchange
17 and associated memory footprint expansion. In many situations this attribute may have performance benefits
18 as many systems are optimized for transporting larger amounts of data. In such applications, a
19 'put/commit/fence/get' pattern is common for efficiently exchanging key-value pairs. For applications where
20 only a small subset of clients access another small subset's key-value pairs this attribute may not be beneficial.
21 As such, applications are not required to use `PMIx_Fence` or `PMIx_Fence_nb` functions nor the
22 associated data collection attribute to ensure correctness of PMIx get/put functionality.

23 8.1 PMIx_Fence

24 Summary

25 Execute a blocking barrier across the processes identified in the specified array, collecting information posted
26 via `PMIx_Put` as directed.

27 Format

PMIx v1.0

C

```
28 pmix_status_t  
29 PMIx_Fence(const pmix_proc_t procs[], size_t nprocs,  
30            const pmix_info_t info[], size_t ninfo);
```

1 **IN** `procs`
 2 Array of `pmix_proc_t` structures (array of handles)
 3 **IN** `nprocs`
 4 Number of elements in the `procs` array (integer)
 5 **IN** `info`
 6 Array of info structures (array of handles)
 7 **IN** `ninfo`
 8 Number of elements in the `info` array (integer)

9 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

10 The following attributes are required to be supported by all PMIx libraries:

11 **PMIX_COLLECT_DATA** "`pmix.collect`" (`bool`)

12 Collect all data posted by the participants using `PMIx_Put` that has been committed via
 13 `PMIx_Commit`, making the collection locally available to each participant at the end of the operation.
 14 By default, this will include all job-level information that was locally generated by PMIx servers unless
 15 excluded using the `PMIX_COLLECT_GENERATED_JOB_INFO` attribute.

16 **PMIX_COLLECT_GENERATED_JOB_INFO** "`pmix.collect.gen`" (`bool`)

17 Collect all job-level information (i.e., reserved keys) that was locally generated by PMIx servers. Some
 18 job-level information (e.g., distance between processes and fabric devices) is best determined on a
 19 distributed basis as it primarily pertains to local processes. Should remote processes need to access the
 20 information, it can either be obtained collectively using the `PMIx_Fence` operation with this
 21 directive, or can be retrieved one peer at a time using `PMIx_Get` without first having performed the
 22 job-wide collection.

Optional Attributes

23 The following attributes are optional for PMIx implementations:

24 **PMIX_ALL_CLONES_PARTICIPATE** "`pmix.clone.part`" (`bool`)

25 All *clones* of the calling process must participate in the collective operation.

26 The following attributes are optional for host environments:

27 **PMIX_TIMEOUT** "`pmix.timeout`" (`int`)

28 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
 29 `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers
 30 (client, server, and host) simultaneously timing the operation.

Description

Passing a **NULL** pointer as the *procs* parameter indicates that the fence is to span all processes in the client's namespace. Each provided `pmix_proc_t` struct can pass `PMIX_RANK_WILDCARD` to indicate that all processes in the given namespace are participating.

The *info* array is used to pass user directives regarding the behavior of the fence operation. Note that for scalability reasons, the default behavior for `PMIx_Fence` is to not collect data posted by the operation's participants.

Advice to PMIx library implementers

`PMIx_Fence` and its non-blocking form are both *collective* operations. Accordingly, the PMIx server library is required to aggregate participation by local clients, passing the request to the host environment once all local participants have executed the API.

Advice to PMIx server hosts

The host will receive a single call for each collective operation. It is the responsibility of the host to identify the nodes containing participating processes, execute the collective across all participating nodes, and notify the local PMIx server library upon completion of the global collective.

8.2 `PMIx_Fence_nb`

Summary

Execute a nonblocking `PMIx_Fence` across the processes identified in the specified array of processes, collecting information posted via `PMIx_Put` as directed.

Format

```
pmix_status_t
PMIx_Fence_nb(const pmix_proc_t procs[], size_t nprocs,
              const pmix_info_t info[], size_t ninfo,
              pmix_op_cbfunc_t cbfunc, void *cbdata);
```

IN procs
Array of [pmix_proc_t](#) structures (array of handles)

IN nprocs
Number of elements in the *procs* array (integer)

IN info
Array of info structures (array of handles)

IN ninfo
Number of elements in the *info* array (integer)

IN cbfunc
Callback function (function reference)

IN cbdata
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when [PMIX_SUCCESS](#) is returned.

Returns [PMIX_SUCCESS](#) or one of the following error codes when the condition described occurs:

- [PMIX_OPERATION_SUCCEEDED](#), indicating that the request was immediately processed and returned *success* - the *cbfunc* will *not* be called. This can occur if the collective involved only processes on the local node.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in [Section 3.1.1](#).

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

[PMIX_COLLECT_DATA](#) "pmix.collect" (bool)

Collect all data posted by the participants using [PMIx_Put](#) that has been committed via [PMIx_Commit](#), making the collection locally available to each participant at the end of the operation. By default, this will include all job-level information that was locally generated by PMIx servers unless excluded using the [PMIX_COLLECT_GENERATED_JOB_INFO](#) attribute.

[PMIX_COLLECT_GENERATED_JOB_INFO](#) "pmix.collect.gen" (bool)

Collect all job-level information (i.e., reserved keys) that was locally generated by PMIx servers. Some job-level information (e.g., distance between processes and fabric devices) is best determined on a distributed basis as it primarily pertains to local processes. Should remote processes need to access the information, it can either be obtained collectively using the [PMIx_Fence](#) operation with this directive, or can be retrieved one peer at a time using [PMIx_Get](#) without first having performed the job-wide collection.

Optional Attributes

The following attributes are optional for PMIx implementations:

PMIX_ALL_CLONES_PARTICIPATE "pmix.clone.part" (bool)

All *clones* of the calling process must participate in the collective operation.

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Nonblocking version of the **PMIX_Fence** routine. See the **PMIX_Fence** description for further details.

8.2.1 Fence-related attributes

The following attributes are defined specifically to support the fence operation:

PMIX_COLLECT_DATA "pmix.collect" (bool)

Collect all data posted by the participants using **PMIX_Put** that has been committed via **PMIX_Commit**, making the collection locally available to each participant at the end of the operation. By default, this will include all job-level information that was locally generated by PMIx servers unless excluded using the **PMIX_COLLECT_GENERATED_JOB_INFO** attribute.

PMIX_LOCAL_COLLECTIVE_STATUS "pmix.loc.col.st" (**pmix_status_t**) (*Provisional*)

Status code for local collective operation being reported to the host by the server library. PMIx servers may aggregate the participation by local client processes in a collective operation - e.g., instead of passing individual client calls to **PMIX_Fence** up to the host environment, the server may pass only a single call to the host when all local participants have executed their **PMIX_Fence** call, thereby reducing the burden placed on the host. However, in cases where the operation locally fails (e.g., if a participating client abnormally terminates prior to calling the operation), the server upcall functions to the host do not include a **pmix_status_t** by which the PMIx server can alert the host to that failure. This attribute resolves that problem by allowing the server to pass the status information regarding the local collective operation.

Advice to PMIx server hosts

The PMIx server is allowed to pass **PMIX_SUCCESS** using this attribute, but is not required to do so. PMIx implementations may choose to only report errors in this manner. The lack of an included status shall therefore be taken to indicate that the collective operation locally succeeded.

1 **PMIX_COLLECT_GENERATED_JOB_INFO** "pmix.collect.gen" (bool)
2 Collect all job-level information (i.e., reserved keys) that was locally generated by PMIx servers. Some
3 job-level information (e.g., distance between processes and fabric devices) is best determined on a
4 distributed basis as it primarily pertains to local processes. Should remote processes need to access the
5 information, it can either be obtained collectively using the **PMIx_Fence** operation with this
6 directive, or can be retrieved one peer at a time using **PMIx_Get** without first having performed the
7 job-wide collection.

8 **PMIX_ALL_CLONES_PARTICIPATE** "pmix.clone.part" (bool)
9 All *clones* of the calling process must participate in the collective operation.

CHAPTER 9

Publish/Lookup Operations

Chapter 6 and Chapter 5.1 discussed how reserved and non-reserved keys dealt with information that either was associated with a specific process (i.e., the retrieving process knew the identifier of the process that posted it) or required a synchronization operation prior to retrieval (e.g., the case of globally unique non-reserved keys). However, another requirement exists for an asynchronous exchange of data where neither the posting nor the retrieving process is known in advance. For example, two separate namespaces may need to rendezvous with each other without knowing in advance the identity of the other namespace or when that namespace might become active.

The APIs defined in this section focus on resolving that specific situation by allowing processes to publish data that can subsequently be retrieved solely by referral to its key. Mechanisms for constraining availability of the information are also provided as a means for better targeting of the eventual recipient(s).

Note that no presumption is made regarding how the published information is to be stored, nor as to the entity (host environment or PMIx implementation) that shall act as the datastore. The descriptions in the remainder of this chapter shall simply refer to that entity as the *datastore*.

9.1 PMIx_Publish

Summary

Publish data for later access via [PMIx_Lookup](#).

Format

```
pmix_status_t
PMIx_Publish(const pmix_info_t info[], size_t ninfo);
```

IN *info*

Array of info structures containing both data to be published and directives (array of handles)

IN *ninfo*

Number of elements in the *info* array (integer)

Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Required Attributes

There are no required attributes for this API. PMIx implementations that do not directly support the operation but are hosted by environments that do support it must pass any attributes that are provided by the client to the host environment for processing. In addition, the PMIx library is required to add the [PMIX_USERID](#) and the [PMIX_GRPID](#) attributes of the client process that published the information to the *info* array passed to the host environment.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

PMIX_RANGE "pmix.range" (pmix_data_range_t)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

PMIX_PERSISTENCE "pmix.persist" (pmix_persistence_t)

Declare how long the datastore shall retain the provided data. The datastore is to delete the data upon reaching the persistence criterion.

PMIX_ACCESS_PERMISSIONS "pmix.aperms" (pmix_data_array_t)

Define access permissions for the published data. The value shall contain an array of **pmix_info_t** structs containing the specified permissions.

Description

Publish the data in the *info* array for subsequent lookup. By default, the data will be published into the **PMIX_RANGE_SESSION** range and with **PMIX_PERSIST_APP** persistence. Changes to those values, and any additional directives, can be included in the **pmix_info_t** array. Attempts to access the data by processes outside of the provided data range shall be rejected. The **PMIX_PERSISTENCE** attribute instructs the datastore holding the published information as to how long that information is to be retained.

The blocking form of this call will block until it has obtained confirmation from the datastore that the data is available for lookup. The *info* array can be released upon return from the blocking function call.

Publishing duplicate keys is permitted provided they are published to different ranges. Duplicate keys being published on the same data range shall return the **PMIX_ERR_DUPLICATE_KEY** error.

9.2 PMIx_Publish_nb

Summary

Nonblocking **PMIx_Publish** routine.

Format

C

```
pmix_status_t
PMIx_Publish_nb(const pmix_info_t info[], size_t ninfo,
                pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

- IN** **info**
Array of info structures containing both data to be published and directives (array of handles)
- IN** **ninfo**
Number of elements in the *info* array (integer)
- IN** **cbfunc**
Callback function [pmix_op_cbfunc_t](#) (function reference)
- IN** **cbdata**
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

There are no required attributes for this API. PMIx implementations that do not directly support the operation but are hosted by environments that do support it must pass any attributes that are provided by the client to the host environment for processing. In addition, the PMIx library is required to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the client process that published the information to the *info* array passed to the host environment.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

PMIX_RANGE "pmix.range" (pmix_data_range_t)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

PMIX_PERSISTENCE "pmix.persist" (pmix_persistence_t)

1 Declare how long the datastore shall retain the provided data. The datastore is to delete the data upon
2 reaching the persistence criterion.
3 **PMIX_ACCESS_PERMISSIONS** "pmix.aperms" (pmix_data_array_t)
4 Define access permissions for the published data. The value shall contain an array of pmix_info_t
5 structs containing the specified permissions.



6 **Description**
7 Nonblocking **PMIx_Publish** routine.

8 9.3 Publish-specific constants

9 The following constants are defined for use with the **PMIx_Publish** APIs:

10 **PMIX_ERR_DUPLICATE_KEY** -53 The provided key has already been published on the same data
11 range.

12 9.4 Publish-specific attributes

13 The following attributes are defined for use with the **PMIx_Publish** APIs:

14 **PMIX_RANGE** "pmix.range" (pmix_data_range_t)
15 Define constraints on the processes that can access the provided data. Only processes that meet the
16 constraints are allowed to access it.
17 **PMIX_PERSISTENCE** "pmix.persist" (pmix_persistence_t)
18 Declare how long the datastore shall retain the provided data. The datastore is to delete the data upon
19 reaching the persistence criterion.
20 **PMIX_ACCESS_PERMISSIONS** "pmix.aperms" (pmix_data_array_t)
21 Define access permissions for the published data. The value shall contain an array of pmix_info_t
22 structs containing the specified permissions.
23 **PMIX_ACCESS_USERIDS** "pmix.auids" (pmix_data_array_t)
24 Array of effective User IDs (UIDs) that are allowed to access the published data.
25 **PMIX_ACCESS_GRPIDS** "pmix.agids" (pmix_data_array_t)
26 Array of effective Group IDs (GIDs) that are allowed to access the published data.

27 9.5 Publish-Lookup Datatypes

28 The following data types are defined for use with the **PMIx_Publish** APIs.

1 9.5.1 Range of Published Data

2 The `pmix_data_range_t` structure is a `uint8_t` type that defines a range for both data *published* via the
3 `PMIx_Publish` API and generated events. The following constants can be used to set a variable of the type
4 `pmix_data_range_t`.

5 `PMIX_RANGE_UNDEF` 0 Undefined range.

6 `PMIX_RANGE_RM` 1 Data is intended for the host environment, or lookup is restricted to data published
7 by the host environment.

8 `PMIX_RANGE_LOCAL` 2 Data is only available to processes on the local node, or lookup is restricted to
9 data published by processes on the local node of the requester.

10 `PMIX_RANGE_NAMESPACE` 3 Data is only available to processes in the same namespace, or lookup is
11 restricted to data published by processes in the same namespace as the requester.

12 `PMIX_RANGE_SESSION` 4 Data is only available to all processes in the session, or lookup is restricted
13 to data published by other processes in the same session as the requester.

14 `PMIX_RANGE_GLOBAL` 5 Data is available to all processes, or lookup is open to data published by
15 anyone.

16 `PMIX_RANGE_CUSTOM` 6 Data is available only to processes as specified in the `pmix_info_t`
17 associated with this call, or lookup is restricted to data published by processes as specified in the
18 `pmix_info_t`.

19 `PMIX_RANGE_PROC_LOCAL` 7 Data is only available to this process, or lookup is restricted to data
20 published by this process.

21 `PMIX_RANGE_INVALID` `UINT8_MAX` Invalid value - typically used to indicate that a range has not yet
22 been set.

23 9.5.2 Data Persistence Structure

24 *PMIx v1.0*

25 The `pmix_persistence_t` structure is a `uint8_t` type that defines the policy for data published by
26 clients via the `PMIx_Publish` API. The following constants can be used to set a variable of the type
27 `pmix_persistence_t`.

28 `PMIX_PERSIST_INDEF` 0 Retain data until specifically deleted.

29 `PMIX_PERSIST_FIRST_READ` 1 Retain data until the first access, then the data is deleted.

30 `PMIX_PERSIST_PROC` 2 Retain data until the publishing process terminates.

31 `PMIX_PERSIST_APP` 3 Retain data until the application terminates.

32 `PMIX_PERSIST_SESSION` 4 Retain data until the session/allocation terminates.

33 `PMIX_PERSIST_INVALID` `UINT8_MAX` Invalid value - typically used to indicate that a persistence
has not yet been set.

34 9.6 PMIx_Lookup

35 Summary

36 Lookup information published by this or another process with `PMIx_Publish` or `PMIx_Publish_nb`.

Format

C

```
pmix_status_t
PMIx_Lookup(pmix_pdata_t data[], size_t ndata,
            const pmix_info_t info[], size_t ninfo);
```

C

INOUT data

Array of publishable data structures (array of `pmix_pdata_t`)

IN `ndata`

Number of elements in the `data` array (integer)

IN `info`

Array of info structures (array of `pmix_info_t`)

IN `ninfo`

Number of elements in the `info` array (integer)

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

- **PMIX_ERR_NOT_FOUND** None of the requested data could be found within the requester's range.
- **PMIX_ERR_PARTIAL_SUCCESS** Some of the requested data was found. Any key that cannot be found will return with a data type of **PMIX_UNDEF** in the associated `value` struct. Note that the specific reason for a particular piece of missing information (e.g., lack of permissions) cannot be communicated back to the requester in this situation.
- **PMIX_ERR_NO_PERMISSIONS** All of the requested data was found and range restrictions were met for each specified key, but none of the matching data could be returned due to lack of access permissions.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host environment for processing, and the PMIx library is required to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the client process that is requesting the info.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (`int`)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

PMIX_RANGE "pmix.range" (`pmix_data_range_t`)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

PMIX_WAIT "pmix.wait" (`int`)

1 Caller requests that the PMIx server wait until at least the specified number of values are found (a value
2 of zero indicates *all* and is the default).



3 **Description**

4 Lookup information published by this or another process. By default, the search will be constrained to
5 publishers that fall within the `PMIX_RANGE_SESSION` range in case duplicate keys exist on different ranges.
6 Changes to the range (e.g., expanding the search to all potential publishers via the `PMIX_RANGE_GLOBAL`
7 constant), and any additional directives, can be provided in the `pmix_info_t` array. Data is returned per the
8 retrieval rules of Section 9.8.

9 The *data* parameter consists of an array of `pmix_pdata_t` structures with the keys specifying the requested
10 information. Data will be returned for each **key** field in the associated **value** field of this structure as per the
11 above description of return values. The **proc** field in each `pmix_pdata_t` structure will contain the
12 namespace/rank of the process that published the data.

▼ Advice to users ▼

13 Although this is a blocking function, it will not wait by default for the requested data to be published. Instead,
14 it will block for the time required by the datastore to lookup its current data and return any found items. Thus,
15 the caller is responsible for either ensuring that data is published prior to executing a lookup, using
16 `PMIX_WAIT` to instruct the datastore to wait for the data to be published, or retrying until the requested data is
17 found.



18 **9.7 PMIx_Lookup_nb**

19 **Summary**

20 Nonblocking version of `PMIx_Lookup`.

Format

C

```
pmix_status_t
PMIx_Lookup_nb(char **keys,
               const pmix_info_t info[], size_t ninfo,
               pmix_lookup_cbfunc_t cbfunc, void *cbdata);
```

C

- IN keys**
NULL-terminated array of keys (array of strings)
- IN info**
Array of info structures (array of handles)
- IN ninfo**
Number of elements in the *info* array (integer)
- IN cbfunc**
Callback function (handle)
- IN cbdata**
Callback data to be provided to the callback function (pointer)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

If executed, the status returned in the provided callback function will be one of the following constants:

- **PMIX_SUCCESS** All data was found and has been returned.
- **PMIX_ERR_NOT_FOUND** None of the requested data was available within the requester's range. The *pdata* array in the callback function shall be **NULL** and the *npdata* parameter set to zero.
- **PMIX_ERR_PARTIAL_SUCCESS** Some of the requested data was found. Only found data will be included in the returned *pdata* array. Note that the specific reason for a particular piece of missing information (e.g., lack of permissions) cannot be communicated back to the requester in this situation.
- **PMIX_ERR_NOT_SUPPORTED** There is no available datastore (either at the host environment or PMIx implementation level) on this system that supports this function.
- **PMIX_ERR_NO_PERMISSIONS** All of the requested data was found and range restrictions were met for each specified key, but none of the matching data could be returned due to lack of access permissions.
- a non-zero PMIx error constant indicating a reason for the request's failure.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host environment for processing, and the PMIx library is required to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the client process that is requesting the info.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

PMIX_RANGE "pmix.range" (pmix_data_range_t)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

PMIX_WAIT "pmix.wait" (int)

Caller requests that the PMIx server wait until at least the specified number of values are found (a value of zero indicates *all* and is the default).

Description

Non-blocking form of the **PMIx_Lookup** function.

9.7.1 Lookup Returned Data Structure

The **pmix_pdata_t** structure is used by **PMIx_Lookup** to describe the data being accessed.

PMIx v1.0

```
typedef struct pmix_pdata {
    pmix_proc_t proc;
    pmix_key_t key;
    pmix_value_t value;
} pmix_pdata_t;
```

where:

- *proc* is the process identifier of the data publisher.
- *key* is the string key of the published data.
- *value* is the value associated with the *key*.

9.7.1.1 Lookup data structure support macros

The following macros are provided to support the **pmix_pdata_t** structure.

1 **Static initializer for the pdata structure**

2 *(Provisional)*

3 Provide a static initializer for the `pmix_pdata_t` fields.



4 **PMIX_LOOKUP_STATIC_INIT**



5 **Initialize the pdata structure**

6 Initialize the `pmix_pdata_t` fields

PMIx v1.0



7 **PMIX_PDATA_CONSTRUCT (m)**



8 **IN m**

9 Pointer to the structure to be initialized (pointer to `pmix_pdata_t`)

10 **Destruct the pdata structure**

11 Destruct the `pmix_pdata_t` fields

PMIx v1.0



12 **PMIX_PDATA_DESTRUCT (m)**



13 **IN m**

14 Pointer to the structure to be destructed (pointer to `pmix_pdata_t`)

15 **Create a pdata array**

16 Allocate and initialize an array of `pmix_pdata_t` structures

PMIx v1.0



17 **PMIX_PDATA_CREATE (m, n)**



18 **INOUT m**

19 Address where the pointer to the array of `pmix_pdata_t` structures shall be stored (handle)

20 **IN n**

21 Number of structures to be allocated (`size_t`)

22 **Free a pdata structure**

23 Release a `pmix_pdata_t` structure

PMIx v4.0



24 **PMIX_PDATA_RELEASE (m)**



25 **IN m**

26 Pointer to a `pmix_pdata_t` structure (handle)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Free a pdata array

Release an array of `pmix_pdata_t` structures

PMIX_PDATA_FREE(m, n)

- IN** m
Pointer to the array of `pmix_pdata_t` structures (handle)
- IN** n
Number of structures in the array (`size_t`)

Load a lookup data structure

This macro simplifies the loading of key, process identifier, and data into a `pmix_pdata_t` by correctly assigning values to the structure's fields.

PMIx v1.0

PMIX_PDATA_LOAD(m, p, k, d, t);

- IN** m
Pointer to the `pmix_pdata_t` structure into which the key and data are to be loaded (pointer to `pmix_pdata_t`)
- IN** p
Pointer to the `pmix_proc_t` structure containing the identifier of the process being referenced (pointer to `pmix_proc_t`)
- IN** k
String key to be loaded - must be less than or equal to `PMIX_MAX_KEYLEN` in length (handle)
- IN** d
Pointer to the data value to be loaded (handle)
- IN** t
Type of the provided data value (`pmix_data_type_t`)

Advice to users

Key, process identifier, and data will all be copied into the `pmix_pdata_t` - thus, the source information can be modified or free'd without affecting the copied data once the macro has completed.

Transfer a lookup data structure

This macro simplifies the transfer of key, process identifier, and data value between two `pmix_pdata_t` structures.

```
PMIX_PDATA_XFER(d, s);
```

IN `d`
Pointer to the destination `pmix_pdata_t` (pointer to `pmix_pdata_t`)

IN `s`
Pointer to the source `pmix_pdata_t` (pointer to `pmix_pdata_t`)

Advice to users

Key, process identifier, and data will all be copied into the destination `pmix_pdata_t` - thus, the source `pmix_pdata_t` may free'd without affecting the copied data once the macro has completed.

9.7.2 Lookup Callback Function

Summary

The `pmix_lookup_cbfunc_t` is used by `PMIx_Lookup_nb` to return data.

PMIx v1.0

```
typedef void (*pmix_lookup_cbfunc_t)  
    (pmix_status_t status,  
     pmix_pdata_t data[], size_t ndata,  
     void *cbdata);
```

IN `status`
Status associated with the operation (handle)

IN `data`
Array of data returned (`pmix_pdata_t`)

IN `ndata`
Number of elements in the `data` array (`size_t`)

IN `cbdata`
Callback data passed to original API call (memory reference)

Description

A callback function for calls to `PMIx_Lookup_nb`. The function will be called upon completion of the `PMIx_Lookup_nb` API with the `status` indicating the success or failure of the request. Any retrieved data will be returned in an array of `pmix_pdata_t` structs. The namespace and rank of the process that provided each data element is also returned.

Note that the `pmix_pdata_t` structures will be released upon return from the callback function, so the receiver must copy/protect the data prior to returning if it needs to be retained.

1 9.8 Retrieval rules for published data

2 The retrieval rules for published data primarily revolve around enforcing data access permissions and range
3 constraints. The datastore shall search its stored information for each specified key according to the following
4 precedence logic:

- 5 1. If the requester specified the range, then the search shall be constrained to data where the publishing
6 process falls within the specified range.
- 7 2. If the key of the stored information does not match the specified key, then the search will continue.
- 8 3. If the requester's identifier does not fall within the range specified by the publisher, then the search will
9 continue.
- 10 4. If the publisher specified access permissions, the effective UID and GID of the requester shall be checked
11 against those permissions, with the datastore rejecting the match if the requester fails to meet the
12 requirements.
- 13 5. If all of the above checks pass, then the value is added to the information that is to be returned.

14 The status returned by the datastore shall be set to:

- 15 • **PMIX_SUCCESS** All data was found and is included in the returned information.
- 16 • **PMIX_ERR_NOT_FOUND** None of the requested data could be found within a requester's range.
- 17 • **PMIX_ERR_PARTIAL_SUCCESS** Some of the requested data was found. Only found data will be
18 included in the returned information. Note that the specific reason for a particular piece of missing
19 information (e.g., lack of permissions) cannot be communicated back to the requester in this situation.
- 20 • a non-zero PMIx error constant indicating a reason for the request's failure.

21 In the case where data was found and range restrictions were met for each specified key, but none of the
22 matching data could be returned due to lack of access permissions, the datastore must return the
23 **PMIX_ERR_NO_PERMISSIONS** error.

▼ Advice to users ▼

24 Note that duplicate keys are allowed to exist on different ranges, and that ranges do overlap each other. Thus, if
25 duplicate keys are published on overlapping ranges, it is possible for the datastore to successfully find multiple
26 responses for a given key should publisher and requester specify sufficiently broad ranges. In this situation, the
27 choice of resolving the duplication is left to the datastore implementation - e.g., it may return the first value
28 found in its search, or the value corresponding to the most limited range of the found values, or it may choose
29 to simply return an error.

30 Users are advised to avoid this ambiguity by careful selection of key values and ranges - e.g., by creating
31 range-specific keys where necessary.

32 9.9 PMIx_Unpublish

33 Summary

34 Unpublish data posted by this process using the given keys.

1 *PMIx v1.0*

Format

C

```
2 pmix_status_t  
3 PMIx_Unpublish(char **keys,  
4               const pmix_info_t info[], size_t ninfo);
```

C

5 **IN** **keys**
6 NULL-terminated array of keys (array of strings)
7 **IN** **info**
8 Array of info structures (array of handles)
9 **IN** **ninfo**
10 Number of elements in the *info* array (integer)

11 Returns **PMIX_SUCCESS** or a negative value indicating the error.

Required Attributes

12 PMIx libraries are not required to directly support any attributes for this function. However, any provided
13 attributes **MUST** be passed to the host environment for processing, and the PMIx library is required to add the
14 **PMIX_USERID** and the **PMIX_GRPID** attributes of the client process that is requesting the operation.

Optional Attributes

15 The following attributes are optional for host environments that support this operation:

16 **PMIX_TIMEOUT** "pmix.timeout" (**int**)

17 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
18 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
19 (client, server, and host) simultaneously timing the operation.

20 **PMIX_RANGE** "pmix.range" (**pmix_data_range_t**)

21 Define constraints on the processes that can access the provided data. Only processes that meet the
22 constraints are allowed to access it.

Description

24 Unpublish data posted by this process using the given *keys*. The function will block until the data has been
25 removed by the server (i.e., it is safe to publish that key again within the specified range). A value of **NULL** for
26 the *keys* parameter instructs the server to remove all data published by this process.

27 By default, the range is assumed to be **PMIX_RANGE_SESSION**. Changes to the range, and any additional
28 directives, can be provided in the *info* array.

9.10 PMIx_Unpublish_nb

Summary

30 Nonblocking version of **PMIx_Unpublish**.
31

Format

```
pmix_status_t
PMIx_Unpublish_nb(char **keys,
                  const pmix_info_t info[], size_t ninfo,
                  pmix_op_cbfunc_t cbfunc, void *cbdata);
```

IN keys
NULL-terminated array of keys (array of strings)

IN info
Array of info structures (array of handles)

IN ninfo
Number of elements in the *info* array (integer)

IN cbfunc
Callback function `pmix_op_cbfunc_t` (function reference)

IN cbdata
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host environment for processing, and the PMIx library is required to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the client process that is requesting the operation.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (**int**)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

PMIX_RANGE "pmix.range" (**pmix_data_range_t**)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

1
2
3

Description

Non-blocking form of the **PMIx_Unpublish** function. The callback function will be executed once the server confirms removal of the specified data. The *info* array must be maintained until the callback is provided.

CHAPTER 10

Event Notification

1 This chapter defines the PMIx event notification system. These interfaces are designed to support the reporting
2 of events to/from clients and servers, and between library layers within a single process.

3 10.1 Notification and Management

4 PMIx event notification provides an asynchronous out-of-band mechanism for communicating events between
5 application processes and/or elements of the SMS. Its uses span a wide range including fault notification,
6 coordination between multiple programming libraries within a single process, and workflow orchestration for
7 non-synchronous programming models. Events can be divided into two distinct classes:

- 8 • *Job-specific events* directly relate to a job executing within the session, such as a debugger attachment,
9 process failure within a related job, or events generated by an application process. Events in this category
10 are to be immediately delivered to the PMIx server library for relay to the related local processes.
- 11 • *Environment events* indirectly relate to a job but do not specifically target the job itself. This category
12 includes SMS-generated events such as Error Check and Correction (ECC) errors, temperature excursions,
13 and other non-job conditions that might directly affect a session's resources, but would never include an
14 event generated by an application process. Note that although these do potentially impact the session's jobs,
15 they are not directly tied to those jobs. Thus, events in this category are to be delivered to the PMIx server
16 library only upon request.

17 Both SMS elements and applications can register for events of either type.

▼ Advice to PMIx library implementers ▼

18 Race conditions can cause the registration to come after events of possible interest (e.g., a memory ECC event
19 that occurs after start of execution but prior to registration, or an application process generating an event prior
20 to another process registering to receive it). SMS vendors are *requested* to cache environment events for some
21 time to mitigate this situation, but are not *required* to do so. However, PMIx implementers are *required* to
22 cache all events received by the PMIx server library and to deliver them to registering clients in the same order
23 in which they were received

▼ Advice to users ▼

24 Applications must be aware that they may not receive environment events that occur prior to registration,
25 depending upon the capabilities of the host SMS.

1 The generator of an event can specify the *target range* for delivery of that event. Thus, the generator can
2 choose to limit notification to processes on the local node, processes within the same job as the generator,
3 processes within the same allocation, other threads within the same process, only the SMS (i.e., not to any
4 application processes), all application processes, or to a custom range based on specific process identifiers.
5 Only processes within the given range that register for the provided event code will be notified. In addition, the
6 generator can use attributes to direct that the event not be delivered to any default event handlers, or to any
7 multi-code handler (as defined below).

8 Event notifications provide the process identifier of the source of the event plus the event code and any
9 additional information provided by the generator. When an event notification is received by a process, the
10 registered handlers are scanned for their event code(s), with matching handlers assembled into an *event chain*
11 for servicing. Note that users can also specify a *source range* when registering an event (using the same range
12 designators described above) to further limit when they are to be invoked. When assembled, PMIx event
13 chains are ordered based on both the specificity of the event handler and user directives at time of handler
14 registration. By default, handlers are grouped into three categories based on the number of event codes that
15 can trigger the callback:

- 16 • *single-code* handlers are serviced first as they are the most specific. These are handlers that are registered
17 against one specific event code.
- 18 • *multi-code* handlers are serviced once all single-code handlers have completed. The handler will be
19 included in the chain upon receipt of an event matching any of the provided codes.
- 20 • *default* handlers are serviced once all multi-code handlers have completed. These handlers are always
21 included in the chain unless the generator specifically excludes them.

22 Users can specify the callback order of a handler within its category at the time of registration. Ordering can
23 be specified by providing the relevant event handler names, if the user specified an event handler name when
24 registering the corresponding event. Thus, users can specify that a given handler be executed before or after
25 another handler should both handlers appear in an event chain (the ordering is ignored if the other handler isn't
26 included). Note that ordering does not imply immediate relationships. For example, multiple handlers
27 registered to be serviced after event handler *A* will all be executed after *A*, but are not guaranteed to be
28 executed in any particular order amongst themselves.

29 In addition, one event handler can be declared as the *first* handler to be executed in the chain. This handler will
30 *always* be called prior to any other handler, regardless of category, provided the incoming event matches both
31 the specified range and event code. Only one handler can be so designated — attempts to designate additional
32 handlers as *first* will return an error. Deregistration of the declared *first* handler will re-open the position for
33 subsequent assignment.

34 Similarly, one event handler can be declared as the *last* handler to be executed in the chain. This handler will
35 *always* be called after all other handlers have executed, regardless of category, provided the incoming event
36 matches both the specified range and event code. Note that this handler will not be called if the chain is
37 terminated by an earlier handler. Only one handler can be designated as *last* — attempts to designate
38 additional handlers as *last* will return an error. Deregistration of the declared *last* handler will re-open the
39 position for subsequent assignment.

Advice to users

Note that the *last* handler is called *after* all registered default handlers that match the specified range of the incoming event unless a handler prior to it terminates the chain. Thus, if the application intends to define a *last* handler, it should ensure that no default handler aborts the process before it.

Upon completing its work and prior to returning, each handler *must* call the event handler completion function provided when it was invoked (including a status code plus any information to be passed to later handlers) so that the chain can continue being progressed. PMIx automatically aggregates the status and any results of each handler (as provided in the completion callback) with status from all prior handlers so that each step in the chain has full knowledge of what preceded it. An event handler can terminate all further progress along the chain by passing the `PMIX_EVENT_ACTION_COMPLETE` status to the completion callback function.

10.1.1 Events versus status constants

Return status constants (see Section 3.1.1) represent values that can be returned from or passed into PMIx APIs. These are distinct from PMIx *events* in that they are not values that can be registered against event handlers. In general, the two types of constants are distinguished by inclusion of an "ERR" in the name of error constants versus an "EVENT" in events, though there are exceptions (e.g, the `PMIX_SUCCESS` constant).

10.1.2 PMIx_Register_event_handler

Summary

Register an event handler.

Format

PMIx v2.0

C

```
pmix_status_t
PMIx_Register_event_handler(pmix_status_t codes[], size_t ncodes,
                           pmix_info_t info[], size_t ninfo,
                           pmix_notification_fn_t evhdlr,
                           pmix_hdlr_reg_cbfunc_t cbfunc,
                           void *cbdata);
```

C

IN codes
Array of status codes (array of `pmix_status_t`)

IN ncodes
Number of elements in the *codes* array (`size_t`)

IN info
Array of info structures (array of handles)

IN ninfo
Number of elements in the *info* array (`size_t`)

IN evhdlr
Event handler to be called `pmix_notification_fn_t` (function reference)

1 **IN cbfunc**
2 Callback function `pmix_hdlr_reg_cbfunc_t` (function reference)
3 **IN cbdata**
4 Data to be passed to the `cbfunc` callback function (memory reference)

5 If `cbfunc` is **NULL**, the function call will be treated as a *blocking* call. In this case, the returned status will be
6 either (a) the event handler reference identifier if the value is greater than or equal to zero, or (b) a negative
7 error code indicative of the reason for the failure.

8 If the `cbfunc` is non-**NULL**, the function call will be treated as a *non-blocking* call and will return the following:

9 A successful return indicates that the request is being processed and the result will be returned in the provided
10 `cbfunc`. Note that the library must not invoke the callback function prior to returning from the API. The
11 callback function, `cbfunc`, is only called when **PMIX_SUCCESS** is returned. The result of the registration
12 operation shall be returned in the provided callback function along with the assigned event handler identifier.

13 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- 14 • **PMIX_ERR_EVENT_REGISTRATION** indicating that the registration has failed for an undetermined
15 reason.

16 If none of the above return codes are appropriate, then an implementation must return either a general PMIX
17 error code or an implementation defined error code as described in Section 3.1.1.

18 The callback function must not be executed prior to returning from the API, and no events corresponding to
19 this registration may be delivered prior to the completion of the registration callback function (`cbfunc`).

▼----- Required Attributes -----▼

20 The following attributes are required to be supported by all PMIX libraries:

21 **PMIX_EVENT_HDLR_NAME** "pmix.evname" (**char***)
22 String name identifying this handler.

23 **PMIX_EVENT_HDLR_FIRST** "pmix.evfirst" (**bool**)
24 Invoke this event handler before any other handlers.

25 **PMIX_EVENT_HDLR_LAST** "pmix.evlast" (**bool**)
26 Invoke this event handler after all other handlers have been called.

27 **PMIX_EVENT_HDLR_FIRST_IN_CATEGORY** "pmix.evfirstcat" (**bool**)
28 Invoke this event handler before any other handlers in this category.

29 **PMIX_EVENT_HDLR_LAST_IN_CATEGORY** "pmix.evlastcat" (**bool**)
30 Invoke this event handler after all other handlers in this category have been called.

31 **PMIX_EVENT_HDLR_BEFORE** "pmix.evbefore" (**char***)
32 Put this event handler immediately before the one specified in the (**char***) value.

33 **PMIX_EVENT_HDLR_AFTER** "pmix.evafter" (**char***)
34 Put this event handler immediately after the one specified in the (**char***) value.

35 **PMIX_EVENT_HDLR_PREPEND** "pmix.evprepend" (**bool**)
36 Prepend this handler to the precedence list within its category.

1 **PMIX_EVENT_HDLR_APPEND** "pmix.evappend" (bool)
 2 Append this handler to the precedence list within its category.

3 **PMIX_EVENT_CUSTOM_RANGE** "pmix.evrangle" (pmix_data_array_t*)
 4 Array of **pmix_proc_t** defining range of event notification.

5 **PMIX_RANGE** "pmix.range" (pmix_data_range_t)
 6 Define constraints on the processes that can access the provided data. Only processes that meet the
 7 constraints are allowed to access it.

8 **PMIX_EVENT_RETURN_OBJECT** "pmix.evobject" (void *)
 9 Object to be returned whenever the registered callback function **cbfunc** is invoked. The object will
 10 only be returned to the process that registered it.

12 Host environments that implement support for PMIx event notification are required to support the following
 13 attributes when registering handlers - these attributes are used to direct that the handler should be invoked only
 14 when the event affects the indicated process(es):

15 **PMIX_EVENT_AFFECTED_PROC** "pmix.evproc" (pmix_proc_t)
 16 The single process that was affected.

17 **PMIX_EVENT_AFFECTED_PROCS** "pmix.evaffected" (pmix_data_array_t*)
 18 Array of **pmix_proc_t** defining affected processes.



19 Description

20 Register an event handler to report events. Note that the codes being registered do *not* need to be PMIx error
 21 constants — any integer value can be registered. This allows for registration of non-PMIx events such as those
 22 defined by a particular SMS vendor or by an application itself.



Advice to users

23 In order to avoid potential conflicts, users are advised to only define codes that lie outside the range of the
 24 PMIx standard's error codes. Thus, SMS vendors and application developers should constrain their definitions
 25 to positive values or negative values beyond the **PMIX_EXTERNAL_ERR_BASE** boundary.



Advice to users

26 As previously stated, upon completing its work, and prior to returning, each handler *must* call the event
 27 handler completion function provided when it was invoked (including a status code plus any information to be
 28 passed to later handlers) so that the chain can continue being progressed. An event handler can terminate all
 29 further progress along the chain by passing the **PMIX_EVENT_ACTION_COMPLETE** status to the completion
 30 callback function. Note that the parameters passed to the event handler (e.g., the *info* and *results* arrays) will
 31 cease to be valid once the completion function has been called - thus, any information in the incoming
 32 parameters that will be referenced following the call to the completion function must be copied.



1 10.1.3 Event registration constants

2 **PMIX_ERR_EVENT_REGISTRATION** -144 Error in event registration.

3 10.1.4 System events

4 **PMIX_EVENT_SYS_BASE** -230 Mark the beginning of a dedicated range of constants for system event
5 reporting.

6 **PMIX_EVENT_NODE_DOWN** -231 A node has gone down - the identifier of the affected node will be
7 included in the notification.

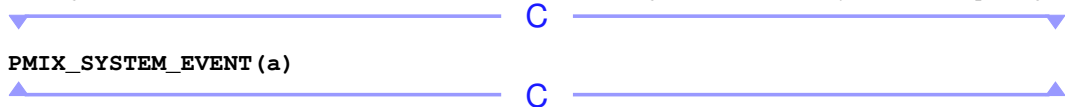
8 **PMIX_EVENT_NODE_OFFLINE** -232 A node has been marked as *offline* - the identifier of the affected
9 node will be included in the notification.

10 **PMIX_EVENT_SYS_OTHER** -330 Mark the end of a dedicated range of constants for system event
11 reporting.

12 Detect system event constant

13 Test a given event constant to see if it falls within the dedicated range of constants for system event reporting.

PMIx v2.2



15 **IN** a

16 Error constant to be checked (**pmix_status_t**)

17 Returns **true** if the provided values falls within the dedicated range of events for system event reporting.

18 10.1.5 Event handler registration and notification attributes

19 Attributes to support event registration and notification.

20 **PMIX_EVENT_HDLR_NAME** "pmix.evname" (**char***)

21 String name identifying this handler.

22 **PMIX_EVENT_HDLR_FIRST** "pmix.evfirst" (**bool**)

23 Invoke this event handler before any other handlers.

24 **PMIX_EVENT_HDLR_LAST** "pmix.evlast" (**bool**)

25 Invoke this event handler after all other handlers have been called.

26 **PMIX_EVENT_HDLR_FIRST_IN_CATEGORY** "pmix.evfirstcat" (**bool**)

27 Invoke this event handler before any other handlers in this category.

28 **PMIX_EVENT_HDLR_LAST_IN_CATEGORY** "pmix.evlastcat" (**bool**)

29 Invoke this event handler after all other handlers in this category have been called.

30 **PMIX_EVENT_HDLR_BEFORE** "pmix.evbefore" (**char***)

31 Put this event handler immediately before the one specified in the (**char***) value.

32 **PMIX_EVENT_HDLR_AFTER** "pmix.evafter" (**char***)

33 Put this event handler immediately after the one specified in the (**char***) value.

34 **PMIX_EVENT_HDLR_PREPEND** "pmix.evprepend" (**bool**)

35 Prepend this handler to the precedence list within its category.

36 **PMIX_EVENT_HDLR_APPEND** "pmix.evappend" (**bool**)

1 Append this handler to the precedence list within its category.

2 **PMIX_EVENT_CUSTOM_RANGE** "pmix.evrange" (pmix_data_array_t*)

3 Array of pmix_proc_t defining range of event notification.

4 **PMIX_EVENT_AFFECTED_PROC** "pmix.evproc" (pmix_proc_t)

5 The single process that was affected.

6 **PMIX_EVENT_AFFECTED_PROCS** "pmix.evaffected" (pmix_data_array_t*)

7 Array of pmix_proc_t defining affected processes.

8 **PMIX_EVENT_NON_DEFAULT** "pmix.evnondef" (bool)

9 Event is not to be delivered to default event handlers.

10 **PMIX_EVENT_RETURN_OBJECT** "pmix.evobject" (void *)

11 Object to be returned whenever the registered callback function cbfunc is invoked. The object will
12 only be returned to the process that registered it.

13 **PMIX_EVENT_DO_NOT_CACHE** "pmix.evnocache" (bool)

14 Instruct the PMIx server not to cache the event.

15 **PMIX_EVENT_PROXY** "pmix.evproxy" (pmix_proc_t*)

16 PMIx server that sourced the event.

17 **PMIX_EVENT_TEXT_MESSAGE** "pmix.evtext" (char*)

18 Text message suitable for output by recipient - e.g., describing the cause of the event.

19 **PMIX_EVENT_TIMESTAMP** "pmix.evtstamp" (time_t)

20 System time when the associated event occurred.

21 10.1.5.1 Fault tolerance event attributes

22 The following attributes may be used by the host environment when providing an event notification as
23 qualifiers indicating the action it intends to take in response to the event:

24 **PMIX_EVENT_TERMINATE_SESSION** "pmix.evterm.sess" (bool)

25 The RM intends to terminate this session.

26 **PMIX_EVENT_TERMINATE_JOB** "pmix.evterm.job" (bool)

27 The RM intends to terminate this job.

28 **PMIX_EVENT_TERMINATE_NODE** "pmix.evterm.node" (bool)

29 The RM intends to terminate all processes on this node.

30 **PMIX_EVENT_TERMINATE_PROC** "pmix.evterm.proc" (bool)

31 The RM intends to terminate just this process.

32 **PMIX_EVENT_ACTION_TIMEOUT** "pmix.evtimeout" (int)

33 The time in seconds before the RM will execute the indicated operation.

34 10.1.5.2 Hybrid programming event attributes

35 The following attributes may be used by programming models to coordinate their use of common resources
36 within a process in conjunction with the **PMIX_OPENMP_PARALLEL_ENTERED** event:

37 **PMIX_MODEL_PHASE_NAME** "pmix.mdl.phase" (char*)

38 User-assigned name for a phase in the application execution (e.g., "cfd reduction").

39 **PMIX_MODEL_PHASE_TYPE** "pmix.mdl.ptype" (char*)

40 Type of phase being executed (e.g., "matrix multiply").

1 10.1.6 Notification Function

2 Summary

3 The `pmix_notification_fn_t` is called by PMIx to deliver notification of an event.

Advice to users

4 The PMIx *ad hoc* v1.0 Standard defined an error notification function with an identical name, but different
5 signature than the v2.0 Standard described below. The *ad hoc* v1.0 version was removed from the v2.0
6 Standard is not included in this document to avoid confusion.

PMIx v2.0

C

```
7 typedef void (*pmix_notification_fn_t)
8     (size_t evhdlr_registration_id,
9      pmix_status_t status,
10     const pmix_proc_t *source,
11     pmix_info_t info[], size_t ninfo,
12     pmix_info_t results[], size_t nresults,
13     pmix_event_notification_cbfunc_fn_t cbfunc,
14     void *cbdata);
```

C

15 **IN** `evhdlr_registration_id`
16 Registration number of the handler being called (`size_t`)

17 **IN** `status`
18 Status associated with the operation (`pmix_status_t`)

19 **IN** `source`
20 Identifier of the process that generated the event (`pmix_proc_t`). If the source is the SMS, then the
21 nspace will be empty and the rank will be `PMIX_RANK_UNDEF`

22 **IN** `info`
23 Information describing the event (`pmix_info_t`). This argument will be `NULL` if no additional
24 information was provided by the event generator.

25 **IN** `ninfo`
26 Number of elements in the info array (`size_t`)

27 **IN** `results`
28 Aggregated results from prior event handlers servicing this event (`pmix_info_t`). This argument will
29 be `NULL` if this is the first handler servicing the event, or if no prior handlers provided results.

30 **IN** `nresults`
31 Number of elements in the results array (`size_t`)

32 **IN** `cbfunc`
33 `pmix_event_notification_cbfunc_fn_t` callback function to be executed upon completion
34 of the handler's operation and prior to handler return (function reference).

35 **IN** `cbdata`
36 Callback data to be passed to `cbfunc` (memory reference)

Description

Note that different RMs may provide differing levels of support for event notification to application processes. Thus, the *info* array may be **NULL** or may contain detailed information of the event. It is the responsibility of the application to parse any provided info array for defined key-values if it so desires.

Advice to users

Possible uses of the *info* array include:

- for the host RM to alert the process as to planned actions, such as aborting the session, in response to the reported event
- provide a timeout for alternative action to occur, such as for the application to request an alternate response to the event

For example, the RM might alert the application to the failure of a node that resulted in termination of several processes, and indicate that the overall session will be aborted unless the application requests an alternative behavior in the next 5 seconds. The application then has time to respond with a checkpoint request, or a request to recover from the failure by obtaining replacement nodes and restarting from some earlier checkpoint.

Support for these options is left to the discretion of the host RM. Info keys are included in the common definitions above but may be augmented by environment vendors.

Advice to PMIx server hosts

On the server side, the notification function is used to inform the PMIx server library's host of a detected event in the PMIx server library. Events generated by PMIx clients are communicated to the PMIx server library, but will be relayed to the host via the `pmix_server_notify_event_fn_t` function pointer, if provided.

10.1.7 PMIx_Deregister_event_handler

Summary

Deregister an event handler.

Format

C

```
pmix_status_t
PMIx_Deregister_event_handler(size_t evhdlr_ref,
                              pmix_op_cbfunc_t cbfunc,
                              void *cbdata);
```

C

IN **evhdlr_ref**
Event handler ID returned by registration (**size_t**)

IN **cbfunc**
Callback function to be executed upon completion of operation **pmix_op_cbfunc_t** (function reference)

IN **cbdata**
Data to be passed to the cbfunc callback function (memory reference)

If *cbfunc* is **NULL**, the function will be treated as a *blocking* call and the result of the operation returned in the status code.

If *cbfunc* is non-**NULL**, the function will be treated as a *non-blocking* call.

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

- **PMIX_OPERATION_SUCCEEDED**, returned when the request was immediately processed successfully - the *cbfunc* will *not* be called.

The returned status code of *cbfunc* will be one of the following:

- **PMIX_SUCCESS** The event handler was successfully deregistered.
- **PMIX_ERR_BAD_PARAM** The provided *evhdlr_ref* was unrecognized.
- **PMIX_ERR_NOT_SUPPORTED** The PMIx implementation does not support event notification.

Description

Deregister an event handler. Note that no events corresponding to the referenced registration may be delivered following completion of the deregistration operation (either return from the API with **PMIX_OPERATION_SUCCEEDED** or execution of the *cbfunc*).

10.1.8 PMIx_Notify_event

Summary

Report an event for notification via any registered event handler.

Format

C

```
pmix_status_t
PMIx_Notify_event(pmix_status_t status,
                 const pmix_proc_t *source,
                 pmix_data_range_t range,
                 pmix_info_t info[], size_t ninfo,
                 pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

IN status
Status code of the event ([pmix_status_t](#))

IN source
Pointer to a [pmix_proc_t](#) identifying the original reporter of the event (handle)

IN range
Range across which this notification shall be delivered ([pmix_data_range_t](#))

IN info
Array of [pmix_info_t](#) structures containing any further info provided by the originator of the event (array of handles)

IN ninfo
Number of elements in the *info* array ([size_t](#))

IN cbfunc
Callback function to be executed upon completion of operation [pmix_op_cbfunc_t](#) (function reference)

IN cbdata
Data to be passed to the *cbfunc* callback function (memory reference)

If *cbfunc* is **NULL**, the function will be treated as a *blocking* call and the result of the operation returned in the status code.

If *cbfunc* is non-**NULL**, the function will be treated as a *non-blocking* call.

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned. Note that a successful call does *not* reflect the success or failure of delivering the event to any recipients.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, returned when the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

PMIX_EVENT_NON_DEFAULT "pmix.evnondef" (**bool**)
Event is not to be delivered to default event handlers.

1 **PMIX_EVENT_CUSTOM_RANGE** "pmix.evrange" (pmix_data_array_t*)
2 Array of **pmix_proc_t** defining range of event notification.
3 **PMIX_EVENT_DO_NOT_CACHE** "pmix.evnocache" (bool)
4 Instruct the PMIx server not to cache the event.
5 **PMIX_EVENT_PROXY** "pmix.evproxy" (pmix_proc_t*)
6 PMIx server that sourced the event.
7 **PMIX_EVENT_TEXT_MESSAGE** "pmix.evtext" (char*)
8 Text message suitable for output by recipient - e.g., describing the cause of the event.

9

10 Host environments that implement support for PMIx event notification are required to provide the following
11 attributes for all events generated by the environment:

12 **PMIX_EVENT_AFFECTED_PROC** "pmix.evproc" (pmix_proc_t)
13 The single process that was affected.
14 **PMIX_EVENT_AFFECTED_PROCS** "pmix.evaffected" (pmix_data_array_t*)
15 Array of **pmix_proc_t** defining affected processes.



▼----- Optional Attributes -----▼

16 Host environments that support PMIx event notification may offer notifications for environmental events
17 impacting the job and for SMS events relating to the job. The following attributes may optionally be included
18 to indicate the host environment's intended response to the event:

19 **PMIX_EVENT_TERMINATE_SESSION** "pmix.evterm.sess" (bool)
20 The RM intends to terminate this session.
21 **PMIX_EVENT_TERMINATE_JOB** "pmix.evterm.job" (bool)
22 The RM intends to terminate this job.
23 **PMIX_EVENT_TERMINATE_NODE** "pmix.evterm.node" (bool)
24 The RM intends to terminate all processes on this node.
25 **PMIX_EVENT_TERMINATE_PROC** "pmix.evterm.proc" (bool)
26 The RM intends to terminate just this process.
27 **PMIX_EVENT_ACTION_TIMEOUT** "pmix.evtimeout" (int)
28 The time in seconds before the RM will execute the indicated operation.



Description

Report an event for notification via any registered event handler. This function can be called by any PMIx process, including application processes, PMIx servers, and SMS elements. The PMIx server calls this API to report events it detected itself so that the host SMS daemon distribute and handle them, and to pass events given to it by its host down to any attached client processes for processing. Examples might include notification of the failure of another process, detection of an impending node failure due to rising temperatures, or an intent to preempt the application. Events may be locally generated or come from anywhere in the system.

Host SMS daemons call the API to pass events down to its embedded PMIx server both for transmittal to local client processes and for the host's own internal processing where the host has registered its own event handlers. The PMIx server library is not allowed to echo any event given to it by its host via this API back to the host through the `pmix_server_notify_event_fn_t` server module function. The host is required to deliver the event to all PMIx servers where the targeted processes either are currently running, or (if they haven't started yet) might be running at some point in the future as the events are required to be cached by the PMIx server library.

Client application processes can call this function to notify the SMS and/or other application processes of an event it encountered. Note that processes are not constrained to report status values defined in the official PMIx standard — any integer value can be used. Thus, applications are free to define their own internal events and use the notification system for their own internal purposes.

Advice to users

The callback function will be called upon completion of the `notify_event` function's actions. At that time, any messages required for executing the operation (e.g., to send the notification to the local PMIx server) will have been queued, but may not yet have been transmitted. The caller is required to maintain the input data until the callback function has been executed — the sole purpose of the callback function is to indicate when the input data is no longer required.

10.1.9 Notification Handler Completion Callback Function

Summary

The `pmix_event_notification_cbfunc_fn_t` is called by event handlers to indicate completion of their operations.

PMIx v2.0

C

```
typedef void (*pmix_event_notification_cbfunc_fn_t)
    (pmix_status_t status,
     pmix_info_t *results, size_t nresults,
     pmix_op_cbfunc_t cbfunc, void *thiscbdata,
     void *notification_cbdata);
```



```

1  IN  status
2      Status returned by the event handler's operation (pmix_status_t)
3  IN  results
4      Results from this event handler's operation on the event (pmix_info_t)
5  IN  nresults
6      Number of elements in the results array (size_t)
7  IN  cbfunc
8      pmix_op_cbfunc_t function to be executed when PMIx completes processing the callback (function
9      reference)
10 IN  thiscbdata
11     Callback data that was passed in to the handler (memory reference)
12 IN  cbdata
13     Callback data to be returned when PMIx executes cbfunc (memory reference)

```

Description

Define a callback by which an event handler can notify the PMIx library that it has completed its response to the notification. The handler is *required* to execute this callback so the library can determine if additional handlers need to be called. The handler shall return `PMIX_EVENT_ACTION_COMPLETE` if no further action is required. The return status of each event handler and any returned `pmix_info_t` structures will be added to the *results* array of `pmix_info_t` passed to any subsequent event handlers to help guide their operation.

If non-`NULL`, the provided callback function will be called to allow the event handler to release the provided info array and execute any other required cleanup operations.

10.1.9.1 Completion Callback Function Status Codes

The following status code may be returned indicating various actions taken by other event handlers.

```

24  PMIX_EVENT_NO_ACTION_TAKEN   -331   Event handler: No action taken.
25  PMIX_EVENT_PARTIAL_ACTION_TAKEN -332   Event handler: Partial action taken.
26  PMIX_EVENT_ACTION_DEFERRED   -333   Event handler: Action deferred.
27  PMIX_EVENT_ACTION_COMPLETE   -334   Event handler: Action complete.

```

CHAPTER 11

Data Packing and Unpacking

1 PMIx intentionally does not include support for internode communications in the standard, instead relying on
2 its host SMS environment to transfer any needed data and/or requests between nodes. These operations
3 frequently involve PMIx-defined public data structures that include binary data. Many HPC clusters are
4 homogeneous, and so transferring the structures can be done rather simply. However, greater effort is required
5 in heterogeneous environments to ensure binary data is correctly transferred. PMIx buffer manipulation
6 functions are provided for this purpose via standardized interfaces to ease adoption.

7 11.1 Data Buffer Type

8 The `pmix_data_buffer_t` structure describes a data buffer used for packing and unpacking.

PMIx v2.0

C

```
9 typedef struct pmix_data_buffer {  
10     /** Start of my memory */  
11     char *base_ptr;  
12     /** Where the next data will be packed to  
13         (within the allocated memory starting  
14         at base_ptr) */  
15     char *pack_ptr;  
16     /** Where the next data will be unpacked  
17         from (within the allocated memory  
18         starting as base_ptr) */  
19     char *unpack_ptr;  
20     /** Number of bytes allocated (starting  
21         at base_ptr) */  
22     size_t bytes_allocated;  
23     /** Number of bytes used by the buffer  
24         (i.e., amount of data - including  
25         overhead - packed in the buffer) */  
26     size_t bytes_used;  
27 } pmix_data_buffer_t;
```

C

28 11.2 Support Macros

29 PMIx provides a set of convenience macros for creating, initiating, and releasing data buffers.

1 **Static initializer for the data buffer structure**

2 *(Provisional)*

3 Provide a static initializer for the `pmix_data_buffer_t` fields.

```
▼ _____ C _____ ▼
```

4 **PMIX_DATA_BUFFER_STATIC_INIT**

```
▲ _____ C _____ ▲
```

5 **PMIX_DATA_BUFFER_CREATE**

6 Allocate memory for a `pmix_data_buffer_t` object and initialize it. This macro uses `calloc` to allocate
7 memory for the buffer and initialize all fields in it

PMIx v2.0

```
▼ _____ C _____ ▼
```

8 **PMIX_DATA_BUFFER_CREATE(buffer);**

```
▲ _____ C _____ ▲
```

9 **OUT buffer**

10 Variable to be assigned the pointer to the allocated `pmix_data_buffer_t` (handle)

11 **PMIX_DATA_BUFFER_RELEASE**

12 Free a `pmix_data_buffer_t` object and the data it contains. Free's the data contained in the buffer, and
13 then free's the buffer itself

PMIx v2.0

```
▼ _____ C _____ ▼
```

14 **PMIX_DATA_BUFFER_RELEASE(buffer);**

```
▲ _____ C _____ ▲
```

15 **IN buffer**

16 Pointer to the `pmix_data_buffer_t` to be released (handle)

17 **PMIX_DATA_BUFFER_CONSTRUCT**

18 Initialize a statically declared `pmix_data_buffer_t` object.

PMIx v2.0

```
▼ _____ C _____ ▼
```

19 **PMIX_DATA_BUFFER_CONSTRUCT(buffer);**

```
▲ _____ C _____ ▲
```

20 **IN buffer**

21 Pointer to the allocated `pmix_data_buffer_t` that is to be initialized (handle)

22 **PMIX_DATA_BUFFER_DESTRUCT**

23 Release the data contained in a `pmix_data_buffer_t` object.

PMIx v2.0

```
▼ _____ C _____ ▼
```

24 **PMIX_DATA_BUFFER_DESTRUCT(buffer);**

```
▲ _____ C _____ ▲
```

25 **IN buffer**

26 Pointer to the `pmix_data_buffer_t` whose data is to be released (handle)

1 **PMIX_DATA_BUFFER_LOAD**

2 Load a blob into a `pmix_data_buffer_t` object. Load the given data into the provided
3 `pmix_data_buffer_t` object, usually done in preparation for unpacking the provided data. Note that the
4 data is *not* copied into the buffer - thus, the blob must not be released until after operations on the buffer have
5 completed.

PMIx v2.0

```
▼ _____ C _____ ▼  
6 PMIX_DATA_BUFFER_LOAD(buffer, data, size);  
▲ _____ C _____ ▲
```

7 **IN** `buffer`
8 Pointer to a pre-allocated `pmix_data_buffer_t` (handle)
9 **IN** `data`
10 Pointer to a blob (`char*`)
11 **IN** `size`
12 Number of bytes in the blob `size_t`

13 **PMIX_DATA_BUFFER_UNLOAD**

14 Unload the data from a `pmix_data_buffer_t` object. Extract the data in a buffer, assigning the pointer to
15 the data (and the number of bytes in the blob) to the provided variables, usually done to transmit the blob to a
16 remote process for unpacking. The buffer's internal pointer will be set to NULL to protect the data upon buffer
17 destruct or release - thus, the user is responsible for releasing the blob when done with it.

PMIx v2.0

```
▼ _____ C _____ ▼  
18 PMIX_DATA_BUFFER_UNLOAD(buffer, data, size);  
▲ _____ C _____ ▲
```

19 **IN** `buffer`
20 Pointer to the `pmix_data_buffer_t` whose data is to be extracted (handle)
21 **OUT** `data`
22 Variable to be assigned the pointer to the extracted blob (`void*`)
23 **OUT** `size`
24 Variable to be assigned the number of bytes in the blob `size_t`

25 **11.3 General Routines**

26 The following routines are provided to support internode transfers in heterogeneous environments.

27 **11.3.1 PMIx_Data_pack**

28 **Summary**

29 Pack one or more values of a specified type into a buffer, usually for transmission to another process.

Format

C

```
1 pmix_status_t
2 PMIx_Data_pack(const pmix_proc_t *target,
3                 pmix_data_buffer_t *buffer,
4                 void *src, int32_t num_vals,
5                 pmix_data_type_t type);
6
```

C

IN target

Pointer to a [pmix_proc_t](#) containing the nspace/rank of the process that will be unpacking the final buffer. A NULL value may be used to indicate that the target is based on the same PMIx version as the caller. Note that only the target's nspace is relevant. (handle)

IN buffer

Pointer to a [pmix_data_buffer_t](#) where the packed data is to be stored (handle)

IN src

Pointer to a location where the data resides. Strings are to be passed as (char **) — i.e., the caller must pass the address of the pointer to the string as the (void*). This allows the caller to pass multiple strings in a single call. (memory reference)

IN num_vals

Number of elements pointed to by the *src* pointer. A string value is counted as a single value regardless of length. The values must be contiguous in memory. Arrays of pointers (e.g., string arrays) should be contiguous, although the data pointed to need not be contiguous across array entries. ([int32_t](#))

IN type

The type of the data to be packed ([pmix_data_type_t](#))

Returns PMIX_SUCCESS or one of the following error codes when the condition described occurs:

[PMIX_ERR_BAD_PARAM](#) The provided buffer or src is NULL

[PMIX_ERR_UNKNOWN_DATA_TYPE](#) The specified data type is not known to this implementation

[PMIX_ERR_OUT_OF_RESOURCE](#) Not enough memory to support the operation

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

The pack function packs one or more values of a specified type into the specified buffer. The buffer must have already been initialized via the [PMIX_DATA_BUFFER_CREATE](#) or [PMIX_DATA_BUFFER_CONSTRUCT](#) macros — otherwise, [PMIx_Data_pack](#) will return an error. Providing an unsupported type flag will likewise be reported as an error.

Note that any data to be packed that is not hard type cast (i.e., not type cast to a specific size) may lose precision when unpacked by a non-homogeneous recipient. The [PMIx_Data_pack](#) function will do its best to deal with heterogeneity issues between the packer and unpacker in such cases. Sending a number larger than can be handled by the recipient will return an error code (generated upon unpacking) — the error cannot be detected during packing.

The namespace of the intended recipient of the packed buffer (i.e., the process that will be unpacking it) is used solely to resolve any data type differences between PMIx versions. The recipient must, therefore, be known to the user prior to calling the pack function so that the PMIx library is aware of the version the

recipient is using. Note that all processes in a given namespace are *required* to use the same PMIx version — thus, the caller must only know at least one process from the target’s namespace.

11.3.2 PMIx_Data_unpack

Summary

Unpack values from a `pmix_data_buffer_t`

Format

C

PMIx v2.0

```
pmix_status_t
PMIx_Data_unpack(const pmix_proc_t *source,
                 pmix_data_buffer_t *buffer, void *dest,
                 int32_t *max_num_values,
                 pmix_data_type_t type);
```

C

IN source

Pointer to a `pmix_proc_t` structure containing the nspace/rank of the process that packed the provided buffer. A NULL value may be used to indicate that the source is based on the same PMIx version as the caller. Note that only the source’s nspace is relevant. (handle)

IN buffer

A pointer to the buffer from which the value will be extracted. (handle)

INOUT dest

A pointer to the memory location into which the data is to be stored. Note that these values will be stored contiguously in memory. For strings, this pointer must be to (char**) to provide a means of supporting multiple string operations. The unpack function will allocate memory for each string in the array - the caller must only provide adequate memory for the array of pointers. (void*)

INOUT max_num_values

The number of values to be unpacked — upon completion, the parameter will be set to the actual number of values unpacked. In most cases, this should match the maximum number provided in the parameters — but in no case will it exceed the value of this parameter. Note that unpacking fewer values than are actually available will leave the buffer in an ununpackable state — the function will return an error code to warn of this condition. (int32_t)

IN type

The type of the data to be unpacked — must be one of the PMIx defined data types (`pmix_data_type_t`)

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

`PMIX_ERR_BAD_PARAM` The provided buffer or dest is `NULL`

`PMIX_ERR_UNKNOWN_DATA_TYPE` The specified data type is not known to this implementation

`PMIX_ERR_OUT_OF_RESOURCE` Not enough memory to support the operation

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

The unpack function unpacks the next value (or values) of a specified type from the given buffer. The buffer must have already been initialized via an `PMIX_DATA_BUFFER_CREATE` or `PMIX_DATA_BUFFER_CONSTRUCT` call (and assumedly filled with some data) — otherwise, the `unpack_value` function will return an error. Providing an unsupported type flag will likewise be reported as an error, as will specifying a data type that *does not* match the type of the next item in the buffer. An attempt to read beyond the end of the stored data held in the buffer will also return an error.

Note that it is possible for the buffer to be corrupted and that PMIx will *think* there is a proper variable type at the beginning of an unpack region — but that the value is bogus (e.g., just a byte field in a string array that so happens to have a value that matches the specified data type flag). Therefore, the data type error check is *not* completely safe.

Unpacking values is a "nondestructive" process — i.e., the values are not removed from the buffer. It is therefore possible for the caller to re-unpack a value from the same buffer by resetting the `unpack_ptr`.

Warning: The caller is responsible for providing adequate memory storage for the requested data. The user must provide a parameter indicating the maximum number of values that can be unpacked into the allocated memory. If more values exist in the buffer than can fit into the memory storage, then the function will unpack what it can fit into that location and return an error code indicating that the buffer was only partially unpacked.

Note that any data that was not hard type cast (i.e., not type cast to a specific size) when packed may lose precision when unpacked by a non-homogeneous recipient. PMIx will do its best to deal with heterogeneity issues between the packer and unpacker in such cases. Sending a number larger than can be handled by the recipient will return an error code generated upon unpacking — these errors cannot be detected during packing.

The namespace of the process that packed the buffer is used solely to resolve any data type differences between PMIx versions. The packer must, therefore, be known to the user prior to calling the pack function so that the PMIx library is aware of the version the packer is using. Note that all processes in a given namespace are *required* to use the same PMIx version — thus, the caller must only know at least one process from the packer's namespace.

11.3.3 PMIx_Data_copy

Summary

Copy a data value from one location to another.

Format

PMIx v2.0

```
pmix_status_t
PMIx_Data_copy(void **dest, void *src,
               pmix_data_type_t type);
```

IN dest

The address of a pointer into which the address of the resulting data is to be stored. (**void****)

IN src

A pointer to the memory location from which the data is to be copied (handle)

1 **IN type**
2 The type of the data to be copied — must be one of the PMIx defined data types.
3 ([pmix_data_type_t](#))

4 Returns PMIX_SUCCESS or one of the following error codes when the condition described occurs:

5 [PMIX_ERR_BAD_PARAM](#) The provided src or dest is **NULL**
6 [PMIX_ERR_UNKNOWN_DATA_TYPE](#) The specified data type is not known to this implementation
7 [PMIX_ERR_OUT_OF_RESOURCE](#) Not enough memory to support the operation

8 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
9 error code or an implementation defined error code as described in Section 3.1.1.

10 **Description**

11 Since registered data types can be complex structures, the system needs some way to know how to copy the
12 data from one location to another (e.g., for storage in the registry). This function, which can call other copy
13 functions to build up complex data types, defines the method for making a copy of the specified data type.

14 **11.3.4 PMIx_Data_print**

15 **Summary**

16 Pretty-print a data value.

17 **Format**

PMIx v2.0

```
18 pmix_status_t  
19 PMIx_Data_print(char **output, char *prefix,  
20                 void *src, pmix_data_type_t type);
```

21 **IN output**
22 The address of a pointer into which the address of the resulting output is to be stored. (**char****)

23 **IN prefix**
24 String to be prepended to the resulting output (**char***)

25 **IN src**
26 A pointer to the memory location of the data value to be printed (handle)

27 **IN type**
28 The type of the data value to be printed — must be one of the PMIx defined data types.
29 ([pmix_data_type_t](#))

30 Returns PMIX_SUCCESS or one of the following error codes when the condition described occurs:

31 [PMIX_ERR_BAD_PARAM](#) The provided data type is not recognized.

32 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
33 error code or an implementation defined error code as described in Section 3.1.1.

34 **Description**

35 Since registered data types can be complex structures, the system needs some way to know how to print them
36 (i.e., convert them to a string representation). Primarily for debug purposes.

1 11.3.5 PMIx_Data_copy_payload

2 Summary

3 Copy a payload from one buffer to another

4 *PMIx v2.0* Format

C

5 `pmix_status_t`

```
6 PMIx_Data_copy_payload(pmix_data_buffer_t *dest,  
7                       pmix_data_buffer_t *src);
```

C

8 **IN** `dest`

9 Pointer to the destination `pmix_data_buffer_t` (handle)

10 **IN** `src`

11 Pointer to the source `pmix_data_buffer_t` (handle)

12 Returns one of the following:

13 `PMIX_SUCCESS` The data has been copied as requested

14 `PMIX_ERR_BAD_PARAM` The `src` and `dest` `pmix_data_buffer_t` types do not match

15 `PMIX_ERR_NOT_SUPPORTED` The PMIx implementation does not support this function.

16 Description

17 This function will append a copy of the payload in one buffer into another buffer. Note that this is *not* a
18 destructive procedure — the source buffer’s payload will remain intact, as will any pre-existing payload in the
19 destination’s buffer. Only the unpacked portion of the source payload will be copied.

20 11.3.6 PMIx_Data_load

21 Summary

22 Load a buffer with the provided payload

23 *Provisional v4.1* Format

C

24 `pmix_status_t`

```
25 PMIx_Data_load(pmix_data_buffer_t *dest,  
26               pmix_byte_object_t *src);
```

C

27 **IN** `dest`

28 Pointer to the destination `pmix_data_buffer_t` (handle)

29 **IN** `src`

30 Pointer to the source `pmix_byte_object_t` (handle)

31 Returns one of the following:

32 `PMIX_SUCCESS` The data has been loaded as requested

33 `PMIX_ERR_BAD_PARAM` The `dest` structure pointer is `NULL`

34 `PMIX_ERR_NOT_SUPPORTED` The PMIx implementation does not support this function.

Description

The load function allows the caller to transfer the contents of the *src* `pmix_byte_object_t` to the *dest* target buffer. If a payload already exists in the buffer, the function will "free" the existing data to release it, and then replace the data payload with the one provided by the caller.

Advice to users

The buffer must be allocated or constructed in advance - failing to do so will cause the load function to return an error code.

The caller is responsible for pre-packing the provided payload. For example, the load function cannot convert to network byte order any data contained in the provided payload.

11.3.7 PMIx_Data_unload

Summary

Unload a buffer into a byte object

Format

Provisional
v4.1

```
pmix_status_t
PMIx_Data_unload(pmix_data_buffer_t *src,
                 pmix_byte_object_t *dest);
```

IN src

Pointer to the source `pmix_data_buffer_t` (handle)

IN dest

Pointer to the destination `pmix_byte_object_t` (handle)

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

`PMIX_ERR_BAD_PARAM` The destination and/or source pointer is `NULL`

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

The unload function provides the caller with a pointer to the portion of the data payload within the buffer that has not yet been unpacked, along with the size of that region. Any portion of the payload that was previously unpacked using the `PMIx_Data_unpack` routine will be ignored. This allows the user to directly access the payload.

Advice to users

This is a destructive operation. While the payload returned in the destination `pmix_byte_object_t` is undisturbed, the function will clear the *src*'s pointers to the payload. Thus, the *src* and the payload are completely separated, leaving the caller able to free or destruct the *src*.

1 11.3.8 PMIx_Data_compress

2 Summary

3 Perform a lossless compression on the provided data

4 *Provisional* Format

C

5 *v4.1*

```
6 bool  
7 PMIx_Data_compress(const uint8_t *inbytes, size_t size,  
8                   uint8_t **outbytes, size_t *nbytes);
```

C

8 **IN** *inbytes*

9 Pointer to the source data (handle)

10 **IN** *size*

11 Number of bytes in the source data region (*size_t*)

12 **OUT** *outbytes*

13 Address where the pointer to the compressed data region is to be returned (handle)

14 **OUT** *nbytes*

15 Address where the number of bytes in the compressed data region is to be returned (handle)

16 Returns one of the following:

- 17 • **True** The data has been compressed as requested
- 18 • **False** The data has not been compressed

19 Description

20 Compress the provided data block. Destination memory will be allocated if operation is successfully
21 concluded. Caller is responsible for release of the allocated region. The input data block will remain unaltered.

22 Note: the compress function will return **False** if the operation would not result in a smaller data block.

23 11.3.9 PMIx_Data_decompress

24 Summary

25 Decompress the provided data

26 *Provisional* Format

v4.1

```

1  bool
2  PMIx_Data_decompress(const uint8_t *inbytes, size_t size,
3                      uint8_t **outbytes, size_t *nbytes);

```

OUT outbytes
Address where the pointer to the decompressed data region is to be returned (handle)

OUT nbytes
Address where the number of bytes in the decompressed data region is to be returned (handle)

IN inbytes
Pointer to the source data (handle)

IN size
Number of bytes in the source data region (**size_t**)

Returns one of the following:

- **True** The data has been decompressed as requested
- **False** The data has not been decompressed

Description
Decompress the provided data block. Destination memory will be allocated if operation is successfully concluded. Caller is responsible for release of the allocated region. The input data block will remain unaltered.

Only data compressed by the **PMIx_Data_compress** API can be decompressed by this function. Passing data that has not been compressed by **PMIx_Data_compress** will lead to unexpected and potentially catastrophic results.

21 11.3.10 PMIx_Data_embed

22 *(Provisional)*

23 **Summary**
24 Embed a data payload into a buffer

25 *PMIx v5.0* **Format**

```

26  pmix_status_t
27  PMIx_Data_embed(pmix_data_buffer_t *buffer,
28                const pmix_byte_object_t *payload);

```

OUT buffer
Address of the buffer where the payload is to be embedded (handle)

IN payload
Address of the **pmix_byte_object_t** structure containing the data to be embedded into the buffer (handle)

1 Returns one of the following:
2 **PMIX_SUCCESS** The data has been embedded as requested
3 **PMIX_ERR_BAD_PARAM** The destination and/or source pointer is **NULL**
4 **PMIX_ERR_NOT_SUPPORTED** The PMIx implementation does not support this function.

5 **Description**

6 The embed function is identical in operation to **PMIx_Data_load** except that it does *not* clear the payload
7 object upon completion.

CHAPTER 12

Process Management

This chapter defines functionality processes can use to abort processes, spawn processes, and determine the relative locality of local processes.

12.1 Abort

PMIx provides a dedicated API by which an application can request that specified processes be aborted by the system.

12.1.1 PMIx_Abort

Summary

Abort the specified processes

Format

PMIx v1.0

C

```
pmix_status_t
PMIx_Abort(int status, const char msg[],
           pmix_proc_t procs[], size_t nprocs)
```

C

- IN status**
Error code to return to invoking environment (integer)
- IN msg**
String message to be returned to user (string)
- IN procs**
Array of `pmix_proc_t` structures (array of handles)
- IN nprocs**
Number of elements in the `procs` array (integer)

A successful return indicates that the requested processes are in a terminated state. Note that the function shall not return in this situation if the caller's own process was included in the request.

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

- **PMIX_ERR_PARAM_VALUE_NOT_SUPPORTED** if the PMIx implementation and host environment support this API, but the request includes processes that the host environment cannot abort - e.g., if the request is to abort subsets of processes from a namespace, or processes outside of the caller's own namespace, and the host environment does not permit such operations. In this case, none of the specified processes will be terminated.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

Request that the host resource manager print the provided message and abort the provided array of *procs*. A Unix or POSIX environment should handle the provided status as a return error code from the main program that launched the application. A **NULL** for the *procs* array indicates that all processes in the caller's namespace are to be aborted, including itself - this is the equivalent of passing a **pmix_proc_t** array element containing the caller's namespace and a rank value of **PMIX_RANK_WILDCARD**. While it is permitted for a caller to request abort of processes from namespaces other than its own, not all environments will support such requests. Passing a **NULL** *msg* parameter is allowed.

The function shall not return until the host environment has carried out the operation on the specified processes. If the caller is included in the array of targets, then the function will not return unless the host is unable to execute the operation.

Advice to users

The response to this request is somewhat dependent on the specific RM and its configuration (e.g., some resource managers will not abort the application if the provided status is zero unless specifically configured to do so, some cannot abort subsets of processes in an application, and some may not permit termination of processes outside of the caller's own namespace), and thus lies outside the control of PMIx itself. However, the PMIx client library shall inform the RM of the request that the specified *procs* be aborted, regardless of the value of the provided status.

Note that race conditions caused by multiple processes calling **PMIx_Abort** are left to the server implementation to resolve with regard to which status is returned and what messages (if any) are printed.

12.2 Process Creation

The **PMIx_Spawn** commands spawn new processes and/or applications in the PMIx universe. This may include requests to extend the existing resource allocation or obtain a new one, depending upon provided and supported attributes.

12.2.1 PMIx_Spawn

Summary

Spawn a new job.

Format

C

```
pmix_status_t
PMIx_Spawn(const pmix_info_t job_info[], size_t ninfo,
           const pmix_app_t apps[], size_t napps,
           char nspace[])
```

C

IN `job_info`
Array of info structures (array of handles)

IN `ninfo`
Number of elements in the `job_info` array (integer)

IN `apps`
Array of `pmix_app_t` structures (array of handles)

IN `napps`
Number of elements in the `apps` array (integer)

OUT `nspace`
Namespace of the new job (string)

Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host environment for processing.

Host environments are required to support the following attributes when present in either the `job_info` or the `info` array of an element of the `apps` array:

PMIX_WDIR "pmix.wdir" (char*)
Working directory for spawned processes.

PMIX_SET_SESSION_CWD "pmix.ssn cwd" (bool)
Set the current working directory to the session working directory assigned by the RM - can be assigned to the entire job (by including attribute in the `job_info` array) or on a per-application basis in the `info` array for each `pmix_app_t`.

PMIX_PREFIX "pmix.prefix" (char*)
Prefix to use for starting spawned processes - i.e., the directory where the executables can be found.

PMIX_HOST "pmix.host" (char*)
Comma-delimited list of hosts to use for spawned processes.

PMIX_HOSTFILE "pmix.hostfile" (char*)
Hostfile to use for spawned processes.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_ADD_HOSTFILE "pmix.addhostfile" (char*)

Hostfile containing hosts to add to existing allocation.

PMIX_ADD_HOST "pmix.addhost" (char*)

Comma-delimited list of hosts to add to the allocation.

PMIX_PRELOAD_BIN "pmix.preloadbin" (bool)

Preload executables onto nodes prior to executing launch procedure.

PMIX_PRELOAD_FILES "pmix.preloadfiles" (char*)

Comma-delimited list of files to pre-position on nodes prior to executing launch procedure.

PMIX_PERSONALITY "pmix.pers" (char*)

Name of personality corresponding to programming model used by application - supported values depend upon PMIx implementation.

PMIX_DISPLAY_MAP "pmix.dispmap" (bool)

Display process mapping upon spawn.

PMIX_PPR "pmix.ppr" (char*)

Number of processes to spawn on each identified resource.

PMIX_MAPBY "pmix.mapby" (char*)

Process mapping policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value for the rank to discover the mapping policy used for the provided namespace. Supported values are launcher specific.

PMIX_RANKBY "pmix.rankby" (char*)

Process ranking policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value for the rank to discover the ranking algorithm used for the provided namespace. Supported values are launcher specific.

PMIX_BINDTO "pmix.bindto" (char*)

Process binding policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value for the rank to discover the binding policy used for the provided namespace. Supported values are launcher specific.

PMIX_STDIN_TGT "pmix.stdin" (uint32_t)

Spawned process rank that is to receive any forwarded **stdin**.

PMIX_TAG_OUTPUT "pmix.tagout" (bool)

Tag **stdout/stderr** with the identity of the source process - can be assigned to the entire job (by including attribute in the *job_info* array) or on a per-application basis in the *info* array for each **pmix_app_t**.

PMIX_TIMESTAMP_OUTPUT "pmix.tsout" (bool)

Timestamp output - can be assigned to the entire job (by including attribute in the *job_info* array) or on a per-application basis in the *info* array for each **pmix_app_t**.

1 **PMIX_MERGE_STDERR_STDOUT** "pmix.mergeerrout" (bool)
2 Merge **stdout** and **stderr** streams - can be assigned to the entire job (by including attribute in the
3 *job_info* array) or on a per-application basis in the *info* array for each **pmix_app_t**.

4 **PMIX_OUTPUT_TO_FILE** "pmix.outfile" (char*)
5 Direct output (both stdout and stderr) into files of form "<filename>.rank" - can be assigned to
6 the entire job (by including attribute in the *job_info* array) or on a per-application basis in the *info* array
7 for each **pmix_app_t**.

8 **PMIX_INDEX_ARGV** "pmix.indxargv" (bool)
9 Mark the **argv** with the rank of the process.

10 **PMIX_CPUS_PER_PROC** "pmix.cputerproc" (uint32_t)
11 Number of PUs to assign to each rank - when accessed using **PMIx_Get**, use the
12 **PMIX_RANK_WILDCARD** value for the rank to discover the PUs/process assigned to the provided
13 namespace.

14 **PMIX_NO_PROCS_ON_HEAD** "pmix.nolocal" (bool)
15 Do not place processes on the head node.

16 **PMIX_NO_OVERSUBSCRIBE** "pmix.noover" (bool)
17 Do not oversubscribe the nodes - i.e., do not place more processes than allocated slots on a node.

18 **PMIX_REPORT_BINDINGS** "pmix.repbinding" (bool)
19 Report bindings of the individual processes.

20 **PMIX_CPU_LIST** "pmix.cpulist" (char*)
21 List of PUs to use for this job - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD**
22 value for the rank to discover the PU list used for the provided namespace.

23 **PMIX_JOB_RECOVERABLE** "pmix.recover" (bool)
24 Application supports recoverable operations.

25 **PMIX_JOB_CONTINUOUS** "pmix.continuous" (bool)
26 Application is continuous, all failed processes should be immediately restarted.

27 **PMIX_MAX_RESTARTS** "pmix.maxrestarts" (uint32_t)
28 Maximum number of times to restart a process - when accessed using **PMIx_Get**, use the
29 **PMIX_RANK_WILDCARD** value for the rank to discover the max restarts for the provided namespace.

30 **PMIX_SET_ENVAR** "pmix.envar.set" (pmix_envar_t*)
31 Set the envar to the given value, overwriting any pre-existing one

32 **PMIX_UNSET_ENVAR** "pmix.envar.unset" (char*)
33 Unset the environment variable specified in the string.

34 **PMIX_ADD_ENVAR** "pmix.envar.add" (pmix_envar_t*)
35 Add the environment variable, but do not overwrite any pre-existing one

36 **PMIX_PREPEND_ENVAR** "pmix.envar.prepnd" (pmix_envar_t*)
37 Prepend the given value to the specified environmental value using the given separator character,
38 creating the variable if it doesn't already exist

1 **PMIX_APPEND_ENVAR** "pmix.envar.append" (pmix_envar_t*)
2 Append the given value to the specified environmental value using the given separator character,
3 creating the variable if it doesn't already exist

4 **PMIX_FIRST_ENVAR** "pmix.envar.first" (pmix_envar_t*)
5 Ensure the given value appears first in the specified envar using the separator character, creating the
6 envar if it doesn't already exist

7 **PMIX_ALLOC_QUEUE** "pmix.alloc.queue" (char*)
8 Name of the WLM queue to which the allocation request is to be directed, or the queue being
9 referenced in a query.

10 **PMIX_ALLOC_TIME** "pmix.alloc.time" (uint32_t)
11 Total session time (in seconds) being requested in an allocation request.

12 **PMIX_ALLOC_NUM_NODES** "pmix.alloc.nnodes" (uint64_t)
13 The number of nodes being requested in an allocation request.

14 **PMIX_ALLOC_NODE_LIST** "pmix.alloc.nlist" (char*)
15 Regular expression of the specific nodes being requested in an allocation request.

16 **PMIX_ALLOC_NUM_CPUS** "pmix.alloc.ncpus" (uint64_t)
17 Number of PUs being requested in an allocation request.

18 **PMIX_ALLOC_NUM_CPU_LIST** "pmix.alloc.ncpulist" (char*)
19 Regular expression of the number of PUs for each node being requested in an allocation request.

20 **PMIX_ALLOC_CPU_LIST** "pmix.alloc.cpulist" (char*)
21 Regular expression of the specific PUs being requested in an allocation request.

22 **PMIX_ALLOC_MEM_SIZE** "pmix.alloc.msize" (float)
23 Number of Megabytes[base2] of memory (per process) being requested in an allocation request.

24 **PMIX_ALLOC_BANDWIDTH** "pmix.alloc.bw" (float)
25 Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

26 **PMIX_ALLOC_FABRIC_QOS** "pmix.alloc.netqos" (char*)
27 Fabric quality of service level for the job being requested in an allocation request.

28 **PMIX_ALLOC_FABRIC_TYPE** "pmix.alloc.nettype" (char*)
29 Type of desired transport (e.g., "tcp", "udp") being requested in an allocation request.

30 **PMIX_ALLOC_FABRIC_PLANE** "pmix.alloc.netplane" (char*)
31 ID string for the *fabric plane* to be used for the requested allocation.

32 **PMIX_ALLOC_FABRIC_ENDPTS** "pmix.alloc.endpts" (size_t)
33 Number of endpoints to allocate per *process* in the job.

34 **PMIX_ALLOC_FABRIC_ENDPTS_NODE** "pmix.alloc.endpts.nd" (size_t)
35 Number of endpoints to allocate per *node* for the job.

36 **PMIX_COSPAWN_APP** "pmix.cospawn" (bool)

1 Designated application is to be spawned as a disconnected job - i.e., the launcher shall not include the
2 application in any of the job-level values (e.g., **PMIX_RANK** within the job) provided to any other
3 application process generated by the same spawn request. Typically used to cospawn debugger
4 daemons alongside an application.

5 **PMIX_SPAWN_TOOL** "pmix.spwn.tool" (bool)

6 Indicate that the job being spawned is a tool.

7 **PMIX_EVENT_SILENT_TERMINATION** "pmix.evsilentterm" (bool)

8 Do not generate an event when this job normally terminates.

9 **PMIX_ENVARS_HARVESTED** "pmix.evar.hvstd" (bool)

10 Environmental parameters have been harvested by the spawn requestor - the server does not need to
11 harvest them.

12 **PMIX_JOB_TIMEOUT** "pmix.job.time" (int)

13 Time in seconds before the spawned job should time out and be terminated (0 => infinite), defined as
14 the total runtime of the job (equivalent to the walltime limit of typical batch schedulers).

15 **PMIX_SPAWN_TIMEOUT** "pmix.sp.time" (int)

16 Time in seconds before spawn operation should time out (0 => infinite). Logically equivalent to
17 passing the **PMIX_TIMEOUT** attribute to the **PMIx_Spawn** API, it is provided as a separate attribute
18 to distinguish it from the **PMIX_JOB_TIMEOUT** attribute



19 Description

20 Spawn a new job. The assigned namespace of the spawned applications is returned in the *nspace* parameter. A
21 **NULL** value in that location indicates that the caller doesn't wish to have the namespace returned. The *nspace*
22 array must be at least of size one more than **PMIX_MAX_NSLEN**.

23 By default, the spawned processes will be PMIx "connected" to the parent process upon successful launch (see
24 Section 12.3 for details). This includes that (a) the parent process will be given a copy of the new job's
25 information so it can query job-level info without incurring any communication penalties, (b) newly spawned
26 child processes will receive a copy of the parent processes job-level info, and (c) both the parent process and
27 members of the child job will receive notification of errors from processes in their combined assemblage.

▼ Advice to users ▼

28 Behavior of individual resource managers may differ, but it is expected that failure of any application process
29 to start will result in termination/cleanup of all processes in the newly spawned job and return of an error code
30 to the caller.

▼ Advice to PMIx library implementers ▼

31 Tools may utilize **PMIx_Spawn** to start intermediate launchers as described in Section 18.2.2. For times
32 where the tool is not attached to a PMIx server, internal support for fork/exec of the specified applications
33 would allow the tool to maintain a single code path for both the connected and disconnected cases. Inclusion
34 of such support is recommended, but not required.

1 12.2.2 PMIx_Spawn_nb

2 Summary

3 Nonblocking version of the [PMIx_Spawn](#) routine.

4 *PMIx v1.0* Format

C

```
5 pmix_status_t  
6 PMIx_Spawn_nb(const pmix_info_t job_info[], size_t ninfo,  
7               const pmix_app_t apps[], size_t napps,  
8               pmix_spawn_cbfunc_t cbfunc, void *cbdata)
```

C

9 **IN** `job_info`
10 Array of info structures (array of handles)
11 **IN** `ninfo`
12 Number of elements in the `job_info` array (integer)
13 **IN** `apps`
14 Array of [pmix_app_t](#) structures (array of handles)
15 **IN** `cbfunc`
16 Callback function [pmix_spawn_cbfunc_t](#) (function reference)
17 **IN** `cbdata`
18 Data to be passed to the callback function (memory reference)

19 A successful return indicates that the request is being processed and the result will be returned in the provided
20 `cbfunc`. Note that the library must not invoke the callback function prior to returning from the API. The
21 callback function, `cbfunc`, is only called when [PMIX_SUCCESS](#) is returned.

22 Returns [PMIX_SUCCESS](#) or one of the following error codes when the condition described occurs:

- 23 • [PMIX_OPERATION_SUCCEEDED](#), indicating that the request was immediately processed and returned
24 *success* - the `cbfunc` will *not* be called

25 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
26 error code or an implementation defined error code as described in Section [3.1.1](#).

Required Attributes

27 PMIx libraries are not required to directly support any attributes for this function. However, any provided
28 attributes must be passed to the host SMS daemon for processing.

29 Host environments are required to support the following attributes when present in either the `job_info` or the
30 `info` array of an element of the `apps` array:

31 **PMIX_WDIR** "`pmix.wdir`" (`char*`)
32 Working directory for spawned processes.

33 **PMIX_SET_SESSION_CWD** "`pmix.ssn cwd`" (`bool`)
34 Set the current working directory to the session working directory assigned by the RM - can be
35 assigned to the entire job (by including attribute in the `job_info` array) or on a per-application basis in
36 the `info` array for each [pmix_app_t](#).

1 **PMIX_PREFIX** "pmix.prefix" (char*)
2 Prefix to use for starting spawned processes - i.e., the directory where the executables can be found.
3 **PMIX_HOST** "pmix.host" (char*)
4 Comma-delimited list of hosts to use for spawned processes.
5 **PMIX_HOSTFILE** "pmix.hostfile" (char*)
6 Hostfile to use for spawned processes.

▲-----▲
▼-----▼ **Optional Attributes** -----▼

7 The following attributes are optional for host environments that support this operation:

8 **PMIX_ADD_HOSTFILE** "pmix.addhostfile" (char*)
9 Hostfile containing hosts to add to existing allocation.
10 **PMIX_ADD_HOST** "pmix.addhost" (char*)
11 Comma-delimited list of hosts to add to the allocation.
12 **PMIX_PRELOAD_BIN** "pmix.preloadbin" (bool)
13 Preload executables onto nodes prior to executing launch procedure.
14 **PMIX_PRELOAD_FILES** "pmix.preloadfiles" (char*)
15 Comma-delimited list of files to pre-position on nodes prior to executing launch procedure.
16 **PMIX_PERSONALITY** "pmix.pers" (char*)
17 Name of personality corresponding to programming model used by application - supported values
18 depend upon PMIx implementation.
19 **PMIX_DISPLAY_MAP** "pmix.dispmap" (bool)
20 Display process mapping upon spawn.
21 **PMIX_PPR** "pmix.ppr" (char*)
22 Number of processes to spawn on each identified resource.
23 **PMIX_MAPBY** "pmix.mapby" (char*)
24 Process mapping policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
25 for the rank to discover the mapping policy used for the provided namespace. Supported values are
26 launcher specific.
27 **PMIX_RANKBY** "pmix.rankby" (char*)
28 Process ranking policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
29 for the rank to discover the ranking algorithm used for the provided namespace. Supported values are
30 launcher specific.
31 **PMIX_BINDTO** "pmix.bindto" (char*)
32 Process binding policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
33 for the rank to discover the binding policy used for the provided namespace. Supported values are
34 launcher specific.
35 **PMIX_STDIN_TGT** "pmix.stdin" (uint32_t)
36 Spawned process rank that is to receive any forwarded **stdin**.

1 **PMIX_TAG_OUTPUT** "pmix.tagout" (bool)
2 Tag **stdout**/**stderr** with the identity of the source process - can be assigned to the entire job (by
3 including attribute in the *job_info* array) or on a per-application basis in the *info* array for each
4 **pmix_app_t**.

5 **PMIX_TIMESTAMP_OUTPUT** "pmix.tsout" (bool)
6 Timestamp output - can be assigned to the entire job (by including attribute in the *job_info* array) or on
7 a per-application basis in the *info* array for each **pmix_app_t**.

8 **PMIX_MERGE_STDERR_STDOUT** "pmix.mergeerrout" (bool)
9 Merge **stdout** and **stderr** streams - can be assigned to the entire job (by including attribute in the
10 *job_info* array) or on a per-application basis in the *info* array for each **pmix_app_t**.

11 **PMIX_OUTPUT_TO_FILE** "pmix.outfile" (char*)
12 Direct output (both **stdout** and **stderr**) into files of form "<filename>.rank" - can be assigned to
13 the entire job (by including attribute in the *job_info* array) or on a per-application basis in the *info* array
14 for each **pmix_app_t**.

15 **PMIX_INDEX_ARGV** "pmix.indxargv" (bool)
16 Mark the **argv** with the rank of the process.

17 **PMIX_CPUS_PER_PROC** "pmix.cputerproc" (uint32_t)
18 Number of PUs to assign to each rank - when accessed using **PMIx_Get**, use the
19 **PMIX_RANK_WILDCARD** value for the rank to discover the PUs/process assigned to the provided
20 namespace.

21 **PMIX_NO_PROCS_ON_HEAD** "pmix.nolocal" (bool)
22 Do not place processes on the head node.

23 **PMIX_NO_OVERSUBSCRIBE** "pmix.noover" (bool)
24 Do not oversubscribe the nodes - i.e., do not place more processes than allocated slots on a node.

25 **PMIX_REPORT_BINDINGS** "pmix.repbind" (bool)
26 Report bindings of the individual processes.

27 **PMIX_CPU_LIST** "pmix.cpulist" (char*)
28 List of PUs to use for this job - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD**
29 value for the rank to discover the PU list used for the provided namespace.

30 **PMIX_JOB_RECOVERABLE** "pmix.recover" (bool)
31 Application supports recoverable operations.

32 **PMIX_JOB_CONTINUOUS** "pmix.continuous" (bool)
33 Application is continuous, all failed processes should be immediately restarted.

34 **PMIX_MAX_RESTARTS** "pmix.maxrestarts" (uint32_t)
35 Maximum number of times to restart a process - when accessed using **PMIx_Get**, use the
36 **PMIX_RANK_WILDCARD** value for the rank to discover the max restarts for the provided namespace.

37 **PMIX_SET_ENVAR** "pmix.envar.set" (pmix_envar_t*)
38 Set the envar to the given value, overwriting any pre-existing one

39 **PMIX_UNSET_ENVAR** "pmix.envar.unset" (char*)

1 Unset the environment variable specified in the string.

2 **PMIX_ADD_ENVAR** "pmix.envar.add" (pmix_envar_t*)

3 Add the environment variable, but do not overwrite any pre-existing one

4 **PMIX_PREPEND_ENVAR** "pmix.envar.prepnd" (pmix_envar_t*)

5 Prepend the given value to the specified environmental value using the given separator character,
6 creating the variable if it doesn't already exist

7 **PMIX_APPEND_ENVAR** "pmix.envar.appnd" (pmix_envar_t*)

8 Append the given value to the specified environmental value using the given separator character,
9 creating the variable if it doesn't already exist

10 **PMIX_FIRST_ENVAR** "pmix.envar.first" (pmix_envar_t*)

11 Ensure the given value appears first in the specified envar using the separator character, creating the
12 envar if it doesn't already exist

13 **PMIX_ALLOC_QUEUE** "pmix.alloc.queue" (char*)

14 Name of the WLM queue to which the allocation request is to be directed, or the queue being
15 referenced in a query.

16 **PMIX_ALLOC_TIME** "pmix.alloc.time" (uint32_t)

17 Total session time (in seconds) being requested in an allocation request.

18 **PMIX_ALLOC_NUM_NODES** "pmix.alloc.nnodes" (uint64_t)

19 The number of nodes being requested in an allocation request.

20 **PMIX_ALLOC_NODE_LIST** "pmix.alloc.nlist" (char*)

21 Regular expression of the specific nodes being requested in an allocation request.

22 **PMIX_ALLOC_NUM_CPUS** "pmix.alloc.ncpus" (uint64_t)

23 Number of PUs being requested in an allocation request.

24 **PMIX_ALLOC_NUM_CPU_LIST** "pmix.alloc.ncpulist" (char*)

25 Regular expression of the number of PUs for each node being requested in an allocation request.

26 **PMIX_ALLOC_CPU_LIST** "pmix.alloc.cpulist" (char*)

27 Regular expression of the specific PUs being requested in an allocation request.

28 **PMIX_ALLOC_MEM_SIZE** "pmix.alloc.msize" (float)

29 Number of Megabytes[base2] of memory (per process) being requested in an allocation request.

30 **PMIX_ALLOC_BANDWIDTH** "pmix.alloc.bw" (float)

31 Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

32 **PMIX_ALLOC_FABRIC_QOS** "pmix.alloc.netqos" (char*)

33 Fabric quality of service level for the job being requested in an allocation request.

34 **PMIX_ALLOC_FABRIC_TYPE** "pmix.alloc.nettype" (char*)

35 Type of desired transport (e.g., "tcp", "udp") being requested in an allocation request.

36 **PMIX_ALLOC_FABRIC_PLANE** "pmix.alloc.netplane" (char*)

37 ID string for the *fabric plane* to be used for the requested allocation.

1 **PMIX_ALLOC_FABRIC_ENDPTS** "pmix.alloc.endpts" (size_t)
2 Number of endpoints to allocate per *process* in the job.

3 **PMIX_ALLOC_FABRIC_ENDPTS_NODE** "pmix.alloc.endpts.nd" (size_t)
4 Number of endpoints to allocate per *node* for the job.

5 **PMIX_COSPAWN_APP** "pmix.cospawn" (bool)
6 Designated application is to be spawned as a disconnected job - i.e., the launcher shall not include the
7 application in any of the job-level values (e.g., **PMIX_RANK** within the job) provided to any other
8 application process generated by the same spawn request. Typically used to cospawn debugger
9 daemons alongside an application.

10 **PMIX_SPAWN_TOOL** "pmix.spwn.tool" (bool)
11 Indicate that the job being spawned is a tool.

12 **PMIX_EVENT_SILENT_TERMINATION** "pmix.evsilentterm" (bool)
13 Do not generate an event when this job normally terminates.

14 **PMIX_ENVARS_HARVESTED** "pmix.evar.hvstd" (bool)
15 Environmental parameters have been harvested by the spawn requestor - the server does not need to
16 harvest them.

17 **PMIX_JOB_TIMEOUT** "pmix.job.time" (int)
18 Time in seconds before the spawned job should time out and be terminated (0 => infinite), defined as
19 the total runtime of the job (equivalent to the walltime limit of typical batch schedulers).

20 **PMIX_SPAWN_TIMEOUT** "pmix.sp.time" (int)
21 Time in seconds before spawn operation should time out (0 => infinite). Logically equivalent to
22 passing the **PMIX_TIMEOUT** attribute to the **PMIx_Spawn** API, it is provided as a separate attribute
23 to distinguish it from the **PMIX_JOB_TIMEOUT** attribute



24 Description

25 Nonblocking version of the **PMIx_Spawn** routine. The provided callback function will be executed upon
26 successful start of *all* specified application processes.



Advice to users

27 Behavior of individual resource managers may differ, but it is expected that failure of any application process
28 to start will result in termination/cleanup of all processes in the newly spawned job and return of an error code
29 to the caller.



12.2.3 Spawn-specific constants

In addition to the generic error constants, the following spawn-specific error constants may be returned by the spawn APIs:

PMIX_ERR_JOB_ALLOC_FAILED	-188	The job request could not be executed due to failure to obtain the specified allocation
PMIX_ERR_JOB_APP_NOT_EXECUTABLE	-177	The specified application executable either could not be found, or lacks execution privileges.
PMIX_ERR_JOB_NO_EXE_SPECIFIED	-178	The job request did not specify an executable.
PMIX_ERR_JOB_FAILED_TO_MAP	-179	The launcher was unable to map the processes for the specified job request.
PMIX_ERR_JOB_FAILED_TO_LAUNCH	-181	One or more processes in the job request failed to launch
PMIX_ERR_JOB_EXE_NOT_FOUND	<i>(Provisional)</i>	Specified executable not found
PMIX_ERR_JOB_INSUFFICIENT_RESOURCES	<i>(Provisional)</i>	Insufficient resources to spawn job
PMIX_ERR_JOB_SYS_OP_FAILED	<i>(Provisional)</i>	System library operation failed
PMIX_ERR_JOB_WDIR_NOT_FOUND	<i>(Provisional)</i>	Specified working directory not found

12.2.4 Spawn attributes

Attributes used to describe **PMIx_Spawn** behavior - they are values passed to the **PMIx_Spawn** API and therefore are not accessed using the **PMIx_Get** APIs when used in that context. However, some of the attributes defined in this section can be provided by the host environment for other purposes - e.g., the host might provide the **PMIX_MAPPBY** attribute in the job-related information so that an application can use **PMIx_Get** to discover the mapping used for determining process locations. Multi-use attributes and their respective access reference rank are denoted below.

PMIX_PERSONALITY	"pmix.pers"	(char*)
		Name of personality corresponding to programming model used by application - supported values depend upon PMIx implementation.
PMIX_HOST	"pmix.host"	(char*)
		Comma-delimited list of hosts to use for spawned processes.
PMIX_HOSTFILE	"pmix.hostfile"	(char*)
		Hostfile to use for spawned processes.
PMIX_ADD_HOST	"pmix.addhost"	(char*)
		Comma-delimited list of hosts to add to the allocation.
PMIX_ADD_HOSTFILE	"pmix.addhostfile"	(char*)
		Hostfile containing hosts to add to existing allocation.
PMIX_PREFIX	"pmix.prefix"	(char*)
		Prefix to use for starting spawned processes - i.e., the directory where the executables can be found.
PMIX_WDIR	"pmix.wdir"	(char*)
		Working directory for spawned processes.
PMIX_DISPLAY_MAP	"pmix.dispmap"	(bool)
		Display process mapping upon spawn.
PMIX_PPR	"pmix.ppr"	(char*)
		Number of processes to spawn on each identified resource.

1 **PMIX_MAPBY** "pmix.mapby" (char*)
2 Process mapping policy - when accessed using [PMIx_Get](#), use the [PMIX_RANK_WILDCARD](#) value
3 for the rank to discover the mapping policy used for the provided namespace. Supported values are
4 launcher specific.

5 **PMIX_RANKBY** "pmix.rankby" (char*)
6 Process ranking policy - when accessed using [PMIx_Get](#), use the [PMIX_RANK_WILDCARD](#) value
7 for the rank to discover the ranking algorithm used for the provided namespace. Supported values are
8 launcher specific.

9 **PMIX_BINDTO** "pmix.bindto" (char*)
10 Process binding policy - when accessed using [PMIx_Get](#), use the [PMIX_RANK_WILDCARD](#) value
11 for the rank to discover the binding policy used for the provided namespace. Supported values are
12 launcher specific.

13 **PMIX_PRELOAD_BIN** "pmix.preloadbin" (bool)
14 Preload executables onto nodes prior to executing launch procedure.

15 **PMIX_PRELOAD_FILES** "pmix.preloadfiles" (char*)
16 Comma-delimited list of files to pre-position on nodes prior to executing launch procedure.

17 **PMIX_STDIN_TGT** "pmix.stdin" (uint32_t)
18 Spawned process rank that is to receive any forwarded **stdin**.

19 **PMIX_SET_SESSION_CWD** "pmix.ssn cwd" (bool)
20 Set the current working directory to the session working directory assigned by the RM - can be
21 assigned to the entire job (by including attribute in the *job_info* array) or on a per-application basis in
22 the *info* array for each [pmix_app_t](#).

23 **PMIX_TAG_OUTPUT** "pmix.tagout" (bool)
24 Tag **stdout/stderr** with the identity of the source process - can be assigned to the entire job (by
25 including attribute in the *job_info* array) or on a per-application basis in the *info* array for each
26 [pmix_app_t](#).

27 **PMIX_TIMESTAMP_OUTPUT** "pmix.tsout" (bool)
28 Timestamp output - can be assigned to the entire job (by including attribute in the *job_info* array) or on
29 a per-application basis in the *info* array for each [pmix_app_t](#).

30 **PMIX_MERGE_STDERR_STDOUT** "pmix.mergeerrout" (bool)
31 Merge **stdout** and **stderr** streams - can be assigned to the entire job (by including attribute in the
32 *job_info* array) or on a per-application basis in the *info* array for each [pmix_app_t](#).

33 **PMIX_OUTPUT_TO_FILE** "pmix.outfile" (char*)
34 Direct output (both **stdout** and **stderr**) into files of form "**<filename>.rank**" - can be assigned to
35 the entire job (by including attribute in the *job_info* array) or on a per-application basis in the *info* array
36 for each [pmix_app_t](#).

37 **PMIX_OUTPUT_TO_DIRECTORY** "pmix.outdir" (char*)
38 Direct output into files of form "**<directory>/<jobid>/rank.<rank>/stdout[err]**" -
39 can be assigned to the entire job (by including attribute in the *job_info* array) or on a per-application
40 basis in the *info* array for each [pmix_app_t](#).

41 **PMIX_INDEX_ARGV** "pmix.indxargv" (bool)
42 Mark the **argv** with the rank of the process.

43 **PMIX_CPUS_PER_PROC** "pmix.cpusperproc" (uint32_t)
44 Number of PUs to assign to each rank - when accessed using [PMIx_Get](#), use the
45 [PMIX_RANK_WILDCARD](#) value for the rank to discover the PUs/process assigned to the provided
46 namespace.

1 **PMIX_NO_PROCS_ON_HEAD** "pmix.nolocal" (bool)
2 Do not place processes on the head node.

3 **PMIX_NO_OVERSUBSCRIBE** "pmix.noover" (bool)
4 Do not oversubscribe the nodes - i.e., do not place more processes than allocated slots on a node.

5 **PMIX_REPORT_BINDINGS** "pmix.repbind" (bool)
6 Report bindings of the individual processes.

7 **PMIX_CPU_LIST** "pmix.cpulist" (char*)
8 List of PUs to use for this job - when accessed using [PMIx_Get](#), use the [PMIX_RANK_WILDCARD](#)
9 value for the rank to discover the PU list used for the provided namespace.

10 **PMIX_JOB_RECOVERABLE** "pmix.recover" (bool)
11 Application supports recoverable operations.

12 **PMIX_JOB_CONTINUOUS** "pmix.continuous" (bool)
13 Application is continuous, all failed processes should be immediately restarted.

14 **PMIX_MAX_RESTARTS** "pmix.maxrestarts" (uint32_t)
15 Maximum number of times to restart a process - when accessed using [PMIx_Get](#), use the
16 [PMIX_RANK_WILDCARD](#) value for the rank to discover the max restarts for the provided namespace.

17 **PMIX_SPAWN_TOOL** "pmix.spwn.tool" (bool)
18 Indicate that the job being spawned is a tool.

19 **PMIX_TIMEOUT_STACKTRACES** "pmix.tim.stack" (bool)
20 Include process stacktraces in timeout report from a job.

21 **PMIX_TIMEOUT_REPORT_STATE** "pmix.tim.state" (bool)
22 Report process states in timeout report from a job.

23 **PMIX_NOTIFY_JOB_EVENTS** "pmix.note.jev" (bool)
24 Requests that the launcher generate the [PMIX_EVENT_JOB_START](#), [PMIX_LAUNCH_COMPLETE](#),
25 and [PMIX_EVENT_JOB_END](#) events. Each event is to include at least the namespace of the
26 corresponding job and a [PMIX_EVENT_TIMESTAMP](#) indicating the time the event occurred. Note
27 that the requester must register for these individual events, or capture and process them by registering a
28 default event handler instead of individual handlers and then process the events based on the returned
29 status code. Another common method is to register one event handler for all job-related events, with a
30 separate handler for non-job events - see [PMIx_Register_event_handler](#) for details.

31 **PMIX_NOTIFY_COMPLETION** "pmix.notecomp" (bool)
32 Requests that the launcher generate the [PMIX_EVENT_JOB_END](#) event for normal or abnormal
33 termination of the spawned job. The event shall include the returned status code
34 ([PMIX_JOB_TERM_STATUS](#)) for the corresponding job; the identity ([PMIX_PROCID](#)) and exit
35 status ([PMIX_EXIT_CODE](#)) of the first failed process, if applicable; and a
36 [PMIX_EVENT_TIMESTAMP](#) indicating the time the termination occurred. Note that the requester
37 must register for the event or capture and process it within a default event handler.

38 **PMIX_NOTIFY_PROC_TERMINATION** "pmix.noteproc" (bool)
39 Requests that the launcher generate the [PMIX_EVENT_PROC_TERMINATED](#) event whenever a
40 process either normally or abnormally terminates.

41 **PMIX_NOTIFY_PROC_ABNORMAL_TERMINATION** "pmix.noteabproc" (bool)
42 Requests that the launcher generate the [PMIX_EVENT_PROC_TERMINATED](#) event only when a
43 process abnormally terminates.

44 **PMIX_LOG_PROC_TERMINATION** "pmix.logproc" (bool)
45 Requests that the launcher log the [PMIX_EVENT_PROC_TERMINATED](#) event whenever a process
46 either normally or abnormally terminates.

1 **PMIX_LOG_PROC_ABNORMAL_TERMINATION** "pmix.logabproc" (bool)
2 Requests that the launcher log the **PMIX_EVENT_PROC_TERMINATED** event only when a process
3 abnormally terminates.

4 **PMIX_LOG_JOB_EVENTS** "pmix.log.jev" (bool)
5 Requests that the launcher log the **PMIX_EVENT_JOB_START**, **PMIX_LAUNCH_COMPLETE**, and
6 **PMIX_EVENT_JOB_END** events using **PMIx_Log**, subject to the logging attributes of Section 13.4.3.

7 **PMIX_LOG_COMPLETION** "pmix.logcomp" (bool)
8 Requests that the launcher log the **PMIX_EVENT_JOB_END** event for normal or abnormal
9 termination of the spawned job using **PMIx_Log**, subject to the logging attributes of Section 13.4.3.
10 The event shall include the returned status code (**PMIX_JOB_TERM_STATUS**) for the corresponding
11 job; the identity (**PMIX_PROCID**) and exit status (**PMIX_EXIT_CODE**) of the first failed process, if
12 applicable; and a **PMIX_EVENT_TIMESTAMP** indicating the time the termination occurred.

13 **PMIX_EVENT_SILENT_TERMINATION** "pmix.evsilentterm" (bool)
14 Do not generate an event when this job normally terminates.

15 **PMIX_ENVARS_HARVESTED** "pmix.evar.hvstd" (bool) *(Provisional)*
16 Environmental parameters have been harvested by the spawn requestor - the server does not need to
17 harvest them.

18 **PMIX_JOB_TIMEOUT** "pmix.job.time" (int) *(Provisional)*
19 Time in seconds before the spawned job should time out and be terminated (0 => infinite), defined as
20 the total runtime of the job (equivalent to the walltime limit of typical batch schedulers).

21 **PMIX_SPAWN_TIMEOUT** "pmix.sp.time" (int) *(Provisional)*
22 Time in seconds before spawn operation should time out (0 => infinite). Logically equivalent to
23 passing the **PMIX_TIMEOUT** attribute to the **PMIx_Spawn** API, it is provided as a separate attribute
24 to distinguish it from the **PMIX_JOB_TIMEOUT** attribute

25 Attributes used to adjust remote environment variables prior to spawning the specified application processes.

26 **PMIX_SET_ENVAR** "pmix.envar.set" (pmix_envar_t*)
27 Set the envar to the given value, overwriting any pre-existing one

28 **PMIX_UNSET_ENVAR** "pmix.envar.unset" (char*)
29 Unset the environment variable specified in the string.

30 **PMIX_ADD_ENVAR** "pmix.envar.add" (pmix_envar_t*)
31 Add the environment variable, but do not overwrite any pre-existing one

32 **PMIX_PREPEND_ENVAR** "pmix.envar.prepnd" (pmix_envar_t*)
33 Prepend the given value to the specified environmental value using the given separator character,
34 creating the variable if it doesn't already exist

35 **PMIX_APPEND_ENVAR** "pmix.envar.appnd" (pmix_envar_t*)
36 Append the given value to the specified environmental value using the given separator character,
37 creating the variable if it doesn't already exist

38 **PMIX_FIRST_ENVAR** "pmix.envar.first" (pmix_envar_t*)
39 Ensure the given value appears first in the specified envar using the separator character, creating the
40 envar if it doesn't already exist

41 12.2.5 Application Structure

42 The **pmix_app_t** structure describes the application context for the **PMIx_Spawn** and **PMIx_Spawn_nb**
43 operations.

PMIx v1.0

```

1 typedef struct pmix_app {
2     /** Executable */
3     char *cmd;
4     /** Argument set, NULL terminated */
5     char **argv;
6     /** Environment set, NULL terminated */
7     char **env;
8     /** Current working directory */
9     char *cwd;
10    /** Maximum processes with this profile */
11    int maxprocs;
12    /** Array of info keys describing this application*/
13    pmix_info_t *info;
14    /** Number of info keys in 'info' array */
15    size_t ninfo;
16 } pmix_app_t;

```

17 12.2.5.1 App structure support macros

18 The following macros are provided to support the `pmix_app_t` structure.

19 **Static initializer for the app structure** 20 *(Provisional)*

21 Provide a static initializer for the `pmix_app_t` fields.

PMIx v5.0

```

22 PMIX_APP_STATIC_INIT

```

23 **Initialize the app structure**

24 Initialize the `pmix_app_t` fields

PMIx v1.0

```

25 PMIX_APP_CONSTRUCT(m)

```

26 **IN** m
27 Pointer to the structure to be initialized (pointer to `pmix_app_t`)

1 **Destruct the app structure**

2 Destruct the `pmix_app_t` fields



3 **PMIX_APP_DESTRUCT (m)**



4 **IN m**

5 Pointer to the structure to be destructed (pointer to `pmix_app_t`)

6 **Create an app array**

7 Allocate and initialize an array of `pmix_app_t` structures

PMIx v1.0



8 **PMIX_APP_CREATE (m, n)**



9 **INOUT m**

10 Address where the pointer to the array of `pmix_app_t` structures shall be stored (handle)

11 **IN n**

12 Number of structures to be allocated (`size_t`)

13 **Free an app structure**

14 Release a `pmix_app_t` structure

PMIx v4.0



15 **PMIX_APP_RELEASE (m)**



16 **IN m**

17 Pointer to a `pmix_app_t` structure (handle)

18 **Free an app array**

19 Release an array of `pmix_app_t` structures

PMIx v1.0



20 **PMIX_APP_FREE (m, n)**



21 **IN m**

22 Pointer to the array of `pmix_app_t` structures (handle)

23 **IN n**

24 Number of structures in the array (`size_t`)

Create the info array of application directives

Create an array of `pmix_info_t` structures for passing application-level directives, updating the `ninfo` field of the `pmix_app_t` structure.

```
PMIX_APP_INFO_CREATE(m, n)
```

```
IN m
    Pointer to the pmix_app_t structure (handle)
IN n
    Number of directives to be allocated (size_t)
```

12.2.5.2 Spawn Callback Function

Summary

The `pmix_spawn_cbfunc_t` is used on the PMIx client side by `PMIx_Spawn_nb` and on the PMIx server side by `pmix_server_spawn_fn_t`.

PMIx v1.0

```
typedef void (*pmix_spawn_cbfunc_t)
    (pmix_status_t status,
     pmix_namespace_t nspace, void *cbdata);
```

```
IN status
    Status associated with the operation (handle)
IN nspace
    Namespace string (pmix_namespace_t)
IN cbdata
    Callback data passed to original API call (memory reference)
```

Description

The callback will be executed upon launch of the specified applications in `PMIx_Spawn_nb`, or upon failure to launch any of them.

The `status` of the callback will indicate whether or not the spawn succeeded. The `nspace` of the spawned processes will be returned, along with any provided callback data. Note that the returned `nspace` value will not be protected upon return from the callback function, so the receiver must copy it if it needs to be retained.

1 12.3 Connecting and Disconnecting Processes

2 This section defines functions to connect and disconnect processes in two or more separate PMIx namespaces.
3 The PMIx definition of *connected* solely implies that the host environment should treat the failure of any
4 process in the assemblage as a reportable event, taking action on the assemblage as if it were a single
5 application. For example, if the environment defaults (in the absence of any application directives) to
6 terminating an application upon failure of any process in that application, then the environment should
7 terminate all processes in the connected assemblage upon failure of any member.

8 The host environment may choose to assign a new namespace to the connected assemblage and/or assign new
9 ranks for its members for its own internal tracking purposes. However, it is not required to communicate such
10 assignments to the participants (e.g., in response to an appropriate call to `PMIx_Query_info_nb`). The
11 host environment is required to generate a `PMIX_ERR_PROC_TERM_WO_SYNC` event should any process in
12 the assemblage terminate or call `PMIx_Finalize` without first *disconnecting* from the assemblage. If the
13 job including the process is terminated as a result of that action, then the host environment is required to also
14 generate the `PMIX_ERR_JOB_TERM_WO_SYNC` for all jobs that were terminated as a result.

▼ Advice to PMIx server hosts ▼

15 The *connect* operation does not require the exchange of job-level information nor the inclusion of information
16 posted by participating processes via `PMIx_Put`. Indeed, the callback function utilized in
17 `pmix_server_connect_fn_t` cannot pass information back into the PMIx server library. However, host
18 environments are advised that collecting such information at the participating daemons represents an
19 optimization opportunity as participating processes are likely to request such information after the connect
20 operation completes.

▼ Advice to users ▼

21 Attempting to *connect* processes solely within the same namespace is essentially a *no-op* operation. While not
22 explicitly prohibited, users are advised that a PMIx implementation or host environment may return an error in
23 such cases.

24 Neither the PMIx implementation nor host environment are required to provide any tracking support for the
25 assemblage. Thus, the application is responsible for maintaining the membership list of the assemblage.

26 12.3.1 PMIx_Connect

27 Summary

28 Connect namespaces.

Format

C

`pmix_status_t`

```
PMIx_Connect(const pmix_proc_t procs[], size_t nprocs,  
             const pmix_info_t info[], size_t ninfo)
```

C

IN `procs`

Array of proc structures (array of handles)

IN `nprocs`

Number of elements in the *procs* array (integer)

IN `info`

Array of info structures (array of handles)

IN `ninfo`

Number of elements in the *info* array (integer)

Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host SMS daemon for processing.

Optional Attributes

The following attributes are optional for PMIx implementations:

`PMIX_ALL_CLONES_PARTICIPATE` "`pmix.clone.part`" (`bool`)

All *clones* of the calling process must participate in the collective operation.

The following attributes are optional for host environments that support this operation:

`PMIX_TIMEOUT` "`pmix.timeout`" (`int`)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Record the processes specified by the *procs* array as *connected* as per the PMIx definition. The function will return once all processes identified in *procs* have called either **PMIx_Connect** or its non-blocking version, and the host environment has completed any supporting operations required to meet the terms of the PMIx definition of *connected* processes.

A process can only engage in one connect operation involving the identical *procs* array at a time. However, a process can be simultaneously engaged in multiple connect operations, each involving a different *procs* array.

As in the case of the **PMIx_Fence** operation, the *info* array can be used to pass user-level directives regarding timeout constraints and other options available from the host RM.

Advice to users

All processes engaged in a given **PMIx_Connect** operation must provide the identical *procs* array as ordering of entries in the array and the method by which those processes are identified (e.g., use of **PMIX_RANK_WILDCARD** versus listing the individual processes) *may* impact the host environment's algorithm for uniquely identifying an operation.

Advice to PMIx library implementers

PMIx_Connect and its non-blocking form are both *collective* operations. Accordingly, the PMIx server library is required to aggregate participation by local clients, passing the request to the host environment once all local participants have executed the API.

Advice to PMIx server hosts

The host will receive a single call for each collective operation. It is the responsibility of the host to identify the nodes containing participating processes, execute the collective across all participating nodes, and notify the local PMIx server library upon completion of the global collective.

12.3.2 PMIx_Connect_nb

Summary

Nonblocking **PMIx_Connect_nb** routine.

Format

C

`pmix_status_t`

```
PMIx_Connect_nb(const pmix_proc_t procs[], size_t nprocs,  
                const pmix_info_t info[], size_t ninfo,  
                pmix_op_cbfunc_t cbfunc, void *cbdata)
```

C

IN `procs`

Array of proc structures (array of handles)

IN `nprocs`

Number of elements in the *procs* array (integer)

IN `info`

Array of info structures (array of handles)

IN `ninfo`

Number of elements in the *info* array (integer)

IN `cbfunc`

Callback function `pmix_op_cbfunc_t` (function reference)

IN `cbdata`

Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will *not* be called

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host SMS daemon for processing.

Optional Attributes

The following attributes are optional for PMIx implementations:

PMIX_ALL_CLONES_PARTICIPATE "`pmix.clone.part`" (**bool**)

All *clones* of the calling process must participate in the collective operation.

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "`pmix.timeout`" (**int**)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.



1
2
3
4
5

Description

Nonblocking version of [PMIx_Connect](#). The callback function is called once all processes identified in *procs* have called either [PMIx_Connect](#) or its non-blocking version, *and* the host environment has completed any supporting operations required to meet the terms of the PMIx definition of *connected* processes. See the advice provided in the description for [PMIx_Connect](#) for more information.

6 12.3.3 PMIx_Disconnect

7
8

Summary

Disconnect a previously connected set of processes.

9 *PMIx v1.0*

Format

C

10
11
12

```
pmix_status_t
PMIx_Disconnect(const pmix_proc_t procs[], size_t nprocs,
                const pmix_info_t info[], size_t ninfo);
```

13
14
15
16
17
18
19
20

- IN procs**
Array of proc structures (array of handles)
- IN nprocs**
Number of elements in the *procs* array (integer)
- IN info**
Array of info structures (array of handles)
- IN ninfo**
Number of elements in the *info* array (integer)

21
22
23
24
25

Returns PMIX_SUCCESS or one of the following error codes when the condition described occurs:

- the [PMIX_ERR_INVALID_OPERATION](#) error indicating that the specified set of *procs* was not previously *connected* via a call to [PMIx_Connect](#) or its non-blocking form.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

26
27

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host SMS daemon for processing.



Optional Attributes

The following attributes are optional for PMIx implementations:

PMIX_ALL_CLONES_PARTICIPATE "`pmix.clone.part`" (`bool`)

All *clones* of the calling process must participate in the collective operation.

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "`pmix.timeout`" (`int`)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Disconnect a previously connected set of processes. The function will return once all processes identified in *procs* have called either **PMIx_Disconnect** or its non-blocking version, *and* the host environment has completed any required supporting operations.

A process can only engage in one disconnect operation involving the identical *procs* array at a time. However, a process can be simultaneously engaged in multiple disconnect operations, each involving a different *procs* array.

As in the case of the **PMIx_Fence** operation, the *info* array can be used to pass user-level directives regarding the algorithm to be used for any collective operation involved in the operation, timeout constraints, and other options available from the host RM.

Advice to users

All processes engaged in a given **PMIx_Disconnect** operation must provide the identical *procs* array as ordering of entries in the array and the method by which those processes are identified (e.g., use of **PMIX_RANK_WILDCARD** versus listing the individual processes) *may* impact the host environment's algorithm for uniquely identifying an operation.

Advice to PMIx library implementers

PMIx_Disconnect and its non-blocking form are both *collective* operations. Accordingly, the PMIx server library is required to aggregate participation by local clients, passing the request to the host environment once all local participants have executed the API.

Advice to PMIx server hosts

The host will receive a single call for each collective operation. The host will receive a single call for each collective operation. It is the responsibility of the host to identify the nodes containing participating processes, execute the collective across all participating nodes, and notify the local PMIx server library upon completion of the global collective.

1 12.3.4 PMIx_Disconnect_nb

2 Summary

3 Nonblocking [PMIx_Disconnect](#) routine.

4 *PMIx v1.0* Format

C

5 `pmix_status_t`

```
6 PMIx_Disconnect_nb(const pmix_proc_t procs[], size_t nprocs,  
7                   const pmix_info_t info[], size_t ninfo,  
8                   pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

9 **IN** `procs`

10 Array of proc structures (array of handles)

11 **IN** `nprocs`

12 Number of elements in the *procs* array (integer)

13 **IN** `info`

14 Array of info structures (array of handles)

15 **IN** `ninfo`

16 Number of elements in the *info* array (integer)

17 **IN** `cbfunc`

18 Callback function [pmix_op_cbfunc_t](#) (function reference)

19 **IN** `cbdata`

20 Data to be passed to the callback function (memory reference)

21 A successful return indicates that the request is being processed and the result will be returned in the provided
22 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
23 callback function, *cbfunc*, is only called when [PMIX_SUCCESS](#) is returned.

24 Returns [PMIX_SUCCESS](#) or one of the following error codes when the condition described occurs:

- 25 • [PMIX_OPERATION_SUCCEEDED](#), indicating that the request was immediately processed and returned
26 *success* - the *cbfunc* will *not* be called

27 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
28 error code or an implementation defined error code as described in Section [3.1.1](#).

Required Attributes

29 PMIx libraries are not required to directly support any attributes for this function. However, any provided
30 attributes must be passed to the host SMS daemon for processing.

Optional Attributes

The following attributes are optional for PMIx implementations:

PMIX_ALL_CLONES_PARTICIPATE "`pmix.clone.part`" (`bool`)

All *clones* of the calling process must participate in the collective operation.

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "`pmix.timeout`" (`int`)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Nonblocking **PMIx_Disconnect** routine. The callback function is called either:

- to return the **PMIX_ERR_INVALID_OPERATION** error indicating that the specified set of *procs* was not previously *connected* via a call to **PMIx_Connect** or its non-blocking form;
- to return a PMIx error constant indicating that the operation failed; or
- once all processes identified in *procs* have called either **PMIx_Disconnect_nb** or its blocking version, *and* the host environment has completed any required supporting operations.

See the advice provided in the description for **PMIx_Disconnect** for more information.

12.4 Process Locality

The relative locality of processes is often used to optimize their interactions with the hardware and other processes. PMIx provides a means by which the host environment can communicate the locality of a given process using the **PMIx_server_generate_locality_string** to generate an abstracted representation of that value. This provides a human-readable format and allows the client to parse the locality string with a method of its choice that may differ from the one used by the server that generated it.

There are times, however, when relative locality and other PMIx-provided information doesn't include some element required by the application. In these instances, the application may need access to the full description of the local hardware topology. PMIx does not itself generate such descriptions - there are multiple third-party libraries that fulfill that role. Instead, PMIx offers an abstraction method by which users can obtain a pointer to the description. This transparently enables support for different methods of sharing the topology between the host environment (which may well have already generated it prior to local start of application processes) and the clients - e.g., through passing of a shared memory region.

12.4.1 PMIx_Load_topology

Summary

Load the local hardware topology description

Format

C

`pmix_status_t`

`PMIx_Load_topology(pmix_topology_t *topo);`

C

INOUT `topo`

Address of a `pmix_topology_t` structure where the topology information is to be loaded (handle)

A successful return indicates that the *topo* was successfully loaded.

Returns `PMIX_SUCCESS` or a negative value indicating the error.

Description

Obtain a pointer to the topology description of the local node. If the *source* field of the provided `pmix_topology_t` is set, then the PMIx library must return a description from the specified implementation or else indicate that the implementation is not available by returning the `PMIX_ERR_NOT_SUPPORTED` error constant.

The returned pointer may point to a shared memory region or an actual instance of the topology description. In either case, the description shall be treated as a "read-only" object - attempts to modify the object are likely to fail and return an error. The PMIx library is responsible for performing any required cleanup when the client library finalizes.

Advice to users

It is the responsibility of the user to ensure that the *topo* argument is properly initialized prior to calling this API, and to check the returned *source* to verify that the returned topology description is compatible with the user's code.

12.4.2 `PMIx_Get_relative_locality`

Summary

Get the relative locality of two local processes given their locality strings.

1 **Format** C

```
2 pmix_status_t  
3 PMIx_Get_relative_locality(const char *locality1,  
4                             const char *locality2,  
5                             pmix_locality_t *locality);
```

6 **IN** `locality1`
7 String returned by the `PMIx_server_generate_locality_string` API (handle)
8 **IN** `locality2`
9 String returned by the `PMIx_server_generate_locality_string` API (handle)
10 **INOUT** `locality`
11 Location where the relative locality bitmask is to be constructed (memory reference)
12 A successful return indicates that the `locality` was successfully loaded.
13 Returns `PMIX_SUCCESS` or a negative value indicating the error.

14 **Description**
15 Parse the locality strings of two processes (as returned by `PMIx_Get` using the `PMIX_LOCALITY_STRING`
16 key) and set the appropriate `pmix_locality_t` locality bits in the provided memory location.

17 12.4.2.1 Topology description

18 The `pmix_topology_t` structure contains a (case-insensitive) string identifying the source of the topology
19 (e.g., "hwloc") and a pointer to the corresponding implementation-specific topology description.

PMIx v4.0

```
20 typedef struct pmix_topology {  
21     char *source;  
22     void *topology;  
23 } pmix_topology_t;
```

24 12.4.2.2 Topology support macros

25 The following macros support the `pmix_topology_t` structure.

26 Static initializer for the topology structure

27 *(Provisional)*

28 Provide a static initializer for the `pmix_topology_t` fields.

PMIx v5.0

```
29 PMIX_TOPOLOGY_STATIC_INIT
```

1 **Initialize the topology structure**

2 Initialize the `pmix_topology_t` fields to `NULL`



3 `PMIX_TOPOLOGY_CONSTRUCT (m)`



4 **IN** `m`

5 Pointer to the structure to be initialized (pointer to `pmix_topology_t`)

6 **Destruct a topology structure**

7 **Summary**

8 Destruct a `pmix_topology_t` fields

9 **Format**

PMIx v5.0



10 `void`

11 `PMIx_Topology_destruct (pmix_topology_t *topo);`



12 **IN** `topo`

13 Pointer to the structure to be destructed (pointer to `pmix_topology_t`)

14 **Description**

15 Release any memory storage held by the `pmix_topology_t` structure

16 **Create a topology array**

17 Allocate and initialize a `pmix_topology_t` array.

PMIx v4.0



18 `PMIX_TOPOLOGY_CREATE (m, n)`



19 **INOUT** `m`

20 Address where the pointer to the array of `pmix_topology_t` structures shall be stored (handle)

21 **IN** `n`

22 Number of structures to be allocated (size_t)

1 12.4.2.3 Relative locality of two processes

2 The `pmix_locality_t` datatype is a `uint16_t` bitmask that defines the relative locality of two processes
3 on a node. The following constants represent specific bits in the mask and can be used to test a locality value
4 using standard bit-test methods.

5 **PMIX_LOCALITY_UNKNOWN** 0x0000 All bits are set to zero, indicating that the relative locality of the
6 two processes is unknown
7 **PMIX_LOCALITY_NONLOCAL** 0x0000 The two processes do not share any common locations
8 **PMIX_LOCALITY_SHARE_HWTHREAD** 0x0001 The two processes share at least one hardware thread
9 **PMIX_LOCALITY_SHARE_CORE** 0x0002 The two processes share at least one core
10 **PMIX_LOCALITY_SHARE_L1CACHE** 0x0004 The two processes share at least an L1 cache
11 **PMIX_LOCALITY_SHARE_L2CACHE** 0x0008 The two processes share at least an L2 cache
12 **PMIX_LOCALITY_SHARE_L3CACHE** 0x0010 The two processes share at least an L3 cache
13 **PMIX_LOCALITY_SHARE_PACKAGE** 0x0020 The two processes share at least a package
14 **PMIX_LOCALITY_SHARE_NUMA** 0x0040 The two processes share at least one Non-Uniform
15 Memory Access (NUMA) region
16 **PMIX_LOCALITY_SHARE_NODE** 0x4000 The two processes are executing on the same node

17 Implementers and vendors may choose to extend these definitions as needed to describe a particular system.

18 12.4.2.4 Locality keys

19 **PMIX_LOCALITY_STRING** "pmix.locstr" (char*)

20 String describing a process's bound location - referenced using the process's rank. The string is
21 prefixed by the implementation that created it (e.g., "hwloc") followed by a colon. The remainder of the
22 string represents the corresponding locality as expressed by the underlying implementation. The entire
23 string must be passed to `PMIx_Get_relative_locality` for processing. Note that hosts are
24 only required to provide locality strings for local client processes - thus, a call to `PMIx_Get` for the
25 locality string of a process that returns `PMIX_ERR_NOT_FOUND` indicates that the process is not
26 executing on the same node.

27 12.4.3 PMIx_Parse_cpuset_string

28 Summary

29 Parse the PU binding bitmap from its string representation.

30 *PMIx v4.0* Format

31 `pmix_status_t`

32 `PMIx_Parse_cpuset_string(const char *cpuset_string,`
33 `pmix_cpuset_t *cpuset);`

34 **IN** `cpuset_string`

35 String returned by the `PMIx_server_generate_cpuset_string` API (handle)

36 **INOUT** `cpuset`

37 Address of an object where the bitmap is to be stored (memory reference)

38 A successful return indicates that the `cpuset` was successfully loaded.

39 Returns `PMIX_SUCCESS` or a negative value indicating the error.

1 **Description**
2 Parse the string representation of the binding bitmap (as returned by [PMIx_Get](#) using the [PMIX_CPUSET](#)
3 key) and set the appropriate PU binding location information in the provided memory location.

4 12.4.4 [PMIx_Get_cpuset](#)

5 **Summary**
6 Get the PU binding bitmap of the current process.

7 *PMIx v4.0* **Format** C
8 `pmix_status_t`
9 `PMIx_Get_cpuset (pmix_cpuset_t *cpuset, pmix_bind_envelope_t ref);`
C

10 **INOUT** `cpuset`
11 Address of an object where the bitmap is to be stored (memory reference)
12 **IN** `ref`
13 The binding envelope to be considered when formulating the bitmap ([pmix_bind_envelope_t](#))
14 A successful return indicates that the *cpuset* was successfully loaded.
15 Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

16 **Description**
17 Obtain and set the appropriate PU binding location information in the provided memory location based on the
18 specified binding envelope.

19 12.4.4.1 [Binding envelope](#)

20 *PMIx v4.0* The [pmix_bind_envelope_t](#) data type defines the envelope of threads within a possibly multi-threaded
21 process that are to be considered when getting the cpuset associated with the process. Valid values include:

22 `PMIX_CPUBIND_PROCESS 0` Use the location of all threads in the possibly multi-threaded process.
23 `PMIX_CPUBIND_THREAD 1` Use only the location of the thread calling the API.

24 12.4.5 [PMIx_Compute_distances](#)

25 **Summary**
26 Compute distances from specified process location to local devices.

Format

C

```
pmix_status_t
PMIx_Compute_distances(pmix_topology_t *topo,
                      pmix_cpuset_t *cpuset,
                      pmix_info_t info[], size_t ninfo[],
                      pmix_device_distance_t *distances[],
                      size_t *ndist);
```

C

IN *topo*
Pointer to the topology description of the node where the process is located (**NULL** indicates the local node) ([pmix_topology_t](#))

IN *cpuset*
Pointer to the location of the process ([pmix_cpuset_t](#))

IN *info*
Array of [pmix_info_t](#) describing the devices whose distance is to be computed (handle)

IN *ninfo*
Number of elements in *info* (integer)

INOUT *distances*
Pointer to an address where the array of [pmix_device_distance_t](#) structures containing the distances from the caller to the specified devices is to be returned (handle)

INOUT *ndist*
Pointer to an address where the number of elements in the *distances* array is to be returned (handle)

Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Description

Both the minimum and maximum distance fields in the elements of the array shall be filled with the respective distances between the current process location and the types of devices or specific device identified in the *info* directives. In the absence of directives, distances to all supported device types shall be returned.

Advice to users

A process whose threads are not all bound to the same location may return inconsistent results from calls to this API by different threads if the [PMIX_CPUBIND_THREAD](#) binding envelope was used when generating the *cpuset*.

12.4.6 PMIx_Compute_distances_nb

Summary

Compute distances from specified process location to local devices.

Format

C

```
pmix_status_t
PMIx_Compute_distances_nb(pmix_topology_t *topo,
                          pmix_cpuset_t *cpuset,
                          pmix_info_t info[], size_t ninfo[],
                          pmix_device_dist_cbfunc_t cbfunc,
                          void *cbdata);
```

C

- IN topo**
Pointer to the topology description of the node where the process is located (**NULL** indicates the local node) ([pmix_topology_t](#))
- IN cpuset**
Pointer to the location of the process ([pmix_cpuset_t](#))
- IN info**
Array of [pmix_info_t](#) describing the devices whose distance is to be computed (handle)
- IN ninfo**
Number of elements in *info* (integer)
- IN cbfunc**
Callback function [pmix_info_cbfunc_t](#) (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general **PMIX** error code or an implementation defined error code as described in Section 3.1.1.

Description

Non-blocking form of the [PMIx_Compute_distances](#) API.

12.4.7 Device Distance Callback Function

Summary

The [pmix_device_dist_cbfunc_t](#) is used to return an array of device distances.

PMIx v4.0

```

1 typedef void (*pmix_device_dist_cbfunc_t)
2     (pmix_status_t status,
3      pmix_device_distance_t *dist,
4      size_t ndist,
5      void *cbdata,
6      pmix_release_cbfunc_t release_fn,
7      void *release_cbdata);

```

IN status
Status associated with the operation ([pmix_status_t](#))

IN dist
Array of [pmix_device_distance_t](#) returned by the operation (pointer)

IN ndist
Number of elements in the *dist* array ([size_t](#))

IN cbdata
Callback data passed to original API call (memory reference)

IN release_fn
Function to be called when done with the *dist* data (function pointer)

IN release_cbdata
Callback data to be passed to *release_fn* (memory reference)

Description

The *status* indicates if requested data was found or not. The array of [pmix_device_distance_t](#) will contain the distance information.

12.4.8 Device type

The [pmix_device_type_t](#) is a [uint64_t](#) bitmask for identifying the type(s) whose distances are being requested, or the type of a specific device being referenced (e.g., in a [pmix_device_distance_t](#) object).

PMIx v1.0

```

26 typedef uint16_t pmix_device_type_t;

```

The following constants can be used to set a variable of the type [pmix_device_type_t](#).

PMIX_DEVTTYPE_UNKNOWN 0x00 The device is of an unknown type - will not be included in returned device distances.

PMIX_DEVTTYPE_BLOCK 0x01 Operating system block device, or non-volatile memory device (e.g., "sda" or "dax2.0" on Linux).

PMIX_DEVTTYPE_GPU 0x02 Operating system Graphics Processing Unit (GPU) device (e.g., "card0" for a Linux Direct Rendering Manager (DRM) device).

PMIX_DEVTTYPE_NETWORK 0x04 Operating system network device (e.g., the "eth0" interface on Linux).

1 **PMIX_DEVTYPE_OPENFABRICS 0x08** Operating system OpenFabrics device (e.g., an "mlx4_0"
 2 InfiniBand Host Channel Adapter (HCA), or "hfi1_0" Omni-Path interface on Linux).
 3 **PMIX_DEVTYPE_DMA 0x10** Operating system Direct Memory Access (DMA) engine device (e.g., the
 4 "dma0chan0" DMA channel on Linux).
 5 **PMIX_DEVTYPE_COPROC 0x20** Operating system co-processor device (e.g., "mic0" for a Xeon Phi on
 6 Linux, "opencl0d0" for an OpenCL device, or "cuda0" for a Compute Unified Device
 7 Architecture (CUDA) device).

8 12.4.9 Device Distance Structure

9 The `pmix_device_distance_t` structure contains the minimum and maximum relative distance from
 10 the caller to a given device.

PMIx v4.0

```
11 typedef struct pmix_device_distance {
12     char *uuid;
13     char *osname;
14     pmix_device_type_t type;
15     uint16_t mindist;
16     uint16_t maxdist;
17 } pmix_device_distance_t;
```

18 The *uuid* is a string identifier guaranteed to be unique within the cluster and is typically assembled from
 19 discovered device attributes (e.g., the Internet Protocol (IP) address of the device). The *osname* is the local
 20 operating system name of the device and is only unique to that node.

21 The two distance fields provide the minimum and maximum relative distance to the device from the specified
 22 location of the process, expressed as a 16-bit integer value where a smaller number indicates that this device is
 23 closer to the process than a device with a larger distance value. Note that relative distance values are not
 24 necessarily correlated to a physical property - e.g., a device at twice the distance from another device does not
 25 necessarily have twice the latency for communication with it.

26 Relative distances only apply to similar devices and cannot be used to compare devices of different types. Both
 27 minimum and maximum distances are provided to support cases where the process may be bound to more than
 28 one location, and the locations are at different distances from the device.

29 A relative distance value of `UINT16_MAX` indicates that the distance from the process to the device could not
 30 be provided. This may be due to lack of available information (e.g., the PMIx library not having access to
 31 device locations) or other factors.

32 12.4.10 Device distance support macros

33 The following macros are provided to support the `pmix_device_distance_t` structure.

1 **Static initializer for the device distance structure**

2 *(Provisional)*

3 Provide a static initializer for the `pmix_device_distance_t` fields.



4 `PMIX_DEVICE_DIST_STATIC_INIT`



5 **Initialize the device distance structure**

6 Initialize the `pmix_device_distance_t` fields.

PMIx v4.0



7 `PMIX_DEVICE_DIST_CONSTRUCT (m)`



8 **IN** `m`

9 Pointer to the structure to be initialized (pointer to `pmix_device_distance_t`)

10 **Destruct the device distance structure**

11 Destruct the `pmix_device_distance_t` fields.

PMIx v4.0



12 `PMIX_DEVICE_DIST_DESTRUCT (m)`



13 **IN** `m`

14 Pointer to the structure to be destructed (pointer to `pmix_device_distance_t`)

15 **Create an device distance array**

16 Allocate and initialize a `pmix_device_distance_t` array.

PMIx v4.0



17 `PMIX_DEVICE_DIST_CREATE (m, n)`



18 **INOUT** `m`

19 Address where the pointer to the array of `pmix_device_distance_t` structures shall be stored
20 (handle)

21 **IN** `n`

22 Number of structures to be allocated (`size_t`)

1 **Release an device distance array**
2 Release an array of `pmix_device_distance_t` structures.

3 **PMIX_DEVICE_DIST_FREE**(m, n)

4 **IN** m
5 Pointer to the array of `pmix_device_distance_t` structures (handle)
6 **IN** n
7 Number of structures in the array (`size_t`)

8 12.4.11 Device distance attributes

9 The following attributes can be used to retrieve device distances from the PMIx data store. Note that distances
10 stored by the host environment are based on the process location at the time of start of execution and may not
11 reflect changes to location imposed by the process itself. **PMIX_DEVICE_DISTANCES**

12 **"pmix.dev.dist"** (`pmix_data_array_t`)

13 Return an array of `pmix_device_distance_t` containing the minimum and maximum distances
14 of the given process location to all devices of the specified type on the local node.

15 **PMIX_DEVICE_TYPE** **"pmix.dev.type"** (`pmix_device_type_t`)

16 Bitmask specifying the type(s) of device(s) whose information is being requested. Only used as a
17 directive/qualifier.

18 **PMIX_DEVICE_ID** **"pmix.dev.id"** (`string`)

19 System-wide Universally Unique Identifier (UUID) or node-local Operating System (OS) name of a
20 particular device.

CHAPTER 13

Job Management and Reporting

1 The job management APIs provide an application with the ability to orchestrate its operation in partnership
2 with the SMS. Members of this category include the [PMIx_Allocation_request](#),
3 [PMIx_Job_control](#), and [PMIx_Process_monitor](#) APIs.

4 13.1 Allocation Requests

5 This section defines functionality to request new allocations from the RM, and request modifications to
6 existing allocations. These are primarily used in the following scenarios:

- 7 • *Evolving* applications that dynamically request and return resources as they execute.
- 8 • *Malleable* environments where the scheduler redirects resources away from executing applications for
9 higher priority jobs or load balancing.
- 10 • *Resilient* applications that need to request replacement resources in the face of failures.
- 11 • *Rigid* jobs where the user has requested a static allocation of resources for a fixed period of time, but
12 realizes that they underestimated their required time while executing.

13 PMIx attempts to address this range of use-cases with a flexible API.

14 13.1.1 PMIx_Allocation_request

15 Summary

16 Request an allocation operation from the host resource manager.

17 *PMIx v3.0* Format

C

```
18 pmix_status_t  
19 PMIx_Allocation_request(pmix_alloc_directive_t directive,  
20                        pmix_info_t info[], size_t ninfo,  
21                        pmix_info_t *results[], size_t *nresults);
```

C

- 22 **IN** `directive`
23 Allocation directive ([pmix_alloc_directive_t](#))
- 24 **IN** `info`
25 Array of [pmix_info_t](#) structures (array of handles)
- 26 **IN** `ninfo`
27 Number of elements in the *info* array (integer)

1 **INOUT results**
2 Address where a pointer to an array of `pmix_info_t` containing the results of the request can be
3 returned (memory reference)

4 **INOUT nresults**
5 Address where the number of elements in *results* can be returned (handle)

6 Returns `PMIX_SUCCESS` or a negative value indicating the error.

▼----- Required Attributes -----▼

7 PMIx libraries are not required to directly support any attributes for this function. However, any provided
8 attributes must be passed to the host SMS daemon for processing, and the PMIx library is *required* to add the
9 `PMIX_USERID` and the `PMIX_GRPID` attributes of the client process making the request.

10 Host environments that implement support for this operation are required to support the following attributes:

11 `PMIX_ALLOC_REQ_ID` "`pmix.alloc.reqid`" (`char*`)
12 User-provided string identifier for this allocation request which can later be used to query status of the
13 request.

14 `PMIX_ALLOC_NUM_NODES` "`pmix.alloc.nnodes`" (`uint64_t`)
15 The number of nodes being requested in an allocation request.

16 `PMIX_ALLOC_NUM_CPUS` "`pmix.alloc.ncpus`" (`uint64_t`)
17 Number of PUs being requested in an allocation request.

18 `PMIX_ALLOC_TIME` "`pmix.alloc.time`" (`uint32_t`)
19 Total session time (in seconds) being requested in an allocation request.



▼----- Optional Attributes -----▼

20 The following attributes are optional for host environments that support this operation:

21 `PMIX_ALLOC_NODE_LIST` "`pmix.alloc.nlist`" (`char*`)
22 Regular expression of the specific nodes being requested in an allocation request.

23 `PMIX_ALLOC_NUM_CPU_LIST` "`pmix.alloc.ncpulist`" (`char*`)
24 Regular expression of the number of PUs for each node being requested in an allocation request.

25 `PMIX_ALLOC_CPU_LIST` "`pmix.alloc.cpulist`" (`char*`)
26 Regular expression of the specific PUs being requested in an allocation request.

27 `PMIX_ALLOC_MEM_SIZE` "`pmix.alloc.msize`" (`float`)
28 Number of Megabytes[base2] of memory (per process) being requested in an allocation request.

29 `PMIX_ALLOC_FABRIC` "`pmix.alloc.net`" (`array`)
30 Array of `pmix_info_t` describing requested fabric resources. This must include at least:
31 `PMIX_ALLOC_FABRIC_ID`, `PMIX_ALLOC_FABRIC_TYPE`, and
32 `PMIX_ALLOC_FABRIC_ENDPTS`, plus whatever other descriptors are desired.

33 `PMIX_ALLOC_FABRIC_ID` "`pmix.alloc.netid`" (`char*`)

1 The key to be used when accessing this requested fabric allocation. The fabric allocation will be
2 returned/stored as a `pmix_data_array_t` of `pmix_info_t` whose first element is composed of
3 this key and the allocated resource description. The type of the included value depends upon the fabric
4 support. For example, a Transmission Control Protocol (TCP) allocation might consist of a
5 comma-delimited string of socket ranges such as "32000-32100, 33005, 38123-38146".
6 Additional array entries will consist of any provided resource request directives, along with their
7 assigned values. Examples include: `PMIX_ALLOC_FABRIC_TYPE` - the type of resources provided;
8 `PMIX_ALLOC_FABRIC_PLANE` - if applicable, what plane the resources were assigned from;
9 `PMIX_ALLOC_FABRIC_QOS` - the assigned QoS; `PMIX_ALLOC_BANDWIDTH` - the allocated
10 bandwidth; `PMIX_ALLOC_FABRIC_SEC_KEY` - a security key for the requested fabric allocation.
11 NOTE: the array contents may differ from those requested, especially if `PMIX_INFO_REQD` was not
12 set in the request.

13 `PMIX_ALLOC_BANDWIDTH` "pmix.alloc.bw" (float)

14 Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

15 `PMIX_ALLOC_FABRIC_QOS` "pmix.alloc.netqos" (char*)

16 Fabric quality of service level for the job being requested in an allocation request.

17 `PMIX_ALLOC_FABRIC_TYPE` "pmix.alloc.nettype" (char*)

18 Type of desired transport (e.g., "tcp", "udp") being requested in an allocation request.

19 `PMIX_ALLOC_FABRIC_PLANE` "pmix.alloc.netplane" (char*)

20 ID string for the *fabric plane* to be used for the requested allocation.

21 `PMIX_ALLOC_FABRIC_ENDPTS` "pmix.alloc.endpts" (size_t)

22 Number of endpoints to allocate per *process* in the job.

23 `PMIX_ALLOC_FABRIC_ENDPTS_NODE` "pmix.alloc.endpts.nd" (size_t)

24 Number of endpoints to allocate per *node* for the job.

25 `PMIX_ALLOC_FABRIC_SEC_KEY` "pmix.alloc.nsec" (pmix_byte_object_t)

26 Request that the allocation include a fabric security key for the spawned job.



27 Description

28 Request an allocation operation from the host resource manager. Several broad categories are envisioned,
29 including the ability to:

- 30 • Request allocation of additional resources, including memory, bandwidth, and compute. This should be
31 accomplished in a non-blocking manner so that the application can continue to progress while waiting for
32 resources to become available. Note that the new allocation will be disjoint from (i.e., not affiliated with)
33 the allocation of the requestor - thus the termination of one allocation will not impact the other.
- 34 • Extend the reservation on currently allocated resources, subject to scheduling availability and priorities.
35 This includes extending the time limit on current resources, and/or requesting additional resources be
36 allocated to the requesting job. Any additional allocated resources will be considered as part of the current
37 allocation, and thus will be released at the same time.
- 38 • Return no-longer-required resources to the scheduler. This includes the "loan" of resources back to the
39 scheduler with a promise to return them upon subsequent request.

1 If successful, the returned results for a request for additional resources must include the host resource
2 manager's identifier ([PMIX_ALLOC_ID](#)) that the requester can use to specify the resources in, for example, a
3 call to [PMIx_Spawn](#).

4 13.1.2 [PMIx_Allocation_request_nb](#)

5 Summary

6 Request an allocation operation from the host resource manager.

7 *PMIx v2.0* Format

C

```
8 pmix_status_t  
9 PMIx_Allocation_request_nb(pmix_alloc_directive_t directive,  
10                          pmix_info_t info[], size_t ninfo,  
11                          pmix_info_cbfunc_t cbfunc, void *cbdata);
```

C

12 **IN** **directive**
13 Allocation directive ([pmix_alloc_directive_t](#))
14 **IN** **info**
15 Array of [pmix_info_t](#) structures (array of handles)
16 **IN** **ninfo**
17 Number of elements in the *info* array (integer)
18 **IN** **cbfunc**
19 Callback function [pmix_info_cbfunc_t](#) (function reference)
20 **IN** **cbdata**
21 Data to be passed to the callback function (memory reference)

22 Returns one of the following:

23 A successful return indicates that the request is being processed and the result will be returned in the provided
24 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
25 callback function, *cbfunc*, is only called when [PMIX_SUCCESS](#) is returned.

26 Returns [PMIX_SUCCESS](#) or one of the following error codes when the condition described occurs:

- 27 • [PMIX_OPERATION_SUCCEEDED](#), indicating that the request was immediately processed and returned
28 *success* - the *cbfunc* will *not* be called

29 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
30 error code or an implementation defined error code as described in Section [3.1.1](#).

Required Attributes

31 PMIx libraries are not required to directly support any attributes for this function. However, any provided
32 attributes must be passed to the host SMS daemon for processing, and the PMIx library is *required* to add the
33 [PMIX_USERID](#) and the [PMIX_GRPID](#) attributes of the client process making the request.

34 Host environments that implement support for this operation are required to support the following attributes:

35 [PMIX_ALLOC_REQ_ID](#) "pmix.alloc.reqid" ([char*](#))

1 User-provided string identifier for this allocation request which can later be used to query status of the
2 request.

3 **PMIX_ALLOC_NUM_NODES** "pmix.alloc.nnodes" (uint64_t)

4 The number of nodes being requested in an allocation request.

5 **PMIX_ALLOC_NUM_CPUS** "pmix.alloc.ncpus" (uint64_t)

6 Number of PUs being requested in an allocation request.

7 **PMIX_ALLOC_TIME** "pmix.alloc.time" (uint32_t)

8 Total session time (in seconds) being requested in an allocation request.



▼----- Optional Attributes -----▼

9 The following attributes are optional for host environments that support this operation:

10 **PMIX_ALLOC_NODE_LIST** "pmix.alloc.nlist" (char*)

11 Regular expression of the specific nodes being requested in an allocation request.

12 **PMIX_ALLOC_NUM_CPU_LIST** "pmix.alloc.ncpulist" (char*)

13 Regular expression of the number of PUs for each node being requested in an allocation request.

14 **PMIX_ALLOC_CPU_LIST** "pmix.alloc.cpulist" (char*)

15 Regular expression of the specific PUs being requested in an allocation request.

16 **PMIX_ALLOC_MEM_SIZE** "pmix.alloc.msize" (float)

17 Number of Megabytes[base2] of memory (per process) being requested in an allocation request.

18 **PMIX_ALLOC_FABRIC** "pmix.alloc.net" (array)

19 Array of **pmix_info_t** describing requested fabric resources. This must include at least:

20 **PMIX_ALLOC_FABRIC_ID**, **PMIX_ALLOC_FABRIC_TYPE**, and

21 **PMIX_ALLOC_FABRIC_ENDPTS**, plus whatever other descriptors are desired.

22 **PMIX_ALLOC_FABRIC_ID** "pmix.alloc.netid" (char*)

23 The key to be used when accessing this requested fabric allocation. The fabric allocation will be
24 returned/stored as a **pmix_data_array_t** of **pmix_info_t** whose first element is composed of
25 this key and the allocated resource description. The type of the included value depends upon the fabric
26 support. For example, a TCP allocation might consist of a comma-delimited string of socket ranges
27 such as "32000-32100, 33005, 38123-38146". Additional array entries will consist of any
28 provided resource request directives, along with their assigned values. Examples include:

29 **PMIX_ALLOC_FABRIC_TYPE** - the type of resources provided; **PMIX_ALLOC_FABRIC_PLANE** -

30 if applicable, what plane the resources were assigned from; **PMIX_ALLOC_FABRIC_QOS** - the

31 assigned QoS; **PMIX_ALLOC_BANDWIDTH** - the allocated bandwidth;

32 **PMIX_ALLOC_FABRIC_SEC_KEY** - a security key for the requested fabric allocation. NOTE: the

33 array contents may differ from those requested, especially if **PMIX_INFO_REQD** was not set in the

34 request.

35 **PMIX_ALLOC_BANDWIDTH** "pmix.alloc.bw" (float)

36 Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

37 **PMIX_ALLOC_FABRIC_QOS** "pmix.alloc.netqos" (char*)

Fabric quality of service level for the job being requested in an allocation request.

PMIX_ALLOC_FABRIC_TYPE "pmix.alloc.nettype" (char*)

Type of desired transport (e.g., "tcp", "udp") being requested in an allocation request.

PMIX_ALLOC_FABRIC_PLANE "pmix.alloc.netplane" (char*)

ID string for the *fabric plane* to be used for the requested allocation.

PMIX_ALLOC_FABRIC_ENDPTS "pmix.alloc.endpts" (size_t)

Number of endpoints to allocate per *process* in the job.

PMIX_ALLOC_FABRIC_ENDPTS_NODE "pmix.alloc.endpts.nd" (size_t)

Number of endpoints to allocate per *node* for the job.

PMIX_ALLOC_FABRIC_SEC_KEY "pmix.alloc.nsec" (pmix_byte_object_t)

Request that the allocation include a fabric security key for the spawned job.



Description

Non-blocking form of the [PMIx_Allocation_request](#) API.

13.1.3 Job Allocation attributes

Attributes used to describe the job allocation - these are values passed to and/or returned by the [PMIx_Allocation_request_nb](#) and [PMIx_Allocation_request](#) APIs and are not accessed using the [PMIx_Get](#) API.

PMIX_ALLOC_REQ_ID "pmix.alloc.reqid" (char*)

User-provided string identifier for this allocation request which can later be used to query status of the request.

PMIX_ALLOC_ID "pmix.alloc.id" (char*)

A string identifier (provided by the host environment) for the resulting allocation which can later be used to reference the allocated resources in, for example, a call to [PMIx_Spawn](#).

PMIX_ALLOC_QUEUE "pmix.alloc.queue" (char*)

Name of the WLM queue to which the allocation request is to be directed, or the queue being referenced in a query.

PMIX_ALLOC_NUM_NODES "pmix.alloc.nnodes" (uint64_t)

The number of nodes being requested in an allocation request.

PMIX_ALLOC_NODE_LIST "pmix.alloc.nlist" (char*)

Regular expression of the specific nodes being requested in an allocation request.

PMIX_ALLOC_NUM_CPUS "pmix.alloc.ncpus" (uint64_t)

Number of PUs being requested in an allocation request.

PMIX_ALLOC_NUM_CPU_LIST "pmix.alloc.ncpulist" (char*)

Regular expression of the number of PUs for each node being requested in an allocation request.

PMIX_ALLOC_CPU_LIST "pmix.alloc.cpulist" (char*)

Regular expression of the specific PUs being requested in an allocation request.

PMIX_ALLOC_MEM_SIZE "pmix.alloc.msize" (float)

Number of Megabytes[base2] of memory (per process) being requested in an allocation request.

PMIX_ALLOC_FABRIC "pmix.alloc.net" (array)

1 Array of `pmix_info_t` describing requested fabric resources. This must include at least:

2 `PMIX_ALLOC_FABRIC_ID`, `PMIX_ALLOC_FABRIC_TYPE`, and

3 `PMIX_ALLOC_FABRIC_ENDPTS`, plus whatever other descriptors are desired.

4 **PMIX_ALLOC_FABRIC_ID** "pmix.alloc.netid" (char*)

5 The key to be used when accessing this requested fabric allocation. The fabric allocation will be
6 returned/stored as a `pmix_data_array_t` of `pmix_info_t` whose first element is composed of
7 this key and the allocated resource description. The type of the included value depends upon the fabric
8 support. For example, a TCP allocation might consist of a comma-delimited string of socket ranges
9 such as "32000–32100, 33005, 38123–38146". Additional array entries will consist of any
10 provided resource request directives, along with their assigned values. Examples include:

11 `PMIX_ALLOC_FABRIC_TYPE` - the type of resources provided; `PMIX_ALLOC_FABRIC_PLANE` -

12 if applicable, what plane the resources were assigned from; `PMIX_ALLOC_FABRIC_QOS` - the

13 assigned QoS; `PMIX_ALLOC_BANDWIDTH` - the allocated bandwidth;

14 `PMIX_ALLOC_FABRIC_SEC_KEY` - a security key for the requested fabric allocation. NOTE: the
15 array contents may differ from those requested, especially if `PMIX_INFO_REQD` was not set in the
16 request.

17 **PMIX_ALLOC_BANDWIDTH** "pmix.alloc.bw" (float)

18 Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

19 **PMIX_ALLOC_FABRIC_QOS** "pmix.alloc.netqos" (char*)

20 Fabric quality of service level for the job being requested in an allocation request.

21 **PMIX_ALLOC_TIME** "pmix.alloc.time" (uint32_t)

22 Total session time (in seconds) being requested in an allocation request.

23 **PMIX_ALLOC_FABRIC_TYPE** "pmix.alloc.nettype" (char*)

24 Type of desired transport (e.g., "tcp", "udp") being requested in an allocation request.

25 **PMIX_ALLOC_FABRIC_PLANE** "pmix.alloc.netplane" (char*)

26 ID string for the *fabric plane* to be used for the requested allocation.

27 **PMIX_ALLOC_FABRIC_ENDPTS** "pmix.alloc.endpts" (size_t)

28 Number of endpoints to allocate per *process* in the job.

29 **PMIX_ALLOC_FABRIC_ENDPTS_NODE** "pmix.alloc.endpts.nd" (size_t)

30 Number of endpoints to allocate per *node* for the job.

31 **PMIX_ALLOC_FABRIC_SEC_KEY** "pmix.alloc.nsec" (pmix_byte_object_t)

32 Request that the allocation include a fabric security key for the spawned job.

33 13.1.4 Job Allocation Directives

34 *PMIx v2.0*

35 The `pmix_alloc_directive_t` structure is a `uint8_t` type that defines the behavior of allocation
36 requests. The following constants can be used to set a variable of the type `pmix_alloc_directive_t`.
All definitions were introduced in version 2 of the standard unless otherwise marked.

37 **PMIX_ALLOC_NEW** 1 A new allocation is being requested. The resulting allocation will be disjoint (i.e.,
38 not connected in a job sense) from the requesting allocation.

39 **PMIX_ALLOC_EXTEND** 2 Extend the existing allocation, either in time or as additional resources.

40 **PMIX_ALLOC_RELEASE** 3 Release part of the existing allocation. Attributes in the accompanying
41 `pmix_info_t` array may be used to specify permanent release of the identified resources, or "lending"
42 of those resources for some period of time.

43 **PMIX_ALLOC_REAQUIRE** 4 Reacquire resources that were previously "lent" back to the scheduler.

44 **PMIX_ALLOC_EXTERNAL** 128 A value boundary above which implementers are free to define their
45 own directive values.

1 13.2 Job Control

2 This section defines APIs that enable the application and host environment to coordinate the response to
3 failures and other events. This can include requesting termination of the entire job or a subset of processes
4 within a job, but can also be used in combination with other PMIx capabilities (e.g., allocation support and
5 event notification) for more nuanced responses. For example, an application notified of an incipient
6 over-temperature condition on a node could use the [PMIx_Allocation_request_nb](#) interface to request
7 replacement nodes while simultaneously using the [PMIx_Job_control_nb](#) interface to direct that a
8 checkpoint event be delivered to all processes in the application. If replacement resources are not available, the
9 application might use the [PMIx_Job_control_nb](#) interface to request that the job continue at a lower
10 power setting, perhaps sufficient to avoid the over-temperature failure.

11 The job control APIs can also be used by an application to register itself as available for preemption when
12 operating in an environment such as a cloud or where incentives, financial or otherwise, are provided to jobs
13 willing to be preempted. Registration can include attributes indicating how many resources are being offered
14 for preemption (e.g., all or only some portion), whether the application will require time to prepare for
15 preemption, etc. Jobs that request a warning will receive an event notifying them of an impending preemption
16 (possibly including information as to the resources that will be taken away, how much time the application will
17 be given prior to being preempted, whether the preemption will be a suspension or full termination, etc.) so
18 they have an opportunity to save their work. Once the application is ready, it calls the provided event
19 completion callback function to indicate that the SMS is free to suspend or terminate it, and can include
20 directives regarding any desired restart.

21 13.2.1 PMIx_Job_control

22 Summary

23 Request a job control action.

24 *PMIx v3.0* Format

C

```
25 pmix_status_t  
26 PMIx_Job_control(const pmix_proc_t targets[], size_t ntargets,  
27                 const pmix_info_t directives[], size_t ndirs,  
28                 pmix_info_t *results[], size_t *nresults);
```

C

29 IN targets

30 Array of proc structures (array of handles)

31 IN ntargets

32 Number of elements in the *targets* array (integer)

33 IN directives

34 Array of info structures (array of handles)

35 IN ndirs

36 Number of elements in the *directives* array (integer)

37 INOUT results

38 Address where a pointer to an array of [pmix_info_t](#) containing the results of the request can be
39 returned (memory reference)

INOUT `nresults`

Address where the number of elements in *results* can be returned (handle)

Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host SMS daemon for processing, and the PMIx library is *required* to add the `PMIX_USERID` and the `PMIX_GRPID` attributes of the client process making the request.

Host environments that implement support for this operation are required to support the following attributes:

`PMIX_JOB_CTRL_ID` "`pmix.jctrl.id`" (`char*`)

Provide a string identifier for this request. The user can provide an identifier for the requested operation, thus allowing them to later request status of the operation or to terminate it. The host, therefore, shall track it with the request for future reference.

`PMIX_JOB_CTRL_PAUSE` "`pmix.jctrl.pause`" (`bool`)

Pause the specified processes.

`PMIX_JOB_CTRL_RESUME` "`pmix.jctrl.resume`" (`bool`)

Resume ("un-pause") the specified processes.

`PMIX_JOB_CTRL_KILL` "`pmix.jctrl.kill`" (`bool`)

Forcibly terminate the specified processes and cleanup.

`PMIX_JOB_CTRL_SIGNAL` "`pmix.jctrl.sig`" (`int`)

Send given signal to specified processes.

`PMIX_JOB_CTRL_TERMINATE` "`pmix.jctrl.term`" (`bool`)

Politely terminate the specified processes.

`PMIX_REGISTER_CLEANUP` "`pmix.reg.cleanup`" (`char*`)

Comma-delimited list of files to be removed upon process termination.

`PMIX_REGISTER_CLEANUP_DIR` "`pmix.reg.cleanupdir`" (`char*`)

Comma-delimited list of directories to be removed upon process termination.

`PMIX_CLEANUP_RECURSIVE` "`pmix.clnup.recurse`" (`bool`)

Recursively cleanup all subdirectories under the specified one(s).

`PMIX_CLEANUP_EMPTY` "`pmix.clnup.empty`" (`bool`)

Only remove empty subdirectories.

`PMIX_CLEANUP_IGNORE` "`pmix.clnup.ignore`" (`char*`)

Comma-delimited list of filenames that are not to be removed.

`PMIX_CLEANUP_LEAVE_TOPDIR` "`pmix.clnup.lvtop`" (`bool`)

When recursively cleaning subdirectories, do not remove the top-level directory (the one given in the cleanup request).

Optional Attributes

The following attributes are optional for host environments that support this operation:

- PMIX_JOB_CTRL_CANCEL** "pmix.jctrl.cancel" (char*)
Cancel the specified request - the provided request ID must match the **PMIX_JOB_CTRL_ID** provided to a previous call to **PMIx_Job_control**. An ID of **NULL** implies cancel all requests from this requestor.
- PMIX_JOB_CTRL_RESTART** "pmix.jctrl.restart" (char*)
Restart the specified processes using the given checkpoint ID.
- PMIX_JOB_CTRL_CHECKPOINT** "pmix.jctrl.ckpt" (char*)
Checkpoint the specified processes and assign the given ID to it.
- PMIX_JOB_CTRL_CHECKPOINT_EVENT** "pmix.jctrl.ckptev" (bool)
Use event notification to trigger a process checkpoint.
- PMIX_JOB_CTRL_CHECKPOINT_SIGNAL** "pmix.jctrl.ckptsig" (int)
Use the given signal to trigger a process checkpoint.
- PMIX_JOB_CTRL_CHECKPOINT_TIMEOUT** "pmix.jctrl.ckptsig" (int)
Time in seconds to wait for a checkpoint to complete.
- PMIX_JOB_CTRL_CHECKPOINT_METHOD** "pmix.jctrl.ckmethod" (pmix_data_array_t)
Array of **pmix_info_t** declaring each method and value supported by this application.
- PMIX_JOB_CTRL_PROVISION** "pmix.jctrl.pvn" (char*)
Regular expression identifying nodes that are to be provisioned.
- PMIX_JOB_CTRL_PROVISION_IMAGE** "pmix.jctrl.pvning" (char*)
Name of the image that is to be provisioned.
- PMIX_JOB_CTRL_PREEMPTIBLE** "pmix.jctrl.preempt" (bool)
Indicate that the job can be pre-empted.

Description

Request a job control action. The *targets* array identifies the processes to which the requested job control action is to be applied. All *clones* of an identified process are to have the requested action applied to them. A **NULL** value can be used to indicate all processes in the caller's namespace. The use of **PMIX_RANK_WILDCARD** can also be used to indicate that all processes in the given namespace are to be included.

The directives are provided as **pmix_info_t** structures in the *directives* array. The returned *status* indicates whether or not the request was granted, and information as to the reason for any denial of the request shall be returned in the *results* array.

13.2.2 PMIx_Job_control_nb

Summary

Request a job control action.

Format

C

```
pmix_status_t
PMIx_Job_control_nb(const pmix_proc_t targets[], size_t ntargets,
                    const pmix_info_t directives[], size_t ndirs,
                    pmix_info_cbfunc_t cbfunc, void *cbdata);
```

C

IN targets
Array of proc structures (array of handles)

IN ntargets
Number of elements in the *targets* array (integer)

IN directives
Array of info structures (array of handles)

IN ndirs
Number of elements in the *directives* array (integer)

IN cbfunc
Callback function `pmix_info_cbfunc_t` (function reference)

IN cbdata
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will *not* be called

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

PMIx libraries are not required to directly support any attributes for this function. However, any provided attributes must be passed to the host SMS daemon for processing, and the PMIx library is *required* to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the client process making the request.

Host environments that implement support for this operation are required to support the following attributes:

PMIX_JOB_CTRL_ID "pmix.jctrl.id" (char*)
Provide a string identifier for this request. The user can provide an identifier for the requested operation, thus allowing them to later request status of the operation or to terminate it. The host, therefore, shall track it with the request for future reference.

PMIX_JOB_CTRL_PAUSE "pmix.jctrl.pause" (bool)
Pause the specified processes.

PMIX_JOB_CTRL_RESUME "pmix.jctrl.resume" (bool)
Resume ("un-pause") the specified processes.

1 **PMIX_JOB_CTRL_KILL** "pmix.jctrl.kill" (bool)
 2 Forcibly terminate the specified processes and cleanup.

3 **PMIX_JOB_CTRL_SIGNAL** "pmix.jctrl.sig" (int)
 4 Send given signal to specified processes.

5 **PMIX_JOB_CTRL_TERMINATE** "pmix.jctrl.term" (bool)
 6 Politely terminate the specified processes.

7 **PMIX_REGISTER_CLEANUP** "pmix.reg.cleanup" (char*)
 8 Comma-delimited list of files to be removed upon process termination.

9 **PMIX_REGISTER_CLEANUP_DIR** "pmix.reg.cleanupdir" (char*)
 10 Comma-delimited list of directories to be removed upon process termination.

11 **PMIX_CLEANUP_RECURSIVE** "pmix.clnup.recurse" (bool)
 12 Recursively cleanup all subdirectories under the specified one(s).

13 **PMIX_CLEANUP_EMPTY** "pmix.clnup.empty" (bool)
 14 Only remove empty subdirectories.

15 **PMIX_CLEANUP_IGNORE** "pmix.clnup.ignore" (char*)
 16 Comma-delimited list of filenames that are not to be removed.

17 **PMIX_CLEANUP_LEAVE_TOPDIR** "pmix.clnup.lvtop" (bool)
 18 When recursively cleaning subdirectories, do not remove the top-level directory (the one given in the
 19 cleanup request).

▲-----▲
 ▼-----▼ **Optional Attributes** -----▼

20 The following attributes are optional for host environments that support this operation:

21 **PMIX_JOB_CTRL_CANCEL** "pmix.jctrl.cancel" (char*)
 22 Cancel the specified request - the provided request ID must match the **PMIX_JOB_CTRL_ID**
 23 provided to a previous call to **PMIx_Job_control**. An ID of **NULL** implies cancel all requests from
 24 this requestor.

25 **PMIX_JOB_CTRL_RESTART** "pmix.jctrl.restart" (char*)
 26 Restart the specified processes using the given checkpoint ID.

27 **PMIX_JOB_CTRL_CHECKPOINT** "pmix.jctrl.ckpt" (char*)
 28 Checkpoint the specified processes and assign the given ID to it.

29 **PMIX_JOB_CTRL_CHECKPOINT_EVENT** "pmix.jctrl.ckptev" (bool)
 30 Use event notification to trigger a process checkpoint.

31 **PMIX_JOB_CTRL_CHECKPOINT_SIGNAL** "pmix.jctrl.ckptsig" (int)
 32 Use the given signal to trigger a process checkpoint.

33 **PMIX_JOB_CTRL_CHECKPOINT_TIMEOUT** "pmix.jctrl.ckptsig" (int)
 34 Time in seconds to wait for a checkpoint to complete.

35 **PMIX_JOB_CTRL_CHECKPOINT_METHOD** "pmix.jctrl.ckmethod" (pmix_data_array_t)

1 Array of `pmix_info_t` declaring each method and value supported by this application.

2 **PMIX_JOB_CTRL_PROVISION** "pmix.jctrl.pvn" (char*)

3 Regular expression identifying nodes that are to be provisioned.

4 **PMIX_JOB_CTRL_PROVISION_IMAGE** "pmix.jctrl.pvning" (char*)

5 Name of the image that is to be provisioned.

6 **PMIX_JOB_CTRL_PREEMPTIBLE** "pmix.jctrl.preempt" (bool)

7 Indicate that the job can be pre-empted.



8 **Description**

9 Non-blocking form of the `PMIx_Job_control` API. The *targets* array identifies the processes to which the
10 requested job control action is to be applied. All *clones* of an identified process are to have the requested
11 action applied to them. A `NULL` value can be used to indicate all processes in the caller's namespace. The use
12 of `PMIX_RANK_WILDCARD` can also be used to indicate that all processes in the given namespace are to be
13 included.

14 The directives are provided as `pmix_info_t` structures in the *directives* array. The callback function
15 provides a *status* to indicate whether or not the request was granted, and to provide some information as to the
16 reason for any denial in the `pmix_info_cbfunc_t` array of `pmix_info_t` structures.

17 **13.2.3 Job control constants**

18 The following constants are specifically defined for return by the job control APIs:

19 **PMIX_ERR_CONFLICTING_CLEANUP_DIRECTIVES** -51 Conflicting directives given for
20 job/process cleanup.

21 **13.2.4 Job control events**

22 The following job control events may be available for registration, depending upon implementation and host
23 environment support:

24 **PMIX_JCTRL_CHECKPOINT** -106 Monitored by PMIx client to trigger a checkpoint operation.

25 **PMIX_JCTRL_CHECKPOINT_COMPLETE** -107 Sent by a PMIx client and monitored by a PMIx
26 server to notify that requested checkpoint operation has completed.

27 **PMIX_JCTRL_PREEMPT_ALERT** -108 Monitored by a PMIx client to detect that an RM intends to
28 preempt the job.

29 **PMIX_ERR_PROC_RESTART** -4 Error in process restart.

30 **PMIX_ERR_PROC_CHECKPOINT** -5 Error in process checkpoint.

31 **PMIX_ERR_PROC_MIGRATE** -6 Error in process migration.

13.2.5 Job control attributes

Attributes used to request control operations on an executing application - these are values passed to the job control APIs and are not accessed using the `PMIx_Get` API.

PMIX_JOB_CTRL_ID "pmix.jctrl.id" (char*)

Provide a string identifier for this request. The user can provide an identifier for the requested operation, thus allowing them to later request status of the operation or to terminate it. The host, therefore, shall track it with the request for future reference.

PMIX_JOB_CTRL_PAUSE "pmix.jctrl.pause" (bool)

Pause the specified processes.

PMIX_JOB_CTRL_RESUME "pmix.jctrl.resume" (bool)

Resume ("un-pause") the specified processes.

PMIX_JOB_CTRL_CANCEL "pmix.jctrl.cancel" (char*)

Cancel the specified request - the provided request ID must match the `PMIX_JOB_CTRL_ID` provided to a previous call to `PMIx_Job_control`. An ID of `NULL` implies cancel all requests from this requestor.

PMIX_JOB_CTRL_KILL "pmix.jctrl.kill" (bool)

Forcibly terminate the specified processes and cleanup.

PMIX_JOB_CTRL_RESTART "pmix.jctrl.restart" (char*)

Restart the specified processes using the given checkpoint ID.

PMIX_JOB_CTRL_CHECKPOINT "pmix.jctrl.ckpt" (char*)

Checkpoint the specified processes and assign the given ID to it.

PMIX_JOB_CTRL_CHECKPOINT_EVENT "pmix.jctrl.ckptev" (bool)

Use event notification to trigger a process checkpoint.

PMIX_JOB_CTRL_CHECKPOINT_SIGNAL "pmix.jctrl.ckptsig" (int)

Use the given signal to trigger a process checkpoint.

PMIX_JOB_CTRL_CHECKPOINT_TIMEOUT "pmix.jctrl.ckptsig" (int)

Time in seconds to wait for a checkpoint to complete.

PMIX_JOB_CTRL_CHECKPOINT_METHOD "pmix.jctrl.ckmethod" (pmix_data_array_t)

Array of `pmix_info_t` declaring each method and value supported by this application.

PMIX_JOB_CTRL_SIGNAL "pmix.jctrl.sig" (int)

Send given signal to specified processes.

PMIX_JOB_CTRL_PROVISION "pmix.jctrl.pvn" (char*)

Regular expression identifying nodes that are to be provisioned.

PMIX_JOB_CTRL_PROVISION_IMAGE "pmix.jctrl.pvning" (char*)

Name of the image that is to be provisioned.

PMIX_JOB_CTRL_PREEMPTIBLE "pmix.jctrl.preempt" (bool)

Indicate that the job can be pre-empted.

PMIX_JOB_CTRL_TERMINATE "pmix.jctrl.term" (bool)

Politely terminate the specified processes.

PMIX_REGISTER_CLEANUP "pmix.reg.cleanup" (char*)

Comma-delimited list of files to be removed upon process termination.

PMIX_REGISTER_CLEANUP_DIR "pmix.reg.cleanupdir" (char*)

Comma-delimited list of directories to be removed upon process termination.

PMIX_CLEANUP_RECURSIVE "pmix.clnup.recurse" (bool)

Recursively cleanup all subdirectories under the specified one(s).

```

1 PMIX_CLEANUP_EMPTY "pmix.clnup.empty" (bool)
2     Only remove empty subdirectories.
3 PMIX_CLEANUP_IGNORE "pmix.clnup.ignore" (char*)
4     Comma-delimited list of filenames that are not to be removed.
5 PMIX_CLEANUP_LEAVE_TOPDIR "pmix.clnup.lvtop" (bool)
6     When recursively cleaning subdirectories, do not remove the top-level directory (the one given in the
7     cleanup request).

```

8 13.3 Process and Job Monitoring

9 In addition to external faults, a common problem encountered in HPC applications is a failure to make
10 progress due to some internal conflict in the computation. These situations can result in a significant waste of
11 resources as the SMS is unaware of the problem, and thus cannot terminate the job. Various watchdog
12 methods have been developed for detecting this situation, including requiring a periodic “heartbeat” from the
13 application and monitoring a specified file for changes in size and/or modification time.

14 The following APIs allow applications to request monitoring, directing what is to be monitored, the frequency
15 of the associated check, whether or not the application is to be notified (via the event notification subsystem) of
16 stall detection, and other characteristics of the operation.

17 13.3.1 PMIx_Process_monitor

18 Summary

19 Request that application processes be monitored.

20 Format

PMIx v3.0

```

21 pmix_status_t
22 PMIx_Process_monitor(const pmix_info_t *monitor,
23                     pmix_status_t error,
24                     const pmix_info_t directives[], size_t ndirs,
25                     pmix_info_t *results[], size_t *nresults);

```

26 **IN monitor**

info (handle)

28 **IN error**

status (integer)

30 **IN directives**

Array of info structures (array of handles)

32 **IN ndirs**

Number of elements in the *directives* array (integer)

34 **INOUT results**

Address where a pointer to an array of `pmix_info_t` containing the results of the request can be
returned (memory reference)

37 **INOUT nresults**

Address where the number of elements in *results* can be returned (handle)

1 A successful return indicates that the results have been placed in the *results* array.

2 Returns **PMIX_SUCCESS** or a negative value indicating the error.

▼----- Optional Attributes -----▼

3 The following attributes may be implemented by a PMIx library or by the host environment. If supported by
4 the PMIx server library, then the library must not pass the supported attributes to the host environment. All
5 attributes not directly supported by the server library must be passed to the host environment if it supports this
6 operation, and the library is *required* to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the
7 requesting process:

8 **PMIX_MONITOR_ID** "pmix.monitor.id" (char*)

9 Provide a string identifier for this request.

10 **PMIX_MONITOR_CANCEL** "pmix.monitor.cancel" (char*)

11 Identifier to be canceled (**NULL** means cancel all monitoring for this process).

12 **PMIX_MONITOR_APP_CONTROL** "pmix.monitor.appctrl" (bool)

13 The application desires to control the response to a monitoring event - i.e., the application is requesting
14 that the host environment not take immediate action in response to the event (e.g., terminating the job).

15
16 **PMIX_MONITOR_HEARTBEAT** "pmix.monitor.mbeat" (void)

17 Register to have the PMIx server monitor the requestor for heartbeats.

18 **PMIX_MONITOR_HEARTBEAT_TIME** "pmix.monitor.btime" (uint32_t)

19 Time in seconds before declaring heartbeat missed.

20 **PMIX_MONITOR_HEARTBEAT_DROPS** "pmix.monitor.bdrop" (uint32_t)

21 Number of heartbeats that can be missed before generating the event.

22 **PMIX_MONITOR_FILE** "pmix.monitor.fmon" (char*)

23 Register to monitor file for signs of life.

24 **PMIX_MONITOR_FILE_SIZE** "pmix.monitor.fsize" (bool)

25 Monitor size of given file is growing to determine if the application is running.

26 **PMIX_MONITOR_FILE_ACCESS** "pmix.monitor.faccess" (char*)

27 Monitor time since last access of given file to determine if the application is running.

28 **PMIX_MONITOR_FILE_MODIFY** "pmix.monitor.fmod" (char*)

29 Monitor time since last modified of given file to determine if the application is running.

30 **PMIX_MONITOR_FILE_CHECK_TIME** "pmix.monitor.ftime" (uint32_t)

31 Time in seconds between checking the file.

32 **PMIX_MONITOR_FILE_DROPS** "pmix.monitor.fdrop" (uint32_t)

33 Number of file checks that can be missed before generating the event.

34 **PMIX_SEND_HEARTBEAT** "pmix.monitor.beat" (void)

35 Send heartbeat to local PMIx server.
▲-----

Description

Request that application processes be monitored via several possible methods. For example, that the server monitor this process for periodic heartbeats as an indication that the process has not become “wedged”. When a monitor detects the specified alarm condition, it will generate an event notification using the provided error code and passing along any available relevant information. It is up to the caller to register a corresponding event handler.

The *monitor* argument is an attribute indicating the type of monitor being requested. For example, `PMIX_MONITOR_FILE` to indicate that the requestor is asking that a file be monitored.

The *error* argument is the status code to be used when generating an event notification alerting that the monitor has been triggered. The range of the notification defaults to `PMIX_RANGE_NAMESPACE`. This can be changed by providing a `PMIX_RANGE` directive.

The *directives* argument characterizes the monitoring request (e.g., monitor file size) and frequency of checking to be done

The returned *status* indicates whether or not the request was granted, and information as to the reason for any denial of the request shall be returned in the *results* array.

13.3.2 PMIx_Process_monitor_nb

Summary

Request that application processes be monitored.

Format

PMIx v2.0

```
pmix_status_t
PMIx_Process_monitor_nb(const pmix_info_t *monitor,
                        pmix_status_t error,
                        const pmix_info_t directives[],
                        size_t ndirs,
                        pmix_info_cbfunc_t cbfunc, void *cbdata);
```

IN monitor
info (handle)

IN error
status (integer)

IN directives
Array of info structures (array of handles)

IN ndirs
Number of elements in the *directives* array (integer)

IN cbfunc
Callback function `pmix_info_cbfunc_t` (function reference)

IN cbdata
Data to be passed to the callback function (memory reference)

1 Returns one of the following:

2 A successful return indicates that the request is being processed and the result will be returned in the provided
3 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
4 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

5 Returns PMIX_SUCCESS or one of the following error codes when the condition described occurs:

- 6 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
7 *success* - the *cbfunc* will *not* be called

8 If none of the above return codes are appropriate, then an implementation must return either a general PMIX
9 error code or an implementation defined error code as described in Section 3.1.1.

▼----- Optional Attributes -----▼

10 The following attributes may be implemented by a PMIX library or by the host environment. If supported by
11 the PMIX server library, then the library must not pass the supported attributes to the host environment. All
12 attributes not directly supported by the server library must be passed to the host environment if it supports this
13 operation, and the library is *required* to add the **PMIX_USERID** and the **PMIX_GRPID** attributes of the
14 requesting process:

- 15 **PMIX_MONITOR_ID** "pmix.monitor.id" (char*)
16 Provide a string identifier for this request.
- 17 **PMIX_MONITOR_CANCEL** "pmix.monitor.cancel" (char*)
18 Identifier to be canceled (**NULL** means cancel all monitoring for this process).
- 19 **PMIX_MONITOR_APP_CONTROL** "pmix.monitor.appctrl" (bool)
20 The application desires to control the response to a monitoring event - i.e., the application is requesting
21 that the host environment not take immediate action in response to the event (e.g., terminating the job).
22
- 23 **PMIX_MONITOR_HEARTBEAT** "pmix.monitor.mbeat" (void)
24 Register to have the PMIX server monitor the requestor for heartbeats.
- 25 **PMIX_MONITOR_HEARTBEAT_TIME** "pmix.monitor.btime" (uint32_t)
26 Time in seconds before declaring heartbeat missed.
- 27 **PMIX_MONITOR_HEARTBEAT_DROPS** "pmix.monitor.bdrop" (uint32_t)
28 Number of heartbeats that can be missed before generating the event.
- 29 **PMIX_MONITOR_FILE** "pmix.monitor.fmon" (char*)
30 Register to monitor file for signs of life.
- 31 **PMIX_MONITOR_FILE_SIZE** "pmix.monitor.fsize" (bool)
32 Monitor size of given file is growing to determine if the application is running.
- 33 **PMIX_MONITOR_FILE_ACCESS** "pmix.monitor.faccess" (char*)
34 Monitor time since last access of given file to determine if the application is running.
- 35 **PMIX_MONITOR_FILE_MODIFY** "pmix.monitor.fmod" (char*)
36 Monitor time since last modified of given file to determine if the application is running.

1 **PMIX_MONITOR_FILE_CHECK_TIME** "pmix.monitor.ftime" (uint32_t)
 2 Time in seconds between checking the file.
 3 **PMIX_MONITOR_FILE_DROPS** "pmix.monitor.fdrop" (uint32_t)
 4 Number of file checks that can be missed before generating the event.
 5 **PMIX_SEND_HEARTBEAT** "pmix.monitor.beat" (void)
 6 Send heartbeat to local PMIx server.



7 **Description**

8 Non-blocking form of the **PMIx_Process_monitor** API. The *cbfunc* function provides a *status* to
 9 indicate whether or not the request was granted, and to provide some information as to the reason for any
 10 denial in the **pmix_info_cbfunc_t** array of **pmix_info_t** structures.

11 **13.3.3 PMIx_Heartbeat**

12 **Summary**

13 Send a heartbeat to the PMIx server library

14 **Format**

PMIx v2.0

15 **PMIx_Heartbeat** ();

16 **Description**

17 A simplified macro wrapping **PMIx_Process_monitor_nb** that sends a heartbeat to the PMIx server
 18 library.

19 **13.3.4 Monitoring events**

20 The following monitoring events may be available for registration, depending upon implementation and host
 21 environment support:

- 22 **PMIX_MONITOR_HEARTBEAT_ALERT** -109 Heartbeat failed to arrive within specified window. The
 23 process that triggered this alert will be identified in the event.
 24 **PMIX_MONITOR_FILE_ALERT** -110 File failed its monitoring detection criteria. The file that
 25 triggered this alert will be identified in the event.

1 13.3.5 Monitoring attributes

2 Attributes used to control monitoring of an executing application- these are values passed to the
3 [PMIx_Process_monitor_nb](#) API and are not accessed using the [PMIx_Get](#) API.

4 **PMIX_MONITOR_ID** "pmix.monitor.id" (char*)

5 Provide a string identifier for this request.

6 **PMIX_MONITOR_CANCEL** "pmix.monitor.cancel" (char*)

7 Identifier to be canceled (NULL means cancel all monitoring for this process).

8 **PMIX_MONITOR_APP_CONTROL** "pmix.monitor.appctrl" (bool)

9 The application desires to control the response to a monitoring event - i.e., the application is requesting
10 that the host environment not take immediate action in response to the event (e.g., terminating the job).

11 **PMIX_MONITOR_HEARTBEAT** "pmix.monitor.mbeat" (void)

12 Register to have the PMIx server monitor the requestor for heartbeats.

13 **PMIX_SEND_HEARTBEAT** "pmix.monitor.beat" (void)

14 Send heartbeat to local PMIx server.

15 **PMIX_MONITOR_HEARTBEAT_TIME** "pmix.monitor.btime" (uint32_t)

16 Time in seconds before declaring heartbeat missed.

17 **PMIX_MONITOR_HEARTBEAT_DROPS** "pmix.monitor.bdrop" (uint32_t)

18 Number of heartbeats that can be missed before generating the event.

19 **PMIX_MONITOR_FILE** "pmix.monitor.fmon" (char*)

20 Register to monitor file for signs of life.

21 **PMIX_MONITOR_FILE_SIZE** "pmix.monitor.fsize" (bool)

22 Monitor size of given file is growing to determine if the application is running.

23 **PMIX_MONITOR_FILE_ACCESS** "pmix.monitor.faccess" (char*)

24 Monitor time since last access of given file to determine if the application is running.

25 **PMIX_MONITOR_FILE_MODIFY** "pmix.monitor.fmod" (char*)

26 Monitor time since last modified of given file to determine if the application is running.

27 **PMIX_MONITOR_FILE_CHECK_TIME** "pmix.monitor.ftime" (uint32_t)

28 Time in seconds between checking the file.

29 **PMIX_MONITOR_FILE_DROPS** "pmix.monitor.fdrop" (uint32_t)

30 Number of file checks that can be missed before generating the event.

31 13.4 Logging

32 The logging interface supports posting information by applications and SMS elements to persistent storage.
33 This function is *not* intended for output of computational results, but rather for reporting status and saving
34 state information such as inserting computation progress reports into the application's SMS job log or error
35 reports to the local syslog.

36 13.4.1 PMIx_Log

37 Summary

38 Log data to a data service.

Format

C

```
pmix_status_t
PMix_Log(const pmix_info_t data[], size_t ndata,
         const pmix_info_t directives[], size_t ndirs);
```

C

IN data
Array of info structures (array of handles)

IN ndata
Number of elements in the *data* array (**size_t**)

IN directives
Array of info structures (array of handles)

IN ndirs
Number of elements in the *directives* array (**size_t**)

Returns **PMIX_SUCCESS** or a negative value indicating the error.

Required Attributes

If the PMIx library does not itself perform this operation, then it is required to pass any attributes provided by the client to the host environment for processing. In addition, it must include the following attributes in the passed *info* array:

PMIX_USERID "pmix.euid" (**uint32_t**)
Effective user ID of the connecting process.

PMIX_GRPID "pmix.egid" (**uint32_t**)
Effective group ID of the connecting process.

Host environments or PMIx libraries that implement support for this operation are required to support the following attributes:

PMIX_LOG_STDERR "pmix.log.stderr" (**char***)
Log string to **stderr**.

PMIX_LOG_STDOUT "pmix.log.stdout" (**char***)
Log string to **stdout**.

PMIX_LOG_SYSLOG "pmix.log.syslog" (**char***)
Log data to syslog. Defaults to **ERROR** priority. Will log to global syslog if available, otherwise to local syslog.

PMIX_LOG_LOCAL_SYSLOG "pmix.log.lsys" (**char***)
Log data to local syslog. Defaults to **ERROR** priority.

PMIX_LOG_GLOBAL_SYSLOG "pmix.log.gsys" (**char***)
Forward data to system "gateway" and log msg to that syslog Defaults to **ERROR** priority.

PMIX_LOG_SYSLOG_PRI "pmix.log.syspri" (**int**)
Syslog priority level.

PMIX_LOG_ONCE "pmix.log.once" (**bool**)

1 Only log this once with whichever channel can first support it, taking the channels in priority order.

Optional Attributes

2 The following attributes are optional for host environments or PMIx libraries that support this operation:

3 **PMIX_LOG_SOURCE** "pmix.log.source" (pmix_proc_t*)

4 ID of source of the log request.

5 **PMIX_LOG_TIMESTAMP** "pmix.log.tstmp" (time_t)

6 Timestamp for log report.

7 **PMIX_LOG_GENERATE_TIMESTAMP** "pmix.log.gtstmp" (bool)

8 Generate timestamp for log.

9 **PMIX_LOG_TAG_OUTPUT** "pmix.log.tag" (bool)

10 Label the output stream with the channel name (e.g., "stdout").

11 **PMIX_LOG_TIMESTAMP_OUTPUT** "pmix.log.tsout" (bool)

12 Print timestamp in output string.

13 **PMIX_LOG_XML_OUTPUT** "pmix.log.xml" (bool)

14 Print the output stream in eXtensible Markup Language (XML) format.

15 **PMIX_LOG_EMAIL** "pmix.log.email" (pmix_data_array_t)

16 Log via email based on **pmix_info_t** containing directives.

17 **PMIX_LOG_EMAIL_ADDR** "pmix.log.emaddr" (char*)

18 Comma-delimited list of email addresses that are to receive the message.

19 **PMIX_LOG_EMAIL_SENDER_ADDR** "pmix.log.emfaddr" (char*)

20 Return email address of sender.

21 **PMIX_LOG_EMAIL_SERVER** "pmix.log.esrvr" (char*)

22 Hostname (or IP address) of SMTP server.

23 **PMIX_LOG_EMAIL_SRVR_PORT** "pmix.log.esrvrprt" (int32_t)

24 Port the email server is listening to.

25 **PMIX_LOG_EMAIL_SUBJECT** "pmix.log.emsub" (char*)

26 Subject line for email.

27 **PMIX_LOG_EMAIL_MSG** "pmix.log.emmsg" (char*)

28 Message to be included in email.

29 **PMIX_LOG_JOB_RECORD** "pmix.log.jrec" (bool)

30 Log the provided information to the host environment's job record.

31 **PMIX_LOG_GLOBAL_DATASTORE** "pmix.log.gstore" (bool)

32 Store the log data in a global data store (e.g., database).

Description

Log data subject to the services offered by the host environment. The data to be logged is provided in the *data* array. The (optional) *directives* can be used to direct the choice of logging channel.

Advice to users

It is strongly recommended that the **PMIx_Log** API not be used by applications for streaming data as it is not a “performant” transport and can perturb the application since it involves the local PMIx server and host SMS daemon. Note that a return of **PMIX_SUCCESS** only denotes that the data was successfully handed to the appropriate system call (for local channels) or the host environment and does not indicate receipt at the final destination.

13.4.2 PMIx_Log_nb

Summary

Log data to a data service.

Format

PMIx v2.0

```
pmix_status_t
PMIx_Log_nb(const pmix_info_t data[], size_t ndata,
            const pmix_info_t directives[], size_t ndirs,
            pmix_op_cbfunc_t cbfunc, void *cbdata);
```

- IN data**
Array of info structures (array of handles)
- IN ndata**
Number of elements in the *data* array (**size_t**)
- IN directives**
Array of info structures (array of handles)
- IN ndirs**
Number of elements in the *directives* array (**size_t**)
- IN cbfunc**
Callback function **pmix_op_cbfunc_t** (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

Return codes are one of the following:

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

PMIX_OPERATION_SUCCEEDED, indicating that the request was immediately processed and returned *success* - the *cbfunc* will *not* be called

1 **PMIX_ERR_BAD_PARAM** The logging request contains at least one incorrect entry that prevents it from
2 being processed. The callback function will not be called.

3 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
4 error code or an implementation defined error code as described in Section 3.1.1.

▼----- Required Attributes -----▼

5 If the PMIx library does not itself perform this operation, then it is required to pass any attributes provided by
6 the client to the host environment for processing. In addition, it must include the following attributes in the
7 passed *info* array:

8 **PMIX_USERID** "pmix.euid" (uint32_t)

9 Effective user ID of the connecting process.

10 **PMIX_GRPID** "pmix.egid" (uint32_t)

11 Effective group ID of the connecting process.

12 Host environments or PMIx libraries that implement support for this operation are required to support the
13 following attributes:

14 **PMIX_LOG_STDERR** "pmix.log.stderr" (char*)

15 Log string to **stderr**.

16 **PMIX_LOG_STDOUT** "pmix.log.stdout" (char*)

17 Log string to **stdout**.

18 **PMIX_LOG_SYSLOG** "pmix.log.syslog" (char*)

19 Log data to syslog. Defaults to **ERROR** priority. Will log to global syslog if available, otherwise to
20 local syslog.

21 **PMIX_LOG_LOCAL_SYSLOG** "pmix.log.lsys" (char*)

22 Log data to local syslog. Defaults to **ERROR** priority.

23 **PMIX_LOG_GLOBAL_SYSLOG** "pmix.log.gsys" (char*)

24 Forward data to system "gateway" and log msg to that syslog Defaults to **ERROR** priority.

25 **PMIX_LOG_SYSLOG_PRI** "pmix.log.syspri" (int)

26 Syslog priority level.

27 **PMIX_LOG_ONCE** "pmix.log.once" (bool)

28 Only log this once with whichever channel can first support it, taking the channels in priority order.



Optional Attributes

The following attributes are optional for host environments or PMIx libraries that support this operation:

- 1 **PMIX_LOG_SOURCE** "pmix.log.source" (pmix_proc_t*)
2 ID of source of the log request.
- 3
- 4 **PMIX_LOG_TIMESTAMP** "pmix.log.tstamp" (time_t)
5 Timestamp for log report.
- 6 **PMIX_LOG_GENERATE_TIMESTAMP** "pmix.log.gtstamp" (bool)
7 Generate timestamp for log.
- 8 **PMIX_LOG_TAG_OUTPUT** "pmix.log.tag" (bool)
9 Label the output stream with the channel name (e.g., "stdout").
- 10 **PMIX_LOG_TIMESTAMP_OUTPUT** "pmix.log.tsout" (bool)
11 Print timestamp in output string.
- 12 **PMIX_LOG_XML_OUTPUT** "pmix.log.xml" (bool)
13 Print the output stream in XML format.
- 14 **PMIX_LOG_EMAIL** "pmix.log.email" (pmix_data_array_t)
15 Log via email based on **pmix_info_t** containing directives.
- 16 **PMIX_LOG_EMAIL_ADDR** "pmix.log.emaddr" (char*)
17 Comma-delimited list of email addresses that are to receive the message.
- 18 **PMIX_LOG_EMAIL_SENDER_ADDR** "pmix.log.emfaddr" (char*)
19 Return email address of sender.
- 20 **PMIX_LOG_EMAIL_SERVER** "pmix.log.esrvr" (char*)
21 Hostname (or IP address) of SMTP server.
- 22 **PMIX_LOG_EMAIL_SRVR_PORT** "pmix.log.esrvrprt" (int32_t)
23 Port the email server is listening to.
- 24 **PMIX_LOG_EMAIL_SUBJECT** "pmix.log.emsub" (char*)
25 Subject line for email.
- 26 **PMIX_LOG_EMAIL_MSG** "pmix.log.emmsg" (char*)
27 Message to be included in email.
- 28 **PMIX_LOG_JOB_RECORD** "pmix.log.jrec" (bool)
29 Log the provided information to the host environment's job record.
- 30 **PMIX_LOG_GLOBAL_DATASTORE** "pmix.log.gstore" (bool)
31 Store the log data in a global data store (e.g., database).

Description

Log data subject to the services offered by the host environment. The data to be logged is provided in the *data* array. The (optional) *directives* can be used to direct the choice of logging channel. The callback function will be executed when the log operation has been completed. The *data* and *directives* arrays must be maintained until the callback is provided.

Advice to users

It is strongly recommended that the `PMIx_Log_nb` API not be used by applications for streaming data as it is not a “performant” transport and can perturb the application since it involves the local PMIx server and host SMS daemon. Note that a return of `PMIX_SUCCESS` only denotes that the data was successfully handed to the appropriate system call (for local channels) or the host environment and does not indicate receipt at the final destination.

13.4.3 Log attributes

Attributes used to describe `PMIx_Log` behavior - these are values passed to the `PMIx_Log` API and therefore are not accessed using the `PMIx_Get` API.

`PMIX_LOG_SOURCE` "pmix.log.source" (`pmix_proc_t*`)

ID of source of the log request.

`PMIX_LOG_STDERR` "pmix.log.stderr" (`char*`)

Log string to `stderr`.

`PMIX_LOG_STDOUT` "pmix.log.stdout" (`char*`)

Log string to `stdout`.

`PMIX_LOG_SYSLOG` "pmix.log.syslog" (`char*`)

Log data to syslog. Defaults to `ERROR` priority. Will log to global syslog if available, otherwise to local syslog.

`PMIX_LOG_LOCAL_SYSLOG` "pmix.log.lsys" (`char*`)

Log data to local syslog. Defaults to `ERROR` priority.

`PMIX_LOG_GLOBAL_SYSLOG` "pmix.log.gsys" (`char*`)

Forward data to system “gateway” and log msg to that syslog Defaults to `ERROR` priority.

`PMIX_LOG_SYSLOG_PRI` "pmix.log.syspri" (`int`)

Syslog priority level.

`PMIX_LOG_TIMESTAMP` "pmix.log.tstamp" (`time_t`)

Timestamp for log report.

`PMIX_LOG_GENERATE_TIMESTAMP` "pmix.log.gtstamp" (`bool`)

Generate timestamp for log.

`PMIX_LOG_TAG_OUTPUT` "pmix.log.tag" (`bool`)

Label the output stream with the channel name (e.g., “stdout”).

`PMIX_LOG_TIMESTAMP_OUTPUT` "pmix.log.tsout" (`bool`)

Print timestamp in output string.

`PMIX_LOG_XML_OUTPUT` "pmix.log.xml" (`bool`)

Print the output stream in XML format.

`PMIX_LOG_ONCE` "pmix.log.once" (`bool`)

Only log this once with whichever channel can first support it, taking the channels in priority order.

```

1  PMIX_LOG_MSG "pmix.log.msg" (pmix_byte_object_t)
2      Message blob to be sent somewhere.
3  PMIX_LOG_EMAIL "pmix.log.email" (pmix_data_array_t)
4      Log via email based on pmix_info_t containing directives.
5  PMIX_LOG_EMAIL_ADDR "pmix.log.emaddr" (char*)
6      Comma-delimited list of email addresses that are to receive the message.
7  PMIX_LOG_EMAIL_SENDER_ADDR "pmix.log.emfaddr" (char*)
8      Return email address of sender.
9  PMIX_LOG_EMAIL_SUBJECT "pmix.log.emsub" (char*)
10     Subject line for email.
11  PMIX_LOG_EMAIL_MSG "pmix.log.emmsg" (char*)
12     Message to be included in email.
13  PMIX_LOG_EMAIL_SERVER "pmix.log.esrvr" (char*)
14     Hostname (or IP address) of SMTP server.
15  PMIX_LOG_EMAIL_SRVR_PORT "pmix.log.esrvrprt" (int32_t)
16     Port the email server is listening to.
17  PMIX_LOG_GLOBAL_DATASTORE "pmix.log.gstore" (bool)
18     Store the log data in a global data store (e.g., database).
19  PMIX_LOG_JOB_RECORD "pmix.log.jrec" (bool)
20     Log the provided information to the host environment's job record.

```

CHAPTER 14

Process Sets and Groups

1 PMIx supports two slightly related, but functionally different concepts known as *process sets* and *process*
2 *groups*. This chapter defines these two concepts and describes how they are utilized, along with their
3 corresponding APIs.

4 14.1 Process Sets

5 A PMIx *Process Set* is a user-provided or host environment assigned label associated with a given set of
6 application processes. Processes can belong to multiple process *sets* at a time. Users may define a PMIx
7 process set at time of application execution. For example, if using the command line parallel launcher "prun",
8 one could specify process sets as follows:

```
9 $ prun -n 4 --pset ocean myoceanapp : -n 3 --pset ice myiceapp
```



10 In this example, the processes in the first application will be labeled with a `PMIX_PSET_NAMES` attribute
11 with a value of *ocean* while those in the second application will be labeled with an *ice* value. During the
12 execution, application processes could lookup the process set for any process using `PMIx_Get`.
13 Alternatively, other executing applications could utilize the `PMIx_Query_info` APIs to obtain the number
14 of declared process sets in the system, a list of their names, and other information about them. In other words,
15 the *process set* identifier provides a label by which an application can derive information about a process and
16 its application - it does *not*, however, confer any operational function.

17 Host environments can create or delete process sets at any time through the
18 `PMIx_server_define_process_set` and `PMIx_server_delete_process_set` APIs. PMIx
19 servers shall notify all local clients of process set operations via the `PMIX_PROCESS_SET_DEFINE` or
20 `PMIX_PROCESS_SET_DELETE` events.

21 Process *sets* differ from process *groups* in several key ways:

- 22 • Process *sets* have no implied relationship between their members - i.e., a process in a process set has no
23 concept of a "pset rank" as it would in a process *group*.
- 24 • Process *set* identifiers are set by the host environment or by the user at time of application submission for
25 execution - there are no PMIx APIs provided by which an application can define a process set or change a
26 process *set* membership. In contrast, PMIx process *groups* can only be defined dynamically by the
27 application.
- 28 • Process *sets* are immutable - members cannot be added or removed once the set has been defined. In
29 contrast, PMIx process *groups* can dynamically change their membership using the appropriate APIs.

- Process *groups* can be used in calls to PMIx operations. Members of process *groups* that are involved in an operation are translated by their PMIx server into their *native* identifier prior to the operation being passed to the host environment. For example, an application can define a process group to consist of ranks 0 and 1 from the host-assigned namespace of *210456*, identified by the group id of *foo*. If the application subsequently calls the **PMIx_Fence** API with a process identifier of `{foo, PMIX_RANK_WILDCARD}`, the PMIx server will replace that identifier with an array consisting of `{210456, 0}` and `{210456, 1}` - the host-assigned identifiers of the participating processes - prior to processing the request.
- Process *groups* can request that the host environment assign a unique `size_t` Process Group Context Identifier (PGCID) to the group at time of group construction. An Message Passing Interface (MPI) library may, for example, use the PGCID as the MPI communicator identifier for the group.

The two concepts do, however, overlap in that they both involve collections of processes. Users desiring to create a process group based on a process set could, for example, obtain the membership array of the process set and use that as input to **PMIx_Group_construct**, perhaps including the process set name as the group identifier for clarity. Note that no linkage between the set and group of the same name is implied nor maintained - e.g., changes in process group membership can not be reflected in the process set using the same identifier.

Advice to PMIx server hosts

The host environment is responsible for ensuring:

- consistent knowledge of process set membership across all involved PMIx servers; and
- that process set names do not conflict with system-assigned namespaces within the scope of the set.

14.1.1 Process Set Constants

PMIx v4.0

The PMIx server is required to send a notification to all local clients upon creation or deletion of process sets. Client processes wishing to receive such notifications must register for the corresponding event:

PMIX_PROCESS_SET_DEFINE -55 The host environment has defined a new process set - the event will include the process set name (**PMIX_PSET_NAME**) and the membership (**PMIX_PSET_MEMBERS**).

PMIX_PROCESS_SET_DELETE -56 The host environment has deleted a process set - the event will include the process set name (**PMIX_PSET_NAME**).

1 14.1.2 Process Set Attributes

2 Several attributes are provided for querying the system regarding process sets using the [PMIx_Query_info](#)
3 APIs.

4 **PMIX_QUERY_NUM_PSETS** "pmix.qry.psetnum" (`size_t`)
5 Return the number of process sets defined in the specified range (defaults to
6 [PMIX_RANGE_SESSION](#)).

7 **PMIX_QUERY_PSET_NAMES** "pmix.qry.psets" (`pmix_data_array_t*`)
8 Return a `pmix_data_array_t` containing an array of strings of the process set names defined in
9 the specified range (defaults to [PMIX_RANGE_SESSION](#)).

10 **PMIX_QUERY_PSET_MEMBERSHIP** "pmix.qry.pmems" (`pmix_data_array_t*`)
11 Return an array of `pmix_proc_t` containing the members of the specified process set.

12 The [PMIX_PROCESS_SET_DEFINE](#) event shall include the name of the newly defined process set and its
13 members: **PMIX_PSET_NAME** "pmix.pset.nm" (`char*`)

14 The name of the newly defined process set.

15 **PMIX_PSET_MEMBERS** "pmix.pset.mems" (`pmix_data_array_t*`)
16 An array of `pmix_proc_t` containing the members of the newly defined process set.

17 In addition, a process can request (via [PMIx_Get](#)) the process sets to which a given process (including itself)
18 belongs:

19 **PMIX_PSET_NAMES** "pmix.pset.nms" (`pmix_data_array_t*`)
20 Returns an array of `char*` string names of the process sets in which the given process is a member.

21 14.2 Process Groups

22 *PMIx Groups* are defined as a collection of processes desiring a common, unique identifier for operational
23 purposes such as passing events or participating in *PMIx fence* operations. As with processes that assemble
24 via [PMIx_Connect](#), each member of the group is provided with both the job-level information of any other
25 namespace represented in the group, and the contact information for all group members.

26 However, members of *PMIx Groups* are *loosely coupled* as opposed to *tightly connected* when constructed via
27 [PMIx_Connect](#). Thus, *groups* differ from [PMIx_Connect](#) assemblages in several key areas, as detailed in
28 the following sections.

29 14.2.1 Relation to the host environment

30 Calls to *PMIx Group* APIs are first processed within the local *PMIx* server. When constructed, the server
31 creates a tracker that associates the specified processes with the user-provided group identifier, and assigns a
32 new *group rank* based on their relative position in the array of processes provided in the call to
33 [PMIx_Group_construct](#). Members of the group can subsequently utilize the group identifier in *PMIx*
34 function calls to address the group's members, using either [PMIX_RANK_WILDCARD](#) to refer to all of them
35 or the group-level rank of specific members. The *PMIx* server will translate the specified processes into their
36 RM-assigned identifiers prior to passing the request up to its host. Thus, the host environment has no visibility
37 into the group's existence or membership.

1 In contrast, calls to **PMIx_Connect** are relayed to the host environment. This means that the host RM should
2 treat the failure of any process in the specified assemblage as a reportable event and take appropriate action.
3 However, the environment is not required to define a new identifier for the connected assemblage or any of its
4 member processes, nor does it define a new rank for each process within that assemblage. In addition, the
5 PMIx server does not provide any tracking support for the assemblage. Thus, the caller is responsible for
6 addressing members of the connected assemblage using their RM-provided identifiers.

Advice to users

7 User-provided group identifiers must be distinct from both other group identifiers within the system and
8 namespaces provided by the RM so as to avoid collisions between group identifiers and RM-assigned
9 namespaces. This can usually be accomplished through the use of an application-specific prefix – e.g.,
10 “myapp-foo”

11 14.2.2 Construction procedure

12 **PMIx_Connect** calls require that every process call the API before completing – i.e., it is modeled upon the
13 bulk synchronous traditional MPI connect/accept methodology. Thus, a given application thread can only be
14 involved in one connect/accept operation at a time, and is blocked in that operation until all specified processes
15 participate. In addition, there is no provision for replacing processes in the assemblage due to failure to
16 participate, nor a mechanism by which a process might decline participation.

17 In contrast, PMIx Groups are designed to be more flexible in their construction procedure by relaxing these
18 constraints. While a standard blocking form of constructing groups is provided, the event notification system is
19 utilized to provide a designated *group leader* with the ability to replace participants that fail to participate
20 within a given timeout period. This provides a mechanism by which the application can, if desired, replace
21 members on-the-fly or allow the group to proceed with partial membership. In such cases, the final group
22 membership is returned to all participants upon completion of the operation.

23 Additionally, PMIx supports dynamic definition of group membership based on an invite/join model. A
24 process can asynchronously initiate construction of a group of any processes via the **PMIx_Group_invite**
25 function call. Invitations are delivered via a PMIx event (using the **PMIX_GROUP_INVITED** event) to the
26 invited processes which can then either accept or decline the invitation using the **PMIx_Group_join** API.
27 The initiating process tracks responses by registering for the events generated by the call to
28 **PMIx_Group_join**, timeouts, or process terminations, optionally replacing processes that decline the
29 invitation, fail to respond in time, or terminate without responding. Upon completion of the operation, the final
30 list of participants is communicated to each member of the new group.

1 14.2.3 Destruct procedure

2 Members of a PMIx Group may depart the group at any time via the **PMIx_Group_leave** API. Other
3 members are notified of the departure via the **PMIX_GROUP_LEFT** event to distinguish such events from
4 those reporting process termination. This leaves the remaining members free to continue group operations.
5 The **PMIx_Group_destruct** operation offers a collective method akin to **PMIx_Disconnect** for
6 deconstructing the entire group.

7 In contrast, processes that assemble via **PMIx_Connect** must all depart the assemblage together – i.e., no
8 member can depart the assemblage while leaving the remaining members in it. Even the non-blocking form of
9 **PMIx_Disconnect** retains this requirement in that members remain a part of the assemblage until all
10 members have called **PMIx_Disconnect_nb**

11 Note that applications supporting dynamic group behaviors such as asynchronous departure take responsibility
12 for ensuring global consistency in the group definition prior to executing group collective operations - i.e., it is
13 the application’s responsibility to either ensure that knowledge of the current group membership is globally
14 consistent across the participants, or to register for appropriate events to deal with the lack of consistency
15 during the operation.

▼ Advice to users ▼

16 The reliance on PMIx events in the PMIx Group concept dictates that processes utilizing these APIs must
17 register for the corresponding events. Failure to do so will likely lead to operational failures. Users are
18 recommended to utilize the **PMIX_TIMEOUT** directive (or retain an internal timer) on calls to PMIx Group
19 APIs (especially the blocking form of those functions) as processes that have not registered for required events
20 will never respond.

21 14.2.4 Process Group Events

22 *PMIx v4.0*

23 Asynchronous process group operations rely heavily on PMIx events. The following events have been defined
24 for that purpose.

25 **PMIX_GROUP_INVITED** -159 The process has been invited to join a PMIx Group - the identifier of
26 the group and the ID’s of other invited (or already joined) members will be included in the notification.

27 **PMIX_GROUP_LEFT** -160 A process has asynchronously left a PMIx Group - the process identifier of
28 the departing process will be included in the notification.

29 **PMIX_GROUP_MEMBER_FAILED** -170 A member of a PMIx Group has abnormally terminated (i.e.,
30 without formally leaving the group prior to termination) - the process identifier of the failed process will
31 be included in the notification.

32 **PMIX_GROUP_INVITE_ACCEPTED** -161 A process has accepted an invitation to join a PMIx Group
33 - the identifier of the group being joined will be included in the notification.

34 **PMIX_GROUP_INVITE_DECLINED** -162 A process has declined an invitation to join a PMIx Group
35 - the identifier of the declined group will be included in the notification.

36 **PMIX_GROUP_INVITE_FAILED** -163 An invited process failed or terminated prior to responding to
37 the invitation - the identifier of the failed process will be included in the notification.

38 **PMIX_GROUP_MEMBERSHIP_UPDATE** -164 The membership of a PMIx group has changed - the
identifiers of the revised membership will be included in the notification.

1 **PMIX_GROUP_CONSTRUCT_ABORT** -165 Any participant in a PMIx group construct operation that
2 returns **PMIX_GROUP_CONSTRUCT_ABORT** from the *leader failed* event handler will cause all
3 participants to receive an event notifying them of that status. Similarly, the leader may elect to abort the
4 procedure by either returning this error code from the handler assigned to the
5 **PMIX_GROUP_INVITE_ACCEPTED** or **PMIX_GROUP_INVITE_DECLINED** codes, or by
6 generating an event for the abort code. Abort events will be sent to all invited or existing members of the
7 group.

8 **PMIX_GROUP_CONSTRUCT_COMPLETE** -166 The group construct operation has completed - the
9 final membership will be included in the notification.

10 **PMIX_GROUP_LEADER_FAILED** -168 The current *leader* of a group including this process has
11 abnormally terminated - the group identifier will be included in the notification.

12 **PMIX_GROUP_LEADER_SELECTED** -167 A new *leader* of a group including this process has been
13 selected - the identifier of the new leader will be included in the notification.

14 **PMIX_GROUP_CONTEXT_ID_ASSIGNED** -169 A new PGCID has been assigned by the host
15 environment to a group that includes this process - the group identifier will be included in the
16 notification.

17 14.2.5 Process Group Attributes

18 *PMIx v4.0*

Attributes for querying the system regarding process groups include:

19 **PMIX_QUERY_NUM_GROUPS** "pmix.qry.pgrpnum" (**size_t**)
20 Return the number of process groups defined in the specified range (defaults to session). OPTIONAL
21 QUALIFIERS: **PMIX_RANGE**.

22 **PMIX_QUERY_GROUP_NAMES** "pmix.qry.pgrp" (**pmix_data_array_t***)
23 Return a **pmix_data_array_t** containing an array of string names of the process groups defined in
24 the specified range (defaults to session). OPTIONAL QUALIFIERS: **PMIX_RANGE**.

25 **PMIX_QUERY_GROUP_MEMBERSHIP** "pmix.qry.pgrpmems" (**pmix_data_array_t***)
26 Return a **pmix_data_array_t** of **pmix_proc_t** containing the members of the specified process
27 group. REQUIRED QUALIFIERS: **PMIX_GROUP_ID**.

28 The following attributes are used as directives in PMIx Group operations:

29 **PMIX_GROUP_ID** "pmix.grp.id" (**char***)
30 User-provided group identifier - as the group identifier may be used in PMIx operations, the user is
31 required to ensure that the provided ID is unique within the scope of the host environment (e.g., by
32 including some user-specific or application-specific prefix or suffix to the string).

33 **PMIX_GROUP_LEADER** "pmix.grp.ldr" (**bool**)
34 This process is the leader of the group.

35 **PMIX_GROUP_OPTIONAL** "pmix.grp.opt" (**bool**)
36 Participation is optional - do not return an error if any of the specified processes terminate without
37 having joined. The default is **false**.

38 **PMIX_GROUP_NOTIFY_TERMINATION** "pmix.grp.notterm" (**bool**)
39 Notify remaining members when another member terminates without first leaving the group.

40 **PMIX_GROUP_FT_COLLECTIVE** "pmix.grp.ftcoll" (**bool**)
41 Adjust internal tracking on-the-fly for terminated processes during a PMIx group collective operation.

42 **PMIX_GROUP_MEMBERSHIP** "pmix.grp.mbrs" (**pmix_data_array_t***)

1 Array `pmix_proc_t` identifiers identifying the members of the specified group.

2 **PMIX_GROUP_ASSIGN_CONTEXT_ID** "`pmix.grp.actxid`" (`bool`)

3 Requests that the RM assign a new context identifier to the newly created group. The identifier is an
4 unsigned, `size_t` value that the RM guarantees to be unique across the range specified in the request.
5 Thus, the value serves as a means of identifying the group within that range. If no range is specified,
6 then the request defaults to **PMIX_RANGE_SESSION**.

7 **PMIX_GROUP_LOCAL_ONLY** "`pmix.grp.lcl`" (`bool`)

8 Group operation only involves local processes. PMIx implementations are *required* to automatically
9 scan an array of group members for local vs remote processes - if only local processes are detected, the
10 implementation need not execute a global collective for the operation unless a context ID has been
11 requested from the host environment. This can result in significant time savings. This attribute can be
12 used to optimize the operation by indicating whether or not only local processes are represented, thus
13 allowing the implementation to bypass the scan.

14 The following attributes are used to return information at the conclusion of a PMIx Group operation and/or in
15 event notifications:

16 **PMIX_GROUP_CONTEXT_ID** "`pmix.grp.ctxid`" (`size_t`)

17 Context identifier assigned to the group by the host RM.

18 **PMIX_GROUP_ENDPT_DATA** "`pmix.grp.endpt`" (`pmix_byte_object_t`)

19 Data collected during group construction to ensure communication between group members is
20 supported upon completion of the operation.

21 In addition, a process can request (via **PMIx_Get**) the process groups to which a given process (including
22 itself) belongs:

23 **PMIX_GROUP_NAMES** "`pmix.pgrp.nm`" (`pmix_data_array_t*`)

24 Returns an array of `char*` string names of the process groups in which the given process is a member.

25 14.2.6 PMIx_Group_construct

26 Summary

27 Construct a PMIx process group.

28 Format

PMIx v4.0

29 `pmix_status_t`

```
30 PMIx_Group_construct(const char grp[],  
31                     const pmix_proc_t procs[], size_t nprocs,  
32                     const pmix_info_t directives[],  
33                     size_t ndirs,  
34                     pmix_info_t **results,  
35                     size_t *nresults);
```

36 **IN** `grp`

37 **NULL**-terminated character array of maximum size **PMIX_MAX_NSLEN** containing the group identifier
38 (string)

1 **IN** **procs**
2 Array of `pmix_proc_t` structures containing the PMIx identifiers of the member processes (array of
3 handles)
4 **IN** **nprocs**
5 Number of elements in the *procs* array (`size_t`)
6 **IN** **directives**
7 Array of `pmix_info_t` structures (array of handles)
8 **IN** **ndirs**
9 Number of elements in the *directives* array (`size_t`)
10 **INOUT** **results**
11 Pointer to a location where the array of `pmix_info_t` describing the results of the operation is to be
12 returned (pointer to handle)
13 **INOUT** **nresults**
14 Pointer to a `size_t` location where the number of elements in *results* is to be returned (memory
15 reference)

16 Returns `PMIX_SUCCESS` or a negative value indicating the error.

▼----- Required Attributes -----▼

17 The following attributes are *required* to be supported by all PMIx libraries that support this operation:

18 `PMIX_GROUP_LEADER` "pmix.grp.ldr" (bool)

19 This process is the leader of the group.

20 `PMIX_GROUP_OPTIONAL` "pmix.grp.opt" (bool)

21 Participation is optional - do not return an error if any of the specified processes terminate without
22 having joined. The default is `false`.

23 `PMIX_GROUP_LOCAL_ONLY` "pmix.grp.lcl" (bool)

24 Group operation only involves local processes. PMIx implementations are *required* to automatically
25 scan an array of group members for local vs remote processes - if only local processes are detected, the
26 implementation need not execute a global collective for the operation unless a context ID has been
27 requested from the host environment. This can result in significant time savings. This attribute can be
28 used to optimize the operation by indicating whether or not only local processes are represented, thus
29 allowing the implementation to bypass the scan.

30 `PMIX_GROUP_FT_COLLECTIVE` "pmix.grp.ftcoll" (bool)

31 Adjust internal tracking on-the-fly for terminated processes during a PMIx group collective operation.

32 Host environments that support this operation are *required* to support the following attributes:

33 `PMIX_GROUP_ASSIGN_CONTEXT_ID` "pmix.grp.actxid" (bool)

34 Requests that the RM assign a new context identifier to the newly created group. The identifier is an
35 unsigned, `size_t` value that the RM guarantees to be unique across the range specified in the request.
36 Thus, the value serves as a means of identifying the group within that range. If no range is specified,
37 then the request defaults to `PMIX_RANGE_SESSION`.

38 `PMIX_GROUP_NOTIFY_TERMINATION` "pmix.grp.notterm" (bool)

39 Notify remaining members when another member terminates without first leaving the group.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Construct a new group composed of the specified processes and identified with the provided group identifier. The group identifier is a user-defined, **NULL**-terminated character array of length less than or equal to **PMIX_MAX_NSLEN**. Only characters accepted by standard string comparison functions (e.g., *strcmp*) are supported. Processes may engage in multiple simultaneous group construct operations so long as each is provided with a unique group ID. The *directives* array can be used to pass user-level directives regarding timeout constraints and other options available from the PMIx server.

If the **PMIX_GROUP_NOTIFY_TERMINATION** attribute is provided and has a value of **true**, then either the construct leader (if **PMIX_GROUP_LEADER** is provided) or all participants who register for the **PMIX_GROUP_MEMBER_FAILED** event will receive events whenever a process fails or terminates prior to calling **PMIx_Group_construct** – i.e. if a *group leader* is declared, *only* that process will receive the event. In the absence of a declared leader, *all* specified group members will receive the event.

The event will contain the identifier of the process that failed to join plus any other information that the host RM provided. This provides an opportunity for the leader or the collective members to react to the event – e.g., to decide to proceed with a smaller group or to abort the operation. The decision is communicated to the PMIx library in the results array at the end of the event handler. This allows PMIx to properly adjust accounting for procedure completion. When construct is complete, the participating PMIx servers will be alerted to any change in participants and each group member will receive an updated group membership (marked with the **PMIX_GROUP_MEMBERSHIP** attribute) as part of the *results* array returned by this API.

Failure of the declared leader at any time will cause a **PMIX_GROUP_LEADER_FAILED** event to be delivered to all participants so they can optionally declare a new leader. A new leader is identified by providing the **PMIX_GROUP_LEADER** attribute in the results array in the return of the event handler. Only one process is allowed to return that attribute, thereby declaring itself as the new leader. Results of the leader selection will be communicated to all participants via a **PMIX_GROUP_LEADER_SELECTED** event identifying the new leader. If no leader was selected, then the **pmix_info_t** provided to that event handler will include that information so the participants can take appropriate action.

Any participant that returns **PMIX_GROUP_CONSTRUCT_ABORT** from either the **PMIX_GROUP_MEMBER_FAILED** or the **PMIX_GROUP_LEADER_FAILED** event handler will cause the construct process to abort, returning from the call with a **PMIX_GROUP_CONSTRUCT_ABORT** status.

If the **PMIX_GROUP_NOTIFY_TERMINATION** attribute is not provided or has a value of **false**, then the **PMIx_Group_construct** operation will simply return an error whenever a proposed group member fails or terminates prior to calling **PMIx_Group_construct**.

1 Providing the `PMIX_GROUP_OPTIONAL` attribute with a value of `true` directs the PMIx library to consider
2 participation by any specified group member as non-required - thus, the operation will return
3 `PMIX_SUCCESS` if all members participate, or `PMIX_ERR_PARTIAL_SUCCESS` if some members fail to
4 participate. The *results* array will contain the final group membership in the latter case. Note that this use-case
5 can cause the operation to hang if the `PMIX_TIMEOUT` attribute is not specified and one or more group
6 members fail to call `PMIx_Group_construct` while continuing to execute. Also, note that no leader or
7 member failed events will be generated during the operation.

8 Processes in a group under construction are not allowed to leave the group until group construction is
9 complete. Upon completion of the construct procedure, each group member will have access to the job-level
10 information of all namespaces represented in the group plus any information posted via `PMIx_Put` (subject to
11 the usual scoping directives) for every group member.

Advice to PMIx library implementers

12 At the conclusion of the construct operation, the PMIx library is *required* to ensure that job-related
13 information from each participating namespace plus any information posted by group members via
14 `PMIx_Put` (subject to scoping directives) is available to each member via calls to `PMIx_Get`.

Advice to PMIx server hosts

15 The collective nature of this API generally results in use of a fence-like operation by the backend host
16 environment. Host environments that utilize the array of process participants as a *signature* for such operations
17 may experience potential conflicts should both a `PMIx_Group_construct` and a `PMIx_Fence` operation
18 involving the same participants be simultaneously executed. As PMIx allows for such use-cases, it is therefore
19 the responsibility of the host environment to resolve any potential conflicts.

14.2.7 `PMIx_Group_construct_nb`

Summary

Non-blocking form of `PMIx_Group_construct`.

Format

C

```
pmix_status_t
PMIx_Group_construct_nb(const char grp[],
                       const pmix_proc_t procs[], size_t nprocs,
                       const pmix_info_t directives[],
                       size_t ndirs,
                       pmix_info_cbfunc_t cbfunc, void *cbdata);
```

C

- IN grp**
NULL-terminated character array of maximum size **PMIX_MAX_NSLEN** containing the group identifier (string)
- IN procs**
Array of **pmix_proc_t** structures containing the PMIx identifiers of the member processes (array of handles)
- IN nprocs**
Number of elements in the *procs* array (**size_t**)
- IN directives**
Array of **pmix_info_t** structures (array of handles)
- IN ndirs**
Number of elements in the *directives* array (**size_t**)
- IN cbfunc**
Callback function **pmix_info_cbfunc_t** (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

PMIX_OPERATION_SUCCEEDED, indicating that the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

If executed, the status returned in the provided callback function will be one of the following constants:

- **PMIX_SUCCESS** The operation succeeded and all specified members participated.
- **PMIX_ERR_PARTIAL_SUCCESS** The operation succeeded but not all specified members participated - the final group membership is included in the callback function.
- **PMIX_ERR_NOT_SUPPORTED** While the PMIx server supports this operation, the host RM does not.
- a non-zero PMIx error constant indicating a reason for the request's failure.

Required Attributes

PMIx libraries that choose not to support this operation *must* return `PMIX_ERR_NOT_SUPPORTED` when the function is called.

The following attributes are *required* to be supported by all PMIx libraries that support this operation:

`PMIX_GROUP_LEADER` "pmix.grp.ldr" (bool)

This process is the leader of the group.

`PMIX_GROUP_OPTIONAL` "pmix.grp.opt" (bool)

Participation is optional - do not return an error if any of the specified processes terminate without having joined. The default is `false`.

`PMIX_GROUP_LOCAL_ONLY` "pmix.grp.lcl" (bool)

Group operation only involves local processes. PMIx implementations are *required* to automatically scan an array of group members for local vs remote processes - if only local processes are detected, the implementation need not execute a global collective for the operation unless a context ID has been requested from the host environment. This can result in significant time savings. This attribute can be used to optimize the operation by indicating whether or not only local processes are represented, thus allowing the implementation to bypass the scan.

`PMIX_GROUP_FT_COLLECTIVE` "pmix.grp.ftcoll" (bool)

Adjust internal tracking on-the-fly for terminated processes during a PMIx group collective operation.

Host environments that support this operation are *required* to provide the following attributes:

`PMIX_GROUP_ASSIGN_CONTEXT_ID` "pmix.grp.actxid" (bool)

Requests that the RM assign a new context identifier to the newly created group. The identifier is an unsigned, `size_t` value that the RM guarantees to be unique across the range specified in the request. Thus, the value serves as a means of identifying the group within that range. If no range is specified, then the request defaults to `PMIX_RANGE_SESSION`.

`PMIX_GROUP_NOTIFY_TERMINATION` "pmix.grp.notterm" (bool)

Notify remaining members when another member terminates without first leaving the group.

Optional Attributes

The following attributes are optional for host environments that support this operation:

`PMIX_TIMEOUT` "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Non-blocking version of the `PMIx_Group_construct` operation. The callback function will be called once all group members have called either `PMIx_Group_construct` or `PMIx_Group_construct_nb`.

1 14.2.8 PMIx_Group_destruct

2 Summary

3 Destruct a PMIx process group.

4 *PMIx v4.0*

Format

C

```

5 pmix_status_t
6 PMIx_Group_destruct(const char grp[],
7                     const pmix_info_t directives[],
8                     size_t ndirs);

```

C

- 9 **IN grp**
10 NULL-terminated character array of maximum size **PMIX_MAX_NSLEN** containing the identifier of the
11 group to be destructed (string)
- 12 **IN directives**
13 Array of **pmix_info_t** structures (array of handles)
- 14 **IN ndirs**
15 Number of elements in the *directives* array (**size_t**)

16 Returns **PMIX_SUCCESS** or a negative value indicating the error.

Required Attributes

17 For implementations and host environments that support the operation, there are no identified required
18 attributes for this API.

Optional Attributes

19 The following attributes are optional for host environments that support this operation:

- 20 **PMIX_TIMEOUT** "pmix.timeout" (**int**)
21 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
22 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
23 (client, server, and host) simultaneously timing the operation.

Description

Destruct a group identified by the provided group identifier. Processes may engage in multiple simultaneous group destruct operations so long as each involves a unique group ID. The *directives* array can be used to pass user-level directives regarding timeout constraints and other options available from the PMIx server.

The destruct API will return an error if any group process fails or terminates prior to calling `PMIx_Group_destruct` or its non-blocking version unless the `PMIX_GROUP_NOTIFY_TERMINATION` attribute was provided (with a value of `false`) at time of group construction. If notification was requested, then the `PMIX_GROUP_MEMBER_FAILED` event will be delivered for each process that fails to call destruct and the destruct tracker updated to account for the lack of participation. The `PMIx_Group_destruct` operation will subsequently return `PMIX_SUCCESS` when the remaining processes have all called destruct – i.e., the event will serve in place of return of an error.

Advice to PMIx server hosts

The collective nature of this API generally results in use of a fence-like operation by the backend host environment. Host environments that utilize the array of process participants as a *signature* for such operations may experience potential conflicts should both a `PMIx_Group_destruct` and a `PMIx_Fence` operation involving the same participants be simultaneously executed. As PMIx allows for such use-cases, it is therefore the responsibility of the host environment to resolve any potential conflicts.

14.2.9 PMIx_Group_destruct_nb

Summary

Non-blocking form of `PMIx_Group_destruct`.

Format

PMIx v4.0

```
pmix_status_t
PMIx_Group_destruct_nb(const char grp[],
                      const pmix_info_t directives[],
                      size_t ndirs,
                      pmix_op_cbfunc_t cbfunc, void *cbdata);
```

- IN** `grp`
NULL-terminated character array of maximum size `PMIX_MAX_NSLEN` containing the identifier of the group to be destructed (string)
- IN** `directives`
Array of `pmix_info_t` structures (array of handles)
- IN** `ndirs`
Number of elements in the *directives* array (`size_t`)
- IN** `cbfunc`
Callback function `pmix_op_cbfunc_t` (function reference)
- IN** `cbdata`
Data to be passed to the callback function (memory reference)

1 A successful return indicates that the request is being processed and the result will be returned in the provided
2 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
3 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

4 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

5 **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed successfully - the
6 *cbfunc* will *not* be called.

7 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
8 error code or an implementation defined error code as described in Section 3.1.1.

9 If executed, the status returned in the provided callback function will be one of the following constants:

- 10 • **PMIX_SUCCESS** The operation was successfully completed.
- 11 • **PMIX_ERR_NOT_SUPPORTED** While the PMIx server supports this operation, the host RM does not.
- 12 • a non-zero PMIx error constant indicating a reason for the request's failure.

Required Attributes

13 PMIx libraries that choose not to support this operation *must* return **PMIX_ERR_NOT_SUPPORTED** when the
14 function is called. For implementations and host environments that support the operation, there are no
15 identified required attributes for this API.

Optional Attributes

16 The following attributes are optional for host environments that support this operation:

17 **PMIX_TIMEOUT** "pmix.timeout" (int)

18 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
19 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
20 (client, server, and host) simultaneously timing the operation.

Description

21 Non-blocking version of the **PMIx_Group_destruct** operation. The callback function will be called once
22 all members of the group have executed either **PMIx_Group_destruct** or
23 **PMIx_Group_destruct_nb**.

14.2.10 PMIx_Group_invite

Summary

26 Asynchronously construct a PMIx process group.
27

Format

C

```
pmix_status_t
PMIx_Group_invite(const char grp[],
                  const pmix_proc_t procs[], size_t nprocs,
                  const pmix_info_t directives[], size_t ndirs,
                  pmix_info_t **results, size_t *nresult);
```

C

IN grp
NULL-terminated character array of maximum size `PMIX_MAX_NSLEN` containing the group identifier (string)

IN procs
Array of `pmix_proc_t` structures containing the PMIx identifiers of the processes to be invited (array of handles)

IN nprocs
Number of elements in the `procs` array (`size_t`)

IN directives
Array of `pmix_info_t` structures (array of handles)

IN ndirs
Number of elements in the `directives` array (`size_t`)

INOUT results
Pointer to a location where the array of `pmix_info_t` describing the results of the operation is to be returned (pointer to handle)

INOUT nresults
Pointer to a `size_t` location where the number of elements in `results` is to be returned (memory reference)

Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

The following attributes are *required* to be supported by all PMIx libraries that support this operation:

PMIX_GROUP_OPTIONAL "pmix.grp.opt" (bool)

Participation is optional - do not return an error if any of the specified processes terminate without having joined. The default is `false`.

PMIX_GROUP_FT_COLLECTIVE "pmix.grp.ftcoll" (bool)

Adjust internal tracking on-the-fly for terminated processes during a PMIx group collective operation.

Host environments that support this operation are *required* to provide the following attributes:

PMIX_GROUP_ASSIGN_CONTEXT_ID "pmix.grp.actxid" (bool)

Requests that the RM assign a new context identifier to the newly created group. The identifier is an unsigned, `size_t` value that the RM guarantees to be unique across the range specified in the request. Thus, the value serves as a means of identifying the group within that range. If no range is specified, then the request defaults to `PMIX_RANGE_SESSION`.

PMIX_GROUP_NOTIFY_TERMINATION "pmix.grp.notterm" (bool)

1 Notify remaining members when another member terminates without first leaving the group.

Optional Attributes

2 The following attributes are optional for host environments that support this operation:

3 **PMIX_TIMEOUT** "pmix.timeout" (int)

4 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
5 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
6 (client, server, and host) simultaneously timing the operation.

Description

7 Explicitly invite the specified processes to join a group. The process making the **PMIX_Group_invite** call
8 is automatically declared to be the *group leader*. Each invited process will be notified of the invitation via the
9 **PMIX_GROUP_INVITED** event - the processes being invited must therefore register for the
10 **PMIX_GROUP_INVITED** event in order to be notified of the invitation. Note that the PMIx event notification
11 system caches events - thus, no ordering of invite versus event registration is required.
12

13 The invitation event will include the identity of the inviting process plus the name of the group. When ready to
14 respond, each invited process provides a response using either the blocking or non-blocking form of
15 **PMIX_Group_join**. This will notify the inviting process that the invitation was either accepted (via the
16 **PMIX_GROUP_INVITE_ACCEPTED** event) or declined (via the **PMIX_GROUP_INVITE_DECLINED**
17 event). The **PMIX_GROUP_INVITE_ACCEPTED** event is captured by the PMIx client library of the inviting
18 process – i.e., the application itself does not need to register for this event. The library will track the number of
19 accepting processes and alert the inviting process (by returning from the blocking form of
20 **PMIX_Group_invite** or calling the callback function of the non-blocking form) when group construction
21 completes.

22 The inviting process should, however, register for the **PMIX_GROUP_INVITE_DECLINED** if the application
23 allows invited processes to decline the invitation. This provides an opportunity for the application to either
24 invite a replacement, declare “abort”, or choose to remove the declining process from the final group. The
25 inviting process should also register to receive **PMIX_GROUP_INVITE_FAILED** events whenever a process
26 fails or terminates prior to responding to the invitation. Actions taken by the inviting process in response to
27 these events must be communicated at the end of the event handler by returning the corresponding result so
28 that the PMIx library can adjust accordingly.

29 Upon completion of the operation, all members of the new group will receive access to the job-level
30 information of each other’s namespaces plus any information posted via **PMIX_Put** by the other members.

31 The inviting process is automatically considered the leader of the asynchronous group construction procedure
32 and will receive all failure or termination events for invited members prior to completion. The inviting process
33 is required to provide a **PMIX_GROUP_CONSTRUCT_COMPLETE** event once the group has been fully
34 assembled – this event is used by the PMIx library as a trigger to release participants from their call to
35 **PMIX_Group_join** and provides information (e.g., the final group membership) to be returned in the
36 *results* array.

37 Failure of the inviting process at any time will cause a **PMIX_GROUP_LEADER_FAILED** event to be
38 delivered to all participants so they can optionally declare a new leader. A new leader is identified by

1 providing the `PMIX_GROUP_LEADER` attribute in the results array in the return of the event handler. Only
2 one process is allowed to return that attribute, declaring itself as the new leader. Results of the leader selection
3 will be communicated to all participants via a `PMIX_GROUP_LEADER_SELECTED` event identifying the
4 new leader. If no leader was selected, then the status code provided in the event handler will provide an error
5 value so the participants can take appropriate action.

Advice to users

6 Applications are not allowed to use the group in any operations until group construction is complete. This is
7 required in order to ensure consistent knowledge of group membership across all participants.

8 14.2.11 `PMIx_Group_invite_nb`

9 Summary

10 Non-blocking form of `PMIx_Group_invite`.

11 Format

PMIx v4.0

```
12 pmix_status_t  
13 PMIx_Group_invite_nb(const char grp[],  
14                     const pmix_proc_t procs[], size_t nprocs,  
15                     const pmix_info_t directives[], size_t ndirs,  
16                     pmix_info_cbfunc_t cbfunc, void *cbdata);
```

17 IN `grp`

18 NULL-terminated character array of maximum size `PMIX_MAX_NSLEN` containing the group identifier
19 (string)

20 IN `procs`

21 Array of `pmix_proc_t` structures containing the PMIx identifiers of the processes to be invited (array
22 of handles)

23 IN `nprocs`

24 Number of elements in the `procs` array (`size_t`)

25 IN `directives`

26 Array of `pmix_info_t` structures (array of handles)

27 IN `ndirs`

28 Number of elements in the `directives` array (`size_t`)

29 IN `cbfunc`

30 Callback function `pmix_info_cbfunc_t` (function reference)

31 IN `cbdata`

32 Data to be passed to the callback function (memory reference)

33 A successful return indicates that the request is being processed and the result will be returned in the provided
34 `cbfunc`. Note that the library must not invoke the callback function prior to returning from the API. The
35 callback function, `cbfunc`, is only called when `PMIX_SUCCESS` is returned.

36 Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

1 **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed successfully - the
2 *cbfunc* will *not* be called.

3 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
4 error code or an implementation defined error code as described in Section 3.1.1.

5 If executed, the status returned in the provided callback function will be one of the following constants:

- 6 • **PMIX_SUCCESS** The operation succeeded and all specified members participated.
- 7 • **PMIX_ERR_PARTIAL_SUCCESS** The operation succeeded but not all specified members participated -
8 the final group membership is included in the callback function.
- 9 • **PMIX_ERR_NOT_SUPPORTED** While the PMIx server supports this operation, the host RM does not.
- 10 • a non-zero PMIx error constant indicating a reason for the request's failure.

Required Attributes

11 The following attributes are *required* to be supported by all PMIx libraries that support this operation:

12 **PMIX_GROUP_OPTIONAL** "pmix.grp.opt" (bool)

13 Participation is optional - do not return an error if any of the specified processes terminate without
14 having joined. The default is **false**.

15 **PMIX_GROUP_FT_COLLECTIVE** "pmix.grp.ftcoll" (bool)

16 Adjust internal tracking on-the-fly for terminated processes during a PMIx group collective operation.

17 Host environments that support this operation are *required* to provide the following attributes:

18 **PMIX_GROUP_ASSIGN_CONTEXT_ID** "pmix.grp.actxid" (bool)

19 Requests that the RM assign a new context identifier to the newly created group. The identifier is an
20 unsigned, **size_t** value that the RM guarantees to be unique across the range specified in the request.
21 Thus, the value serves as a means of identifying the group within that range. If no range is specified,
22 then the request defaults to **PMIX_RANGE_SESSION**.

23 **PMIX_GROUP_NOTIFY_TERMINATION** "pmix.grp.notterm" (bool)

24 Notify remaining members when another member terminates without first leaving the group.

Optional Attributes

25 The following attributes are optional for host environments that support this operation:

26 **PMIX_TIMEOUT** "pmix.timeout" (int)

27 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
28 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
29 (client, server, and host) simultaneously timing the operation.

Description

30 Non-blocking version of the **PMIx_Group_invite** operation. The callback function will be called once all
31 invited members of the group (or their substitutes) have executed either **PMIx_Group_join** or
32 **PMIx_Group_join_nb**.

1 14.2.12 PMIx_Group_join

2 Summary

3 Accept an invitation to join a PMIx process group.

4 Format

PMIx v4.0

C

```

5 pmix_status_t
6 PMIx_Group_join(const char grp[],
7                 const pmix_proc_t *leader,
8                 pmix_group_opt_t opt,
9                 const pmix_info_t directives[], size_t ndirs,
10                pmix_info_t **results, size_t *nresult);

```

C

- 11 **IN grp**
12 NULL-terminated character array of maximum size **PMIX_MAX_NSLEN** containing the group identifier (string)
- 14 **IN leader**
15 Process that generated the invitation (handle)
- 16 **IN opt**
17 Accept or decline flag (**pmix_group_opt_t**)
- 18 **IN directives**
19 Array of **pmix_info_t** structures (array of handles)
- 20 **IN ndirs**
21 Number of elements in the *directives* array (**size_t**)
- 22 **INOUT results**
23 Pointer to a location where the array of **pmix_info_t** describing the results of the operation is to be returned (pointer to handle)
- 25 **INOUT nresults**
26 Pointer to a **size_t** location where the number of elements in *results* is to be returned (memory reference)

28 Returns **PMIX_SUCCESS** or a negative value indicating the error.

▼----- Required Attributes -----▼

29 There are no identified required attributes for implementers.

▲-----

▼----- Optional Attributes -----▼

30 The following attributes are optional for host environments that support this operation:

31 **PMIX_TIMEOUT** "pmix.timeout" (**int**)

32 Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers

33 (client, server, and host) simultaneously timing the operation.

▲-----

Description

Respond to an invitation to join a group that is being asynchronously constructed. The process must have registered for the `PMIX_GROUP_INVITED` event in order to be notified of the invitation. When called, the event information will include the `pmix_proc_t` identifier of the process that generated the invitation along with the identifier of the group being constructed. When ready to respond, the process provides a response using either form of `PMIx_Group_join`.

Advice to users

Since the process is alerted to the invitation in a PMIx event handler, the process *must not* use the blocking form of this call unless it first “thread shifts” out of the handler and into its own thread context. Likewise, while it is safe to call the non-blocking form of the API from the event handler, the process *must not* block in the handler while waiting for the callback function to be called.

Calling this function causes the inviting process (aka the *group leader*) to be notified that the process has either accepted or declined the request. The blocking form of the API will return once the group has been completely constructed or the group’s construction has failed (as described below) – likewise, the callback function of the non-blocking form will be executed upon the same conditions.

Failure of the leader during the call to `PMIx_Group_join` will cause a `PMIX_GROUP_LEADER_FAILED` event to be delivered to all invited participants so they can optionally declare a new leader. A new leader is identified by providing the `PMIX_GROUP_LEADER` attribute in the results array in the return of the event handler. Only one process is allowed to return that attribute, declaring itself as the new leader. Results of the leader selection will be communicated to all participants via a `PMIX_GROUP_LEADER_SELECTED` event identifying the new leader. If no leader was selected, then the status code provided in the event handler will provide an error value so the participants can take appropriate action.

Any participant that returns `PMIX_GROUP_CONSTRUCT_ABORT` from the leader failed event handler will cause all participants to receive an event notifying them of that status. Similarly, the leader may elect to abort the procedure by either returning `PMIX_GROUP_CONSTRUCT_ABORT` from the handler assigned to the `PMIX_GROUP_INVITE_ACCEPTED` or `PMIX_GROUP_INVITE_DECLINED` codes, or by generating an event for the abort code. Abort events will be sent to all invited participants.

14.2.13 `PMIx_Group_join_nb`

Summary

Non-blocking form of `PMIx_Group_join`

Format

```
pmix_status_t
PMIx_Group_join_nb(const char grp[],
                  const pmix_proc_t *leader,
                  pmix_group_opt_t opt,
                  const pmix_info_t directives[], size_t ndirs,
                  pmix_info_cbfunc_t cbfunc, void *cbdata);
```

- IN grp**
NULL-terminated character array of maximum size **PMIX_MAX_NSLEN** containing the group identifier (string)
- IN leader**
Process that generated the invitation (handle)
- IN opt**
Accept or decline flag (**pmix_group_opt_t**)
- IN directives**
Array of **pmix_info_t** structures (array of handles)
- IN ndirs**
Number of elements in the *directives* array (**size_t**)
- IN cbfunc**
Callback function **pmix_info_cbfunc_t** (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

PMIX_OPERATION_SUCCEEDED, indicating that the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

If executed, the status returned in the provided callback function will be one of the following constants:

- **PMIX_SUCCESS** The operation succeeded and group membership is in the callback function parameters.
- **PMIX_ERR_NOT_SUPPORTED** While the PMIx server supports this operation, the host RM does not.
- a non-zero PMIx error constant indicating a reason for the request's failure.

Required Attributes

There are no identified required attributes for implementers.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Non-blocking version of the **PMIx_Group_join** operation. The callback function will be called once all invited members of the group (or their substitutes) have executed either **PMIx_Group_join** or **PMIx_Group_join_nb**.

14.2.13.1 Group accept/decline directives

PMIx v4.0

The **pmix_group_opt_t** type is a **uint8_t** value used with the **PMIx_Group_join** API to indicate *accept* or *decline* of the invitation - these are provided for readability of user code:

PMIX_GROUP_DECLINE 0 Decline the invitation.

PMIX_GROUP_ACCEPT 1 Accept the invitation.

14.2.14 PMIx_Group_leave

Summary

Leave a PMIx process group.

Format

PMIx v4.0

```
pmix_status_t
PMIx_Group_leave(const char grp[],
                 const pmix_info_t directives[],
                 size_t ndirs);
```

IN grp

NULL-terminated character array of maximum size **PMIX_MAX_NSLEN** containing the group identifier (string)

IN directives

Array of **pmix_info_t** structures (array of handles)

IN ndirs

Number of elements in the *directives* array (**size_t**)

Returns **PMIX_SUCCESS** or a negative value indicating the error.

Required Attributes

There are no identified required attributes for implementers.

Description

Calls to `PMIx_Group_leave` (or its non-blocking form) will cause a `PMIX_GROUP_LEFT` event to be generated notifying all members of the group of the caller's departure. The function will return (or the non-blocking function will execute the specified callback function) once the event has been locally generated and is not indicative of remote receipt.

Advice to users

The `PMIx_Group_leave` API is intended solely for asynchronous departures of individual processes from a group as it is not a scalable operation – i.e., when a process determines it should no longer be a part of a defined group, but the remainder of the group retains a valid reason to continue in existence. Developers are advised to use `PMIx_Group_destruct` (or its non-blocking form) for all other scenarios as it represents a more scalable operation.

14.2.15 `PMIx_Group_leave_nb`

Summary

Non-blocking form of `PMIx_Group_leave`.

Format

PMIx v4.0

```
pmix_status_t
PMIx_Group_leave_nb(const char grp[],
                   const pmix_info_t directives[],
                   size_t ndirs,
                   pmix_op_cbfunc_t cbfunc,
                   void *cbdata);
```

- IN** `grp`
NULL-terminated character array of maximum size `PMIX_MAX_NSLEN` containing the group identifier (string)
- IN** `directives`
Array of `pmix_info_t` structures (array of handles)
- IN** `ndirs`
Number of elements in the `directives` array (`size_t`)
- IN** `cbfunc`
Callback function `pmix_op_cbfunc_t` (function reference)
- IN** `cbdata`
Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided `cbfunc`. Note that the library must not invoke the callback function prior to returning from the API. The callback function, `cbfunc`, is only called when `PMIX_SUCCESS` is returned.

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

1 **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed successfully - the
2 *cbfunc* will *not* be called.

3 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
4 error code or an implementation defined error code as described in Section 3.1.1.

5 If executed, the status returned in the provided callback function will be one of the following constants:

- 6 • **PMIX_SUCCESS** The operation succeeded - i.e., the **PMIX_GROUP_LEFT** event was generated.
- 7 • **PMIX_ERR_NOT_SUPPORTED** While the PMIx library supports this operation, the host RM does not.
- 8 • a non-zero PMIx error constant indicating a reason for the request's failure.

9  **Required Attributes**
There are no identified required attributes for implementers.

10 **Description**

11 Non-blocking version of the **PMIx_Group_leave** operation. The callback function will be called once the
12 event has been locally generated and is not indicative of remote receipt.

CHAPTER 15

Fabric Support Definitions

1 As the drive for performance continues, interest has grown in scheduling algorithms that take into account
2 network locality of the allocated resources and in optimizing collective communication patterns by structuring
3 them to follow fabric topology. In addition, concerns over the time required to initiate execution of parallel
4 applications and enable communication across them have grown as the size of those applications extends into
5 the hundreds of thousands of individual processes spanning tens of thousands of nodes.

6 PMIx supports the communication part of these efforts by defining data types and attributes by which fabric
7 endpoints and coordinates for processes and devices can be obtained from the host environment. When used in
8 conjunction with other PMIx methods described in Chapter 17, this results in the ability of a process to obtain
9 the fabric endpoint and coordinate of all other processes without incurring additional overhead associated with
10 a global exchange of that information. This includes:

- 11 • Defining several interfaces specifically intended to support WLMs by providing access to information of
12 potential use to scheduling algorithms - e.g., information on communication costs between different points
13 on the fabric.
- 14 • Supporting hierarchical collective operations by providing the fabric coordinates for all devices on
15 participating nodes as well as a list of the peers sharing each fabric switch. This enables one, for example,
16 to aggregate the contribution from all processes on a node, then again across all nodes on a common switch,
17 and finally across all switches based on detailed knowledge of the fabric location of each participant.
- 18 • Enabling the "*instant on*" paradigm to mitigate the scalable launch problem by providing each process with
19 a rich set of information about the environment and the application, including everything required for
20 communication between peers within the application, at time of process start of execution.

21 Meeting these needs in the case where only a single fabric device exists on each node is relatively
22 straightforward - PMIx and the host environment provide a single endpoint for each process plus a coordinate
23 for the device on each node, and there is no uncertainty regarding the endpoint each process will use.
24 Extending this to the multiple device per node case is more difficult as the choice of endpoint by any given
25 process cannot be known in advance, and questions arise regarding reachability between devices on different
26 nodes. Resolving these ambiguities without requiring a global operation requires that PMIx provide both (a)
27 an endpoint for each application process on each of its local devices; and (b) the fabric coordinates of all
28 remote and local devices on participating nodes. It also requires that each process open all of its assigned
29 endpoints as the endpoint selected for contact by a remote peer cannot be known in advance.

30 While these steps ensure the ability of a process to connect to a remote peer, it leaves unanswered the question
31 of selecting the *preferred* device for that communication. If multiple devices are present on a node, then the
32 application can benefit from having each process utilize its "closest" fabric device (i.e., the device that
33 minimizes the communication distance between the process' location and that device) for messaging
34 operations. In some cases, messaging libraries prefer to also retain the ability to use non-nearest devices,
35 prioritizing the devices based on distance to support multi-device operations (e.g., for large message
36 transmission in parallel).

1 PMIx supports this requirement by providing the array of process-to-device distance information for each
2 process and local fabric device at start of execution. Both minimum and maximum distances are provided
3 since a single process can occupy multiple processor locations. In addition, since processes can relocate
4 themselves by changing their processor bindings, PMIx provides an API that allows the process to dynamically
5 request an update to its distance array.

6 However, while these measures assist a process in selecting its own best endpoint, they do not resolve the
7 uncertainty over the choice of preferred device by a remote peer. There are two methods by which this
8 ambiguity can be resolved:

- 9 a) A process can select a remote endpoint to use based on its own preferred device and reachability of the
10 peer's remote devices. Once the initial connection has been made, the two processes can exchange
11 information and mutually determine their desired communication path going forward.
- 12 b) The application can use knowledge of both the local and remote distance arrays to compute the best
13 communication path and establish that connection. In some instances (e.g., a homogeneous system), a
14 PMIx server may provide distance information for both local and remote devices. Alternatively, when this
15 isn't available, an application can opt to collect the information using the
16 `PMIX_COLLECT_GENERATED_JOB_INFO` with the `PMIx_Fence` API, or can obtain it on a one
17 peer-at-a-time basis using the `PMIx_Get` API on systems where the host environment supports the *Direct*
18 *Modex* operation.

19 Information on fabric coordinates, endpoints, and device distances are provided as *reserved keys* as detailed in
20 Chapter 6 - i.e., they are to be available at client start of execution and are subject to the retrieval rules of
21 Section 6.2. Examples for retrieving fabric-related information include retrieval of:

- 22 • An array of information on fabric devices for a node by passing `PMIX_FABRIC_DEVICES` as the key to
23 `PMIx_Get` along with the `PMIX_HOSTNAME` of the node as a directive
- 24 • An array of information on a specific fabric device by passing `PMIX_FABRIC_DEVICE` as the key to
25 `PMIx_Get` along with the `PMIX_DEVICE_ID` of the device as a directive
- 26 • An array of information on a specific fabric device by passing `PMIX_FABRIC_DEVICE` as the key to
27 `PMIx_Get` along with both `PMIX_FABRIC_DEVICE_NAME` of the device and the `PMIX_HOSTNAME` of
28 the node as directives

29 When requesting data on a device, returned data must include at least the following attributes:

- 30 • `PMIX_HOSTNAME` `"pmix.hname"` (`char*`)
31 Name of the host, as returned by the `gethostname` utility or its equivalent. The
32 `PMIX_NODEID` may be returned in its place, or in addition to the hostname.
- 33 • `PMIX_DEVICE_ID` `"pmix.dev.id"` (`string`)
34 System-wide UUID or node-local OS name of a particular device.
- 35 • `PMIX_FABRIC_DEVICE_NAME` `"pmix.fabdev.nm"` (`string`)
36 The operating system name associated with the device. This may be a logical fabric interface name
37 (e.g. "eth0" or "eno1") or an absolute filename.
- 38 • `PMIX_FABRIC_DEVICE_VENDOR` `"pmix.fabdev.vndr"` (`string`)
39 Indicates the name of the vendor that distributes the device.
- 40 • `PMIX_FABRIC_DEVICE_BUS_TYPE` `"pmix.fabdev.btyp"` (`string`)

1 The type of bus to which the device is attached (e.g., "PCI", "GEN-Z").

- 2 ● **PMIX_FABRIC_DEVICE_PCI_DEVID** "pmix.fabdev.pcidevid" (string)
3 A node-level unique identifier for a Peripheral Component Interconnect (PCI) device. Provided only
4 if the device is located on a PCI bus. The identifier is constructed as a four-part tuple delimited by
5 colons comprised of the PCI 16-bit domain, 8-bit bus, 8-bit device, and 8-bit function IDs, each
6 expressed in zero-extended hexadecimal form. Thus, an example identifier might be
7 "abc1:0f:23:01". The combination of node identifier (**PMIX_HOSTNAME** or **PMIX_NODEID**) and
8 **PMIX_FABRIC_DEVICE_PCI_DEVID** shall be unique within the overall system. This item
9 should be included if the device bus type is PCI - the equivalent should be provided for any other
10 bus type.

11 The returned array may optionally contain one or more of the following in addition to the above list:

- 12 ● **PMIX_FABRIC_DEVICE_INDEX** "pmix.fabdev.idx" (uint32_t)
13 Index of the device within an associated communication cost matrix.
- 14 ● **PMIX_FABRIC_DEVICE_VENDORID** "pmix.fabdev.vendid" (string)
15 This is a vendor-provided identifier for the device or product.
- 16 ● **PMIX_FABRIC_DEVICE_DRIVER** "pmix.fabdev.driver" (string)
17 The name of the driver associated with the device.
- 18 ● **PMIX_FABRIC_DEVICE_FIRMWARE** "pmix.fabdev.fmwr" (string)
19 The device's firmware version.
- 20 ● **PMIX_FABRIC_DEVICE_ADDRESS** "pmix.fabdev.addr" (string)
21 The primary link-level address associated with the device, such as a Media Access Control (MAC)
22 address. If multiple addresses are available, only one will be reported.
- 23 ● **PMIX_FABRIC_DEVICE_COORDINATES** "pmix.fab.coord" (pmix_geometry_t)
24 The **pmix_geometry_t** fabric coordinates for the device, including values for all supported
25 coordinate views.
- 26 ● **PMIX_FABRIC_DEVICE_MTU** "pmix.fabdev.mtu" (size_t)
27 The maximum transfer unit of link level frames or packets, in bytes.
- 28 ● **PMIX_FABRIC_DEVICE_SPEED** "pmix.fabdev.speed" (size_t)
29 The active link data rate, given in bits per second.
- 30 ● **PMIX_FABRIC_DEVICE_STATE** "pmix.fabdev.state" (pmix_link_state_t)
31 The last available physical port state for the specified device. Possible values are
32 **PMIX_LINK_STATE_UNKNOWN**, **PMIX_LINK_DOWN**, and **PMIX_LINK_UP**, to indicate if the
33 port state is unknown or not applicable (unknown), inactive (down), or active (up).
- 34 ● **PMIX_FABRIC_DEVICE_TYPE** "pmix.fabdev.type" (string)
35 Specifies the type of fabric interface currently active on the device, such as Ethernet or InfiniBand.

36 The remainder of this chapter details the events, data types, attributes, and APIs associated with fabric-related
37 operations.

1 **Initialize the endpoint structure**

2 Initialize the `pmix_endpoint_t` fields.



3 **PMIX_ENDPOINT_CONSTRUCT (m)**



4 **IN m**

5 Pointer to the structure to be initialized (pointer to `pmix_endpoint_t`)

6 **Destruct the endpoint structure**

7 Destruct the `pmix_endpoint_t` fields.

PMIx v4.0



8 **PMIX_ENDPOINT_DESTRUCT (m)**



9 **IN m**

10 Pointer to the structure to be destructed (pointer to `pmix_endpoint_t`)

11 **Create an endpoint array**

12 Allocate and initialize a `pmix_endpoint_t` array.

PMIx v4.0



13 **PMIX_ENDPOINT_CREATE (m, n)**



14 **INOUT m**

15 Address where the pointer to the array of `pmix_endpoint_t` structures shall be stored (handle)

16 **IN n**

17 Number of structures to be allocated (`size_t`)

18 **Release an endpoint array**

19 Release an array of `pmix_endpoint_t` structures.

PMIx v4.0



20 **PMIX_ENDPOINT_FREE (m, n)**



21 **IN m**

22 Pointer to the array of `pmix_endpoint_t` structures (handle)

23 **IN n**

24 Number of structures in the array (`size_t`)

1 15.2.3 Fabric Coordinate Structure

2 The `pmix_coord_t` structure describes the fabric coordinates of a specified device in a given view.

```
3 typedef struct pmix_coord {  
4     pmix_coord_view_t view;  
5     uint32_t *coord;  
6     size_t dims;  
7 } pmix_coord_t;
```

8 All coordinate values shall be expressed as unsigned integers due to their units being defined in fabric devices
9 and not physical distances. The coordinate is therefore an indicator of connectivity and not relative
10 communication distance.

Advice to PMIx library implementers

11 Note that the `pmix_coord_t` structure does not imply nor mandate any requirement on how the coordinate
12 data is to be stored within the PMIx library. Implementers are free to store the coordinate in whatever format
13 they choose.

14 A fabric coordinate is associated with a given fabric device and must be unique within a given view. Fabric
15 devices are associated with the operating system which hosts them - thus, fabric coordinates are logically
16 grouped within the *node* realm (as described in Section 6.1) and can be retrieved per the rules detailed in
17 Section 6.1.5.

18 15.2.4 Fabric coordinate support macros

19 The following macros are provided to support the `pmix_coord_t` structure.

20 Static initializer for the coord structure

21 *(Provisional)*

22 Provide a static initializer for the `pmix_coord_t` fields.

PMIx v5.0

23 `PMIX_COORD_STATIC_INIT`

24 Initialize the coord structure

25 Initialize the `pmix_coord_t` fields.

PMIx v4.0

26 `PMIX_COORD_CONSTRUCT(m)`

27 **IN** `m`

28 Pointer to the structure to be initialized (pointer to `pmix_coord_t`)

1 Destruct the coord structure

2 Destruct the `pmix_coord_t` fields.

▼  ▼

3 `PMIX_COORD_DESTRUCT (m)`


▲  ▲

4 **IN** `m`

5 Pointer to the structure to be destructed (pointer to `pmix_coord_t`)

6 Create a coord array

7 Allocate and initialize a `pmix_coord_t` array.

PMIx v4.0 ▼  ▼

8 `PMIX_COORD_CREATE (m, n)`

▲  ▲

9 **INOUT** `m`

10 Address where the pointer to the array of `pmix_coord_t` structures shall be stored (handle)

11 **IN** `n`

12 Number of structures to be allocated (`size_t`)

13 Release a coord array

14 Release an array of `pmix_coord_t` structures.

PMIx v4.0 ▼  ▼

15 `PMIX_COORD_FREE (m, n)`

▲  ▲

16 **IN** `m`

17 Pointer to the array of `pmix_coord_t` structures (handle)

18 **IN** `n`

19 Number of structures in the array (`size_t`)

20 15.2.5 Fabric Geometry Structure

21 The `pmix_geometry_t` structure describes the fabric coordinates of a specified device.

PMIx v4.0 ▼  ▼

```
22 typedef struct pmix_geometry {
23     size_t fabric;
24     char *uuid;
25     char *osname;
26     pmix_coord_t *coordinates;
27     size_t ncoords;
28 } pmix_geometry_t;
```

C

1 All coordinate values shall be expressed as unsigned integers due to their units being defined in fabric devices
2 and not physical distances. The coordinate is therefore an indicator of connectivity and not relative
3 communication distance.

Advice to PMIx library implementers

4 Note that the `pmix_coord_t` structure does not imply nor mandate any requirement on how the coordinate
5 data is to be stored within the PMIx library. Implementers are free to store the coordinate in whatever format
6 they choose.

7 A fabric coordinate is associated with a given fabric device and must be unique within a given view. Fabric
8 devices are associated with the operating system which hosts them - thus, fabric coordinates are logically
9 grouped within the *node* realm (as described in Section 6.1) and can be retrieved per the rules detailed in
10 Section 6.1.5.

15.2.6 Fabric geometry support macros

The following macros are provided to support the `pmix_geometry_t` structure.

Static initializer for the geometry structure

(Provisional)

Provide a static initializer for the `pmix_geometry_t` fields.

PMIx v5.0

`PMIX_GEOMETRY_STATIC_INIT`

Initialize the geometry structure

Initialize the `pmix_geometry_t` fields.

PMIx v4.0

`PMIX_GEOMETRY_CONSTRUCT (m)`

IN *m*

Pointer to the structure to be initialized (pointer to `pmix_geometry_t`)

Destruct the geometry structure

Destruct the `pmix_geometry_t` fields.

PMIx v4.0

`PMIX_GEOMETRY_DESTRUCT (m)`

IN *m*

Pointer to the structure to be destructed (pointer to `pmix_geometry_t`)

1 **Create a geometry array**
2 Allocate and initialize a `pmix_geometry_t` array.

3 `PMIX_GEOMETRY_CREATE(m, n)`

4 **INOUT** `m`
5 Address where the pointer to the array of `pmix_geometry_t` structures shall be stored (handle)
6 **IN** `n`
7 Number of structures to be allocated (`size_t`)

8 **Release a geometry array**
9 Release an array of `pmix_geometry_t` structures.

PMIx v4.0

10 `PMIX_GEOMETRY_FREE(m, n)`

11 **IN** `m`
12 Pointer to the array of `pmix_geometry_t` structures (handle)
13 **IN** `n`
14 Number of structures in the array (`size_t`)

15 15.2.7 Fabric Coordinate Views

PMIx v4.0

```
16 typedef uint8_t pmix_coord_view_t;  
17 #define PMIX_COORD_VIEW_UNDEF 0x00  
18 #define PMIX_COORD_LOGICAL_VIEW 0x01  
19 #define PMIX_COORD_PHYSICAL_VIEW 0x02
```

20 Fabric coordinates can be reported based on different *views* according to user preference at the time of request.
21 The following views have been defined:

22 **PMIX_COORD_VIEW_UNDEF 0x00** The coordinate view has not been defined.

23 **PMIX_COORD_LOGICAL_VIEW 0x01** The coordinates are provided in a *logical* view, typically given
24 in Cartesian (x,y,z) dimensions, that describes the data flow in the fabric as defined by the arrangement
25 of the hierarchical addressing scheme, fabric segmentation, routing domains, and other similar factors
26 employed by that fabric.

27 **PMIX_COORD_PHYSICAL_VIEW 0x02** The coordinates are provided in a *physical* view based on the
28 actual wiring diagram of the fabric - i.e., values along each axis reflect the relative position of that
29 interface on the specific fabric cabling.

30 If the requester does not specify a view, coordinates shall default to the *logical* view.

15.2.8 Fabric Link State

The `pmix_link_state_t` is a `uint32_t` type for fabric link states.

```
typedef uint8_t pmix_link_state_t;
```

The following constants can be used to set a variable of the type `pmix_link_state_t`. All definitions were introduced in version 4 of the standard unless otherwise marked. Valid link state values start at zero.

`PMIX_LINK_STATE_UNKNOWN` 0 The port state is unknown or not applicable.

`PMIX_LINK_DOWN` 1 The port is inactive.

`PMIX_LINK_UP` 2 The port is active.

15.2.9 Fabric Operation Constants

PMIx v4.0

The `pmix_fabric_operation_t` data type is an enumerated type for specifying fabric operations used in the PMIx server module's `pmix_server_fabric_fn_t` API.

`PMIX_FABRIC_REQUEST_INFO` 0 Request information on a specific fabric - if the fabric isn't specified as per `PMIx_Fabric_register`, then return information on the default fabric of the overall system. Information to be returned is described in `pmix_fabric_t`.

`PMIX_FABRIC_UPDATE_INFO` 1 Update information on a specific fabric - the index of the fabric (`PMIX_FABRIC_INDEX`) to be updated must be provided.

15.2.10 Fabric registration structure

The `pmix_fabric_t` structure is used by a WLM to interact with fabric-related PMIx interfaces, and to provide information about the fabric for use in scheduling algorithms or other purposes.

PMIx v4.0

```
typedef struct pmix_fabric_s {
    char *name;
    size_t index;
    pmix_info_t *info;
    size_t ninfo;
    void *module;
} pmix_fabric_t;;
```

Note that in this structure:

- *name* is an optional user-supplied string name identifying the fabric being referenced by this struct. If provided, the field must be a **NULL**-terminated string composed of standard alphanumeric values supported by common utilities such as `strcmp`;
- *index* is a PMIx-provided number identifying this object;
- *info* is an array of `pmix_info_t` containing information (provided by the PMIx library) about the fabric;

1 • *ninfo* is the number of elements in the *info* array;

2 • *module* points to an opaque object reserved for use by the PMIx server library.

3 Note that only the *name* field is provided by the user - all other fields are provided by the PMIx library and
4 must not be modified by the user. The *info* array contains a varying amount of information depending upon
5 both the PMIx implementation and information available from the fabric vendor. At a minimum, it must
6 contain (ordering is arbitrary):

▼----- Required Attributes -----▼

7 **PMIX_FABRIC_VENDOR** "pmix.fab.vndr" (**string**)

8 Name of the vendor (e.g., Amazon, Mellanox, HPE, Intel) for the specified fabric.

9 **PMIX_FABRIC_IDENTIFIER** "pmix.fab.id" (**string**)

10 An identifier for the specified fabric (e.g., MgmtEthernet, Slingshot-11, OmniPath-1).

11 **PMIX_FABRIC_NUM_DEVICES** "pmix.fab.nverts" (**size_t**)

12 Total number of fabric devices in the overall system - corresponds to the number of rows or columns in
13 the cost matrix.

▲-----

14 and may optionally contain one or more of the following:

▼----- Optional Attributes -----▼

15 **PMIX_FABRIC_COST_MATRIX** "pmix.fab.cm" (**pointer**)

16 Pointer to a two-dimensional square array of point-to-point relative communication costs expressed as
17 **uint16_t** values.

18 **PMIX_FABRIC_GROUPS** "pmix.fab.grps" (**string**)

19 A string delineating the group membership of nodes in the overall system, where each fabric group
20 consists of the group number followed by a colon and a comma-delimited list of nodes in that group,
21 with the groups delimited by semi-colons (e.g., **0:node000,node002,node004,node006;**
22 **1:node001,node003,node005,node007**)

23 **PMIX_FABRIC_DIMS** "pmix.fab.dims" (**uint32_t**)

24 Number of dimensions in the specified fabric plane/view. If no plane is specified in a request, then the
25 dimensions of all planes in the overall system will be returned as a **pmix_data_array_t**
26 containing an array of **uint32_t** values. Default is to provide dimensions in *logical* view.

27 **PMIX_FABRIC_PLANE** "pmix.fab.plane" (**string**)

28 ID string of a fabric plane (e.g., CIDR for Ethernet). When used as a modifier in a request for
29 information, specifies the plane whose information is to be returned. When used directly as a key in a
30 request, returns a **pmix_data_array_t** of string identifiers for all fabric planes in the overall
31 system.

32 **PMIX_FABRIC_SHAPE** "pmix.fab.shape" (**pmix_data_array_t***)

1 The size of each dimension in the specified fabric plane/view, returned in a `pmix_data_array_t`
2 containing an array of `uint32_t` values. The size is defined as the number of elements present in
3 that dimension - e.g., the number of devices in one dimension of a physical view of a fabric plane. If no
4 plane is specified, then the shape of each plane in the overall system will be returned in a
5 `pmix_data_array_t` array where each element is itself a two-element array containing the
6 `PMIX_FABRIC_PLANE` followed by that plane's fabric shape. Default is to provide the shape in
7 *logical* view.

8 **PMIX_FABRIC_SHAPE_STRING** "pmix.fab.shapestr" (string)

9 Network shape expressed as a string (e.g., "10x12x2"). If no plane is specified, then the shape of
10 each plane in the overall system will be returned in a `pmix_data_array_t` array where each
11 element is itself a two-element array containing the `PMIX_FABRIC_PLANE` followed by that plane's
12 fabric shape string. Default is to provide the shape in *logical* view.

13 While unusual due to scaling issues, implementations may include an array of `PMIX_FABRIC_DEVICE`
14 elements describing the device information for each device in the overall system. Each element shall contain a
15 `pmix_data_array_t` of `pmix_info_t` values describing the device. Each array may contain one or
16 more of the following (ordering is arbitrary):

17 **PMIX_FABRIC_DEVICE_NAME** "pmix.fabdev.nm" (string)

18 The operating system name associated with the device. This may be a logical fabric interface name
19 (e.g. "eth0" or "eno1") or an absolute filename.

20 **PMIX_FABRIC_DEVICE_VENDOR** "pmix.fabdev.vndr" (string)

21 Indicates the name of the vendor that distributes the device.

22 **PMIX_DEVICE_ID** "pmix.dev.id" (string)

23 System-wide UUID or node-local OS name of a particular device.

24 **PMIX_HOSTNAME** "pmix.hname" (char*)

25 Name of the host, as returned by the `gethostname` utility or its equivalent.

26 **PMIX_FABRIC_DEVICE_DRIVER** "pmix.fabdev.driver" (string)

27 The name of the driver associated with the device.

28 **PMIX_FABRIC_DEVICE_FIRMWARE** "pmix.fabdev.fmwr" (string)

29 The device's firmware version.

30 **PMIX_FABRIC_DEVICE_ADDRESS** "pmix.fabdev.addr" (string)

31 The primary link-level address associated with the device, such as a MAC address. If multiple
32 addresses are available, only one will be reported.

33 **PMIX_FABRIC_DEVICE_MTU** "pmix.fabdev.mtu" (size_t)

34 The maximum transfer unit of link level frames or packets, in bytes.

35 **PMIX_FABRIC_DEVICE_SPEED** "pmix.fabdev.speed" (size_t)

36 The active link data rate, given in bits per second.

37 **PMIX_FABRIC_DEVICE_STATE** "pmix.fabdev.state" (`pmix_link_state_t`)

38 The last available physical port state for the specified device. Possible values are
39 `PMIX_LINK_STATE_UNKNOWN`, `PMIX_LINK_DOWN`, and `PMIX_LINK_UP`, to indicate if the port
40 state is unknown or not applicable (unknown), inactive (down), or active (up).

1 **PMIX_FABRIC_DEVICE_TYPE** "pmix.fabdev.type" (string)
 2 Specifies the type of fabric interface currently active on the device, such as Ethernet or InfiniBand.

3 **PMIX_FABRIC_DEVICE_BUS_TYPE** "pmix.fabdev.btyp" (string)
 4 The type of bus to which the device is attached (e.g., "PCI", "GEN-Z").

5 **PMIX_FABRIC_DEVICE_PCI_DEVID** "pmix.fabdev.pcidevid" (string)
 6 A node-level unique identifier for a PCI device. Provided only if the device is located on a PCI bus.
 7 The identifier is constructed as a four-part tuple delimited by colons comprised of the PCI 16-bit
 8 domain, 8-bit bus, 8-bit device, and 8-bit function IDs, each expressed in zero-extended hexadecimal
 9 form. Thus, an example identifier might be "abc1:0f:23:01". The combination of node identifier
 10 (**PMIX_HOSTNAME** or **PMIX_NODEID**) and **PMIX_FABRIC_DEVICE_PCI_DEVID** shall be
 11 unique within the overall system.

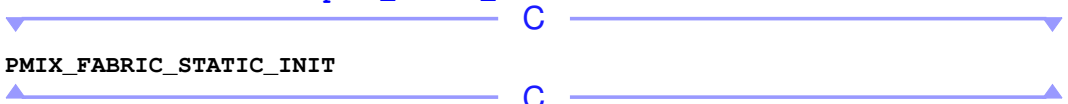


12 15.2.10.1 Static initializer for the fabric structure

13 *(Provisional)*

14 Provide a static initializer for the **pmix_fabric_t** fields.

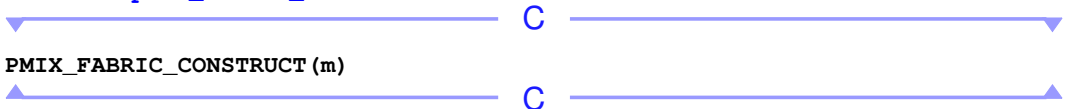
PMIx v5.0



16 15.2.10.2 Initialize the fabric structure

17 Initialize the **pmix_fabric_t** fields.

PMIx v4.0



19 **IN** m

20 Pointer to the structure to be initialized (pointer to **pmix_fabric_t**)

21 15.3 Fabric Support Attributes

22 The following attribute is used by the PMIx server library supporting the system's WLM to indicate that it
 23 wants access to the fabric support functions:

24 **PMIX_SERVER_SCHEDULER** "pmix.srv.sched" (bool)

25 Server is supporting system scheduler and desires access to appropriate WLM-supporting features.

26 Indicates that the library is to be initialized for scheduler support.

27 The following attributes may be returned in response to fabric-specific APIs or queries (e.g., **PMIx_Get** or
 28 **PMIx_Query_info**). These attributes are not related to a specific *data realm* (as described in Section 6.1) -
 29 the **PMIx_Get** function shall therefore ignore the value in its *proc* process identifier argument when
 30 retrieving these values.

31 **PMIX_FABRIC_COST_MATRIX** "pmix.fab.cm" (pointer)

1 Pointer to a two-dimensional square array of point-to-point relative communication costs expressed as
2 `uint16_t` values.

3 **PMIX_FABRIC_GROUPS** "`pmix.fab.grps`" (**string**)

4 A string delineating the group membership of nodes in the overall system, where each fabric group
5 consists of the group number followed by a colon and a comma-delimited list of nodes in that group,
6 with the groups delimited by semi-colons (e.g., `0:node000,node002,node004,node006;`
7 `1:node001,node003,node005,node007`)

8 **PMIX_FABRIC_PLANE** "`pmix.fab.plane`" (**string**)

9 ID string of a fabric plane (e.g., CIDR for Ethernet). When used as a modifier in a request for
10 information, specifies the plane whose information is to be returned. When used directly as a key in a
11 request, returns a `pmix_data_array_t` of string identifiers for all fabric planes in the overall
12 system.

13 **PMIX_FABRIC_SWITCH** "`pmix.fab.switch`" (**string**)

14 ID string of a fabric switch. When used as a modifier in a request for information, specifies the switch
15 whose information is to be returned. When used directly as a key in a request, returns a
16 `pmix_data_array_t` of string identifiers for all fabric switches in the overall system.

17 The following attributes may be returned in response to queries (e.g., `PMIx_Get` or `PMIx_Query_info`).
18 A qualifier (e.g., `PMIX_FABRIC_INDEX`) identifying the fabric whose value is being referenced must be
19 provided for queries on systems supporting more than one fabric when values for the non-default fabric are
20 requested. These attributes are not related to a specific *data realm* (as described in Section 6.1) - the
21 `PMIx_Get` function shall therefore ignore the value in its *proc* process identifier argument when retrieving
22 these values.

23 **PMIX_FABRIC_VENDOR** "`pmix.fab.vndr`" (**string**)

24 Name of the vendor (e.g., Amazon, Mellanox, HPE, Intel) for the specified fabric.

25 **PMIX_FABRIC_IDENTIFIER** "`pmix.fab.id`" (**string**)

26 An identifier for the specified fabric (e.g., MgmtEthernet, Slingshot-11, OmniPath-1).

27 **PMIX_FABRIC_INDEX** "`pmix.fab.idx`" (**size_t**)

28 The index of the fabric as returned in `pmix_fabric_t`.

29 **PMIX_FABRIC_NUM_DEVICES** "`pmix.fab.nverts`" (**size_t**)

30 Total number of fabric devices in the overall system - corresponds to the number of rows or columns in
31 the cost matrix.

32 **PMIX_FABRIC_DIMS** "`pmix.fab.dims`" (**uint32_t**)

33 Number of dimensions in the specified fabric plane/view. If no plane is specified in a request, then the
34 dimensions of all planes in the overall system will be returned as a `pmix_data_array_t`
35 containing an array of `uint32_t` values. Default is to provide dimensions in *logical* view.

36 **PMIX_FABRIC_SHAPE** "`pmix.fab.shape`" (**pmix_data_array_t***)

37 The size of each dimension in the specified fabric plane/view, returned in a `pmix_data_array_t`
38 containing an array of `uint32_t` values. The size is defined as the number of elements present in
39 that dimension - e.g., the number of devices in one dimension of a physical view of a fabric plane. If no
40 plane is specified, then the shape of each plane in the overall system will be returned in a
41 `pmix_data_array_t` array where each element is itself a two-element array containing the
42 `PMIX_FABRIC_PLANE` followed by that plane's fabric shape. Default is to provide the shape in
43 *logical* view.

44 **PMIX_FABRIC_SHAPE_STRING** "`pmix.fab.shapestr`" (**string**)

1 Network shape expressed as a string (e.g., "10x12x2"). If no plane is specified, then the shape of
2 each plane in the overall system will be returned in a `pmix_data_array_t` array where each
3 element is itself a two-element array containing the `PMIX_FABRIC_PLANE` followed by that plane's
4 fabric shape string. Default is to provide the shape in *logical* view.

5 The following attributes are related to the *node realm* (as described in Section 6.1.5) and are retrieved
6 according to those rules.

7 **PMIX_FABRIC_DEVICES** "pmix.fab.devs" (`pmix_data_array_t`)

8 Array of `pmix_info_t` containing information for all devices on the specified node. Each element of
9 the array will contain a `PMIX_FABRIC_DEVICE` entry, which in turn will contain an array of
10 information on a given device.

11 **PMIX_FABRIC_COORDINATES** "pmix.fab.coords" (`pmix_data_array_t`)

12 Array of `pmix_geometry_t` fabric coordinates for devices on the specified node. The array will
13 contain the coordinates of all devices on the node, including values for all supported coordinate views.
14 The information for devices on the local node shall be provided if the node is not specified in the
15 request.

16 **PMIX_FABRIC_DEVICE** "pmix.fabdev" (`pmix_data_array_t`)

17 An array of `pmix_info_t` describing a particular fabric device using one or more of the attributes
18 defined below. The first element in the array shall be the `PMIX_DEVICE_ID` of the device.

19 **PMIX_FABRIC_DEVICE_INDEX** "pmix.fabdev.idx" (`uint32_t`)

20 Index of the device within an associated communication cost matrix.

21 **PMIX_FABRIC_DEVICE_NAME** "pmix.fabdev.nm" (`string`)

22 The operating system name associated with the device. This may be a logical fabric interface name
23 (e.g. "eth0" or "eno1") or an absolute filename.

24 **PMIX_FABRIC_DEVICE_VENDOR** "pmix.fabdev.vndr" (`string`)

25 Indicates the name of the vendor that distributes the device.

26 **PMIX_FABRIC_DEVICE_BUS_TYPE** "pmix.fabdev.btyp" (`string`)

27 The type of bus to which the device is attached (e.g., "PCI", "GEN-Z").

28 **PMIX_FABRIC_DEVICE_VENDORID** "pmix.fabdev.vendid" (`string`)

29 This is a vendor-provided identifier for the device or product.

30 **PMIX_FABRIC_DEVICE_DRIVER** "pmix.fabdev.driver" (`string`)

31 The name of the driver associated with the device.

32 **PMIX_FABRIC_DEVICE_FIRMWARE** "pmix.fabdev.fmwr" (`string`)

33 The device's firmware version.

34 **PMIX_FABRIC_DEVICE_ADDRESS** "pmix.fabdev.addr" (`string`)

35 The primary link-level address associated with the device, such as a MAC address. If multiple
36 addresses are available, only one will be reported.

37 **PMIX_FABRIC_DEVICE_COORDINATES** "pmix.fab.coord" (`pmix_geometry_t`)

38 The `pmix_geometry_t` fabric coordinates for the device, including values for all supported
39 coordinate views.

40 **PMIX_FABRIC_DEVICE_MTU** "pmix.fabdev.mtu" (`size_t`)

41 The maximum transfer unit of link level frames or packets, in bytes.

42 **PMIX_FABRIC_DEVICE_SPEED** "pmix.fabdev.speed" (`size_t`)

43 The active link data rate, given in bits per second.

44 **PMIX_FABRIC_DEVICE_STATE** "pmix.fabdev.state" (`pmix_link_state_t`)

The last available physical port state for the specified device. Possible values are `PMIX_LINK_STATE_UNKNOWN`, `PMIX_LINK_DOWN`, and `PMIX_LINK_UP`, to indicate if the port state is unknown or not applicable (unknown), inactive (down), or active (up).

PMIX_FABRIC_DEVICE_TYPE "`pmix.fabdev.type`" (**string**)

Specifies the type of fabric interface currently active on the device, such as Ethernet or InfiniBand.

PMIX_FABRIC_DEVICE_PCI_DEVID "`pmix.fabdev.pcidevid`" (**string**)

A node-level unique identifier for a PCI device. Provided only if the device is located on a PCI bus. The identifier is constructed as a four-part tuple delimited by colons comprised of the PCI 16-bit domain, 8-bit bus, 8-bit device, and 8-bit function IDs, each expressed in zero-extended hexadecimal form. Thus, an example identifier might be "abc1:0f:23:01". The combination of node identifier (`PMIX_HOSTNAME` or `PMIX_NODEID`) and `PMIX_FABRIC_DEVICE_PCI_DEVID` shall be unique within the overall system.

The following attributes are related to the *process realm* (as described in Section 6.1.4) and are retrieved according to those rules.

PMIX_FABRIC_ENDPT "`pmix.fab.endpt`" (**pmix_data_array_t**)

Fabric endpoints for a specified process. As multiple endpoints may be assigned to a given process (e.g., in the case where multiple devices are associated with a package to which the process is bound), the returned values will be provided in a `pmix_data_array_t` of `pmix_endpoint_t` elements.

The following attributes are related to the *job realm* (as described in Section 6.1.2) and are retrieved according to those rules. Note that distances to fabric devices are retrieved using the `PMIX_DEVICE_DISTANCES` key with the appropriate `pmix_device_type_t` qualifier.

PMIX_SWITCH_PEERS "`pmix.speers`" (**pmix_data_array_t**)

Peer ranks that share the same switch as the process specified in the call to `PMIx_Get`. Returns a `pmix_data_array_t` array of `pmix_info_t` results, each element containing the `PMIX_SWITCH_PEERS` key with a three-element `pmix_data_array_t` array of `pmix_info_t` containing the `PMIX_DEVICE_ID` of the local fabric device, the `PMIX_FABRIC_SWITCH` identifying the switch to which it is connected, and a comma-delimited string of peer ranks sharing the switch to which that device is connected.

15.4 Fabric Support Functions

The following APIs allow the WLM to request specific services from the fabric subsystem via the PMIx library.

Advice to PMIx server hosts

Due to their high cost in terms of execution, memory consumption, and interactions with other SMS components (e.g., a fabric manager), it is strongly advised that the underlying implementation of these APIs be restricted to a single PMIx server in a system that is supporting the SMS component responsible for the scheduling of allocations (i.e., the system *scheduler*). The `PMIX_SERVER_SCHEDULER` attribute can be used for this purpose to control the execution path. Clients, tools, and other servers utilizing these functions are advised to have their requests forwarded to the server supporting the scheduler using the `pmix_server_fabric_fn_t` server module function, as needed.

1 15.4.1 PMIx_Fabric_register

2 Summary

3 Register for access to fabric-related information.

4 *PMIx v4.0* Format

C

```

5 pmix_status_t
6 PMIx_Fabric_register(pmix_fabric_t *fabric,
7                     const pmix_info_t directives[],
8                     size_t ndirs);

```

C

9 INOUT fabric

10 address of a `pmix_fabric_t` (backed by storage). User may populate the "name" field at will - PMIx
11 does not utilize this field (handle)

12 IN directives

13 an optional array of values indicating desired behaviors and/or fabric to be accessed. If `NULL`, then the
14 highest priority available fabric will be used (array of handles)

15 IN ndirs

16 Number of elements in the *directives* array (integer)

17 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

18 The following directives are required to be supported by all PMIx libraries to aid users in identifying the fabric
19 whose data is being sought:

20 `PMIX_FABRIC_PLANE` "pmix.fab.plane" (string)

21 ID string of a fabric plane (e.g., CIDR for Ethernet). When used as a modifier in a request for
22 information, specifies the plane whose information is to be returned. When used directly as a key in a
23 request, returns a `pmix_data_array_t` of string identifiers for all fabric planes in the overall
24 system.

25 `PMIX_FABRIC_IDENTIFIER` "pmix.fab.id" (string)

26 An identifier for the specified fabric (e.g., MgmtEthernet, Slingshot-11, OmniPath-1).

27 `PMIX_FABRIC_VENDOR` "pmix.fab.vndr" (string)

28 Name of the vendor (e.g., Amazon, Mellanox, HPE, Intel) for the specified fabric.

Description

Register for access to fabric-related information, including the communication cost matrix. This call must be made prior to requesting information from a fabric. The caller may request access to a particular fabric using the vendor, type, or identifier, or to a specific *fabric plane* via the `PMIX_FABRIC_PLANE` attribute - otherwise, information for the default fabric will be returned. Upon successful completion of the call, information will have been filled into the fields of the provided *fabric* structure.

For performance reasons, the PMIx library does not provide thread protection for accessing the information in the `pmix_fabric_t` structure. Instead, the PMIx implementation shall provide two methods for coordinating updates to the provided fabric information:

- Users may periodically poll for updates using the `PMIx_Fabric_update` API
- Users may register for `PMIX_FABRIC_UPDATE_PENDING` events indicating that an update to the cost matrix is pending. When received, users are required to terminate or pause any actions involving access to the cost matrix before returning from the event. Completion of the `PMIX_FABRIC_UPDATE_PENDING` event handler indicates to the PMIx library that the fabric object's entries are available for updating. This may include releasing and re-allocating memory as the number of vertices may have changed (e.g., due to addition or removal of one or more devices). When the update has been completed, the PMIx library will generate a `PMIX_FABRIC_UPDATED` event indicating that it is safe to begin using the updated fabric object(s).

There is no requirement that the caller exclusively use either one of these options. For example, the user may choose to both register for fabric update events, but poll for an update prior to some critical operation.

15.4.2 PMIx_Fabric_register_nb

Summary

Register for access to fabric-related information.

Format

PMIx v4.0

C

```
pmix_status_t
PMIx_Fabric_register_nb(pmix_fabric_t *fabric,
                       const pmix_info_t directives[],
                       size_t ndirs,
                       pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

INOUT fabric

address of a `pmix_fabric_t` (backed by storage). User may populate the "name" field at will - PMIx does not utilize this field (handle)

IN directives

an optional array of values indicating desired behaviors and/or fabric to be accessed. If `NULL`, then the highest priority available fabric will be used (array of handles)

IN ndirs

Number of elements in the *directives* array (integer)

IN cbfunc

Callback function `pmix_op_cbfunc_t` (function reference)

1 **IN** **cbdata**
 2 Data to be passed to the callback function (memory reference)

3 A successful return indicates that the request is being processed and the result will be returned in the provided
 4 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
 5 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

6 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

7

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
 8 *success* - the *cbfunc* will not be called

9 If none of the above return codes are appropriate, then an implementation must return either a general PMIX
 10 error code or an implementation defined error code as described in Section 3.1.1.

11 **Description**

12 Non-blocking form of **PMIx_Fabric_register**. The caller is not allowed to access the provided
 13 **pmix_fabric_t** until the callback function has been executed, at which time the fabric information will
 14 have been loaded into the provided structure.

15 **15.4.3 PMIx_Fabric_update**

16 **Summary**

17 Update fabric-related information.

18 **Format**

18 *PMIx v4.0*

19 **pmix_status_t**
 20 **PMIx_Fabric_update**(**pmix_fabric_t** **fabric*);

21 **INOUT fabric**

22 address of a **pmix_fabric_t** (backed by storage) (handle)

23 Returns **PMIX_SUCCESS** or a negative value indicating the error.

24 **Description**

25 Update fabric-related information. This call can be made at any time to request an update of the fabric
 26 information contained in the provided **pmix_fabric_t** object. The caller is not allowed to access the
 27 provided **pmix_fabric_t** until the call has returned. Upon successful return, the information fields in the
 28 *fabric* structure will have been updated.

29 **15.4.4 PMIx_Fabric_update_nb**

30 **Summary**

31 Update fabric-related information.

Format

C

`pmix_status_t`

```
PMIx_Fabric_update_nb(pmix_fabric_t *fabric,  
                      pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

INOUT `fabric`

address of a `pmix_fabric_t` (handle)

IN `cbfunc`

Callback function `pmix_op_cbfunc_t` (function reference)

IN `cbdata`

Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided `cbfunc`. Note that the library must not invoke the callback function prior to returning from the API. The callback function, `cbfunc`, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the `cbfunc` will not be called

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

Non-blocking form of **PMIx_Fabric_update**. The caller is not allowed to access the provided `pmix_fabric_t` until the callback function has been executed, at which time the fields in the provided `fabric` structure will have been updated.

15.4.5 PMIx_Fabric_deregister

Summary

Deregister a fabric object.

Format

C

`pmix_status_t`

```
PMIx_Fabric_deregister(pmix_fabric_t *fabric);
```

C

IN `fabric`

address of a `pmix_fabric_t` (handle)

Returns **PMIX_SUCCESS** or a negative value indicating the error.

Description

Deregister a fabric object, providing an opportunity for the PMIx library to cleanup any information (e.g., cost matrix) associated with it. Contents of the provided `pmix_fabric_t` will be invalidated upon function return.

1 15.4.6 PMIx_Fabric_deregister_nb

2 Summary

3 Deregister a fabric object.

4 Format

PMIx v4.0

C

```
5 pmix_status_t PMIx_Fabric_deregister_nb(pmix_fabric_t *fabric,  
6                                         pmix_op_cbfunc_t cbfunc,  
7                                         void *cbdata);
```

C

8 **IN fabric**

9 address of a [pmix_fabric_t](#) (handle)

10 **IN cbfunc**

11 Callback function [pmix_op_cbfunc_t](#) (function reference)

12 **IN cbdata**

13 Data to be passed to the callback function (memory reference)

14 A successful return indicates that the request is being processed and the result will be returned in the provided
15 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
16 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

17 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- 18 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
19 *success* - the *cbfunc* will not be called

20 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
21 error code or an implementation defined error code as described in Section 3.1.1.

22 Description

23 Non-blocking form of [PMIx_Fabric_deregister](#). Provided *fabric* must not be accessed until after
24 callback function has been executed.

CHAPTER 16

Security

1 PMIx utilizes a multi-layered approach toward security that differs for client versus tool processes. By
2 definition, *client* processes must be preregistered with the PMIx server library via the
3 `PMIx_server_register_client` API before they are spawned. This API requires that the host pass the
4 expected effective UID/GID of the client process.

5 When the client attempts to connect to the PMIx server, the server shall use available standard OS methods to
6 determine the effective UID/GID of the process requesting the connection. PMIx implementations shall not
7 rely on any values reported by the client process itself. The effective UID/GID reported by the OS is compared
8 to the values provided by the host during registration - if the values fail to match, the PMIx server is required
9 to drop the connection request. This ensures that the PMIx server does not allow connection from a client that
10 doesn't at least meet some minimal security requirement.

11 Once the requesting client passes the initial test, the PMIx server can, at the choice of the implementor,
12 perform additional security checks. This may involve a variety of methods such as exchange of a
13 system-provided key or credential. At the conclusion of that process, the PMIx server reports the client
14 connection request to the host via the `pmix_server_client_connected2_fn_t` interface, if
15 provided. The host may perform any additional checks and operations before responding with either
16 `PMIX_SUCCESS` to indicate that the connection is approved, or a PMIx error constant indicating that the
17 connection request is refused. In this latter case, the PMIx server is required to drop the connection.

18 Tools started by the host environment are classed as a subgroup of client processes and follow the client
19 process procedure. However, tools that are not started by the host environment must be handled differently as
20 registration information is not available prior to the connection request. In these cases, the PMIx server library
21 is required to use available standard OS methods to get the effective UID/GID of the tool and report them
22 upwards as part of invoking the `pmix_server_tool_connection_fn_t` interface, deferring initial
23 security screening to the host. Host environments willing to accept tool connections must therefore both
24 explicitly enable them via the `PMIX_SERVER_TOOL_SUPPORT` attribute, thereby confirming acceptance of
25 the authentication and authorization burden, and provide the `pmix_server_tool_connection_fn_t`
26 server module function pointer.

16.1 Obtaining Credentials

27 Applications and tools often interact with the host environment in ways that require security beyond just
28 verifying the user's identity - e.g., access to that user's relevant authorizations. This is particularly important
29 when tools connect directly to a system-level PMIx server that may be operating at a privileged level. A
30 variety of system management software packages provide authorization services, but the lack of standardized
31 interfaces makes portability problematic.
32

33 This section defines two PMIx client-side APIs for this purpose. These are most likely to be used by
34 user-space applications/tools, but are not restricted to that realm.

1 16.1.1 PMIx_Get_credential

2 Summary

3 Request a credential from the PMIx server library or the host environment.

4 *PMIx v3.0* Format C

```
5 pmix_status_t
6 PMIx_Get_credential(const pmix_info_t info[], size_t ninfo,
7                     pmix_byte_object_t *credential);
C
```

- 8 **IN info**
9 Array of `pmix_info_t` structures (array of handles)
- 10 **IN ninfo**
11 Number of elements in the *info* array (`size_t`)
- 12 **IN credential**
13 Address of a `pmix_byte_object_t` within which to return credential (handle)

14 A successful return indicates that the credential has been returned in the provided `pmix_byte_object_t`.

15 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

16 There are no required attributes for this API. Note that implementations may choose to internally execute
17 integration for some security environments (e.g., directly contacting a *munge* server).

18 Implementations that support the operation but cannot directly process the client's request must pass any
19 attributes that are provided by the client to the host environment for processing. In addition, the following
20 attributes are required to be included in the *info* array passed from the PMIx library to the host environment:

21 **PMIX_USERID** "pmix.euid" (`uint32_t`)
22 Effective user ID of the connecting process.

23 **PMIX_GRPID** "pmix.egid" (`uint32_t`)
24 Effective group ID of the connecting process.



Optional Attributes

25 The following attributes are optional for host environments that support this operation:

26 **PMIX_TIMEOUT** "pmix.timeout" (`int`)
27 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
28 `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers
29 (client, server, and host) simultaneously timing the operation.



Description

Request a credential from the PMIx server library or the host environment. The credential is returned as a [pmix_byte_object_t](#) to support potential binary formats - it is therefore opaque to the caller. No information as to the source of the credential is provided.

16.1.2 PMIx_Get_credential_nb

Summary

Request a credential from the PMIx server library or the host environment.

Format

PMIx v3.0

```
pmix_status_t
PMIx_Get_credential_nb(const pmix_info_t info[], size_t ninfo,
                      pmix_credential_cbfunc_t cbfunc,
                      void *cbdata);
```

IN info

Array of [pmix_info_t](#) structures (array of handles)

IN ninfo

Number of elements in the *info* array (**size_t**)

IN cbfunc

Callback function to return credential ([pmix_credential_cbfunc_t](#) function reference)

IN cbdata

Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

There are no required attributes for this API. Note that implementations may choose to internally execute integration for some security environments (e.g., directly contacting a *munge* server).

Implementations that support the operation but cannot directly process the client's request must pass any attributes that are provided by the client to the host environment for processing. In addition, the following attributes are required to be included in the *info* array passed from the PMIx library to the host environment:

PMIX_USERID "pmix.euid" (**uint32_t**)

Effective user ID of the connecting process.

PMIX_GRPID "pmix.egid" (**uint32_t**)

1 Effective group ID of the connecting process.

Optional Attributes

2 The following attributes are optional for host environments that support this operation:

3 **PMIX_TIMEOUT** "pmix.timeout" (int)

4 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
5 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
6 (client, server, and host) simultaneously timing the operation.

Description

7 Request a credential from the PMIx server library or the host environment. This version of the API is generally
8 preferred in scenarios where the host environment may have to contact a remote credential service. Thus,
9 provision is made for the system to return additional information (e.g., the identity of the issuing agent) outside
10 of the credential itself and visible to the application.
11

16.1.3 Credential Attributes

12 The following attributes are defined to support credential operations:

13 **PMIX_CRED_TYPE** "pmix.sec.ctype" (char*)

14 When passed in **PMIx_Get_credential**, a prioritized, comma-delimited list of desired credential
15 types for use in environments where multiple authentication mechanisms may be available. When
16 returned in a callback function, a string identifier of the credential type.

17 **PMIX_CRYPT_KEY** "pmix.sec.key" (pmix_byte_object_t)

18 Blob containing crypto key.
19

16.2 Validating Credentials

20 Given a credential, PMIx provides two methods by which a caller can request that the system validate it,
21 returning any additional information (e.g., authorizations) conveyed within the credential.
22

16.2.1 PMIx_Validate_credential

Summary

23 Request validation of a credential by the PMIx server library or the host environment.
24
25

Format

C

```
pmix_status_t
PMIx_Validate_credential(const pmix_byte_object_t *cred,
                        const pmix_info_t info[], size_t ninfo,
                        pmix_info_t **results, size_t *nresults);
```

C

IN cred

Pointer to `pmix_byte_object_t` containing the credential (handle)

IN info

Array of `pmix_info_t` structures (array of handles)

IN ninfo

Number of elements in the `info` array (`size_t`)

INOUT results

Address where a pointer to an array of `pmix_info_t` containing the results of the request can be returned (memory reference)

INOUT nresults

Address where the number of elements in `results` can be returned (handle)

A successful return indicates that the credential was valid and any information it contained was successfully processed. Details of the result will be returned in the `results` array.

Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

There are no required attributes for this API. Note that implementations may choose to internally execute integration for some security environments (e.g., directly contacting a `munge` server).

Implementations that support the operation but cannot directly process the client's request must pass any attributes that are provided by the client to the host environment for processing. In addition, the following attributes are required to be included in the `info` array passed from the PMIx library to the host environment:

PMIX_USERID "pmix.euid" (`uint32_t`)

Effective user ID of the connecting process.

PMIX_GRPID "pmix.egid" (`uint32_t`)

Effective group ID of the connecting process.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (`int`)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Request validation of a credential by the PMIx server library or the host environment.

16.2.2 PMIx_Validate_credential_nb

Summary

Request validation of a credential by the PMIx server library or the host environment. Provision is made for the system to return additional information regarding possible authorization limitations beyond simple authentication.

Format

PMIx v3.0

C

`pmix_status_t`

```
PMIx_Validate_credential_nb(const pmix_byte_object_t *cred,  
                           const pmix_info_t info[], size_t ninfo,  
                           pmix_validation_cbfunc_t cbfunc,  
                           void *cbdata);
```

C

IN `cred`

Pointer to `pmix_byte_object_t` containing the credential (handle)

IN `info`

Array of `pmix_info_t` structures (array of handles)

IN `ninfo`

Number of elements in the `info` array (`size_t`)

IN `cbfunc`

Callback function to return result (`pmix_validation_cbfunc_t` function reference)

IN `cbdata`

Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided `cbfunc`. Note that the library must not invoke the callback function prior to returning from the API. The callback function, `cbfunc`, is only called when `PMIX_SUCCESS` is returned.

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

`PMIX_OPERATION_SUCCEEDED`, indicating that the request was immediately processed successfully - the `cbfunc` will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

There are no required attributes for this API. Note that implementations may choose to internally execute integration for some security environments (e.g., directly contacting a `munge` server).

Implementations that support the operation but cannot directly process the client's request must pass any attributes that are provided by the client to the host environment for processing. In addition, the following attributes are required to be included in the `info` array passed from the PMIx library to the host environment:

`PMIX_USERID` "pmix.euid" (`uint32_t`)

1 Effective user ID of the connecting process.

2 **PMIX_GRPID** "pmix.egid" (uint32_t)

3 Effective group ID of the connecting process.



4 Optional Attributes

4 The following attributes are optional for host environments that support this operation:

5 **PMIX_TIMEOUT** "pmix.timeout" (int)

6 Time in seconds before the specified operation should time out (zero indicating infinite) and return the

7 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers

8 (client, server, and host) simultaneously timing the operation.



9 **Description**

10 Request validation of a credential by the PMIx server library or the host environment. This version of the API
11 is generally preferred in scenarios where the host environment may have to contact a remote credential service.
12 Provision is made for the system to return additional information (e.g., possible authorization limitations)
13 beyond simple authentication.

CHAPTER 17

Server-Specific Interfaces

1 The process that hosts the PMIx server library interacts with that library in two distinct manners. First, PMIx
2 provides a set of APIs by which the host can request specific services from its library. This includes:

- 3 • collecting inventory to support scheduling algorithms,
- 4 • providing subsystems with an opportunity to precondition their resources for optimized application support,
- 5 • generating regular expressions,
- 6 • registering information to be passed to client processes, and
- 7 • requesting information on behalf of a remote process.

8 Note that the host always has access to all PMIx client APIs - the functions listed below are in addition to those
9 available to a PMIx client.

10 Second, the host can provide a set of callback functions by which the PMIx server library can pass requests
11 upward for servicing by the host. These include notifications of client connection and finalize, as well as
12 requests by clients for information and/or services that the PMIx server library does not itself provide.

13 17.1 Server Initialization and Finalization

14 Initialization and finalization routines for PMIx servers.

15 17.1.1 PMIx_server_init

16 Summary

17 Initialize the PMIx server.

18 Format

PMIx v1.0

C

19 `pmix_status_t`

20 `PMIx_server_init(pmix_server_module_t *module,`
21 `pmix_info_t info[], size_t ninfo);`

C

22 INOUT module

23 `pmix_server_module_t` structure (handle)

24 IN info

25 Array of `pmix_info_t` structures (array of handles)

26 IN ninfo

27 Number of elements in the `info` array (`size_t`)

28 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

- PMIX_SERVER_NAMESPACE** "pmix.srv.namespace" (char*)
Name of the namespace to use for this PMIx server.
- PMIX_SERVER_RANK** "pmix.srv.rank" (pmix_rank_t)
Rank of this PMIx server.
- PMIX_SERVER_TMPDIR** "pmix.srvr.tmpdir" (char*)
Top-level temporary directory for all client processes connected to this server, and where the PMIx server will place its tool rendezvous point and contact information.
- PMIX_SYSTEM_TMPDIR** "pmix.sys.tmpdir" (char*)
Temporary directory for this system, and where a PMIx server that declares itself to be a system-level server will place a tool rendezvous point and contact information.
- PMIX_SERVER_TOOL_SUPPORT** "pmix.srvr.tool" (bool)
The host RM wants to declare itself as willing to accept tool connection requests.
- PMIX_SERVER_SYSTEM_SUPPORT** "pmix.srvr.sys" (bool)
The host RM wants to declare itself as being the local system server for PMIx connection requests.
- PMIX_SERVER_SESSION_SUPPORT** "pmix.srvr.sess" (bool)
The host RM wants to declare itself as being the local session server for PMIx connection requests.
- PMIX_SERVER_GATEWAY** "pmix.srv.gway" (bool)
Server is acting as a gateway for PMIx requests that cannot be serviced on backend nodes (e.g., logging to email).
- PMIX_SERVER_SCHEDULER** "pmix.srv.sched" (bool)
Server is supporting system scheduler and desires access to appropriate WLM-supporting features. Indicates that the library is to be initialized for scheduler support.

Optional Attributes

The following attributes are optional for implementers of PMIx libraries:

- PMIX_USOCK_DISABLE** "pmix.usock.disable" (bool)
Disable legacy UNIX socket (usock) support. If the library supports Unix socket connections, this attribute may be supported for disabling it.
- PMIX_SOCKET_MODE** "pmix.sockmode" (uint32_t)
POSIX *mode_t* (9 bits valid). If the library supports socket connections, this attribute may be supported for setting the socket mode.
- PMIX_SINGLE_LISTENER** "pmix.sing.listnr" (bool)
Use only one rendezvous socket, letting priorities and/or environment parameters select the active transport.
- PMIX_TCP_REPORT_URI** "pmix.tcp.repuri" (char*)

1 If provided, directs that the TCP URI be reported and indicates the desired method of reporting: '-'
2 for stdout, '+' for stderr, or filename. If the library supports TCP socket connections, this attribute
3 may be supported for reporting the URI.

4 **PMIX_TCP_IF_INCLUDE** "pmix.tcp.ifinclude" (char*)
5 Comma-delimited list of devices and/or CIDR notation to include when establishing the TCP
6 connection. If the library supports TCP socket connections, this attribute may be supported for
7 specifying the interfaces to be used.

8 **PMIX_TCP_IF_EXCLUDE** "pmix.tcp.ifexclude" (char*)
9 Comma-delimited list of devices and/or CIDR notation to exclude when establishing the TCP
10 connection. If the library supports TCP socket connections, this attribute may be supported for
11 specifying the interfaces that are *not* to be used.

12 **PMIX_TCP_IPV4_PORT** "pmix.tcp.ipv4" (int)
13 The IPv4 port to be used. If the library supports IPV4 connections, this attribute may be supported
14 for specifying the port to be used.

15 **PMIX_TCP_IPV6_PORT** "pmix.tcp.ipv6" (int)
16 The IPv6 port to be used. If the library supports IPV6 connections, this attribute may be supported
17 for specifying the port to be used.

18 **PMIX_TCP_DISABLE_IPV4** "pmix.tcp.disipv4" (bool)
19 Set to **true** to disable IPv4 family of addresses. If the library supports IPV4 connections, this
20 attribute may be supported for disabling it.

21 **PMIX_TCP_DISABLE_IPV6** "pmix.tcp.disipv6" (bool)
22 Set to **true** to disable IPv6 family of addresses. If the library supports IPV6 connections, this
23 attribute may be supported for disabling it.

24 **PMIX_SERVER_REMOTE_CONNECTIONS** "pmix.srvr.remote" (bool)
25 Allow connections from remote tools. Forces the PMIx server to not exclusively use loopback device.
26 If the library supports connections from remote tools, this attribute may be supported for enabling or
27 disabling it.

28 **PMIX_EXTERNAL_PROGRESS** "pmix.evext" (bool)
29 The host shall progress the PMIx library via calls to **PMIx_Progress**

30 **PMIX_EVENT_BASE** "pmix.evbase" (void*)
31 Pointer to an **event_base** to use in place of the internal progress thread. All PMIx library events are
32 to be assigned to the provided event base. The event base *must* be compatible with the event library
33 used by the PMIx implementation - e.g., either both the host and PMIx library must use libevent, or
34 both must use libev. Cross-matches are unlikely to work and should be avoided - it is the responsibility
35 of the host to ensure that the PMIx implementation supports (and was built with) the appropriate event
36 library.

37 **PMIX_TOPOLOGY2** "pmix.topo2" (pmix_topology_t)
38 Provide a pointer to an implementation-specific description of the local node topology.

39 **PMIX_SERVER_SHARE_TOPOLOGY** "pmix.srvr.share" (bool)

1 The PMIx server is to share its copy of the local node topology (whether given to it or self-discovered)
2 with any clients. The PMIx server will perform the necessary actions to scalably expose the
3 description to the local clients. This includes creating any required shared memory backing stores and/
4 or XML representations, plus ensuring that all necessary key-value pairs for clients to access the
5 description are included in the job-level information provided to each client. All required files are to be
6 installed under the effective `PMIX_SERVER_TMPDIR` directory. The PMIx server library is
7 responsible for cleaning up any artifacts (e.g., shared memory backing files or cached key-value pairs)
8 at library finalize.

9 `PMIX_SERVER_ENABLE_MONITORING` "pmix.srv.monitor" (bool)

10 Enable PMIx internal monitoring by the PMIx server.

11 `PMIX_HOMOGENEOUS_SYSTEM` "pmix.homo" (bool)

12 The nodes comprising the session are homogeneous - i.e., they each contain the same number of
13 identical packages, fabric interfaces, GPUs, and other devices.

14 `PMIX_SINGLETON` "pmix.singleton" (char*)

15 String representation (nspace.rank) of proc ID for the singleton the server was started to support

16 `PMIX_IOF_LOCAL_OUTPUT` "pmix.iof.local" (bool)

17 Write output streams to local stdout/err



18 Description

19 Initialize the PMIx server support library, and provide a pointer to a `pmix_server_module_t` structure
20 containing the caller's callback functions. The array of `pmix_info_t` structs is used to pass additional info
21 that may be required by the server when initializing. For example, it may include the
22 `PMIX_SERVER_TOOL_SUPPORT` attribute, thereby indicating that the daemon is willing to accept
23 connection requests from tools.

Advice to PMIx server hosts

24 Providing a value of `NULL` for the *module* argument is permitted, as is passing an empty *module* structure.
25 Doing so indicates that the host environment will not provide support for multi-node operations such as
26 `PMIx_Fence`, but does intend to support local clients access to information.

27 17.1.2 PMIx_server_finalize

28 Summary

29 Finalize the PMIx server library.

30 Format

31 `pmix_status_t`
32 `PMIx_server_finalize(void);`

33 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Description

Finalize the PMIx server support library, terminating all connections to attached tools and any local clients. All memory usage is released.

17.1.3 Server Initialization Attributes

These attributes are used to direct the configuration and operation of the PMIx server library by passing them into `PMIx_server_init`.

- PMIX_TOPOLOGY2** "`pmix.topo2`" (`pmix_topology_t`)
Provide a pointer to an implementation-specific description of the local node topology.
- PMIX_SERVER_SHARE_TOPOLOGY** "`pmix.srvr.share`" (`bool`)
The PMIx server is to share its copy of the local node topology (whether given to it or self-discovered) with any clients.
- PMIX_USOCK_DISABLE** "`pmix.usock.disable`" (`bool`)
Disable legacy UNIX socket (usock) support.
- PMIX_SOCKET_MODE** "`pmix.sockmode`" (`uint32_t`)
POSIX `mode_t` (9 bits valid).
- PMIX_SINGLE_LISTENER** "`pmix.sing.listnr`" (`bool`)
Use only one rendezvous socket, letting priorities and/or environment parameters select the active transport.
- PMIX_SERVER_TOOL_SUPPORT** "`pmix.srvr.tool`" (`bool`)
The host RM wants to declare itself as willing to accept tool connection requests.
- PMIX_SERVER_REMOTE_CONNECTIONS** "`pmix.srvr.remote`" (`bool`)
Allow connections from remote tools. Forces the PMIx server to not exclusively use loopback device.
- PMIX_SERVER_SYSTEM_SUPPORT** "`pmix.srvr.sys`" (`bool`)
The host RM wants to declare itself as being the local system server for PMIx connection requests.
- PMIX_SERVER_SESSION_SUPPORT** "`pmix.srvr.sess`" (`bool`)
The host RM wants to declare itself as being the local session server for PMIx connection requests.
- PMIX_SERVER_START_TIME** "`pmix.srvr.strtime`" (`char*`)
Time when the server started - i.e., when the server created its rendezvous file (given in ctime string format).
- PMIX_SERVER_TMPDIR** "`pmix.srvr.tmpdir`" (`char*`)
Top-level temporary directory for all client processes connected to this server, and where the PMIx server will place its tool rendezvous point and contact information.
- PMIX_SYSTEM_TMPDIR** "`pmix.sys.tmpdir`" (`char*`)
Temporary directory for this system, and where a PMIx server that declares itself to be a system-level server will place a tool rendezvous point and contact information.
- PMIX_SERVER_ENABLE_MONITORING** "`pmix.srv.monitor`" (`bool`)
Enable PMIx internal monitoring by the PMIx server.
- PMIX_SERVER_NAMESPACE** "`pmix.srv.namespace`" (`char*`)
Name of the namespace to use for this PMIx server.
- PMIX_SERVER_RANK** "`pmix.srv.rank`" (`pmix_rank_t`)
Rank of this PMIx server.
- PMIX_SERVER_GATEWAY** "`pmix.srv.gway`" (`bool`)
Server is acting as a gateway for PMIx requests that cannot be serviced on backend nodes (e.g., logging to email).


```

1  PMIX_SERVER_SCHEDULER "pmix.srv.sched" (bool)
2      Server is supporting system scheduler and desires access to appropriate WLM-supporting features.
3      Indicates that the library is to be initialized for scheduler support.
4  PMIX_EXTERNAL_PROGRESS "pmix.evext" (bool)
5      The host shall progress the PMIx library via calls to PMIx_Progress
6  PMIX_HOMOGENEOUS_SYSTEM "pmix.homo" (bool)
7      The nodes comprising the session are homogeneous - i.e., they each contain the same number of
8      identical packages, fabric interfaces, GPUs, and other devices.
9  PMIX_SINGLETON "pmix.singleton" (char*) (Provisional)
10     String representation (nspace.rank) of proc ID for the singleton the server was started to support

```

11 17.2 Server Support Functions

12 The following APIs allow the RM daemon that hosts the PMIx server library to request specific services from
13 the PMIx library.

14 17.2.1 PMIx_generate_regex

15 Summary

16 Generate a compressed representation of the input string.

17 *PMIx v1.0* Format

```

18 pmix_status_t
19 PMIx_generate_regex(const char *input, char **output);

```

20 IN input

21 String to process (string)

22 OUT output

23 Compressed representation of *input* (array of bytes)

24 Returns **PMIX_SUCCESS** or a negative value indicating the error.

25 Description

26 Given a comma-separated list of *input* values, generate a reduced size representation of the input that can be
27 passed down to the PMIx server library's **PMIx_server_register_nspace** API for parsing. The order
28 of the individual values in the *input* string is preserved across the operation. The caller is responsible for
29 releasing the returned data.

30 The precise compressed representations will be implementation specific. The regular expression itself is not
31 required to be a printable string nor to obey typical string constraints (e.g., include a **NULL** terminator byte).
32 However, all PMIx implementations are required to include a colon-delimited **NULL**-terminated string at the
33 beginning of the output representation that can be printed for diagnostic purposes and identifies the method
34 used to generate the representation. The following identifiers are reserved by the PMIx Standard:

- 35 • "raw:\0" - indicates that the expression following the identifier is simply the comma-delimited input
36 string (no processing was performed).

- "**pmix:\0**" - a PMIx-unique regular expression represented as a **NULL**-terminated string following the identifier.
- "**blob:\0**" - a PMIx-unique regular expression that is not represented as a **NULL**-terminated string following the identifier. Additional implementation-specific metadata may follow the identifier along with the data itself. For example, a compressed binary array format based on the *zlib* compression package, with the size encoded in the space immediately following the identifier.

Communicating the resulting output should be done by first packing the returned expression using the [PMIx_Data_pack](#), declaring the input to be of type [PMIX_REGEX](#), and then obtaining the resulting blob to be communicated using the [PMIX_DATA_BUFFER_UNLOAD](#) macro. The reciprocal method can be used on the remote end prior to passing the regex into [PMIx_server_register_namespace](#). The pack/unpack routines will ensure proper handling of the data based on the regex prefix.

17.2.2 PMIx_generate_ppn

Summary

Generate a compressed representation of the input identifying the processes on each node.

Format

PMIx v1.0

```
pmix_status_t
PMIx_generate_ppn(const char *input, char **ppn);
```

IN input

String to process (string)

OUT ppn

Compressed representation of *input* (array of bytes)

Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Description

The input shall consist of a semicolon-separated list of ranges representing the ranks of processes on each node of the job - e.g., "1-4; 2-5; 8, 10, 11, 12; 6, 7, 9". Each field of the input must correspond to the node name provided at that position in the input to [PMIx_generate_regex](#). Thus, in the example, ranks 1-4 would be located on the first node of the comma-separated list of names provided to [PMIx_generate_regex](#), and ranks 2-5 would be on the second name in the list.

Rules governing the format of the returned regular expression are the same as those specified for [PMIx_generate_regex](#), as detailed [here](#).

17.2.3 PMIx_server_register_namespace

Summary

Setup the data about a particular namespace.

Format

C

```
pmix_status_t
PMIx_server_register_namespace(const pmix_namespace_t nspace,
                               int nlocalprocs,
                               pmix_info_t info[], size_t ninfo,
                               pmix_op_cbfunc_t cbfunc,
                               void *cbdata);
```

C

IN nspace

Character array of maximum size **PMIX_MAX_NSLEN** containing the namespace identifier (string)

IN nlocalprocs

number of local processes (integer)

IN info

Array of info structures (array of handles)

IN ninfo

Number of elements in the *info* array (integer)

IN cbfunc

Callback function **pmix_op_cbfunc_t** to be executed upon completion of the operation. A **NULL** function reference indicates that the function is to be executed as a blocking operation (function reference)

IN cbdata

Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

PMIX_REGISTER_NODATA "pmix.reg.nodata" (bool)

Registration is for this namespace only, do not copy job data.

PMIX_SESSION_INFO_ARRAY "pmix.ssn.arr" (pmix_data_array_t)

Provide an array of **pmix_info_t** containing session-realm information. The **PMIX_SESSION_ID** attribute is required to be included in the array.

PMIX_JOB_INFO_ARRAY "pmix.job.arr" (pmix_data_array_t)

1 Provide an array of `pmix_info_t` containing job-realm information. The `PMIX_SESSION_ID`
2 attribute of the *session* containing the *job* is required to be included in the array whenever the PMIx
3 server library may host multiple sessions (e.g., when executing with a host RM daemon). As
4 information is registered one job (aka namespace) at a time via the
5 `PMIx_server_register_namespace` API, there is no requirement that the array contain either the
6 `PMIX_NAMESPACE` or `PMIX_JOBID` attributes when used in that context (though either or both of them
7 may be included). At least one of the job identifiers must be provided in all other contexts where the
8 job being referenced is ambiguous.

9 **`PMIX_APP_INFO_ARRAY`** "pmix.app.arr" (`pmix_data_array_t`)

10 Provide an array of `pmix_info_t` containing application-realm information. The `PMIX_NAMESPACE`
11 or `PMIX_JOBID` attributes of the *job* containing the application, plus its `PMIX_APPNUM` attribute,
12 must be included in the array when the array is *not* included as part of a call to
13 `PMIx_server_register_namespace` - i.e., when the job containing the application is ambiguous.
14 The job identification is otherwise optional.

15 **`PMIX_PROC_INFO_ARRAY`** "pmix.pdata" (`pmix_data_array_t`)

16 Provide an array of `pmix_info_t` containing process-realm information. The `PMIX_RANK` and
17 `PMIX_NAMESPACE` attributes, or the `PMIX_PROCID` attribute, are required to be included in the array
18 when the array is not included as part of a call to `PMIx_server_register_namespace` - i.e., when
19 the job containing the process is ambiguous. All three may be included if desired. When the array is
20 included in some broader structure that identifies the job, then only the `PMIX_RANK` or the
21 `PMIX_PROCID` attribute must be included (the others are optional).

22 **`PMIX_NODE_INFO_ARRAY`** "pmix.node.arr" (`pmix_data_array_t`)

23 Provide an array of `pmix_info_t` containing node-realm information. At a minimum, either the
24 `PMIX_NODEID` or `PMIX_HOSTNAME` attribute is required to be included in the array, though both
25 may be included.

26
27 Host environments are required to provide a wide range of session-, job-, application-, node-, and
28 process-realm information, and may choose to provide a similarly wide range of optional information. The
29 information is broadly separated into categories based on the *data realm* definitions explained in Section 6.1,
30 and retrieved according to the rules detailed in Section 6.2.

31 Session-realm information may be passed as individual `pmix_info_t` entries, or as part of a
32 `pmix_data_array_t` using the `PMIX_SESSION_INFO_ARRAY` attribute. The list of data referenced in
33 this way shall include:

34 • **`PMIX_UNIV_SIZE`** "pmix.univ.size" (`uint32_t`)

35 Maximum number of process that can be simultaneously executing in a session. Note that this
36 attribute is equivalent to the `PMIX_MAX_PROCS` attribute for the *session* realm - it is included in
37 the PMIx Standard for historical reasons.

38 • **`PMIX_MAX_PROCS`** "pmix.max.size" (`uint32_t`)

39 Maximum number of processes that can be simultaneously executed in the specified realm.
40 Typically, this is a constraint imposed by a scheduler or by user settings in a hostfile or other
41 resource description. Defaults to the *job* realm. Must be provided if `PMIX_UNIV_SIZE` is not
42 given. Requires use of the `PMIX_SESSION_INFO` attribute to avoid ambiguity when retrieving it.

- 1 • **PMIX_SESSION_ID** "pmix.session.id" (uint32_t)

2 Session identifier assigned by the scheduler.

3 plus the following optional information:

- 4 • **PMIX_CLUSTER_ID** "pmix.clid" (char*)

5 A string name for the cluster this allocation is on. As this information is not related to the
6 namespace, it is best passed using the **PMIx_server_register_resources** API.

- 7 • **PMIX_ALLOCATED_NODELIST** "pmix.alist" (char*)

8 Comma-delimited list or regular expression of all nodes in the specified realm regardless of whether
9 or not they currently host processes. Defaults to the *job* realm.

- 10 • **PMIX_RM_NAME** "pmix.rm.name" (char*)

11 String name of the RM. As this information is not related to the namespace, it is best passed using
12 the **PMIx_server_register_resources** API.

- 13 • **PMIX_RM_VERSION** "pmix.rm.version" (char*)

14 RM version string. As this information is not related to the namespace, it is best passed using the
15 **PMIx_server_register_resources** API.

- 16 • **PMIX_SERVER_HOSTNAME** "pmix.srvr.host" (char*)

17 Host where target PMIx server is located. As this information is not related to the namespace, it is
18 best passed using the **PMIx_server_register_resources** API.

19 Job-realm information may be passed as individual **pmix_info_t** entries, or as part of a
20 **pmix_data_array_t** using the **PMIX_JOB_INFO_ARRAY** attribute. The list of data referenced in this
21 way shall include:

- 22 • **PMIX_SERVER_NAMESPACE** "pmix.srv.namespace" (char*)

23 Name of the namespace to use for this PMIx server. Identifies the namespace of the PMIx server
24 itself

- 25 • **PMIX_SERVER_RANK** "pmix.srv.rank" (pmix_rank_t)

26 Rank of this PMIx server. Identifies the rank of the PMIx server itself.

- 27 • **PMIX_NAMESPACE** "pmix.namespace" (char*)

28 Namespace of the job - may be a numerical value expressed as a string, but is often an alphanumeric
29 string carrying information solely of use to the system. Required to be unique within the scope of
30 the host environment. One cannot retrieve the namespace of an arbitrary process since that would
31 require already knowing the namespace of that process. However, a process' own namespace can be
32 retrieved by passing a NULL value of *proc* to **PMIx_Get**. Identifies the namespace of the job
33 being registered.

- 34 • **PMIX_JOBID** "pmix.jobid" (char*)

35 Job identifier assigned by the scheduler to the specified job - may be identical to the namespace, but
36 is often a numerical value expressed as a string (e.g., "12345.3").

- 37 • **PMIX_JOB_SIZE** "pmix.job.size" (uint32_t)

38 Total number of processes in the specified job across all contained applications. Note that this value
39 can be different from **PMIX_MAX_PROCS**. For example, users may choose to subdivide an
40 allocation (running several jobs in parallel within it), and dynamic programming models may

1 support adding and removing processes from a running *job* on-the-fly. In the latter case, PMIx
2 events may be used to notify processes within the job that the job size has changed.

- 3 ● **PMIX_MAX_PROCS** "pmix.max.size" (uint32_t)
4 Maximum number of processes that can be simultaneously executed in the specified realm.
5 Typically, this is a constraint imposed by a scheduler or by user settings in a hostfile or other
6 resource description. Defaults to the *job* realm. Retrieval of this attribute defaults to the job level
7 unless an appropriate specification is given (e.g., **PMIX_SESSION_INFO**).
- 8 ● **PMIX_NODE_MAP** "pmix.nmap" (char*)
9 Regular expression of nodes currently hosting processes in the specified realm - see 17.2.3.2 for an
10 explanation of its generation. Defaults to the *job* realm.
- 11 ● **PMIX_PROC_MAP** "pmix.pmap" (char*)
12 Regular expression describing processes on each node in the specified realm - see 17.2.3.2 for an
13 explanation of its generation. Defaults to the *job* realm.

14 plus the following optional information:

- 15 ● **PMIX_NPROC_OFFSET** "pmix.offset" (pmix_rank_t)
16 Starting global rank of the specified job. The returned value is the same as the value of
17 **PMIX_GLOBAL_RANK** of rank 0 of the specified job.
- 18 ● **PMIX_JOB_NUM_APPS** "pmix.job.napps" (uint32_t)
19 Number of applications in the specified job. This is a required attribute if more than one application
20 is included in the job.
- 21 ● **PMIX_MAPBY** "pmix.mapby" (char*)
22 Process mapping policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD**
23 value for the rank to discover the mapping policy used for the provided namespace. Supported
24 values are launcher specific.
- 25 ● **PMIX_RANKBY** "pmix.rankby" (char*)
26 Process ranking policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
27 for the rank to discover the ranking algorithm used for the provided namespace. Supported values
28 are launcher specific.
- 29 ● **PMIX_BINDTO** "pmix.bindto" (char*)
30 Process binding policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
31 for the rank to discover the binding policy used for the provided namespace. Supported values are
32 launcher specific.
- 33 ● **PMIX_HOSTNAME_KEEP_FQDN** "pmix.fqdn" (bool)
34 FQDNs are being retained by the PMIx library.
- 35 ● **PMIX_ANL_MAP** "pmix.anlmap" (char*)
36 Process map equivalent to **PMIX_PROC_MAP** expressed in Argonne National Laboratory's
37 PMI-1/PMI-2 notation. Defaults to the *job* realm.
- 38 ● **PMIX_TDIR_RMCLEAN** "pmix.tdir.rmclean" (bool)
39 The Resource Manager will remove any directories or files it creates in **PMIX_TMPDIR**.

- 1 • **PMIX_CRYPTO_KEY** "pmix.sec.key" (**pmix_byte_object_t**)

2 Blob containing crypto key.

3 If more than one application is included in the namespace, then the host environment is also required to supply
4 data consisting of the following items for each application in the job, passed as a **pmix_data_array_t**
5 using the **PMIX_APP_INFO_ARRAY** attribute:

- 6 • **PMIX_APPNUM** "pmix.appnum" (**uint32_t**)

7 The application number within the job in which the specified process is a member. This attribute
8 must appear at the beginning of the array.

- 9 • **PMIX_APP_SIZE** "pmix.app.size" (**uint32_t**)

10 Number of processes in the specified application, regardless of their execution state - i.e., this
11 number may include processes that either failed to start or have already terminated.

- 12 • **PMIX_MAX_PROCS** "pmix.max.size" (**uint32_t**)

13 Maximum number of processes that can be simultaneously executed in the specified realm.
14 Typically, this is a constraint imposed by a scheduler or by user settings in a hostfile or other
15 resource description. Defaults to the *job* realm. Requires use of the **PMIX_APP_INFO** attribute to
16 avoid ambiguity when retrieving it.

- 17 • **PMIX_APPLDR** "pmix.aldr" (**pmix_rank_t**)

18 Lowest rank in the specified application.

- 19 • **PMIX_WDIR** "pmix.wdir" (**char***)

20 Working directory for spawned processes. This attribute is required for all registrations, but may be
21 provided as an individual **pmix_info_t** entry if only one application is included in the
22 namespace.

- 23 • **PMIX_APP_ARGV** "pmix.app.argv" (**char***)

24 Consolidated argv passed to the spawn command for the given application (e.g., "/myapp arg1 arg2
25 arg3"). This attribute is required for all registrations, but may be provided as an individual
26 **pmix_info_t** entry if only one application is included in the namespace.

27 plus the following optional information:

- 28 • **PMIX_PSET_NAMES** "pmix.pset.nms" (**pmix_data_array_t***)

29 Returns an array of **char*** string names of the process sets in which the given process is a member.
30

- 31 • **PMIX_APP_MAP_TYPE** "pmix.apmap.type" (**char***)

32 Type of mapping used to layout the application (e.g., **cyclic**). This attribute may be provided as
33 an individual **pmix_info_t** entry if only one application is included in the namespace.

- 34 • **PMIX_APP_MAP_REGEX** "pmix.apmap.regex" (**char***)

35 Regular expression describing the result of the process mapping. This attribute may be provided as
36 an individual **pmix_info_t** entry if only one application is included in the namespace.

37 The data may also include attributes provided by the host environment that identify the programming model
38 (as specified by the user) being executed within the application. The PMIx server library may utilize this
39 information to customize the environment to fit that model (e.g., adding environmental variables specified by
40 the corresponding standard for that model):

- 1 • **PMIX_PROGRAMMING_MODEL** "pmix.pgm.model" (char*)
2 Programming model being initialized (e.g., "MPI" or "OpenMP").
- 3 • **PMIX_MODEL_LIBRARY_NAME** "pmix.mdl.name" (char*)
4 Programming model implementation ID (e.g., "OpenMPI" or "MPICH").
- 5 • **PMIX_MODEL_LIBRARY_VERSION** "pmix.mld.vrs" (char*)
6 Programming model version string (e.g., "2.1.1").

7 Node-realm information may be passed as individual **pmix_info_t** entries if only one node will host
8 processes from the job being registered, or as part of a **pmix_data_array_t** using the
9 **PMIX_NODE_INFO_ARRAY** attribute when multiple nodes are involved in the job. The list of data referenced
10 in this way shall include:

- 11 • **PMIX_NODEID** "pmix.nodeid" (uint32_t)
12 Node identifier expressed as the node's index (beginning at zero) in an array of nodes within the
13 active session. The value must be unique and directly correlate to the **PMIX_HOSTNAME** of the
14 node - i.e., users can interchangeably reference the same location using either the
15 **PMIX_HOSTNAME** or corresponding **PMIX_NODEID**.
- 16 • **PMIX_HOSTNAME** "pmix.hname" (char*)
17 Name of the host, as returned by the **gethostname** utility or its equivalent. As this information is
18 not related to the namespace, it can be passed using the
19 **PMIx_server_register_resources** API. However, either it or the **PMIX_NODEID** must
20 be included in the array to properly identify the node.
- 21 • **PMIX_HOSTNAME_ALIASES** "pmix.alias" (char*)
22 Comma-delimited list of names by which the target node is known. As this information is not
23 related to the namespace, it is best passed using the **PMIx_server_register_resources**
24 API.
- 25 • **PMIX_LOCAL_SIZE** "pmix.local.size" (uint32_t)
26 Number of processes in the specified job or application on the caller's node. Defaults to job unless
27 the **PMIX_APP_INFO** and the **PMIX_APPNUM** qualifiers are given.
- 28 • **PMIX_NODE_SIZE** "pmix.node.size" (uint32_t)
29 Number of processes across all jobs that are executing upon the node.
- 30 • **PMIX_LOCALLDR** "pmix.lldr" (pmix_rank_t)
31 Lowest rank within the specified job on the node (defaults to current node in absence of
32 **PMIX_HOSTNAME** or **PMIX_NODEID** qualifier).
- 33 • **PMIX_LOCAL_PEERS** "pmix.lpeers" (char*)
34 Comma-delimited list of ranks that are executing on the local node within the specified namespace –
35 shortcut for **PMIx_Resolve_peers** for the local node.
- 36 • **PMIX_NODE_OVERSUBSCRIBED** "pmix.ndosub" (bool)
37 True if the number of processes from this job on this node exceeds the number of slots allocated to it
38

39 plus the following information for the server's own node:

- 1 • **PMIX_TMPDIR** "pmix.tmpdir" (char*)
2 Full path to the top-level temporary directory assigned to the session.
- 3 • **PMIX_NSDIR** "pmix.nmdir" (char*)
4 Full path to the temporary directory assigned to the specified job, under **PMIX_TMPDIR**.
- 5 • **PMIX_LOCAL_PROCS** "pmix.lprocs" (pmix_proc_t array)
6 Array of **pmix_proc_t** of all processes executing on the local node – shortcut for
7 **PMIX_Resolve_peers** for the local node and a **NULL** namespace argument. The process
8 identifier is ignored for this attribute.

9 The data may also include the following optional information for the server's own node:

- 10 • **PMIX_LOCAL_CPUSSETS** "pmix.lcpus" (pmix_data_array_t)
11 A **pmix_data_array_t** array of string representations of the PU binding bitmaps applied to
12 each local *peer* on the caller's node upon launch. Each string shall begin with the name of the
13 library that generated it (e.g., "hwloc") followed by a colon and the bitmap string itself. The array
14 shall be in the same order as the processes returned by **PMIX_LOCAL_PEERS** for that namespace.
- 15 • **PMIX_AVAIL_PHYS_MEMORY** "pmix.pmem" (uint64_t)
16 Total available physical memory on a node. As this information is not related to the namespace, it
17 can be passed using the **PMIX_server_register_resources** API.

18 and the following optional information for other nodes:

- 19 • **PMIX_MAX_PROCS** "pmix.max.size" (uint32_t)
20 Maximum number of processes that can be simultaneously executed in the specified realm.
21 Typically, this is a constraint imposed by a scheduler or by user settings in a hostfile or other
22 resource description. Defaults to the *job* realm. Requires use of the **PMIX_NODE_INFO** attribute
23 to avoid ambiguity when retrieving it.

24 Process-realm information shall include the following data for each process in the job, passed as a
25 **pmix_data_array_t** using the **PMIX_PROC_INFO_ARRAY** attribute:

- 26 • **PMIX_RANK** "pmix.rank" (pmix_rank_t)
27 Process rank within the job, starting from zero.
- 28 • **PMIX_APPNUM** "pmix.appnum" (uint32_t)
29 The application number within the job in which the specified process is a member. This attribute
30 may be omitted if only one application is present in the namespace.
- 31 • **PMIX_APP_RANK** "pmix.apprank" (pmix_rank_t)
32 Rank of the specified process within its application. This attribute may be omitted if only one
33 application is present in the namespace.
- 34 • **PMIX_GLOBAL_RANK** "pmix.grank" (pmix_rank_t)
35 Rank of the specified process spanning across all jobs in this session, starting with zero. Note that
36 no ordering of the jobs is implied when computing this value. As jobs can start and end at random
37 times, this is defined as a continually growing number - i.e., it is not dynamically adjusted as
38 individual jobs and processes are started or terminated.
- 39 • **PMIX_LOCAL_RANK** "pmix.lrank" (uint16_t)

1 Rank of the specified process on its node - refers to the numerical location (starting from zero) of
2 the process on its node when counting only those processes from the same job that share the node,
3 ordered by their overall rank within that job.

- 4 • **PMIX_NODE_RANK** "pmix.nrank" (uint16_t)

5 Rank of the specified process on its node spanning all jobs- refers to the numerical location (starting
6 from zero) of the process on its node when counting all processes (regardless of job) that share the
7 node, ordered by their overall rank within the job. The value represents a snapshot in time when the
8 specified process was started on its node and is not dynamically adjusted as processes from other
9 jobs are started or terminated on the node.

- 10 • **PMIX_NODEID** "pmix.nodeid" (uint32_t)

11 Node identifier expressed as the node's index (beginning at zero) in an array of nodes within the
12 active session. The value must be unique and directly correlate to the **PMIX_HOSTNAME** of the
13 node - i.e., users can interchangeably reference the same location using either the
14 **PMIX_HOSTNAME** or corresponding **PMIX_NODEID**.

- 15 • **PMIX_REINCARNATION** "pmix.reinc" (uint32_t)

16 Number of times this process has been re-instantiated - i.e, a value of zero indicates that the process
17 has never been restarted.

- 18 • **PMIX_SPAWNED** "pmix.spawned" (bool)

19 **true** if this process resulted from a call to **PMIx_Spawn**. Lack of inclusion (i.e., a return status of
20 **PMIX_ERR_NOT_FOUND**) corresponds to a value of **false** for this attribute.

21 plus the following information for processes that are local to the server:

- 22 • **PMIX_LOCALITY_STRING** "pmix.locstr" (char*)

23 String describing a process's bound location - referenced using the process's rank. The string is
24 prefixed by the implementation that created it (e.g., "hwloc") followed by a colon. The remainder of
25 the string represents the corresponding locality as expressed by the underlying implementation. The
26 entire string must be passed to **PMIx_Get_relative_locality** for processing. Note that
27 hosts are only required to provide locality strings for local client processes - thus, a call to
28 **PMIx_Get** for the locality string of a process that returns **PMIX_ERR_NOT_FOUND** indicates that
29 the process is not executing on the same node.

- 30 • **PMIX_PROCDIR** "pmix.pdir" (char*)

31 Full path to the subdirectory under **PMIX_NSDIR** assigned to the specified process.

- 32 • **PMIX_PACKAGE_RANK** "pmix.pkgrank" (uint16_t)

33 Rank of the specified process on the *package* where this process resides - refers to the numerical
34 location (starting from zero) of the process on its package when counting only those processes from
35 the same job that share the package, ordered by their overall rank within that job. Note that
36 processes that are not bound to PUs within a single specific package cannot have a package rank.

37 and the following optional information - note that some of this information can be derived from information
38 already provided by other attributes, but it may be included here for ease of retrieval by users:

- 39 • **PMIX_HOSTNAME** "pmix.hname" (char*)

40 Name of the host, as returned by the **gethostname** utility or its equivalent.

- 41 • **PMIX_CPuset** "pmix.cpuset" (char*)

1 A string representation of the PU binding bitmap applied to the process upon launch. The string
2 shall begin with the name of the library that generated it (e.g., "hwloc") followed by a colon and the
3 bitmap string itself.

- 4 • **PMIX_CPUSSET_BITMAP** "pmix.bitmap" (`pmix_cpuset_t*`)
5 Bitmap applied to the process upon launch.
- 6 • **PMIX_DEVICE_DISTANCES** "pmix.dev.dist" (`pmix_data_array_t`)
7 Return an array of `pmix_device_distance_t` containing the minimum and maximum
8 distances of the given process location to all devices of the specified type on the local node.

9
10 Attributes not directly provided by the host environment may be derived by the PMIx server library from other
11 required information and included in the data made available to the server library's clients.

12 **Description**

13 Pass job-related information to the PMIx server library for distribution to local client processes.

14 **Advice to PMIx server hosts**

15 Host environments are required to execute this operation prior to starting any local application process within
16 the given namespace.

17 The PMIx server must register all namespaces that will participate in collective operations with local
18 processes. This means that the server must register a namespace even if it will not host any local processes
19 from within that namespace if any local process of another namespace might at some point perform an
20 operation involving one or more processes from the new namespace. This is necessary so that the collective
operation can identify the participants and know when it is locally complete.

21 The caller must also provide the number of local processes that will be launched within this namespace. This
22 is required for the PMIx server library to correctly handle collectives as a collective operation call can occur
23 before all the local processes have been started.

24 A **NULL** *cbfunc* reference indicates that the function is to be executed as a blocking operation.

25 **Advice to users**

26 The number of local processes for any given namespace is generally fixed at the time of application launch.
27 Calls to **PMIx_Spawn** result in processes launched in their own namespace, not that of their parent. However,
28 it is possible for processes to *migrate* to another node via a call to **PMIx_Job_control_nb**, thus resulting
29 in a change to the number of local processes on both the initial node and the node to which the process moved.
30 It is therefore critical that applications not migrate processes without first ensuring that PMIx-based collective
operations are not in progress, and that no such operations be initiated until process migration has completed.

17.2.3.1 Namespace registration attributes

The following attributes are defined specifically for use with the `PMIx_server_register_nspace` API:

`PMIX_REGISTER_NODATA` "pmix.reg.nodata" (bool)

Registration is for this namespace only, do not copy job data.

The following attributes are used to assemble information according to its data realm (*session*, *job*, *application*, *node*, or *process* as defined in Section 6.1) for registration where ambiguity may exist - see 17.2.3.2 for examples of their use.

`PMIX_SESSION_INFO_ARRAY` "pmix.ssn.arr" (pmix_data_array_t)

Provide an array of `pmix_info_t` containing session-realm information. The `PMIX_SESSION_ID` attribute is required to be included in the array.

`PMIX_JOB_INFO_ARRAY` "pmix.job.arr" (pmix_data_array_t)

Provide an array of `pmix_info_t` containing job-realm information. The `PMIX_SESSION_ID` attribute of the *session* containing the *job* is required to be included in the array whenever the PMIx server library may host multiple sessions (e.g., when executing with a host RM daemon). As information is registered one job (aka namespace) at a time via the `PMIx_server_register_nspace` API, there is no requirement that the array contain either the `PMIX_NAMESPACE` or `PMIX_JOBID` attributes when used in that context (though either or both of them may be included). At least one of the job identifiers must be provided in all other contexts where the job being referenced is ambiguous.

`PMIX_APP_INFO_ARRAY` "pmix.app.arr" (pmix_data_array_t)

Provide an array of `pmix_info_t` containing application-realm information. The `PMIX_NAMESPACE` or `PMIX_JOBID` attributes of the *job* containing the application, plus its `PMIX_APPNUM` attribute, must to be included in the array when the array is *not* included as part of a call to `PMIx_server_register_nspace` - i.e., when the job containing the application is ambiguous. The job identification is otherwise optional.

`PMIX_PROC_INFO_ARRAY` "pmix.pdata" (pmix_data_array_t)

Provide an array of `pmix_info_t` containing process-realm information. The `PMIX_RANK` and `PMIX_NAMESPACE` attributes, or the `PMIX_PROCID` attribute, are required to be included in the array when the array is not included as part of a call to `PMIx_server_register_nspace` - i.e., when the job containing the process is ambiguous. All three may be included if desired. When the array is included in some broader structure that identifies the job, then only the `PMIX_RANK` or the `PMIX_PROCID` attribute must be included (the others are optional).

`PMIX_NODE_INFO_ARRAY` "pmix.node.arr" (pmix_data_array_t)

Provide an array of `pmix_info_t` containing node-realm information. At a minimum, either the `PMIX_NODEID` or `PMIX_HOSTNAME` attribute is required to be included in the array, though both may be included.

Note that these assemblages can be used hierarchically:

- a `PMIX_JOB_INFO_ARRAY` might contain multiple `PMIX_APP_INFO_ARRAY` elements, each describing values for a specific application within the job.
- a `PMIX_JOB_INFO_ARRAY` could contain a `PMIX_NODE_INFO_ARRAY` for each node hosting processes from that job, each array describing job-level values for that node.
- a `PMIX_SESSION_INFO_ARRAY` might contain multiple `PMIX_JOB_INFO_ARRAY` elements, each describing a job executing within the session. Each job array could, in turn, contain both application and

node arrays, thus providing a complete picture of the active operations within the allocation.

Advice to PMIx library implementers

PMIx implementations must be capable of properly parsing and storing any hierarchical depth of information arrays. The resulting stored values are must to be accessible via both `PMIx_Get` and `PMIx_Query_info_nb` APIs, assuming appropriate directives are provided by the caller.

17.2.3.2 Assembling the registration information

The following description is not intended to represent the actual layout of information in a given PMIx library. Instead, it describes how information provided in the *info* parameter of the `PMIx_server_register_namespace` shall be organized for proper processing by a PMIx server library. The ordering of the various information elements is arbitrary - they are presented in a top-down hierarchical form solely for clarity in reading.

Advice to PMIx server hosts

Creating the *info* array of data requires knowing in advance the number of elements required for the array. This can be difficult to compute and somewhat fragile in practice. One method for resolving the problem is to create a linked list of objects, each containing a single `pmix_info_t` structure. Allocation and manipulation of the list can then be accomplished using existing standard methods. Upon completion, the final *info* array can be allocated based on the number of elements on the list, and then the values in the list object `pmix_info_t` structures transferred to the corresponding array element utilizing the `PMIx_Info_xfer` API.

A common building block used in several areas is the construction of a regular expression identifying the nodes involved in that area - e.g., the nodes in a *session* or *job*. PMIx provides several tools to facilitate this operation, beginning by constructing an argv-like array of node names. This array is then passed to the `PMIx_generate_regex` function to create a regular expression parseable by the PMIx server library, as shown below:

```

1  char **nodes = NULL;
2  char *nodelist;
3  char *regex;
4  size_t n;
5  pmix_status_t rc;
6  pmix_info_t info;
7
8  /* loop over an array of nodes, adding each
9   * name to the array */
10 for (n=0; n < num_nodes; n++) {
11     /* filter the nodes to ignore those not included
12      * in the target range (session, job, etc.). In
13      * this example, all nodes are accepted */
14     PMIX_ARGV_APPEND(&nodes, node[n]->name);
15 }
16
17 /* join into a comma-delimited string */
18 nodelist = PMIX_ARGV_JOIN(nodes, ',');
19
20 /* release the array */
21 PMIX_ARGV_FREE(nodes);
22
23 /* generate regex */
24 rc = PMIx_generate_regex(nodelist, &regex);
25
26 /* release list */
27 free(nodelist);
28
29 /* pass the regex as the value to the PMIX_NODE_MAP key */
30 PMIx_Info_load(&info, PMIX_NODE_MAP, regex, PMIX_REGEX);
31 /* release the regex */
32 free(regex);

```

33 Changing the filter criteria allows the construction of node maps for any level of information. A description of
34 the returned regular expression is provided [here](#).

35 A similar method is used to construct the map of processes on each node from the namespace being registered.
36 This may be done for each information level of interest (e.g., to identify the process map for the entire *job* or
37 for each *application* in the job) by changing the search criteria. An example is shown below for the case of
38 creating the process map for a *job*:

```

1  char **ndppn;
2  char rank[30];
3  char **ppnarray = NULL;
4  char *ppn;
5  char *localranks;
6  char *regex;
7  size_t n, m;
8  pmix_status_t rc;
9  pmix_info_t info;
10
11  /* loop over an array of nodes */
12  for (n=0; n < num_nodes; n++) {
13      /* for each node, construct an array of ranks on that node */
14      ndppn = NULL;
15      for (m=0; m < node[n]->num_procs; m++) {
16          /* ignore processes that are not part of the target job */
17          if (!PMIX_CHECK_NAMESPACE(targetjob,node[n]->proc[m].nspace)) {
18              continue;
19          }
20          snprintf(rank, 30, "%d", node[n]->proc[m].rank);
21          PMIX_ARGV_APPEND(&ndppn, rank);
22      }
23      /* convert the array into a comma-delimited string of ranks */
24      localranks = PMIX_ARGV_JOIN(ndppn, ',');
25      /* release the local array */
26      PMIX_ARGV_FREE(ndppn);
27      /* add this node's contribution to the overall array */
28      PMIX_ARGV_APPEND(&ppnarray, localranks);
29      /* release the local list */
30      free(localranks);
31  }
32
33  /* join into a semicolon-delimited string */
34  ppn = PMIX_ARGV_JOIN(ppnarray, ';');
35
36  /* release the array */
37  PMIX_ARGV_FREE(ppnarray);
38
39  /* generate ppn regex */
40  rc = PMIx_generate_ppn(ppn, &regex);
41
42  /* release list */
43  free(ppn);
44
45  /* pass the regex as the value to the PMIX_PROC_MAP key */
46  PMIx_Info_load(&info, PMIX_PROC_MAP, regex, PMIX_REGEX);

```

```

1  /* release the regex */
2  free(regex);

```



3 Note that the **PMIX_NODE_MAP** and **PMIX_PROC_MAP** attributes are linked in that the order of entries in the
4 process map must match the ordering of nodes in the node map - i.e., there is no provision in the PMIx process
5 map regular expression generator/parser pair supporting an out-of-order node or a node that has no
6 corresponding process map entry (e.g., a node with no processes on it). Armed with these tools, the
7 registration *info* array can be constructed as follows:

- 8 • Session-level information includes all session-specific values. In many cases, only two values
9 (**PMIX_SESSION_ID** and **PMIX_UNIV_SIZE**) are included in the registration array. Since both of these
10 values are session-specific, they can be specified independently - i.e., in their own **pmix_info_t** elements
11 of the *info* array. Alternatively, they can be provided as a **pmix_data_array_t** array of **pmix_info_t**
12 using the **PMIX_SESSION_INFO_ARRAY** attribute and identified by including the **PMIX_SESSION_ID**
13 attribute in the array - this is required in cases where non-specific attributes (e.g., **PMIX_NUM_NODES** or
14 **PMIX_NODE_MAP**) are passed to describe aspects of the session. Note that the node map can include
15 nodes not used by the job being registered as no corresponding process map is specified.

16 The *info* array at this point might look like (where the labels identify the corresponding attribute - e.g.,
17 “Session ID” corresponds to the **PMIX_SESSION_ID** attribute):

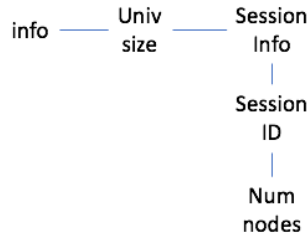


Figure 17.1.: Session-level information elements

- 18 • Job-level information includes all job-specific values such as **PMIX_JOB_SIZE**, **PMIX_JOB_NUM_APPS**,
19 and **PMIX_JOBID**. Since each invocation of **PMIx_server_register_namespace** describes a single
20 *job*, job-specific values can be specified independently - i.e., in their own **pmix_info_t** elements of the
21 *info* array. Alternatively, they can be provided as a **pmix_data_array_t** array of **pmix_info_t**
22 identified by the **PMIX_JOB_INFO_ARRAY** attribute - this is required in cases where non-specific
23 attributes (e.g., **PMIX_NODE_MAP**) are passed to describe aspects of the job. Note that since the invocation
24 only involves a single namespace, there is no need to include the **PMIX_NAMESPACE** attribute in the array.

25 Upon conclusion of this step, the *info* array might look like:

26 Note that in this example, **PMIX_NUM_NODES** is not required as that information is contained in the
27 **PMIX_NODE_MAP** attribute. Similarly, **PMIX_JOB_SIZE** is not technically required as that information
28 is contained in the **PMIX_PROC_MAP** when combined with the corresponding node map - however, there is
29 no issue with including the job size as a separate entry.

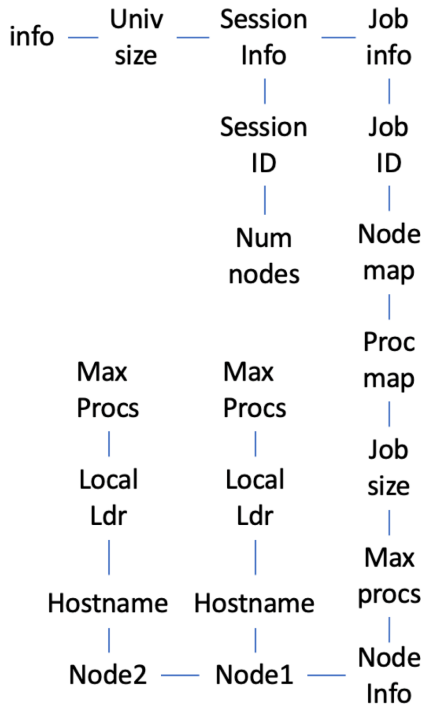


Figure 17.2.: Job-level information elements

The example also illustrates the hierarchical use of the `PMIX_NODE_INFO_ARRAY` attribute. In this case, we have chosen to pass several job-related values for each node - since those values are non-unique across the job, they must be passed in a node-info container. Note that the choice of what information to pass into the PMIx server library versus what information to derive from other values at time of request is left to the host environment. PMIx implementors in turn may, if they choose, pre-parse registration data to create expanded views (thus enabling faster response to requests at the expense of memory footprint) or to compress views into tighter representations (thus trading minimized footprint for longer response times).

- Application-level information includes all application-specific values such as `PMIX_APP_SIZE` and `PMIX_APPLDR`. If the *job* contains only a single *application*, then the application-specific values can be specified independently - i.e., in their own `pmix_info_t` elements of the *info* array - or as a `pmix_data_array_t` array of `pmix_info_t` using the `PMIX_APP_INFO_ARRAY` attribute and identified by including the `PMIX_APPNUM` attribute in the array. Use of the array format is must in cases where non-specific attributes (e.g., `PMIX_NODE_MAP`) are passed to describe aspects of the application.

However, in the case of a job consisting of multiple applications, all application-specific values for each application must be provided using the `PMIX_APP_INFO_ARRAY` format, each identified by its `PMIX_APPNUM` value.

Upon conclusion of this step, the *info* array might look like that shown in 17.3, assuming there are two

1 applications in the job being registered:

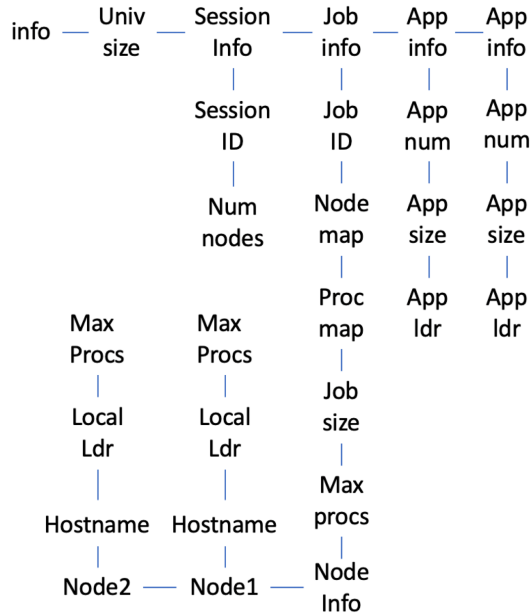


Figure 17.3.: Application-level information elements

- 2
- 3
- 4
- 5
- Process-level information includes an entry for each process in the job being registered, each entry marked with the `PMIX_PROC_INFO_ARRAY` attribute. The *rank* of the process must be the first entry in the array - this provides efficiency when storing the data. Upon conclusion of this step, the *info* array might look like the diagram in 17.4:
- 6
- For purposes of this example, node-level information only includes values describing the local node - i.e., it does not include information about other nodes in the job or session. In many cases, the values included in this level are unique to it and can be specified independently - i.e., in their own `pmix_info_t` elements of the *info* array. Alternatively, they can be provided as a `pmix_data_array_t` array of `pmix_info_t` using the `PMIX_NODE_INFO_ARRAY` attribute - this is required in cases where non-specific attributes are passed to describe aspects of the node, or where values for multiple nodes are being provided.
- 7
- 8
- 9
- 10
- 11

12 The node-level information requires two elements that must be constructed in a manner similar to that used

13 for the node map. The `PMIX_LOCAL_PEERS` value is computed based on the processes on the local node,

14 filtered to select those from the job being registered, as shown below using the tools provided by PMIx:

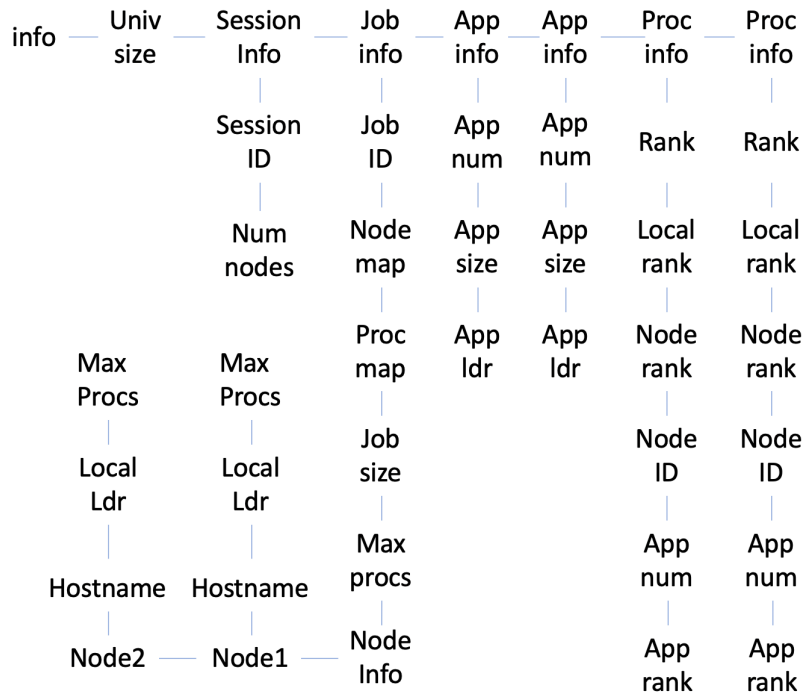


Figure 17.4.: Process-level information elements

C

```

1  char **ndppn = NULL;
2  char rank[30];
3  char *localranks;
4  size_t m;
5  pmix_info_t info;
6
7  for (m=0; m < mynode->num_procs; m++) {
8      /* ignore processes that are not part of the target job */
9      if (!PMIX_CHECK_NAMESPACE(targetjob,mynode->proc[m].namespace)) {
10         continue;
11     }
12     snprintf(rank, 30, "%d", mynode->proc[m].rank);
13     PMIX_ARGV_APPEND(&ndppn, rank);
14 }
15 /* convert the array into a comma-delimited string of ranks */
16 localranks = PMIX_ARGV_JOIN(ndppn, ',');

```

```

1      /* release the local array */
2      PMIX_ARGV_FREE(ndppn);
3
4      /* pass the string as the value to the PMIX_LOCAL_PEERS key */
5      PMIx_Info_load(&info, PMIX_LOCAL_PEERS, localranks, PMIX_STRING);
6
7      /* release the list */
8      free(localranks);

```

C

9 The **PMIX_LOCAL_CPUSSETS** value is constructed in a similar manner. In the provided example, it is assumed that an Hardware Locality (HWLOC) cpuset representation (a comma-delimited string of processor IDs) of the processors assigned to each process has previously been generated and stored on the process description. Thus, the value can be constructed as shown below:

```

13     char **ndcpus = NULL;
14     char *localcpus;
15     size_t m;
16     pmix_info_t info;
17
18     for (m=0; m < mynode->num_procs; m++) {
19         /* ignore processes that are not part of the target job */
20         if (!PMIX_CHECK_NAMESPACE(targetjob, mynode->proc[m].nspace)) {
21             continue;
22         }
23         PMIX_ARGV_APPEND(&ndcpus, mynode->proc[m].cpuset);
24     }
25     /* convert the array into a colon-delimited string */
26     localcpus = PMIX_ARGV_JOIN(ndcpus, ':');
27     /* release the local array */
28     PMIX_ARGV_FREE(ndcpus);
29
30     /* pass the string as the value to the PMIX_LOCAL_CPUSSETS key */
31     PMIx_Info_load(&info, PMIX_LOCAL_CPUSSETS, localcpus, PMIX_STRING);
32
33     /* release the list */
34     free(localcpus);

```

C

35 Note that for efficiency, these two values can be computed at the same time.

36 The final *info* array might therefore look like the diagram in 17.5:

37 17.2.4 PMIx_server_deregister_namespace

38 Summary

39 Deregister a namespace.

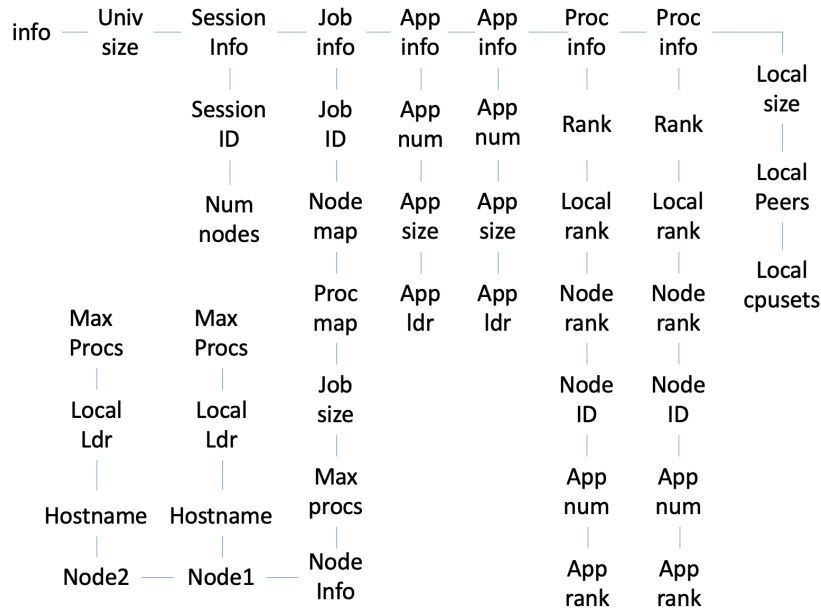


Figure 17.5.: Final information array

Format

```

1 void PMIx_server_deregister_namespace(const pmix_namespace_t nspace,
2                                     pmix_op_cbfunc_t cbfunc, void *cbdata);
3

```

IN nspace

Namespace (string)

IN cbfunc

Callback function [pmix_op_cbfunc_t](#). A **NULL** function reference indicates that the function is to be executed as a blocking operation. (function reference)

IN cbdata

Data to be passed to the callback function (memory reference)

Description

Deregister the specified *nspace* and purge all objects relating to it, including any client information from that namespace. This is intended to support persistent PMIx servers by providing an opportunity for the host RM to tell the PMIx server library to release all memory for a completed job. Note that the library must not invoke the callback function prior to returning from the API, and that a **NULL** *cbfunc* reference indicates that the function is to be executed as a blocking operation.

1 17.2.5 PMIx_server_register_resources

2 Summary

3 Register non-namespace related information with the local PMIx server library.

4 *PMIx v4.0* Format

C

```
5 pmix_status_t
6 PMIx_server_register_resources(pmix_info_t info[], size_t ninfo,
7                               pmix_op_cbfunc_t cbfunc,
8                               void *cbdata);
```

C

9 **IN** *info*

10 Array of info structures (array of handles)

11 **IN** *ninfo*

12 Number of elements in the *info* array (integer)

13 **IN** *cbfunc*

14 Callback function [pmix_op_cbfunc_t](#). A **NULL** function reference indicates that the function is to
15 be executed as a blocking operation (function reference)

16 **IN** *cbdata*

17 Data to be passed to the callback function (memory reference)

18 Description

19 Pass information about resources not associated with a given namespace to the PMIx server library for
20 distribution to local client processes. This includes information on fabric devices, GPUs, and other resources.
21 All information provided through this API shall be made available to each job as part of its job-level
22 information. Duplicate information provided with the [PMIx_server_register_namespace](#) API shall
23 override any information provided by this function for that namespace, but only for that specific namespace.

24 Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Advice to PMIx server hosts

25 Note that information passed in this manner could also have been included in a call to
26 [PMIx_server_register_namespace](#) - e.g., as part of a [PMIX_NODE_INFO_ARRAY](#) array. This API is
27 provided as a logical alternative for code clarity, especially where multiple jobs may be supported by a single
28 PMIx server library instance, to avoid multiple registration of static resource information.

29 A **NULL** *cbfunc* reference indicates that the function is to be executed as a blocking operation.

30 17.2.6 PMIx_server_deregister_resources

31 Summary

32 Remove specified non-namespace related information from the local PMIx server library.

Format

C

```
pmix_status_t
PMIx_server_deregister_resources(pmix_info_t info[], size_t ninfo,
                                pmix_op_cbfunc_t cbfunc,
                                void *cbdata);
```

C

- IN** **info**
Array of info structures (array of handles)
- IN** **ninfo**
Number of elements in the *info* array (integer)
- IN** **cbfunc**
Callback function [pmix_op_cbfunc_t](#). A **NULL** function reference indicates that the function is to be executed as a blocking operation (function reference)
- IN** **cbdata**
Data to be passed to the callback function (memory reference)

Description

Remove information about resources not associated with a given namespace from the PMIx server library. Only the *key* fields of the provided *info* array shall be used for the operation - the associated values shall be ignored except where they serve as qualifiers to the request. For example, to remove a specific fabric device from a given node, the *info* array might include a [PMIX_NODE_INFO_ARRAY](#) containing the [PMIX_NODEID](#) or [PMIX_HOSTNAME](#) identifying the node hosting the device, and the [PMIX_FABRIC_DEVICE_NAME](#) specifying the device to be removed. Alternatively, the device could be removed using only the [PMIX_DEVICE_ID](#) as this is unique across the overall system.

Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Advice to PMIx server hosts

As information not related to namespaces is considered *static*, there is no requirement that the host environment deregister resources prior to finalizing the PMIx server library. The server library shall properly cleanup as part of its normal finalize operations. Deregistration of resources is only required, therefore, when the host environment determines that client processes should no longer have access to that information.

A **NULL** *cbfunc* reference indicates that the function is to be executed as a blocking operation.

17.2.7 PMIx_server_register_client

Summary

Register a client process with the PMIx server library.

Format

C

```
1 pmix_status_t  
2 PMIx_server_register_client(const pmix_proc_t *proc,  
3                               uid_t uid, gid_t gid,  
4                               void *server_object,  
5                               pmix_op_cbfunc_t cbfunc, void *cbdata);  
6
```

C

7 **IN** **proc**
8 [pmix_proc_t](#) structure (handle)
9 **IN** **uid**
10 user id (integer)
11 **IN** **gid**
12 group id (integer)
13 **IN** **server_object**
14 (memory reference)
15 **IN** **cbfunc**
16 Callback function [pmix_op_cbfunc_t](#). A **NULL** function reference indicates that the function is to
17 be executed as a blocking operation (function reference)
18 **IN** **cbdata**
19 Data to be passed to the callback function (memory reference)

20 A successful return indicates that the request is being processed and the result will be returned in the provided
21 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
22 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

23 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- 24 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
25 *success* - the *cbfunc* will not be called

26 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
27 error code or an implementation defined error code as described in Section 3.1.1.

Description

28 Register a client process with the PMIx server library.

29
30 The host server can also, if it desires, provide an object it wishes to be returned when a server function is called
31 that relates to a specific process. For example, the host server may have an object that tracks the specific client.
32 Passing the object to the library allows the library to provide that object to the host server during subsequent
33 calls related to that client, such as a [pmix_server_client_connected2_fn_t](#) function. This allows
34 the host server to access the object without performing a lookup based on the client's namespace and rank.

Advice to PMIx server hosts

35 Host environments are required to execute this operation prior to starting the client process. The expected user
36 ID and group ID of the child process allows the server library to properly authenticate clients as they connect
37 by requiring the two values to match. Accordingly, the detected user and group ID's of the connecting process
38 are not included in the [pmix_server_client_connected2_fn_t](#) server module function.

Advice to PMIx library implementers

For security purposes, the PMIx server library should check the user and group ID's of a connecting process against those provided for the declared client process identifier via the `PMIx_server_register_client` prior to completing the connection.

17.2.8 `PMIx_server_deregister_client`

Summary

Deregister a client and purge all data relating to it.

Format

PMIx v1.0

C

```
void
PMIx_server_deregister_client(const pmix_proc_t *proc,
                             pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

IN `proc`

`pmix_proc_t` structure (handle)

IN `cbfunc`

Callback function `pmix_op_cbfunc_t`. A `NULL` function reference indicates that the function is to be executed as a blocking operation (function reference)

IN `cbdata`

Data to be passed to the callback function (memory reference)

Description

The `PMIx_server_deregister_namespace` API will delete all client information for that namespace. The PMIx server library will automatically perform that operation upon disconnect of all local clients. This API is therefore intended primarily for use in exception cases, but can be called in non-exception cases if desired. Note that the library must not invoke the callback function prior to returning from the API.

17.2.9 `PMIx_server_setup_fork`

Summary

Setup the environment of a child process to be forked by the host.

Format

C

```
pmix_status_t
PMIx_server_setup_fork(const pmix_proc_t *proc,
                      char ***env);
```

C

IN `proc`
 [pmix_proc_t](#) structure (handle)

IN `env`
 Environment array (array of strings)

Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Description

Setup the environment of a child process to be forked by the host so it can correctly interact with the PMIx server.

The PMIx client needs some setup information so it can properly connect back to the server. This function will set appropriate environmental variables for this purpose, and will also provide any environmental variables that were specified in the launch command (e.g., via [PMIx_Spawn](#)) plus other values (e.g., variables required to properly initialize the client's fabric library).

Advice to PMIx server hosts

Host environments are required to execute this operation prior to starting the client process.

17.2.10 PMIx_server_dmodex_request

Summary

Define a function by which the host server can request modex data from the local PMIx server.

Format

C

PMIx v1.0

```
pmix_status_t
PMIx_server_dmodex_request(const pmix_proc_t *proc,
                          pmix_dmodex_response_fn_t cbfunc,
                          void *cbdata);
```

C

IN `proc`
 [pmix_proc_t](#) structure (handle)

IN `cbfunc`
 Callback function [pmix_dmodex_response_fn_t](#) (function reference)

IN `cbdata`
 Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when [PMIX_SUCCESS](#) is returned.

Description

Define a function by which the host server can request modex data from the local PMIx server. Traditional wireup procedures revolve around the per-process posting of data (e.g., location and endpoint information) via the `PMIx_Put` and `PMIx_Commit` functions followed by a `PMIx_Fence` barrier that globally exchanges the posted information. However, the barrier operation represents a significant time impact at large scale.

PMIx supports an alternative wireup method known as *Direct Modex* that replaces the barrier-based exchange of all process-posted information with on-demand fetch of a peer's data. In place of the barrier operation, data posted by each process is cached on the local PMIx server. When a process requests the information posted by a particular peer, it first checks the local cache to see if the data is already available. If not, then the request is passed to the local PMIx server, which subsequently requests that its RM host request the data from the RM daemon on the node where the specified peer process is located. Upon receiving the request, the RM daemon passes the request into its PMIx server library using the `PMIx_server_dmodex_request` function, receiving the response in the provided `cbfunc` once the indicated process has posted its information. The RM daemon then returns the data to the requesting daemon, who subsequently passes the data to its PMIx server library for transfer to the requesting client.

Advice to users

While direct modex allows for faster launch times by eliminating the barrier operation, per-peer retrieval of posted information is less efficient. Optimizations can be implemented - e.g., by returning posted information from all processes on a node upon first request - but in general direct modex remains best suited for sparsely connected applications.

17.2.10.1 Server Direct Modex Response Callback Function

The `PMIx_server_dmodex_request` callback function.

Summary

Provide a function by which the local PMIx server library can return connection and other data posted by local application processes to the host resource manager.

Format

PMIx v1.0

```
typedef void (*pmix_dmodex_response_fn_t) (  
    pmix_status_t status,  
    char *data, size_t sz,  
    void *cbdata);
```

IN status

Returned status of the request (`pmix_status_t`)

IN data

Pointer to a data "blob" containing the requested information (handle)

IN sz

Number of bytes in the *data* blob (integer)

IN cbdata

Data passed into the initial call to `PMIx_server_dmodex_request` (memory reference)

Description

Define a function to be called by the PMIx server library for return of information posted by a local application process (via `PMIx_Put` with subsequent `PMIx_Commit`) in response to a request from the host RM. The returned *data* blob is owned by the PMIx server library and will be free'd upon return from the function.

17.2.11 PMIx_server_setup_application

Summary

Provide a function by which a launcher can request application-specific setup data prior to launch of a *job*.

Format

PMIx v2.0

C

```
pmix_status_t
PMIx_server_setup_application(const pmix_namespace_t nspace,
                             pmix_info_t info[], size_t ninfo,
                             pmix_setup_application_cbfunc_t cbfunc,
                             void *cbdata);
```

C

IN nspace
namespace (string)

IN info
Array of info structures (array of handles)

IN ninfo
Number of elements in the *info* array (integer)

IN cbfunc
Callback function `pmix_setup_application_cbfunc_t` (function reference)

IN cbdata
Data to be passed to the *cbfunc* callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when `PMIX_SUCCESS` is returned.

Required Attributes

PMIx libraries that support this operation are required to support the following:

`PMIX_SETUP_APP_ENVARS` "pmix.setup.env" (bool)
Harvest and include relevant environmental variables.

`PMIX_SETUP_APP_NONENVARS` "pmix.setup.nenv" (bool)
Include all relevant data other than environmental variables.

`PMIX_SETUP_APP_ALL` "pmix.setup.all" (bool)
Include all relevant data.

`PMIX_ALLOC_FABRIC` "pmix.alloc.net" (array)

1 Array of `pmix_info_t` describing requested fabric resources. This must include at least:
2 `PMIX_ALLOC_FABRIC_ID`, `PMIX_ALLOC_FABRIC_TYPE`, and
3 `PMIX_ALLOC_FABRIC_ENDPTS`, plus whatever other descriptors are desired.

4 `PMIX_ALLOC_FABRIC_ID` "pmix.alloc.netid" (char*)

5 The key to be used when accessing this requested fabric allocation. The fabric allocation will be
6 returned/stored as a `pmix_data_array_t` of `pmix_info_t` whose first element is composed of
7 this key and the allocated resource description. The type of the included value depends upon the fabric
8 support. For example, a TCP allocation might consist of a comma-delimited string of socket ranges
9 such as "32000–32100, 33005, 38123–38146". Additional array entries will consist of any
10 provided resource request directives, along with their assigned values. Examples include:
11 `PMIX_ALLOC_FABRIC_TYPE` - the type of resources provided; `PMIX_ALLOC_FABRIC_PLANE` -
12 if applicable, what plane the resources were assigned from; `PMIX_ALLOC_FABRIC_QOS` - the
13 assigned QoS; `PMIX_ALLOC_BANDWIDTH` - the allocated bandwidth;
14 `PMIX_ALLOC_FABRIC_SEC_KEY` - a security key for the requested fabric allocation. NOTE: the
15 array contents may differ from those requested, especially if `PMIX_INFO_REQD` was not set in the
16 request.

17 `PMIX_ALLOC_FABRIC_SEC_KEY` "pmix.alloc.nsec" (pmix_byte_object_t)

18 Request that the allocation include a fabric security key for the spawned job.

19 `PMIX_ALLOC_FABRIC_TYPE` "pmix.alloc.nettype" (char*)

20 Type of desired transport (e.g., "tcp", "udp") being requested in an allocation request.

21 `PMIX_ALLOC_FABRIC_PLANE` "pmix.alloc.netplane" (char*)

22 ID string for the *fabric plane* to be used for the requested allocation.

23 `PMIX_ALLOC_FABRIC_ENDPTS` "pmix.alloc.endpts" (size_t)

24 Number of endpoints to allocate per *process* in the job.

25 `PMIX_ALLOC_FABRIC_ENDPTS_NODE` "pmix.alloc.endpts.nd" (size_t)

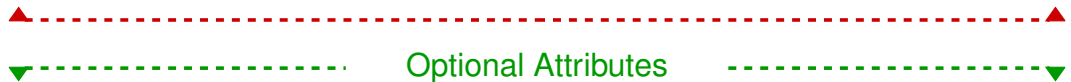
26 Number of endpoints to allocate per *node* for the job.

27 `PMIX_PROC_MAP` "pmix.pmap" (char*)

28 Regular expression describing processes on each node in the specified realm - see 17.2.3.2 for an
29 explanation of its generation. Defaults to the *job* realm.

30 `PMIX_NODE_MAP` "pmix.nmap" (char*)

31 Regular expression of nodes currently hosting processes in the specified realm - see 17.2.3.2 for an
32 explanation of its generation. Defaults to the *job* realm.



Optional Attributes

33 PMIx libraries that support this operation may support the following:

34 `PMIX_ALLOC_BANDWIDTH` "pmix.alloc.bw" (float)

35 Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

36 `PMIX_ALLOC_FABRIC_QOS` "pmix.alloc.netqos" (char*)

37 Fabric quality of service level for the job being requested in an allocation request.

38 `PMIX_SESSION_INFO` "pmix.ssn.info" (bool)

1 Return information regarding the session realm of the target process. In this context, indicates that the
2 information provided in the [PMIX_NODE_MAP](#) is for the entire session and not just the indicated
3 namespace. Thus, subsequent calls to this API may omit node-level information - e.g., the library may
4 not need to include information on the devices on each node in a subsequent call.

5 The following optional attributes may be provided by the host environment to identify the programming model
6 (as specified by the user) being executed within the application. The PMIx server library may utilize this
7 information to harvest/forward model-specific environmental variables, record the programming model
8 associated with the application, etc.

- 9 • [PMIX_PROGRAMMING_MODEL](#) "`pmix.pgm.model`" (`char*`)
10 Programming model being initialized (e.g., "MPI" or "OpenMP").
 - 11 • [PMIX_MODEL_LIBRARY_NAME](#) "`pmix.mdl.name`" (`char*`)
12 Programming model implementation ID (e.g., "OpenMPI" or "MPICH").
 - 13 • [PMIX_MODEL_LIBRARY_VERSION](#) "`pmix.mld.vrs`" (`char*`)
14 Programming model version string (e.g., "2.1.1").
-

15 Description

16 Provide a function by which the RM can request application-specific setup data (e.g., environmental variables,
17 fabric configuration and security credentials) from supporting PMIx server library subsystems prior to
18 initiating launch of a job.

19 This is defined as a non-blocking operation in case contributing subsystems need to perform some potentially
20 time consuming action (e.g., query a remote service) before responding. The returned data must be distributed
21 by the host environment and subsequently delivered to the local PMIx server on each node where application
22 processes will execute, prior to initiating execution of those processes.

▼ Advice to PMIx server hosts ▼

23 Host environments are required to execute this operation prior to launching a job. In addition to supported
24 directives, the *info* array must include a description of the *job* using the [PMIX_NODE_MAP](#) and
25 [PMIX_PROC_MAP](#) attributes.

26 Note that the function can be called on a per-application basis if the [PMIX_PROC_MAP](#) and
27 [PMIX_NODE_MAP](#) are provided only for the corresponding application (as opposed to the entire job) each
28 time.

▼ Advice to PMIx library implementers ▼

29 Support for harvesting of environmental variables and providing of local configuration information by the
30 PMIx implementation is optional.

31 17.2.11.1 Server Setup Application Callback Function

32 The [PMIx_server_setup_application](#) callback function.

Summary

Provide a function by which the resource manager can receive application-specific environmental variables and other setup data prior to launch of an application.

Format

```
typedef void (*pmix_setup_application_cbfunc_t) (  
    pmix_status_t status,  
    pmix_info_t info[], size_t ninfo,  
    void *provided_cbdata,  
    pmix_op_cbfunc_t cbfunc, void *cbdata);
```

IN status

returned status of the request ([pmix_status_t](#))

IN info

Array of info structures (array of handles)

IN ninfo

Number of elements in the *info* array (integer)

IN provided_cbdata

Data originally passed to call to [PMIx_server_setup_application](#) (memory reference)

IN cbfunc

[pmix_op_cbfunc_t](#) function to be called when processing completed (function reference)

IN cbdata

Data to be passed to the *cbfunc* callback function (memory reference)

Description

Define a function to be called by the PMIx server library for return of application-specific setup data in response to a request from the host RM. The returned *info* array is owned by the PMIx server library and will be free'd when the provided *cbfunc* is called.

17.2.11.2 Server Setup Application Attributes

Attributes specifically defined for controlling contents of application setup data.

PMIX_SETUP_APP_ENVARS "pmix.setup.env" (bool)

Harvest and include relevant environmental variables.

PMIX_SETUP_APP_NONENVARS "pmix.setup.nenv" (bool)

Include all relevant data other than environmental variables.

PMIX_SETUP_APP_ALL "pmix.setup.all" (bool)

Include all relevant data.

17.2.12 PMIx_Register_attributes

Summary

Register host environment attribute support for a function.

Format

C

```
pmix_status_t
PMIx_Register_attributes(char *function,
                        pmix_regattr_t attrs[],
                        size_t nattrs);
```

C

IN function
String name of function (string)

IN attrs
Array of [pmix_regattr_t](#) describing the supported attributes (handle)

IN nattrs
Number of elements in *attrs* ([size_t](#))

Returns [PMIX_SUCCESS](#) or a negative value indicating the error.

Description

The [PMIx_Register_attributes](#) function is used by the host environment to register with its PMIx server library the attributes it supports for each [pmix_server_module_t](#) function. The *function* is the string name of the server module function (e.g., "register_events", "validate_credential", or "allocate") whose attributes are being registered. See the [pmix_regattr_t](#) entry for a description of the *attrs* array elements.

Note that the host environment can also query the library (using the [PMIx_Query_info_nb](#) API) for its attribute support both at the server, client, and tool levels once the host has executed [PMIx_server_init](#) since the server will internally register those values.

Advice to PMIx server hosts

Host environments are strongly encouraged to register all supported attributes immediately after initializing the library to ensure that user requests are correctly serviced.

Advice to PMIx library implementers

PMIx implementations are *required* to register all internally supported attributes for each API during initialization of the library (i.e., when the process calls their respective PMIx init function). Specifically, the implementation *must not* register supported attributes upon first call to a given API as this would prevent users from discovering supported attributes prior to first use of an API.

It is the implementation's responsibility to associate registered attributes for a given [pmix_server_module_t](#) function with their corresponding user-facing API. Supported attributes *must* be reported to users in terms of their support for user-facing APIs, broken down by the level (see Section 7.1.5) at which the attribute is supported.

Note that attributes can/will be registered on an API for each level. It is *required* that the implementation support user queries for supported attributes on a per-level basis. Duplicate registrations at the *same* level for a function *shall* return an error - however, duplicate registrations at *different* levels *shall* be independently tracked.

1 17.2.12.1 Attribute registration constants

2 Constants supporting attribute registration.

3 **PMIX_ERR_REPEAT_ATTR_REGISTRATION** -171 The attributes for an identical function have
4 already been registered at the specified level (host, server, or client).

5 17.2.12.2 Attribute registration structure

6 The `pmix_regattr_t` structure is used to register attribute support for a PMIx function.

PMIx v4.0

```
7 typedef struct pmix_regattr {  
8     char *name;  
9     pmix_key_t *string;  
10    pmix_data_type_t type;  
11    pmix_info_t *info;  
12    size_t ninfo;  
13    char **description;  
14 } pmix_regattr_t;
```

15 Note that in this structure:

- 16 • the *name* is the actual name of the attribute - e.g., "PMIX_MAX_PROCS"
- 17 • the *string* is the literal string value of the attribute - e.g., "pmix.max.size" for the **PMIX_MAX_PROCS**
18 attribute
- 19 • *type* must be a PMIx data type identifying the type of data associated with this attribute.
- 20 • the *info* array contains machine-usuable information regarding the range of accepted values. This may
21 include entries for **PMIX_MIN_VALUE**, **PMIX_MAX_VALUE**, **PMIX_ENUM_VALUE**, or a combination of
22 them. For example, an attribute that supports all positive integers might delineate it by including a
23 **pmix_info_t** with a key of **PMIX_MIN_VALUE**, type of **PMIX_INT**, and value of zero. The lack of an
24 entry for **PMIX_MAX_VALUE** indicates that there is no ceiling to the range of accepted values.
- 25 • *ninfo* indicates the number of elements in the *info* array
- 26 • The *description* field consists of a **NULL**-terminated array of strings describing the attribute, optionally
27 including a human-readable description of the range of accepted values - e.g., "ALL POSITIVE
28 INTEGERS", or a comma-delimited list of enum value names. No correlation between the number of
29 entries in the *description* and the number of elements in the *info* array is implied or required.

30 The attribute *name* and *string* fields must be **NULL**-terminated strings composed of standard alphanumeric
31 values supported by common utilities such as *strcmp*.

32 Although not strictly required, both PMIx library implementers and host environments are strongly
33 encouraged to provide both human-readable and machine-parsable descriptions of supported attributes when
34 registering them.

1 17.2.12.3 Attribute registration structure descriptive attributes

2 The following attributes relate to the nature of the values being reported in the `pmix_regattr_t` structures.

3 **PMIX_MAX_VALUE** "pmix.descr.maxval" (varies)

4 Used in `pmix_regattr_t` to describe the maximum valid value for the associated attribute.

5 **PMIX_MIN_VALUE** "pmix.descr.minval" (varies)

6 Used in `pmix_regattr_t` to describe the minimum valid value for the associated attribute.

7 **PMIX_ENUM_VALUE** "pmix.descr.enum" (char*)

8 Used in `pmix_regattr_t` to describe accepted values for the associated attribute. Numerical values
9 shall be presented in a form convertible to the attribute's declared data type. Named values (i.e., values
10 defined by constant names via a typical C-language enum declaration) must be provided as their
11 numerical equivalent.

12 17.2.12.4 Attribute registration structure support macros

13 The following macros are provided to support the `pmix_regattr_t` structure.

14 **Static initializer for the regattr structure**

15 *(Provisional)*

16 Provide a static initializer for the `pmix_regattr_t` fields.

17 *PMIx v5.0* 

18 **PMIX_REGATTR_STATIC_INIT**

19 

20 **Initialize the regattr structure**

21 Initialize the `pmix_regattr_t` fields

22 *PMIx v4.0* 

23 **PMIX_REGATTR_CONSTRUCT(m)**

24 

25 **IN m**

26 Pointer to the structure to be initialized (pointer to `pmix_regattr_t`)

27 **Destruct the regattr structure**

28 Destruct the `pmix_regattr_t` fields, releasing all strings.

29 *PMIx v4.0* 

30 **PMIX_REGATTR_DESTRUCT(m)**

31 

32 **IN m**

33 Pointer to the structure to be destructed (pointer to `pmix_regattr_t`)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Create a regattr array

Allocate and initialize an array of `pmix_regattr_t` structures.

```
▼ _____ C _____ ▼  
PMIX_REGATTR_CREATE(m, n)  
▲ _____ C _____ ▲
```

INOUT `m`

Address where the pointer to the array of `pmix_regattr_t` structures shall be stored (handle)

IN `n`

Number of structures to be allocated (`size_t`)

Free a regattr array

Release an array of `pmix_regattr_t` structures.

PMIx v4.0

```
▼ _____ C _____ ▼  
PMIX_REGATTR_FREE(m, n)  
▲ _____ C _____ ▲
```

INOUT `m`

Pointer to the array of `pmix_regattr_t` structures (handle)

IN `n`

Number of structures in the array (`size_t`)

Load a regattr structure

Load values into a `pmix_regattr_t` structure. The macro can be called multiple times to add as many strings as desired to the same structure by passing the same address and a **NULL** key to the macro. Note that the `t` type value must be given each time.

PMIx v4.0

```
▼ _____ C _____ ▼  
PMIX_REGATTR_LOAD(a, n, k, t, ni, v)  
▲ _____ C _____ ▲
```

IN `a`

Pointer to the structure to be loaded (pointer to `pmix_proc_t`)

IN `n`

String name of the attribute (string)

IN `k`

Key value to be loaded (`pmix_key_t`)

IN `t`

Type of data associated with the provided key (`pmix_data_type_t`)

IN `ni`

Number of `pmix_info_t` elements to be allocated in *info* (`size_t`)

IN `v`

One-line description to be loaded (more can be added separately) (string)

Transfer a regattr to another regattr

Non-destructively transfer the contents of a `pmix_regattr_t` structure to another one.

PMIX_REGATTR_XFER(*m*, *n*)

INOUT *m*

Pointer to the destination `pmix_regattr_t` structure (handle)

IN *m*

Pointer to the source `pmix_regattr_t` structure (handle)

17.2.13 PMIx_server_setup_local_support

Summary

Provide a function by which the local PMIx server can perform any application-specific operations prior to spawning local clients of a given application.

Format

PMIx v2.0

`pmix_status_t`

```
PMIx_server_setup_local_support(const pmix_namespace_t nspace,  
                               pmix_info_t info[], size_t ninfo,  
                               pmix_op_cbfunc_t cbfunc,  
                               void *cbdata);
```

IN *nspace*

Namespace (string)

IN *info*

Array of info structures (array of handles)

IN *ninfo*

Number of elements in the *info* array (**size_t**)

IN *cbfunc*

Callback function `pmix_op_cbfunc_t`. A `NULL` function reference indicates that the function is to be executed as a blocking operation (function reference)

IN *cbdata*

Data to be passed to the callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

Provide a function by which the local PMIx server can perform any application-specific operations prior to spawning local clients of a given application. For example, a fabric library might need to setup the local driver for “instant on” addressing. The data provided in the *info* array is the data returned to the host RM by the callback function executed as a result of a call to `PMIx_server_setup_application`.

Advice to PMIx server hosts

Host environments are required to execute this operation prior to starting any local application processes from the specified namespace if information was obtained from a call to `PMIx_server_setup_application`.

Host environments must register the *nspace* using `PMIx_server_register_nspace` prior to calling this API to ensure that all namespace-related information required to support this function is available to the library. This eliminates the need to include any of the registration information in the *info* array passed to this API.

17.2.14 PMIx_server_IOF_deliver

Summary

Provide a function by which the host environment can pass forwarded Input/Output (IO) to the PMIx server library for distribution to its clients.

Format

PMIx v3.0

C

```
pmix_status_t
PMIx_server_IOF_deliver(const pmix_proc_t *source,
                        pmix_iof_channel_t channel,
                        const pmix_byte_object_t *bo,
                        const pmix_info_t info[], size_t ninfo,
                        pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

- IN source**
Pointer to `pmix_proc_t` identifying source of the IO (handle)
- IN channel**
IO channel of the data (`pmix_iof_channel_t`)
- IN bo**
Pointer to `pmix_byte_object_t` containing the payload to be delivered (handle)
- IN info**
Array of `pmix_info_t` metadata describing the data (array of handles)
- IN ninfo**
Number of elements in the *info* array (`size_t`)
- IN cbfunc**
Callback function `pmix_op_cbfunc_t`. A `NULL` function reference indicates that the function is to be executed as a blocking operation (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

1 A successful return indicates that the request is being processed and the result will be returned in the provided
2 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
3 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

4 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- 5 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
6 *success* - the *cbfunc* will not be called

7 If none of the above return codes are appropriate, then an implementation must return either a general PMIx
8 error code or an implementation defined error code as described in Section 3.1.1.

9 **Description**

10 Provide a function by which the host environment can pass forwarded IO to the PMIx server library for
11 distribution to its clients. The PMIx server library is responsible for determining which of its clients have
12 actually registered for the provided data and delivering it. The *cbfunc* callback function will be called once the
13 PMIx server library no longer requires access to the provided data.

14 **17.2.15 PMIx_server_collect_inventory**

15 **Summary**

16 Collect inventory of resources on a node.

17 **Format**

PMIx v3.0

C

```
18 pmix_status_t  
19 PMIx_server_collect_inventory(const pmix_info_t directives[],  
20                               size_t ndirs,  
21                               pmix_info_cbfunc_t cbfunc,  
22                               void *cbdata);
```

C

23 **IN directives**

24 Array of **pmix_info_t** directing the request (array of handles)

25 **IN ndirs**

26 Number of elements in the *directives* array (**size_t**)

27 **IN cbfunc**

28 Callback function to return collected data (**pmix_info_cbfunc_t** function reference)

29 **IN cbdata**

30 Data to be passed to the callback function (memory reference)

31 A successful return indicates that the request is being processed and the result will be returned in the provided
32 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
33 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Description

Provide a function by which the host environment can request its PMIx server library collect an inventory of local resources. Supported resources depends upon the PMIx implementation, but may include the local node topology and fabric interfaces.

Advice to PMIx server hosts

This is a non-blocking API as it may involve somewhat lengthy operations to obtain the requested information. Inventory collection is expected to be a rare event – at system startup and upon command from a system administrator. Inventory updates are expected to initiate a smaller operation involving only the changed information. For example, replacement of a node would generate an event to notify the scheduler with an inventory update without invoking a global inventory operation.

17.2.16 PMIx_server_deliver_inventory

Summary

Pass collected inventory to the PMIx server library for storage.

Format

PMIx v3.0

```
pmix_status_t
PMIx_server_deliver_inventory(const pmix_info_t info[],
                             size_t ninfo,
                             const pmix_info_t directives[],
                             size_t ndirs,
                             pmix_op_cbfunc_t cbfunc,
                             void *cbdata);
```

IN info

Array of [pmix_info_t](#) containing the inventory (array of handles)

IN ninfo

Number of elements in the *info* array ([size_t](#))

IN directives

Array of [pmix_info_t](#) directing the request (array of handles)

IN ndirs

Number of elements in the *directives* array ([size_t](#))

IN cbfunc

Callback function [pmix_op_cbfunc_t](#). A **NULL** function reference indicates that the function is to be executed as a blocking operation (function reference)

IN cbdata

Data to be passed to the callback function (memory reference)

1 Returns one of the following:

2 A successful return indicates that the request is being processed and the result will be returned in the provided
3 *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The
4 callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

5 Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

- 6 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
7 *success* - the *cbfunc* will not be called

8 If none of the above return codes are appropriate, then an implementation must return either a general PMIX
9 error code or an implementation defined error code as described in Section 3.1.1.

10 Description

11 Provide a function by which the host environment can pass inventory information obtained from a node (as a
12 result of a call to **PMIX_server_collect_inventory**) to the PMIX server library for storage. Inventory
13 data is subsequently used by the PMIX server library for allocations in response to
14 **PMIX_server_setup_application**, and may be available to the library's host via the **PMIX_Get** API
15 (depending upon PMIX implementation). The *cbfunc* callback function will be called once the PMIX server
16 library no longer requires access to the provided data.

17 17.2.17 PMIX_server_generate_locality_string

18 Summary

19 Generate a PMIX locality string from a given cpuset.

20 Format

PMIX v4.0

```
21 pmix_status_t
22 PMIX_server_generate_locality_string(const pmix_cpuset_t *cpuset,
23                                     char **locality);
```

24 IN cpuset

25 Pointer to a **pmix_cpuset_t** containing the bitmap of assigned PUs (handle)

26 OUT locality

27 String representation of the PMIX locality corresponding to the input bitmap (**char***)

28 A successful return indicates that the returned string contains the generated locality string.

29 Returns **PMIX_SUCCESS** or a negative value indicating the error.

30 Description

31 Provide a function by which the host environment can generate a PMIX locality string for inclusion in the call
32 to **PMIX_server_register_nspace**. This function shall only be called for local client processes, with
33 the returned locality included in the job-level information (via the **PMIX_LOCALITY_STRING** attribute)
34 provided to local clients. Local clients can use these strings as input to determine the relative locality of their
35 local peers via the **PMIX_Get_relative_locality** API.

36 The function is required to return a string prefixed by the *source* field of the provided *cpuset* followed by a
37 colon. The remainder of the string shall represent the corresponding locality as expressed by the underlying
38 implementation.

1 17.2.18 PMIx_server_generate_cpuset_string

2 Summary

3 Generate a PMIx string representation of the provided cpuset.

4 *PMIx v4.0* Format

C

5 `pmix_status_t`

6 `PMIx_server_generate_cpuset_string(const pmix_cpuset_t *cpuset,`
7 `char **cpuset_string);`

C

8 **IN** `cpuset`

9 Pointer to a `pmix_cpuset_t` containing the bitmap of assigned PUs (handle)

10 **OUT** `cpuset_string`

11 String representation of the input bitmap (`char*`)

12 A successful return indicates that the returned string contains the generated cpuset representation string.

13 Returns `PMIX_SUCCESS` or a negative value indicating the error.

14 Description

15 Provide a function by which the host environment can generate a string representation of the cpuset bitmap for
16 inclusion in the call to `PMIx_server_register_nspace`. This function shall only be called for local
17 client processes, with the returned string included in the job-level information (via the `PMIX_CPUSSET`
18 attribute) provided to local clients. Local clients can use these strings as input to obtain their PU bindings via
19 the `PMIx_Parse_cpuset_string` API.

20 The function is required to return a string prefixed by the *source* field of the provided *cpuset* followed by a
21 colon. The remainder of the string shall represent the PUs to which the process is bound as expressed by the
22 underlying implementation.

23 17.2.18.1 Cpuset Structure

24 The `pmix_cpuset_t` structure contains a character string identifying the source of the bitmap (e.g.,
25 "hwloc") and a pointer to the corresponding implementation-specific structure (e.g., `hwloc_cpuset_t`).

PMIx v4.0

C

26 `typedef struct pmix_cpuset {`
27 `char *source;`
28 `void *bitmap;`
29 `} pmix_cpuset_t;`

C

30 17.2.18.2 Cpuset support macros

31 The following macros support the `pmix_cpuset_t` structure.

1 **Static initializer for the cpuset structure**

2 *(Provisional)*

3 Provide a static initializer for the `pmix_cpuset_t` fields.



4 `PMIX_CPUSSET_STATIC_INIT`



5 **Initialize the cpuset structure**

6 Initialize the `pmix_cpuset_t` fields.

PMIx v4.0



7 `PMIX_CPUSSET_CONSTRUCT (m)`



8 **IN** `m`

9 Pointer to the structure to be initialized (pointer to `pmix_cpuset_t`)

10 **Destruct the cpuset structure**

11 Destruct the `pmix_cpuset_t` fields.

PMIx v4.0



12 `PMIX_CPUSSET_DESTRUCT (m)`



13 **IN** `m`

14 Pointer to the structure to be destructed (pointer to `pmix_cpuset_t`)

15 **Create a cpuset array**

16 Allocate and initialize a `pmix_cpuset_t` array.

PMIx v4.0



17 `PMIX_CPUSSET_CREATE (m, n)`



18 **INOUT** `m`

19 Address where the pointer to the array of `pmix_cpuset_t` structures shall be stored (handle)

20 **IN** `n`

21 Number of structures to be allocated (size_t)

22 **Release a cpuset array**

23 Deconstruct and free a `pmix_cpuset_t` array.

PMIx v4.0



24 `PMIX_CPUSSET_FREE (m, n)`



25 **INOUT** `m`

26 Address the array of `pmix_cpuset_t` structures to be released (handle)

27 **IN** `n`

28 Number of structures in the array (size_t)

1 17.2.19 PMIx_server_define_process_set

2 Summary

3 Define a PMIx process set.

4 *PMIx v4.0* Format

C

5 `pmix_status_t`

```
6 PMIx_server_define_process_set(const pmix_proc_t members[],  
7                               size_t nmembers,  
8                               char *pset_name);
```

C

9 **IN** `members`

10 Pointer to an array of `pmix_proc_t` containing the identifiers of the processes in the process set
11 (handle)

12 **IN** `nmembers`

13 Number of elements in `members` (integer)

14 **IN** `pset_name`

15 String name of the process set being defined (`char*`)

16 Returns `PMIX_SUCCESS` or a negative value indicating the error.

17 Description

18 Provide a function by which the host environment can create a process set. The PMIx server shall alert all
19 local clients of the new process set (including process set name and membership) via the
20 `PMIX_PROCESS_SET_DEFINE` event.

Advice to PMIx server hosts

21 The host environment is responsible for ensuring:

- 22 • consistent knowledge of process set membership across all involved PMIx servers; and
- 23 • that process set names do not conflict with system-assigned namespaces within the scope of the set

24 17.2.20 PMIx_server_delete_process_set

25 Summary

26 Delete a PMIx process set name

27 *PMIx v4.0* Format

C

28 `pmix_status_t`

```
29 PMIx_server_delete_process_set(char *pset_name);
```

C

30 **IN** `pset_name`

31 String name of the process set being deleted (`char*`)

32 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Description

Provide a function by which the host environment can delete a process set name. The PMIx server shall alert all local clients of the process set name being deleted via the `PMIX_PROCESS_SET_DELETE` event. Deletion of the name has no impact on the member processes.

Advice to PMIx server hosts

The host environment is responsible for ensuring consistent knowledge of process set membership across all involved PMIx servers.

17.3 Server Function Pointers

PMIx utilizes a "function-shipping" approach to support for implementing the server-side of the protocol. This method allows RMs to implement the server without being burdened with PMIx internal details. When a request is received from the client, the corresponding server function will be called with the information.

Any functions not supported by the RM can be indicated by a `NULL` for the function pointer. PMIx implementations are required to return a `PMIX_ERR_NOT_SUPPORTED` status to all calls to functions that require host environment support and are not backed by a corresponding server module entry. Host environments may, if they choose, include a function pointer for operations they have not yet implemented and simply return `PMIX_ERR_NOT_SUPPORTED`.

Functions that accept directives (i.e., arrays of `pmix_info_t` structures) must check any provided directives for those marked as *required* via the `PMIX_INFO_REQD` flag. PMIx client and server libraries are required to mark any such directives with the `PMIX_INFO_REQD_PROCESSED` flag should they have handled the request. Any required directive that has not been marked therefore becomes the responsibility of the host environment. If a required directive that hasn't been processed by a lower level cannot be supported by the host, then the `PMIX_ERR_NOT_SUPPORTED` error constant must be returned. If the directive can be processed by the host, then the host shall do so and mark the attribute with the `PMIX_INFO_REQD_PROCESSED` flag.

The host RM will provide the function pointers in a `pmix_server_module_t` structure passed to `PMIx_server_init`. The module structure and associated function references are defined in this section.

Advice to PMIx server hosts

For performance purposes, the host server is required to return as quickly as possible from all functions. Execution of the function is thus to be done asynchronously so as to allow the PMIx server support library to handle multiple client requests as quickly and scalably as possible.

All data passed to the host server functions is "owned" by the PMIx server support library and must not be free'd. Data returned by the host server via callback function is owned by the host server, which is free to release it upon return from the callback

17.3.1 `pmix_server_module_t` Module

Summary

List of function pointers that a PMIx server passes to `PMIx_server_init` during startup.

Format

C

```
1
2 typedef struct pmix_server_module_4_0_0_t {
3     /* v1x interfaces */
4     pmix_server_client_connected_fn_t    client_connected; // DEPRECATED
5     pmix_server_client_finalized_fn_t    client_finalized;
6     pmix_server_abort_fn_t               abort;
7     pmix_server_fence_nb_fn_t            fence_nb;
8     pmix_server_dmodex_req_fn_t          direct_modex;
9     pmix_server_publish_fn_t             publish;
10    pmix_server_lookup_fn_t               lookup;
11    pmix_server_unpublish_fn_t            unpublish;
12    pmix_server_spawn_fn_t                spawn;
13    pmix_server_connect_fn_t              connect;
14    pmix_server_disconnect_fn_t            disconnect;
15    pmix_server_register_events_fn_t       register_events;
16    pmix_server_deregister_events_fn_t     deregister_events;
17    pmix_server_listener_fn_t              listener;
18    /* v2x interfaces */
19    pmix_server_notify_event_fn_t          notify_event;
20    pmix_server_query_fn_t                 query;
21    pmix_server_tool_connection_fn_t       tool_connected;
22    pmix_server_log_fn_t                   log;
23    pmix_server_alloc_fn_t                 allocate;
24    pmix_server_job_control_fn_t           job_control;
25    pmix_server_monitor_fn_t               monitor;
26    /* v3x interfaces */
27    pmix_server_get_cred_fn_t              get_credential;
28    pmix_server_validate_cred_fn_t         validate_credential;
29    pmix_server_iof_fn_t                   iof_pull;
30    pmix_server_stdin_fn_t                 push_stdin;
31    /* v4x interfaces */
32    pmix_server_grp_fn_t                   group;
33    pmix_server_fabric_fn_t                fabric;
34    pmix_server_client_connected2_fn_t     client_connected2;
35 } pmix_server_module_t;
```

C

Advice to PMIx server hosts

36 Note that some PMIx implementations *require* the use of C99-style designated initializers to clearly correlate
37 each provided function pointer with the correct member of the `pmix_server_module_t` structure as the
38 location/ordering of struct members may change over time.

1 17.3.2 pmix_server_client_connected_fn_t

2 Summary

3 Notify the host server that a client connected to this server. This function module entry has been
4 DEPRECATED in favor of [pmix_server_client_connected2_fn_t](#).

5 Format

PMIx v1.0

C

```
6 typedef pmix_status_t (*pmix_server_client_connected_fn_t) (  
7     const pmix_proc_t *proc,  
8     void* server_object,  
9     pmix_op_cbfunc_t cbfunc,  
10    void *cbdata);
```

C

11 **IN** `proc`
12 [pmix_proc_t](#) structure (handle)
13 **IN** `server_object`
14 object reference (memory reference)
15 **IN** `cbfunc`
16 Callback function [pmix_op_cbfunc_t](#) (function reference)
17 **IN** `cbdata`
18 Data to be passed to the callback function (memory reference)

19 Returns one of the following:

- 20 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
21 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
22 from the API.
- 23 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
24 *success* - the *cbfunc* will not be called
- 25 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
26 and failed - the *cbfunc* will not be called

27 Description

28 This function module entry has been DEPRECATED in favor of
29 [pmix_server_client_connected2_fn_t](#). If both functions are provided, the PMIx library will
30 ignore this function module entry in favor of its replacement.

31 17.3.3 pmix_server_client_connected2_fn_t

32 Summary

33 Notify the host server that a client connected to this server - this version of the original function definition has
34 been extended to include an array of [pmix_info_t](#), thereby allowing the PMIx server library to pass
35 additional information identifying the client to the host environment.

Format

```
typedef pmix_status_t (*pmix_server_client_connected2_fn_t) (  
    const pmix_proc_t *proc,  
    void* server_object,  
    pmix_info_t info[], size_t ninfo,  
    pmix_op_cbfunc_t cbfunc,  
    void *cbdata)
```

IN `proc`
 [pmix_proc_t](#) structure (handle)

IN `server_object`
 object reference (memory reference)

IN `info`
 Array of info structures (array of handles)

IN `ninfo`
 Number of elements in the *info* array (integer)

IN `cbfunc`
 Callback function [pmix_op_cbfunc_t](#) (function reference)

IN `cbdata`
 Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the *cbfunc* will not be called. The PMIx server library is to immediately terminate the connection.

Description

Notify the host environment that a client has called [PMIx_Init](#). Note that the client will be in a blocked state until the host server executes the callback function, thus allowing the PMIx server support library to release the client. The `server_object` parameter will be the value of the `server_object` parameter passed to [PMIx_server_register_client](#) by the host server when registering the connecting client. A host server can choose to not be notified when clients connect by setting [pmix_server_client_connected2_fn_t](#) to `NULL`.

It is possible that only a subset of the clients in a namespace call [PMIx_Init](#). The server's [pmix_server_client_connected2_fn_t](#) implementation should therefore not depend on being called once per rank in a namespace or delay calling the callback function until all ranks have connected. However, the host may rely on the [pmix_server_client_connected2_fn_t](#) function module entry being called for a given rank prior to any other function module entries being executed on behalf of that rank.

1 17.3.4 pmix_server_client_finalized_fn_t

2 Summary

3 Notify the host environment that a client called `PMIx_Finalize`.

4 Format

PMIx v1.0

C

```
5 typedef pmix_status_t (*pmix_server_client_finalized_fn_t) (  
6     const pmix_proc_t *proc,  
7     void* server_object,  
8     pmix_op_cbfunc_t cbfunc,  
9     void *cbdata);
```

C

- 10 **IN** `proc`
11 `pmix_proc_t` structure (handle)
- 12 **IN** `server_object`
13 object reference (memory reference)
- 14 **IN** `cbfunc`
15 Callback function `pmix_op_cbfunc_t` (function reference)
- 16 **IN** `cbdata`
17 Data to be passed to the callback function (memory reference)

18 Returns one of the following:

- 19 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
20 returned in the provided `cbfunc`. Note that the host must not invoke the callback function prior to returning
21 from the API.
- 22 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
23 `success` - the `cbfunc` will not be called
- 24 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
25 and failed - the `cbfunc` will not be called

26 Description

27 Notify the host environment that a client called `PMIx_Finalize`. Note that the client will be in a blocked
28 state until the host server executes the callback function, thus allowing the PMIx server support library to
29 release the client. The `server_object` parameter will be the value of the `server_object` parameter passed to
30 `PMIx_server_register_client` by the host server when registering the connecting client. If provided,
31 an implementation of `pmix_server_client_finalized_fn_t` is only required to call the callback
32 function designated. A host server can choose to not be notified when clients finalize by setting
33 `pmix_server_client_finalized_fn_t` to `NULL`.

34 Note that the host server is only being informed that the client has called `PMIx_Finalize`. The client might
35 not have exited. If a client exits without calling `PMIx_Finalize`, the server support library will not call the
36 `pmix_server_client_finalized_fn_t` implementation.

Advice to PMIx server hosts

This operation is an opportunity for a host server to update the status of the tasks it manages. It is also a convenient and well defined time to release resources used to support that client.

17.3.5 `pmix_server_abort_fn_t`

Summary

Notify the host environment that a local client called `PMIx_Abort`.

Format

C

```
typedef pmix_status_t (*pmix_server_abort_fn_t) (  
    const pmix_proc_t *proc,  
    void *server_object,  
    int status,  
    const char msg[],  
    pmix_proc_t procs[],  
    size_t nprocs,  
    pmix_op_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN** `proc`
`pmix_proc_t` structure identifying the process requesting the abort (handle)
- IN** `server_object`
object reference (memory reference)
- IN** `status`
exit status (integer)
- IN** `msg`
exit status message (string)
- IN** `procs`
Array of `pmix_proc_t` structures identifying the processes to be terminated (array of handles)
- IN** `nprocs`
Number of elements in the `procs` array (integer)
- IN** `cbfunc`
Callback function `pmix_op_cbfunc_t` (function reference)
- IN** `cbdata`
Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided `cbfunc`. Note that the host must not invoke the callback function prior to returning from the API.

- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called
- **PMIX_ERR_PARAM_VALUE_NOT_SUPPORTED** indicating that the host environment supports this API, but the request includes processes that the host environment cannot abort - e.g., if the request is to abort subsets of processes from a namespace, or processes outside of the caller's own namespace, and the host environment does not permit such operations. In this case, none of the specified processes will be terminated - the *cbfunc* will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the *cbfunc* will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the *cbfunc* will not be called

Description

A local client called **PMIx_Abort**. Note that the client will be in a blocked state until the host server executes the callback function, thus allowing the PMIx server library to release the client. The array of *procs* indicates which processes are to be terminated. A **NULL** for the *procs* array indicates that all processes in the caller's namespace are to be aborted, including itself - this is the equivalent of passing a **pmix_proc_t** array element containing the caller's namespace and a rank value of **PMIX_RANK_WILDCARD**.

17.3.6 pmix_server_fenceb_fn_t

Summary

At least one client called either **PMIx_Fence** or **PMIx_Fence_nb**.

Format

C

```
typedef pmix_status_t (*pmix_server_fenceb_fn_t) (
    const pmix_proc_t procs[],
    size_t nprocs,
    const pmix_info_t info[],
    size_t ninfo,
    char *data, size_t ndata,
    pmix_modex_cbfunc_t cbfunc,
    void *cbdata);
```

C

- IN procs**
Array of **pmix_proc_t** structures identifying operation participants(array of handles)
- IN nprocs**
Number of elements in the *procs* array (integer)
- IN info**
Array of info structures (array of handles)
- IN ninfo**
Number of elements in the *info* array (integer)
- IN data**
(string)

1 **IN** **ndata**
2 (integer)
3 **IN** **cbfunc**
4 Callback function [pmix_modex_cbfunc_t](#) (function reference)
5 **IN** **cbdata**
6 Data to be passed to the callback function (memory reference)

7 Returns one of the following:

- 8 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
9 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
10 from the API.
- 11 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
12 though the function entry was provided in the server module - the *cbfunc* will not be called
- 13 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
14 and failed - the *cbfunc* will not be called

▼----- Required Attributes -----▼

15 PMIx libraries are required to pass any provided attributes to the host environment for processing.

16 The following attributes are required to be supported by all host environments:

17 **PMIX_COLLECT_DATA** "pmix.collect" (bool)

18 Collect all data posted by the participants using [PMIx_Put](#) that has been committed via
19 [PMIx_Commit](#), making the collection locally available to each participant at the end of the operation.
20 By default, this will include all job-level information that was locally generated by PMIx servers unless
21 excluded using the [PMIX_COLLECT_GENERATED_JOB_INFO](#) attribute.

22 **PMIX_LOCAL_COLLECTIVE_STATUS** "pmix.loc.col.st" (pmix_status_t)

23 Status code for local collective operation being reported to the host by the server library. PMIx servers
24 may aggregate the participation by local client processes in a collective operation - e.g., instead of
25 passing individual client calls to [PMIx_Fence](#) up to the host environment, the server may pass only a
26 single call to the host when all local participants have executed their [PMIx_Fence](#) call, thereby
27 reducing the burden placed on the host. However, in cases where the operation locally fails (e.g., if a
28 participating client abnormally terminates prior to calling the operation), the server upcall functions to
29 the host do not include a [pmix_status_t](#) by which the PMIx server can alert the host to that failure.
30 This attribute resolves that problem by allowing the server to pass the status information regarding the
31 local collective operation.

▲-----

▼----- Optional Attributes -----▼

32 The following attributes are optional for host environments:

33 **PMIX_TIMEOUT** "pmix.timeout" (int)

34 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
35 [PMIX_ERR_TIMEOUT](#) error. Care should be taken to avoid race conditions caused by multiple layers
36 (client, server, and host) simultaneously timing the operation.

Advice to PMIx server hosts

Host environment are required to return `PMIX_ERR_NOT_SUPPORTED` if passed an attributed marked as `PMIX_INFO_REQD` that they do not support, even if support for that attribute is optional.

Description

All local clients in the provided array of *procs* called either `PMIx_Fence` or `PMIx_Fence_nb`. In either case, the host server will be called via a non-blocking function to execute the specified operation once all participating local processes have contributed. All processes in the specified *procs* array are required to participate in the `PMIx_Fence/PMIx_Fence_nb` operation. The callback is to be executed once every daemon hosting at least one participant has called the host server's `pmix_server_fence_nb_fn_t` function.

The provided data is to be collectively shared with all PMIx servers involved in the fence operation, and returned in the modex *cbfunc*. A `NULL` data value indicates that the local processes had no data to contribute.

The array of *info* structs is used to pass user-requested options to the server. This can include directives as to the algorithm to be used to execute the fence operation. The directives are optional unless the `PMIX_INFO_REQD` flag has been set - in such cases, the host RM is required to return an error if the directive cannot be met.

Advice to PMIx library implementers

The PMIx server library is required to aggregate participation by local clients, passing the request to the host environment once all local participants have executed the API.

Advice to PMIx server hosts

The host will receive a single call for each collective operation. It is the responsibility of the host to identify the nodes containing participating processes, execute the collective across all participating nodes, and notify the local PMIx server library upon completion of the global collective. Data received from each node must be simply concatenated to form an aggregated unit, as shown in the following example:

```
uint8_t *blob1, *blob2, *total;
size_t sz_blob1, sz_blob2, sz_total;

sz_total = sz_blob1 + sz_blob2;
total = (uint8_t*)malloc(sz_total);
memcpy(total, blob1, sz_blob1);
memcpy(&total[sz_blob1], blob2, sz_blob2);
```

Note that the ordering of the data blobs does not matter. The host is responsible for free'ing the *data* object passed to it by the PMIx server library.

1 17.3.6.1 Modex Callback Function

2 Summary

3 The `pmix_modex_cbfunc_t` is used by the `pmix_server_fencefn_t` and
4 `pmix_server_dmodex_req_fn_t` PMIx server operations to return modex Business Card
5 Exchange (BCX) data.

PMIx v1.0

C

```
6 typedef void (*pmix_modex_cbfunc_t)
7     (pmix_status_t status,
8      const char *data, size_t ndata,
9      void *cbdata,
10     pmix_release_cbfunc_t release_fn,
11     void *release_cbdata);
```

C

12 **IN status**
13 Status associated with the operation (handle)

14 **IN data**
15 Data to be passed (pointer)

16 **IN ndata**
17 size of the data (`size_t`)

18 **IN cbdata**
19 Callback data passed to original API call (memory reference)

20 **IN release_fn**
21 Callback for releasing *data* (function pointer)

22 **IN release_cbdata**
23 Pointer to be passed to *release_fn* (memory reference)

24 Description

25 A callback function that is solely used by PMIx servers, and not clients, to return modex BCX data in response
26 to “fence” and “get” operations. The returned blob contains the data collected from each server participating in
27 the operation.

28 17.3.7 pmix_server_dmodex_req_fn_t

29 Summary

30 Used by the PMIx server to request its local host contact the PMIx server on the remote node that hosts the
31 specified process to obtain and return a direct modex blob for that process.

32 Format

PMIx v1.0

C

```
33 typedef pmix_status_t (*pmix_server_dmodex_req_fn_t) (
34     const pmix_proc_t *proc,
35     const pmix_info_t info[],
36     size_t ninfo,
37     pmix_modex_cbfunc_t cbfunc,
38     void *cbdata);
```

1 **IN** `proc`
 2 `pmix_proc_t` structure identifying the process whose data is being requested (handle)
 3 **IN** `info`
 4 Array of info structures (array of handles)
 5 **IN** `ninfo`
 6 Number of elements in the `info` array (integer)
 7 **IN** `cbfunc`
 8 Callback function `pmix_modex_cbfunc_t` (function reference)
 9 **IN** `cbdata`
 10 Data to be passed to the callback function (memory reference)

11 Returns one of the following:

- 12 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
 13 returned in the provided `cbfunc`. Note that the host must not invoke the callback function prior to returning
 14 from the API.
- 15 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
 16 though the function entry was provided in the server module - the `cbfunc` will not be called
- 17 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
 18 and failed - the `cbfunc` will not be called

Required Attributes

19 PMIx libraries are required to pass any provided attributes to the host environment for processing.

20 All host environments are required to support the following attributes:

21 **PMIX_REQUIRED_KEY** "`pmix.req.key`" (`char*`)
 22 Identifies a key that must be included in the requested information. If the specified key is not already
 23 available, then the PMIx servers are required to delay response to the dmodex request until either the
 24 key becomes available or the request times out.

Optional Attributes

25 The following attributes are optional for host environments that support this operation:

26 **PMIX_TIMEOUT** "`pmix.timeout`" (`int`)
 27 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
 28 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
 29 (client, server, and host) simultaneously timing the operation.

Description

Used by the PMIx server to request its local host contact the PMIx server on the remote node that hosts the specified *proc* to obtain and return any information that process posted via calls to [PMIx_Put](#) and [PMIx_Commit](#).

The array of *info* structs is used to pass user-requested options to the server. This can include a timeout to preclude an indefinite wait for data that may never become available. The directives are optional unless the *mandatory* flag has been set - in such cases, the host RM is required to return an error if the directive cannot be met.

17.3.7.1 Dmodex attributes

PMIX_REQUIRED_KEY "pmix.req.key" (char*)

Identifies a key that must be included in the requested information. If the specified key is not already available, then the PMIx servers are required to delay response to the dmodex request until either the key becomes available or the request times out.

17.3.8 pmix_server_publish_fn_t

Summary

Publish data per the PMIx API specification.

Format

PMIx v1.0

```
typedef pmix_status_t (*pmix_server_publish_fn_t) (
    const pmix_proc_t *proc,
    const pmix_info_t info[],
    size_t ninfo,
    pmix_op_cbfunc_t cbfunc,
    void *cbdata);
```

IN proc
[pmix_proc_t](#) structure of the process publishing the data (handle)

IN info
Array of info structures (array of handles)

IN ninfo
Number of elements in the *info* array (integer)

IN cbfunc
Callback function [pmix_op_cbfunc_t](#) (function reference)

IN cbdata
Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning from the API.

- 1 ● **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
2 *success* - the *cbfunc* will not be called
- 3 ● **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
4 though the function entry was provided in the server module - the *cbfunc* will not be called
- 5 ● a PMIx error constant indicating either an error in the input or that the request was immediately processed
6 and failed - the *cbfunc* will not be called

Required Attributes

7 PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition,
8 the following attributes are required to be included in the passed *info* array:

9 **PMIX_USERID** "pmix.euid" (uint32_t)

10 Effective user ID of the connecting process.

11 **PMIX_GRPID** "pmix.egid" (uint32_t)

12 Effective group ID of the connecting process.

14 Host environments that implement this entry point are required to support the following attributes:

15 **PMIX_RANGE** "pmix.range" (pmix_data_range_t)

16 Define constraints on the processes that can access the provided data. Only processes that meet the
17 constraints are allowed to access it.

18 **PMIX_PERSISTENCE** "pmix.persist" (pmix_persistence_t)

19 Declare how long the datastore shall retain the provided data. The datastore is to delete the data upon
20 reaching the persistence criterion.

Optional Attributes

21 The following attributes are optional for host environments that support this operation:

22 **PMIX_TIMEOUT** "pmix.timeout" (int)

23 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
24 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
25 (client, server, and host) simultaneously timing the operation.

Description

Publish data per the [PMIx_Publish](#) specification. The callback is to be executed upon completion of the operation. The default data range is left to the host environment, but expected to be [PMIX_RANGE_SESSION](#), and the default persistence [PMIX_PERSIST_SESSION](#) or their equivalent. These values can be specified by including the respective attributed in the *info* array.

The persistence indicates how long the server should retain the data.

Advice to PMIx server hosts

The host environment is not required to guarantee support for any specific range - i.e., the environment does not need to return an error if the data store doesn't support a specified range so long as it is covered by some internally defined range. However, the server must return an error (a) if the key is duplicative within the storage range, and (b) if the server does not allow overwriting of published info by the original publisher - it is left to the discretion of the host environment to allow info-key-based flags to modify this behavior.

The [PMIX_USERID](#) and [PMIX_GRPID](#) of the publishing process will be provided to support authorization-based access to published information and must be returned on any subsequent lookup request.

17.3.9 pmix_server_lookup_fn_t

Summary

Lookup published data.

Format

PMIx v1.0

```
typedef pmix_status_t (*pmix_server_lookup_fn_t) (
    const pmix_proc_t *proc,
    char **keys,
    const pmix_info_t info[],
    size_t ninfo,
    pmix_lookup_cbfunc_t cbfunc,
    void *cbdata);
```

- IN** **proc**
[pmix_proc_t](#) structure of the process seeking the data (handle)
- IN** **keys**
(array of strings)
- IN** **info**
Array of info structures (array of handles)
- IN** **ninfo**
Number of elements in the *info* array (integer)
- IN** **cbfunc**
Callback function [pmix_lookup_cbfunc_t](#) (function reference)
- IN** **cbdata**
Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the *cbfunc* will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the *cbfunc* will not be called

Required Attributes

PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition, the following attributes are required to be included in the passed *info* array:

PMIX_USERID "pmix.euid" (**uint32_t**)

Effective user ID of the connecting process.

PMIX_GRPID "pmix.egid" (**uint32_t**)

Effective group ID of the connecting process.

Host environments that implement this entry point are required to support the following attributes:

PMIX_RANGE "pmix.range" (**pmix_data_range_t**)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

PMIX_WAIT "pmix.wait" (**int**)

Caller requests that the PMIx server wait until at least the specified number of values are found (a value of zero indicates *all* and is the default).

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (**int**)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Lookup published data. The host server will be passed a **NULL**-terminated array of string keys identifying the data being requested.

The array of *info* structs is used to pass user-requested options to the server. The default data range is left to the host environment, but expected to be **PMIX_RANGE_SESSION**. This can include a wait flag to indicate that the server should wait for all data to become available before executing the callback function, or should immediately callback with whatever data is available. In addition, a timeout can be specified on the wait to preclude an indefinite wait for data that may never be published.

Advice to PMIx server hosts

The **PMIX_USERID** and **PMIX_GRPID** of the requesting process will be provided to support authorization-based access to published information. The host environment is not required to guarantee support for any specific range - i.e., the environment does not need to return an error if the data store doesn't support a specified range so long as it is covered by some internally defined range.

17.3.10 pmix_server_unpublish_fn_t

Summary

Delete data from the data store.

Format

C

```
typedef pmix_status_t (*pmix_server_unpublish_fn_t) (  
    const pmix_proc_t *proc,  
    char **keys,  
    const pmix_info_t info[],  
    size_t ninfo,  
    pmix_op_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN proc**
[pmix_proc_t](#) structure identifying the process making the request (handle)
- IN keys**
(array of strings)
- IN info**
Array of info structures (array of handles)
- IN ninfo**
Number of elements in the *info* array (integer)
- IN cbfunc**
Callback function [pmix_op_cbfunc_t](#) (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

Returns one of the following:

- 1 ● **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
2 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
3 from the API.
- 4 ● **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
5 *success* - the *cbfunc* will not be called
- 6 ● **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
7 though the function entry was provided in the server module - the *cbfunc* will not be called
- 8 ● a PMIx error constant indicating either an error in the input or that the request was immediately processed
9 and failed - the *cbfunc* will not be called

▼----- Required Attributes -----▼

10 PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition,
11 the following attributes are required to be included in the passed *info* array:

12 **PMIX_USERID** "**pmix.euid**" (**uint32_t**)
13 Effective user ID of the connecting process.

14 **PMIX_GRPID** "**pmix.egid**" (**uint32_t**)
15 Effective group ID of the connecting process.

17 Host environments that implement this entry point are required to support the following attributes:

18 **PMIX_RANGE** "**pmix.range**" (**pmix_data_range_t**)
19 Define constraints on the processes that can access the provided data. Only processes that meet the
20 constraints are allowed to access it.

▲----- Optional Attributes -----▼

21 The following attributes are optional for host environments that support this operation:

22 **PMIX_TIMEOUT** "**pmix.timeout**" (**int**)
23 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
24 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
25 (client, server, and host) simultaneously timing the operation.

Description

Delete data from the data store. The host server will be passed a **NULL**-terminated array of string keys, plus potential directives such as the data range within which the keys should be deleted. The default data range is left to the host environment, but expected to be **PMIX_RANGE_SESSION**. The callback is to be executed upon completion of the delete procedure.

Advice to PMIx server hosts

The **PMIX_USERID** and **PMIX_GRPID** of the requesting process will be provided to support authorization-based access to published information. The host environment is not required to guarantee support for any specific range - i.e., the environment does not need to return an error if the data store doesn't support a specified range so long as it is covered by some internally defined range.

17.3.11 pmix_server_spawn_fn_t

Summary

Spawn a set of applications/processes as per the **PMIx_Spawn** API.

Format

PMIx v1.0

C

```
typedef pmix_status_t (*pmix_server_spawn_fn_t) (  
    const pmix_proc_t *proc,  
    const pmix_info_t job_info[],  
    size_t ninfo,  
    const pmix_app_t apps[],  
    size_t napps,  
    pmix_spawn_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN proc**
pmix_proc_t structure of the process making the request (handle)
- IN job_info**
Array of info structures (array of handles)
- IN ninfo**
Number of elements in the *jobinfo* array (integer)
- IN apps**
Array of **pmix_app_t** structures (array of handles)
- IN napps**
Number of elements in the *apps* array (integer)
- IN cbfunc**
Callback function **pmix_spawn_cbfunc_t** (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

Returns one of the following:

- 1 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
2 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
3 from the API.
- 4 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
5 *success* - the *cbfunc* will not be called
- 6 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
7 though the function entry was provided in the server module - the *cbfunc* will not be called
- 8 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
9 and failed - the *cbfunc* will not be called

▼ ----- Required Attributes ----- ▼

10 PMIx server libraries are required to pass any provided attributes to the host environment for processing. In
11 addition, the following attributes are required to be included in the passed *info* array:

12 **PMIX_USERID** "pmix.euid" (uint32_t)
13 Effective user ID of the connecting process.

14 **PMIX_GRPID** "pmix.egid" (uint32_t)
15 Effective group ID of the connecting process.

16 **PMIX_SPAWNED** "pmix.spawned" (bool)
17 **true** if this process resulted from a call to **PMIx_Spawn**. Lack of inclusion (i.e., a return status of
18 **PMIX_ERR_NOT_FOUND**) corresponds to a value of **false** for this attribute.

19 **PMIX_PARENT_ID** "pmix.parent" (pmix_proc_t)
20 Process identifier of the parent process of the specified process - typically used to identify the
21 application process that caused the job containing the specified process to be spawned (e.g., the process
22 that called **PMIx_Spawn**). This attribute is only provided for a process if it was created by a call to
23 **PMIx_Spawn** or **PMIx_Spawn_nb**.

24 **PMIX_REQUESTOR_IS_TOOL** "pmix.req.tool" (bool)
25 The requesting process is a PMIx tool.

26 **PMIX_REQUESTOR_IS_CLIENT** "pmix.req.client" (bool)
27 The requesting process is a PMIx client.

28
29 Host environments that provide this module entry point are required to pass the **PMIX_SPAWNED** and
30 **PMIX_PARENT_ID** attributes to all PMIx servers launching new child processes so those values can be
31 returned to clients upon connection to the PMIx server. In addition, they are required to support the following
32 attributes when present in either the *job_info* or the *info* array of an element of the *apps* array:

33 **PMIX_WDIR** "pmix.wdir" (char*)
34 Working directory for spawned processes.

35 **PMIX_SET_SESSION_CWD** "pmix.ssn cwd" (bool)
36 Set the current working directory to the session working directory assigned by the RM - can be
37 assigned to the entire job (by including attribute in the *job_info* array) or on a per-application basis in
38 the *info* array for each **pmix_app_t**.

1 **PMIX_PREFIX** "pmix.prefix" (char*)
2 Prefix to use for starting spawned processes - i.e., the directory where the executables can be found.
3 **PMIX_HOST** "pmix.host" (char*)
4 Comma-delimited list of hosts to use for spawned processes.
5 **PMIX_HOSTFILE** "pmix.hostfile" (char*)
6 Hostfile to use for spawned processes.

▲-----▲
▼-----▼ **Optional Attributes** -----▼

7 The following attributes are optional for host environments that support this operation:

8 **PMIX_ADD_HOSTFILE** "pmix.addhostfile" (char*)
9 Hostfile containing hosts to add to existing allocation.
10 **PMIX_ADD_HOST** "pmix.addhost" (char*)
11 Comma-delimited list of hosts to add to the allocation.
12 **PMIX_PRELOAD_BIN** "pmix.preloadbin" (bool)
13 Preload executables onto nodes prior to executing launch procedure.
14 **PMIX_PRELOAD_FILES** "pmix.preloadfiles" (char*)
15 Comma-delimited list of files to pre-position on nodes prior to executing launch procedure.
16 **PMIX_PERSONALITY** "pmix.pers" (char*)
17 Name of personality corresponding to programming model used by application - supported values
18 depend upon PMIx implementation.
19 **PMIX_DISPLAY_MAP** "pmix.dispmap" (bool)
20 Display process mapping upon spawn.
21 **PMIX_PPR** "pmix.ppr" (char*)
22 Number of processes to spawn on each identified resource.
23 **PMIX_MAPBY** "pmix.mapby" (char*)
24 Process mapping policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
25 for the rank to discover the mapping policy used for the provided namespace. Supported values are
26 launcher specific.
27 **PMIX_RANKBY** "pmix.rankby" (char*)
28 Process ranking policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
29 for the rank to discover the ranking algorithm used for the provided namespace. Supported values are
30 launcher specific.
31 **PMIX_BINDTO** "pmix.bindto" (char*)
32 Process binding policy - when accessed using **PMIx_Get**, use the **PMIX_RANK_WILDCARD** value
33 for the rank to discover the binding policy used for the provided namespace. Supported values are
34 launcher specific.
35 **PMIX_STDIN_TGT** "pmix.stdin" (uint32_t)
36 Spawned process rank that is to receive any forwarded **stdin**.

1 **PMIX_FWD_STDIN** "pmix.fwd.stdin" (**pmix_rank_t**)
2 The requester intends to push information from its **stdin** to the indicated process. The local spawn
3 agent should, therefore, ensure that the **stdin** channel to that process remains available. A rank of
4 **PMIX_RANK_WILDCARD** indicates that all processes in the spawned job are potential recipients. The
5 requester will issue a call to **PMIx_IOF_push** to initiate the actual forwarding of information to
6 specified targets - this attribute simply requests that the IL retain the ability to forward the information
7 to the designated targets.

8 **PMIX_FWD_STDOUT** "pmix.fwd.stdout" (**bool**)
9 Requests that the ability to forward the **stdout** of the spawned processes be maintained. The
10 requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for
11 delivery of the forwarded output.

12 **PMIX_FWD_STDERR** "pmix.fwd.stderr" (**bool**)
13 Requests that the ability to forward the **stderr** of the spawned processes be maintained. The
14 requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for
15 delivery of the forwarded output.

16 **PMIX_DEBUGGER_DAEMONS** "pmix.debugger" (**bool**)
17 Included in the **pmix_info_t** array of a **pmix_app_t**, this attribute declares that the application
18 consists of debugger daemons and shall be governed accordingly. If used as the sole **pmix_app_t** in
19 a **PMIx_Spawn** request, then the **PMIX_DEBUG_TARGET** attribute must also be provided (in either
20 the *job_info* or in the *info* array of the **pmix_app_t**) to identify the namespace to be debugged so that
21 the launcher can determine where to place the spawned daemons. If neither
22 **PMIX_DEBUG_DAEMONS_PER_PROC** nor **PMIX_DEBUG_DAEMONS_PER_NODE** is specified, then
23 the launcher shall default to a placement policy of one daemon per process in the target job.

24 **PMIX_TAG_OUTPUT** "pmix.tagout" (**bool**)
25 Tag **stdout/stderr** with the identity of the source process - can be assigned to the entire job (by
26 including attribute in the *job_info* array) or on a per-application basis in the *info* array for each
27 **pmix_app_t**.

28 **PMIX_TIMESTAMP_OUTPUT** "pmix.tsout" (**bool**)
29 Timestamp output - can be assigned to the entire job (by including attribute in the *job_info* array) or on
30 a per-application basis in the *info* array for each **pmix_app_t**.

31 **PMIX_MERGE_STDERR_STDOUT** "pmix.mergeerrout" (**bool**)
32 Merge **stdout** and **stderr** streams - can be assigned to the entire job (by including attribute in the
33 *job_info* array) or on a per-application basis in the *info* array for each **pmix_app_t**.

34 **PMIX_OUTPUT_TO_FILE** "pmix.outfile" (**char***)
35 Direct output (both **stdout** and **stderr**) into files of form "**<filename>.rank**" - can be assigned to
36 the entire job (by including attribute in the *job_info* array) or on a per-application basis in the *info* array
37 for each **pmix_app_t**.

38 **PMIX_INDEX_ARGV** "pmix.indxargv" (**bool**)
39 Mark the **argv** with the rank of the process.

40 **PMIX_CPUS_PER_PROC** "pmix.cputerproc" (**uint32_t**)

1 Number of PUs to assign to each rank - when accessed using `PMIx_Get`, use the
2 `PMIX_RANK_WILDCARD` value for the rank to discover the PUs/process assigned to the provided
3 namespace.

4 `PMIX_NO_PROCS_ON_HEAD` "pmix.nolocal" (bool)

5 Do not place processes on the head node.

6 `PMIX_NO_OVERSUBSCRIBE` "pmix.noover" (bool)

7 Do not oversubscribe the nodes - i.e., do not place more processes than allocated slots on a node.

8 `PMIX_REPORT_BINDINGS` "pmix.repbind" (bool)

9 Report bindings of the individual processes.

10 `PMIX_CPU_LIST` "pmix.cpulist" (char*)

11 List of PUs to use for this job - when accessed using `PMIx_Get`, use the `PMIX_RANK_WILDCARD`
12 value for the rank to discover the PU list used for the provided namespace.

13 `PMIX_JOB_RECOVERABLE` "pmix.recover" (bool)

14 Application supports recoverable operations.

15 `PMIX_JOB_CONTINUOUS` "pmix.continuous" (bool)

16 Application is continuous, all failed processes should be immediately restarted.

17 `PMIX_MAX_RESTARTS` "pmix.maxrestarts" (uint32_t)

18 Maximum number of times to restart a process - when accessed using `PMIx_Get`, use the
19 `PMIX_RANK_WILDCARD` value for the rank to discover the max restarts for the provided namespace.

20 `PMIX_TIMEOUT` "pmix.timeout" (int)

21 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
22 `PMIX_ERR_TIMEOUT` error. Care should be taken to avoid race conditions caused by multiple layers
23 (client, server, and host) simultaneously timing the operation.

24 `PMIX_JOB_TIMEOUT` "pmix.job.time" (int)

25 Time in seconds before the spawned job should time out and be terminated (0 => infinite), defined as
26 the total runtime of the job (equivalent to the walltime limit of typical batch schedulers).

27 `PMIX_SPAWN_TIMEOUT` "pmix.sp.time" (int)

28 Time in seconds before spawn operation should time out (0 => infinite). Logically equivalent to
29 passing the `PMIX_TIMEOUT` attribute to the `PMIx_Spawn` API, it is provided as a separate attribute
30 to distinguish it from the `PMIX_JOB_TIMEOUT` attribute



31 Description

32 Spawn a set of applications/processes as per the `PMIx_Spawn` API. Note that applications are not required to
33 be MPI or any other programming model. Thus, the host server cannot make any assumptions as to their
34 required support. The callback function is to be executed once all processes have been started. An error in
35 starting any application or process in this request shall cause all applications and processes in the request to be
36 terminated, and an error returned to the originating caller.

37 Note that a timeout can be specified in the `job_info` array to indicate that failure to start the requested job
38 within the given time should result in termination to avoid hangs.

1 17.3.11.1 Server spawn attributes

- 2 **PMIX_REQUESTOR_IS_TOOL** "pmix.req.tool" (bool)
3 The requesting process is a PMIx tool.
- 4 **PMIX_REQUESTOR_IS_CLIENT** "pmix.req.client" (bool)
5 The requesting process is a PMIx client.

6 17.3.12 pmix_server_connect_fn_t

7 Summary

8 Record the specified processes as *connected*.

9 Format

PMIx v1.0

C

```
10 typedef pmix_status_t (*pmix_server_connect_fn_t) (  
11     const pmix_proc_t procs[],  
12     size_t nprocs,  
13     const pmix_info_t info[],  
14     size_t ninfo,  
15     pmix_op_cbfunc_t cbfunc,  
16     void *cbdata);
```

C

- 17 **IN procs**
18 Array of [pmix_proc_t](#) structures identifying participants (array of handles)
- 19 **IN nprocs**
20 Number of elements in the *procs* array (integer)
- 21 **IN info**
22 Array of info structures (array of handles)
- 23 **IN ninfo**
24 Number of elements in the *info* array (integer)
- 25 **IN cbfunc**
26 Callback function [pmix_op_cbfunc_t](#) (function reference)
- 27 **IN cbdata**
28 Data to be passed to the callback function (memory reference)

29 Returns one of the following:

- 30 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
31 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
32 from the API.
- 33 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
34 *success* - the *cbfunc* will not be called
- 35 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
36 though the function entry was provided in the server module - the *cbfunc* will not be called
- 37 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
38 and failed - the *cbfunc* will not be called

Required Attributes

PMIX_LOCAL_COLLECTIVE_STATUS "pmix.loc.col.st" (pmix_status_t)

Status code for local collective operation being reported to the host by the server library. PMIx servers may aggregate the participation by local client processes in a collective operation - e.g., instead of passing individual client calls to **PMIx_Fence** up to the host environment, the server may pass only a single call to the host when all local participants have executed their **PMIx_Fence** call, thereby reducing the burden placed on the host. However, in cases where the operation locally fails (e.g., if a participating client abnormally terminates prior to calling the operation), the server upcall functions to the host do not include a **pmix_status_t** by which the PMIx server can alert the host to that failure. This attribute resolves that problem by allowing the server to pass the status information regarding the local collective operation.

PMIx libraries are required to pass any provided attributes to the host environment for processing.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Record the processes specified by the *procs* array as *connected* as per the PMIx definition. The callback is to be executed once every daemon hosting at least one participant has called the host server's **pmix_server_connect_fn_t** function, and the host environment has completed any supporting operations required to meet the terms of the PMIx definition of *connected* processes.

Advice to PMIx library implementers

The PMIx server library is required to aggregate participation by local clients, passing the request to the host environment once all local participants have executed the API.

Advice to PMIx server hosts

The host will receive a single call for each collective operation. It is the responsibility of the host to identify the nodes containing participating processes, execute the collective across all participating nodes, and notify the local PMIx server library upon completion of the global collective.

17.3.13 pmix_server_disconnect_fn_t

Summary

Disconnect a previously connected set of processes.

Format

C

```
typedef pmix_status_t (*pmix_server_disconnect_fn_t) (  
    const pmix_proc_t procs[],  
    size_t nprocs,  
    const pmix_info_t info[],  
    size_t ninfo,  
    pmix_op_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN** `procs`
Array of `pmix_proc_t` structures identifying participants (array of handles)
- IN** `nprocs`
Number of elements in the `procs` array (integer)
- IN** `info`
Array of info structures (array of handles)
- IN** `ninfo`
Number of elements in the `info` array (integer)
- IN** `cbfunc`
Callback function `pmix_op_cbfunc_t` (function reference)
- IN** `cbdata`
Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided `cbfunc`. Note that the host must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned `success` - the `cbfunc` will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the `cbfunc` will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the `cbfunc` will not be called

Required Attributes

PMIX_LOCAL_COLLECTIVE_STATUS "pmix.loc.col.st" (`pmix_status_t`)

Status code for local collective operation being reported to the host by the server library. PMIx servers may aggregate the participation by local client processes in a collective operation - e.g., instead of passing individual client calls to **PMIx_Fence** up to the host environment, the server may pass only a single call to the host when all local participants have executed their **PMIx_Fence** call, thereby reducing the burden placed on the host. However, in cases where the operation locally fails (e.g., if a participating client abnormally terminates prior to calling the operation), the server upcall functions to the host do not include a `pmix_status_t` by which the PMIx server can alert the host to that failure.

1 This attribute resolves that problem by allowing the server to pass the status information regarding the
2 local collective operation.

3 PMIx libraries are required to pass any provided attributes to the host environment for processing.



4 **Optional Attributes**

5 The following attributes are optional for host environments that support this operation:

6 **PMIX_TIMEOUT** "pmix.timeout" (int)

7 Time in seconds before the specified operation should time out (zero indicating infinite) and return the
8 **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers
(client, server, and host) simultaneously timing the operation.



9 **Description**

10 Disconnect a previously connected set of processes. The callback is to be executed once every daemon hosting
11 at least one participant has called the host server's has called the **pmix_server_disconnect_fn_t**
12 function, and the host environment has completed any required supporting operations.

13 **Advice to PMIx library implementers**

14 The PMIx server library is required to aggregate participation by local clients, passing the request to the host
environment once all local participants have executed the API.



15 **Advice to PMIx server hosts**

16 The host will receive a single call for each collective operation. It is the responsibility of the host to identify
17 the nodes containing participating processes, execute the collective across all participating nodes, and notify
the local PMIx server library upon completion of the global collective.

18 A **PMIX_ERR_INVALID_OPERATION** error must be returned if the specified set of *procs* was not
19 previously *connected* via a call to the **pmix_server_connect_fn_t** function.



20 **17.3.14 pmix_server_register_events_fn_t**

21 **Summary**

22 Register to receive notifications for the specified events.

Format

```
typedef pmix_status_t (*pmix_server_register_events_fn_t) (  
    pmix_status_t *codes,  
    size_t ncodes,  
    const pmix_info_t info[],  
    size_t ninfo,  
    pmix_op_cbfunc_t cbfunc,  
    void *cbdata);
```

- IN codes**
Array of `pmix_status_t` values (array of handles)
- IN ncodes**
Number of elements in the `codes` array (integer)
- IN info**
Array of info structures (array of handles)
- IN ninfo**
Number of elements in the `info` array (integer)
- IN cbfunc**
Callback function `pmix_op_cbfunc_t` (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided `cbfunc`. Note that the host must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the `cbfunc` will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the `cbfunc` will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the `cbfunc` will not be called

Required Attributes

PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition, the following attributes are required to be included in the passed `info` array:

- PMIX_USERID** "pmix.euid" (`uint32_t`)
Effective user ID of the connecting process.
- PMIX_GRPID** "pmix.egid" (`uint32_t`)
Effective group ID of the connecting process.

Description

Register to receive notifications for the specified status codes. The *info* array included in this API is reserved for possible future directives to further steer notification.

Advice to PMIx library implementers

The PMIx server library must track all client registrations for subsequent notification. This module function shall only be called when:

- the client has requested notification of an environmental code (i.e., a PMIx codes in the range between `PMIX_EVENT_SYS_BASE` and `PMIX_EVENT_SYS_OTHER`, inclusive) or codes that lies outside the defined PMIx range of constants; and
- the PMIx server library has not previously requested notification of that code - i.e., the host environment is to be contacted only once a given unique code value

Advice to PMIx server hosts

The host environment is required to pass to its PMIx server library all non-environmental events that directly relate to a registered namespace without the PMIx server library explicitly requesting them. Environmental events are to be translated to their nearest PMIx equivalent code as defined in the range between `PMIX_EVENT_SYS_BASE` and `PMIX_EVENT_SYS_OTHER` (inclusive).

17.3.15 `pmix_server_deregister_events_fn_t`

Summary

Deregister to receive notifications for the specified events.

Format

```
typedef pmix_status_t (*pmix_server_deregister_events_fn_t) (  
    pmix_status_t *codes,  
    size_t ncodes,  
    pmix_op_cbfunc_t cbfunc,  
    void *cbdata);
```

IN codes

Array of `pmix_status_t` values (array of handles)

IN ncodes

Number of elements in the `codes` array (integer)

IN cbfunc

Callback function `pmix_op_cbfunc_t` (function reference)

IN cbdata

Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided `cbfunc`. Note that the host must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the `cbfunc` will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the `cbfunc` will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the `cbfunc` will not be called

Description

Deregister to receive notifications for the specified events to which the PMIx server has previously registered.

Advice to PMIx library implementers

The PMIx server library must track all client registrations. This module function shall only be called when:

- the library is deregistering environmental codes (i.e., a PMIx codes in the range between **PMIX_EVENT_SYS_BASE** and **PMIX_EVENT_SYS_OTHER**, inclusive) or codes that lies outside the defined PMIx range of constants; and
- no client (including the server library itself) remains registered for notifications on any included code - i.e., a code should be included in this call only when no registered notifications against it remain.

1 17.3.16 pmix_server_notify_event_fn_t

2 Summary

3 Notify the specified processes of an event.

4 Format

PMIx v2.0

C

```
5 typedef pmix_status_t (*pmix_server_notify_event_fn_t) (  
6     pmix_status_t code,  
7     const pmix_proc_t *source,  
8     pmix_data_range_t range,  
9     pmix_info_t info[],  
10    size_t ninfo,  
11    pmix_op_cbfunc_t cbfunc,  
12    void *cbdata);
```

C

13 **IN code**
14 The [pmix_status_t](#) event code being referenced structure (handle)

15 **IN source**
16 [pmix_proc_t](#) of process that generated the event (handle)

17 **IN range**
18 [pmix_data_range_t](#) range over which the event is to be distributed (handle)

19 **IN info**
20 Optional array of [pmix_info_t](#) structures containing additional information on the event (array of
21 handles)

22 **IN ninfo**
23 Number of elements in the *info* array (integer)

24 **IN cbfunc**
25 Callback function [pmix_op_cbfunc_t](#) (function reference)

26 **IN cbdata**
27 Data to be passed to the callback function (memory reference)

28 Returns one of the following:

- 29 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
30 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
31 from the API.
- 32 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
33 *success* - the *cbfunc* will not be called
- 34 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
35 though the function entry was provided in the server module - the *cbfunc* will not be called
- 36 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
37 and failed - the *cbfunc* will not be called

Required Attributes

PMIx libraries are required to pass any provided attributes to the host environment for processing.
Host environments that provide this module entry point are required to support the following attributes:

PMIX_RANGE "pmix.range" (**pmix_data_range_t**)

Define constraints on the processes that can access the provided data. Only processes that meet the constraints are allowed to access it.

Description

Notify the specified processes (described through a combination of *range* and attributes provided in the *info* array) of an event generated either by the PMIx server itself or by one of its local clients. The process generating the event is provided in the *source* parameter, and any further descriptive information is included in the *info* array.

Note that the PMIx server library is not allowed to echo any event given to it by its host via the **PMIx_Notify_event** API back to the host through the **pmix_server_notify_event_fn_t** server module function.

Advice to PMIx server hosts

The callback function is to be executed once the host environment no longer requires that the PMIx server library maintain the provided data structures. It does not necessarily indicate that the event has been delivered to any process, nor that the event has been distributed for delivery

17.3.17 pmix_server_listener_fn_t

Summary

Register a socket the host server can monitor for connection requests.

Format

C

```
typedef pmix_status_t (*pmix_server_listener_fn_t) (  
    int listening_sd,  
    pmix_connection_cbfunc_t cbfunc,  
    void *cbdata);
```

C

IN **incoming_sd**

(integer)

IN **cbfunc**

Callback function **pmix_connection_cbfunc_t** (function reference)

IN **cbdata**

(memory reference)

Returns **PMIX_SUCCESS** indicating that the request is accepted, or a negative value corresponding to a PMIx error constant indicating that the request has been rejected.

Description

Register a socket the host environment can monitor for connection requests, harvest them, and then call the PMIx server library's internal callback function for further processing. A listener thread is essential to efficiently harvesting connection requests from large numbers of local clients such as occur when running on large SMPs. The host server listener is required to call `accept` on the incoming connection request, and then pass the resulting socket to the provided `cbfunc`. A `NULL` for this function will cause the internal PMIx server to spawn its own listener thread.

17.3.17.1 PMIx Client Connection Callback Function

Summary

Callback function for incoming connection request from a local client.

Format

PMIx v1.0

```
typedef void (*pmix_connection_cbfunc_t) (  
    int incoming_sd, void *cbdata);
```

IN `incoming_sd`
(integer)
IN `cbdata`
(memory reference)

Description

Callback function for incoming connection requests from local clients - only used by host environments that wish to directly handle socket connection requests.

17.3.18 `pmix_server_query_fn_t`

Summary

Query information from the resource manager.

Format

PMIx v2.0

```
typedef pmix_status_t (*pmix_server_query_fn_t) (  
    pmix_proc_t *proct,  
    pmix_query_t *queries,  
    size_t nqueries,  
    pmix_info_cbfunc_t cbfunc,  
    void *cbdata);
```

IN `proct`
`pmix_proc_t` structure of the requesting process (handle)
IN `queries`
Array of `pmix_query_t` structures (array of handles)

1 **IN** **nqueries**
 2 Number of elements in the *queries* array (integer)
 3 **IN** **cbfunc**
 4 Callback function `pmix_info_cbfunc_t` (function reference)
 5 **IN** **cbdata**
 6 Data to be passed to the callback function (memory reference)

7 Returns one of the following:

- 8 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning from the API.
- 11 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called
- 13 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the *cbfunc* will not be called
- 15 • a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the *cbfunc* will not be called

▼----- Required Attributes -----▼

17 PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition, the following attributes are required to be included in the passed *info* array:

19 **PMIX_USERID** "pmix.euid" (uint32_t)
 20 Effective user ID of the connecting process.

21 **PMIX_GRPID** "pmix.egid" (uint32_t)
 22 Effective group ID of the connecting process.

▲-----

▼----- Optional Attributes -----▼

23 The following attributes are optional for host environments that support this operation:

24 **PMIX_QUERY_NAMESPACES** "pmix.qry.ns" (char*)
 25 Request a comma-delimited list of active namespaces. NO QUALIFIERS.

26 **PMIX_QUERY_JOB_STATUS** "pmix.qry.jst" (pmix_status_t)
 27 Status of a specified, currently executing job. REQUIRED QUALIFIER: **PMIX_NAMESPACE** indicating the namespace whose status is being queried.

29 **PMIX_QUERY_QUEUE_LIST** "pmix.qry.qlst" (char*)
 30 Request a comma-delimited list of scheduler queues. NO QUALIFIERS.

31 **PMIX_QUERY_QUEUE_STATUS** "pmix.qry.qst" (char*)
 32 Returns status of a specified scheduler queue, expressed as a string. OPTIONAL QUALIFIERS:
 33 **PMIX_ALLOC_QUEUE** naming specific queue whose status is being requested.

34 **PMIX_QUERY_PROC_TABLE** "pmix.qry.ptable" (char*)

1 Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in
2 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: `PMIX_NAMESPACE`
3 indicating the namespace whose process table is being queried.

4 `PMIX_QUERY_LOCAL_PROC_TABLE` "`pmix.qry.lptable`" (`char*`)

5 Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in
6 the specified namespace executing on the same node as the requester, ordered by process job rank.
7 REQUIRED QUALIFIER: `PMIX_NAMESPACE` indicating the namespace whose local process table is
8 being queried. OPTIONAL QUALIFIER: `PMIX_HOSTNAME` indicating the host whose local process
9 table is being queried. By default, the query assumes that the host upon which the request was made is
10 to be used.

11 `PMIX_QUERY_SPAWN_SUPPORT` "`pmix.qry.spawn`" (`bool`)

12 Return a comma-delimited list of supported spawn attributes. NO QUALIFIERS.

13 `PMIX_QUERY_DEBUG_SUPPORT` "`pmix.qry.debug`" (`bool`)

14 Return a comma-delimited list of supported debug attributes. NO QUALIFIERS.

15 `PMIX_QUERY_MEMORY_USAGE` "`pmix.qry.mem`" (`bool`)

16 Return information on memory usage for the processes indicated in the qualifiers. OPTIONAL
17 QUALIFIERS: `PMIX_NAMESPACE` and `PMIX_RANK`, or `PMIX_PROCID` of specific process(es) whose
18 memory usage is being requested.

19 `PMIX_QUERY_LOCAL_ONLY` "`pmix.qry.local`" (`bool`)

20 Constrain the query to local information only. NO QUALIFIERS.

21 `PMIX_QUERY_REPORT_AVG` "`pmix.qry.avg`" (`bool`)

22 Report only average values for sampled information. NO QUALIFIERS.

23 `PMIX_QUERY_REPORT_MINMAX` "`pmix.qry.minmax`" (`bool`)

24 Report minimum and maximum values. NO QUALIFIERS.

25 `PMIX_QUERY_ALLOC_STATUS` "`pmix.query.alloc`" (`char*`)

26 String identifier of the allocation whose status is being requested. NO QUALIFIERS.

27 `PMIX_TIME_REMAINING` "`pmix.time.remaining`" (`char*`)

28 Query number of seconds (`uint32_t`) remaining in allocation for the specified namespace.
29 OPTIONAL QUALIFIERS: `PMIX_NAMESPACE` of the namespace whose info is being requested (defaults
30 to allocation containing the caller).



31 Description

32 Query information from the host environment. The query will include the namespace/rank of the process that
33 is requesting the info, an array of `pmix_query_t` describing the request, and a callback function/data for the
34 return.



Advice to PMIx library implementers

35 The PMIx server library should not block in this function as the host environment may, depending upon the
36 information being requested, require significant time to respond.



1 17.3.19 pmix_server_tool_connection_fn_t

2 Summary

3 Register that a tool has connected to the server.

4 Format

PMIx v2.0

C

```
5 typedef void (*pmix_server_tool_connection_fn_t) (  
6             pmix_info_t info[], size_t ninfo,  
7             pmix_tool_connection_cbfnc_t cbfunc,  
8             void *cbdata);
```

C

9 **IN info**

10 Array of `pmix_info_t` structures (array of handles)

11 **IN ninfo**

12 Number of elements in the *info* array (integer)

13 **IN cbfunc**

14 Callback function `pmix_tool_connection_cbfnc_t` (function reference)

15 **IN cbdata**

16 Data to be passed to the callback function (memory reference)

Required Attributes

17 PMIx libraries are required to pass the following attributes in the *info* array:

18 **PMIX_USERID** "pmix.euid" (uint32_t)

19 Effective user ID of the connecting process.

20 **PMIX_GRPID** "pmix.egid" (uint32_t)

21 Effective group ID of the connecting process.

22 **PMIX_TOOL_NAMESPACE** "pmix.tool.namespace" (char*)

23 Name of the namespace to use for this tool. This must be included only if the tool already has an
24 assigned namespace.

25 **PMIX_TOOL_RANK** "pmix.tool.rank" (uint32_t)

26 Rank of this tool. This must be included only if the tool already has an assigned rank.

27 **PMIX_CREDENTIAL** "pmix.cred" (char*)

28 Security credential assigned to the process.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_FWD_STDOUT "pmix.fwd.stdout" (bool)

Requests that the ability to forward the **stdout** of the spawned processes be maintained. The requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for delivery of the forwarded output.

PMIX_FWD_STDERR "pmix.fwd.stderr" (bool)

Requests that the ability to forward the **stderr** of the spawned processes be maintained. The requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for delivery of the forwarded output.

PMIX_FWD_STDIN "pmix.fwd.stdin" (pmix_rank_t)

The requester intends to push information from its **stdin** to the indicated process. The local spawn agent should, therefore, ensure that the **stdin** channel to that process remains available. A rank of **PMIX_RANK_WILDCARD** indicates that all processes in the spawned job are potential recipients. The requester will issue a call to **PMIx_IOF_push** to initiate the actual forwarding of information to specified targets - this attribute simply requests that the IL retain the ability to forward the information to the designated targets.

PMIX_VERSION_INFO "pmix.version" (char*)

PMIx version of the library being used by the connecting process.

Description

Register that a tool has connected to the server, possibly requesting that the tool be assigned a namespace/rank identifier for further interactions. The **pmix_info_t** array is used to pass qualifiers for the connection request, including the effective uid and gid of the calling tool for authentication purposes.

If the tool already has an assigned process identifier, then this must be indicated in the *info* array. The host is responsible for checking that the provided namespace does not conflict with any currently known assignments, returning an appropriate error in the callback function if a conflict is found.

The host environment is solely responsible for authenticating and authorizing the connection using whatever means it deems appropriate. If certificates or other authentication information are required, then the tool must provide them. The conclusion of those operations shall be communicated back to the PMIx server library via the callback function.

Approval or rejection of the connection request shall be returned in the *status* parameter of the **pmix_tool_connection_cbfunc_t**. If the connection is refused, the PMIx server library must terminate the connection attempt. The host must not execute the callback function prior to returning from the API.

1 17.3.19.1 Tool connection attributes

2 Attributes associated with tool connections.

3 **PMIX_USERID** "pmix.euid" (uint32_t)

4 Effective user ID of the connecting process.

5 **PMIX_GRPID** "pmix.egid" (uint32_t)

6 Effective group ID of the connecting process.

7 **PMIX_VERSION_INFO** "pmix.version" (char*)

8 PMIx version of the library being used by the connecting process.

9 17.3.19.2 PMIx Tool Connection Callback Function

10 Summary

11 Callback function for incoming tool connections.

12 Format

PMIx v2.0

C

```
13 typedef void (*pmix_tool_connection_cbfunc_t) (  
14             pmix_status_t status,  
15             pmix_proc_t *proc, void *cbdata);
```

C

16 **IN** status

17 [pmix_status_t](#) value (handle)

18 **IN** proc

19 [pmix_proc_t](#) structure containing the identifier assigned to the tool (handle)

20 **IN** cbdata

21 Data to be passed (memory reference)

22 Description

23 Callback function for incoming tool connections. The host environment shall provide a namespace/rank
24 identifier for the connecting tool.

Advice to PMIx server hosts

25 It is assumed that **rank=0** will be the normal assignment, but allow for the future possibility of a parallel set
26 of tools connecting, and thus each process requiring a unique rank.

27 17.3.20 pmix_server_log_fn_t

28 Summary

29 Log data on behalf of a client.

Format

```
typedef void (*pmix_server_log_fn_t) (  
    const pmix_proc_t *client,  
    const pmix_info_t data[], size_t ndata,  
    const pmix_info_t directives[], size_t ndirs,  
    pmix_op_cbfunc_t cbfunc, void *cbdata);
```

IN client
[pmix_proc_t](#) structure (handle)

IN data
Array of info structures (array of handles)

IN ndata
Number of elements in the *data* array (integer)

IN directives
Array of info structures (array of handles)

IN ndirs
Number of elements in the *directives* array (integer)

IN cbfunc
Callback function [pmix_op_cbfunc_t](#) (function reference)

IN cbdata
Data to be passed to the callback function (memory reference)

Required Attributes

PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition, the following attributes are required to be included in the passed *info* array:

PMIX_USERID "pmix.euid" ([uint32_t](#))
Effective user ID of the connecting process.

PMIX_GRPID "pmix.egid" ([uint32_t](#))
Effective group ID of the connecting process.

Host environments that provide this module entry point are required to support the following attributes:

PMIX_LOG_STDERR "pmix.log.stderr" ([char*](#))
Log string to *stderr*.

PMIX_LOG_STDOUT "pmix.log.stdout" ([char*](#))
Log string to *stdout*.

PMIX_LOG_SYSLOG "pmix.log.syslog" ([char*](#))
Log data to syslog. Defaults to **ERROR** priority. Will log to global syslog if available, otherwise to local syslog.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_LOG_MSG "pmix.log.msg" (`pmix_byte_object_t`)

Message blob to be sent somewhere.

PMIX_LOG_EMAIL "pmix.log.email" (`pmix_data_array_t`)

Log via email based on `pmix_info_t` containing directives.

PMIX_LOG_EMAIL_ADDR "pmix.log.emaddr" (`char*`)

Comma-delimited list of email addresses that are to receive the message.

PMIX_LOG_EMAIL_SUBJECT "pmix.log.emsub" (`char*`)

Subject line for email.

PMIX_LOG_EMAIL_MSG "pmix.log.emmsg" (`char*`)

Message to be included in email.

Description

Log data on behalf of a client. This function is not intended for output of computational results, but rather for reporting status and error messages. The host must not execute the callback function prior to returning from the API.

17.3.21 pmix_server_alloc_fn_t

Summary

Request allocation operations on behalf of a client.

Format

PMIx v2.0

```
typedef pmix_status_t (*pmix_server_alloc_fn_t) (  
    const pmix_proc_t *client,  
    pmix_alloc_directive_t directive,  
    const pmix_info_t data[],  
    size_t ndata,  
    pmix_info_cbfunc_t cbfunc,  
    void *cbdata);
```

- IN client**
`pmix_proc_t` structure of process making request (handle)
- IN directive**
Specific action being requested (`pmix_alloc_directive_t`)
- IN data**
Array of info structures (array of handles)
- IN ndata**
Number of elements in the *data* array (integer)

1 **IN** **cbfunc**
2 Callback function `pmix_info_cbfunc_t` (function reference)
3 **IN** **cbdata**
4 Data to be passed to the callback function (memory reference)

5 Returns one of the following:

- 6 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
7 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
8 from the API.
- 9 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
10 *success* - the *cbfunc* will not be called
- 11 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
12 though the function entry was provided in the server module - the *cbfunc* will not be called
- 13 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
14 and failed - the *cbfunc* will not be called

▼----- Required Attributes -----▼

15 PMIx libraries are required to pass any provided attributes to the host environment for processing. In addition,
16 the following attributes are required to be included in the passed *info* array:

17 **PMIX_USERID** "`pmix.euid`" (`uint32_t`)
18 Effective user ID of the connecting process.
19 **PMIX_GRPID** "`pmix.egid`" (`uint32_t`)
20 Effective group ID of the connecting process.

22 Host environments that provide this module entry point are required to support the following attributes:

23 **PMIX_ALLOC_ID** "`pmix.alloc.id`" (`char*`)
24 A string identifier (provided by the host environment) for the resulting allocation which can later be
25 used to reference the allocated resources in, for example, a call to **PMIx_Spawn**.
26 **PMIX_ALLOC_NUM_NODES** "`pmix.alloc.nnodes`" (`uint64_t`)
27 The number of nodes being requested in an allocation request.
28 **PMIX_ALLOC_NUM_CPUS** "`pmix.alloc.ncpus`" (`uint64_t`)
29 Number of PUs being requested in an allocation request.
30 **PMIX_ALLOC_TIME** "`pmix.alloc.time`" (`uint32_t`)
31 Total session time (in seconds) being requested in an allocation request.

▲-----

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_ALLOC_NODE_LIST "pmix.alloc.nlist" (char*)

Regular expression of the specific nodes being requested in an allocation request.

PMIX_ALLOC_NUM_CPU_LIST "pmix.alloc.ncpulist" (char*)

Regular expression of the number of PUs for each node being requested in an allocation request.

PMIX_ALLOC_CPU_LIST "pmix.alloc.cpulist" (char*)

Regular expression of the specific PUs being requested in an allocation request.

PMIX_ALLOC_MEM_SIZE "pmix.alloc.msize" (float)

Number of Megabytes[base2] of memory (per process) being requested in an allocation request.

PMIX_ALLOC_FABRIC "pmix.alloc.net" (array)

Array of `pmix_info_t` describing requested fabric resources. This must include at least:

PMIX_ALLOC_FABRIC_ID, **PMIX_ALLOC_FABRIC_TYPE**, and

PMIX_ALLOC_FABRIC_ENDPTS, plus whatever other descriptors are desired.

PMIX_ALLOC_FABRIC_ID "pmix.alloc.netid" (char*)

The key to be used when accessing this requested fabric allocation. The fabric allocation will be returned/stored as a `pmix_data_array_t` of `pmix_info_t` whose first element is composed of this key and the allocated resource description. The type of the included value depends upon the fabric support. For example, a TCP allocation might consist of a comma-delimited string of socket ranges such as "32000-32100, 33005, 38123-38146". Additional array entries will consist of any provided resource request directives, along with their assigned values. Examples include:

PMIX_ALLOC_FABRIC_TYPE - the type of resources provided; **PMIX_ALLOC_FABRIC_PLANE** -

if applicable, what plane the resources were assigned from; **PMIX_ALLOC_FABRIC_QOS** - the

assigned QoS; **PMIX_ALLOC_BANDWIDTH** - the allocated bandwidth;

PMIX_ALLOC_FABRIC_SEC_KEY - a security key for the requested fabric allocation. NOTE: the

array contents may differ from those requested, especially if **PMIX_INFO_REQD** was not set in the request.

PMIX_ALLOC_BANDWIDTH "pmix.alloc.bw" (float)

Fabric bandwidth (in Megabits[base2]/sec) for the job being requested in an allocation request.

PMIX_ALLOC_FABRIC_QOS "pmix.alloc.netqos" (char*)

Fabric quality of service level for the job being requested in an allocation request.

Description

Request new allocation or modifications to an existing allocation on behalf of a client. Several broad categories are envisioned, including the ability to:

- Request allocation of additional resources, including memory, bandwidth, and compute for an existing allocation. Any additional allocated resources will be considered as part of the current allocation, and thus will be released at the same time.
- Request a new allocation of resources. Note that the new allocation will be disjoint from (i.e., not affiliated with) the allocation of the requestor - thus the termination of one allocation will not impact the other.
- Extend the reservation on currently allocated resources, subject to scheduling availability and priorities.
- Return no-longer-required resources to the scheduler. This includes the *loan* of resources back to the scheduler with a promise to return them upon subsequent request.

The callback function provides a *status* to indicate whether or not the request was granted, and to provide some information as to the reason for any denial in the `pmix_info_cbfunc_t` array of `pmix_info_t` structures.

17.3.22 pmix_server_job_control_fn_t

Summary

Execute a job control action on behalf of a client.

Format

PMIx v2.0

C

```
typedef pmix_status_t (*pmix_server_job_control_fn_t) (  
    const pmix_proc_t *requestor,  
    const pmix_proc_t targets[],  
    size_t ntargets,  
    const pmix_info_t directives[],  
    size_t ndirs,  
    pmix_info_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN requestor**
`pmix_proc_t` structure of requesting process (handle)
- IN targets**
Array of proc structures (array of handles)
- IN ntargets**
Number of elements in the *targets* array (integer)
- IN directives**
Array of info structures (array of handles)
- IN ndirs**
Number of elements in the *info* array (integer)
- IN cbfunc**
Callback function `pmix_info_cbfunc_t` (function reference)
- IN cbdata**
Data to be passed to the callback function (memory reference)

1 Returns one of the following:

- 2 ● **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
3 returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning
4 from the API.
- 5 ● **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
6 *success* - the *cbfunc* will not be called
- 7 ● **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
8 though the function entry was provided in the server module - the *cbfunc* will not be called
- 9 ● a PMIx error constant indicating either an error in the input or that the request was immediately processed
10 and failed - the *cbfunc* will not be called

Required Attributes

11 PMIx libraries are required to pass any attributes provided by the client to the host environment for processing.
12 In addition, the following attributes are required to be included in the passed *info* array:

13 **PMIX_USERID** "pmix.euid" (uint32_t)

14 Effective user ID of the connecting process.

15 **PMIX_GRPID** "pmix.egid" (uint32_t)

16 Effective group ID of the connecting process.

17
18 Host environments that provide this module entry point are required to support the following attributes:

19 **PMIX_JOB_CTRL_ID** "pmix.jctrl.id" (char*)

20 Provide a string identifier for this request. The user can provide an identifier for the requested
21 operation, thus allowing them to later request status of the operation or to terminate it. The host,
22 therefore, shall track it with the request for future reference.

23 **PMIX_JOB_CTRL_PAUSE** "pmix.jctrl.pause" (bool)

24 Pause the specified processes.

25 **PMIX_JOB_CTRL_RESUME** "pmix.jctrl.resume" (bool)

26 Resume ("un-pause") the specified processes.

27 **PMIX_JOB_CTRL_KILL** "pmix.jctrl.kill" (bool)

28 Forcibly terminate the specified processes and cleanup.

29 **PMIX_JOB_CTRL_SIGNAL** "pmix.jctrl.sig" (int)

30 Send given signal to specified processes.

31 **PMIX_JOB_CTRL_TERMINATE** "pmix.jctrl.term" (bool)

32 Politely terminate the specified processes.

Optional Attributes

The following attributes are optional for host environments that support this operation:

- PMIX_JOB_CTRL_CANCEL** "pmix.jctrl.cancel" (char*)
Cancel the specified request - the provided request ID must match the **PMIX_JOB_CTRL_ID** provided to a previous call to **PMIx_Job_control**. An ID of **NULL** implies cancel all requests from this requestor.
- PMIX_JOB_CTRL_RESTART** "pmix.jctrl.restart" (char*)
Restart the specified processes using the given checkpoint ID.
- PMIX_JOB_CTRL_CHECKPOINT** "pmix.jctrl.ckpt" (char*)
Checkpoint the specified processes and assign the given ID to it.
- PMIX_JOB_CTRL_CHECKPOINT_EVENT** "pmix.jctrl.ckptev" (bool)
Use event notification to trigger a process checkpoint.
- PMIX_JOB_CTRL_CHECKPOINT_SIGNAL** "pmix.jctrl.ckptsig" (int)
Use the given signal to trigger a process checkpoint.
- PMIX_JOB_CTRL_CHECKPOINT_TIMEOUT** "pmix.jctrl.ckptsig" (int)
Time in seconds to wait for a checkpoint to complete.
- PMIX_JOB_CTRL_CHECKPOINT_METHOD** "pmix.jctrl.ckmethod" (pmix_data_array_t)
Array of **pmix_info_t** declaring each method and value supported by this application.
- PMIX_JOB_CTRL_PROVISION** "pmix.jctrl.pvn" (char*)
Regular expression identifying nodes that are to be provisioned.
- PMIX_JOB_CTRL_PROVISION_IMAGE** "pmix.jctrl.pvning" (char*)
Name of the image that is to be provisioned.
- PMIX_JOB_CTRL_PREEMPTIBLE** "pmix.jctrl.preempt" (bool)
Indicate that the job can be pre-empted.

Description

Execute a job control action on behalf of a client. The *targets* array identifies the processes to which the requested job control action is to be applied. A **NULL** value can be used to indicate all processes in the caller's namespace. The use of **PMIX_RANK_WILDCARD** can also be used to indicate that all processes in the given namespace are to be included.

The directives are provided as **pmix_info_t** structures in the *directives* array. The callback function provides a *status* to indicate whether or not the request was granted, and to provide some information as to the reason for any denial in the **pmix_info_cbfunc_t** array of **pmix_info_t** structures.

17.3.23 pmix_server_monitor_fn_t

Summary

Request that a client be monitored for activity.

Format

C

```
typedef pmix_status_t (*pmix_server_monitor_fn_t) (  
    const pmix_proc_t *requestor,  
    const pmix_info_t *monitor,  
    pmix_status_t error,  
    const pmix_info_t directives[],  
    size_t ndirs,  
    pmix_info_cbfunc_t cbfunc,  
    void *cbdata);
```

C

IN requestor
 pmix_proc_t structure of requesting process (handle)

IN monitor
 pmix_info_t identifying the type of monitor being requested (handle)

IN error
 Status code to use in generating event if alarm triggers (integer)

IN directives
 Array of info structures (array of handles)

IN ndirs
 Number of elements in the *info* array (integer)

IN cbfunc
 Callback function **pmix_info_cbfunc_t** (function reference)

IN cbdata
 Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the host must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the *cbfunc* will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the *cbfunc* will not be called

This entry point is only called for monitoring requests that are not directly supported by the PMIx server library itself.

Required Attributes

1 If supported by the PMI_x server library, then the library must not pass any supported attributes to the host
2 environment. Any attributes provided by the client that are not directly supported by the server library must be
3 passed to the host environment if it provides this module entry. In addition, the following attributes are
4 required to be included in the passed *info* array:

5 **PMIX_USERID** "pmix.euid" (uint32_t)

6 Effective user ID of the connecting process.

7 **PMIX_GRPID** "pmix.egid" (uint32_t)

8 Effective group ID of the connecting process.

9 Host environments are not required to support any specific monitoring attributes.

Optional Attributes

10 The following attributes may be implemented by a host environment.

11 **PMIX_MONITOR_ID** "pmix.monitor.id" (char*)

12 Provide a string identifier for this request.

13 **PMIX_MONITOR_CANCEL** "pmix.monitor.cancel" (char*)

14 Identifier to be canceled (**NULL** means cancel all monitoring for this process).

15 **PMIX_MONITOR_APP_CONTROL** "pmix.monitor.appctrl" (bool)

16 The application desires to control the response to a monitoring event - i.e., the application is requesting
17 that the host environment not take immediate action in response to the event (e.g., terminating the job).
18

19 **PMIX_MONITOR_HEARTBEAT** "pmix.monitor.mbeat" (void)

20 Register to have the PMI_x server monitor the requestor for heartbeats.

21 **PMIX_MONITOR_HEARTBEAT_TIME** "pmix.monitor.btime" (uint32_t)

22 Time in seconds before declaring heartbeat missed.

23 **PMIX_MONITOR_HEARTBEAT_DROPS** "pmix.monitor.bdrop" (uint32_t)

24 Number of heartbeats that can be missed before generating the event.

25 **PMIX_MONITOR_FILE** "pmix.monitor.fmon" (char*)

26 Register to monitor file for signs of life.

27 **PMIX_MONITOR_FILE_SIZE** "pmix.monitor.fsize" (bool)

28 Monitor size of given file is growing to determine if the application is running.

29 **PMIX_MONITOR_FILE_ACCESS** "pmix.monitor.faccess" (char*)

30 Monitor time since last access of given file to determine if the application is running.

31 **PMIX_MONITOR_FILE_MODIFY** "pmix.monitor.fmod" (char*)

32 Monitor time since last modified of given file to determine if the application is running.

33 **PMIX_MONITOR_FILE_CHECK_TIME** "pmix.monitor.ftime" (uint32_t)

34 Time in seconds between checking the file.

1 **PMIX_MONITOR_FILE_DROPS** "pmix.monitor.fdrop" (uint32_t)
2 Number of file checks that can be missed before generating the event.



3 **Description**

4 Request that a client be monitored for activity.

5 **17.3.24 pmix_server_get_cred_fn_t**

6 **Summary**

7 Request a credential from the host environment.

8 **Format**

PMIx v3.0

C

```
9 typedef pmix_status_t (*pmix_server_get_cred_fn_t) (  
10     const pmix_proc_t *proc,  
11     const pmix_info_t directives[],  
12     size_t ndirs,  
13     pmix_credential_cbfunc_t cbfunc,  
14     void *cbdata);
```

C

15 **IN proc**

pmix_proc_t structure of requesting process (handle)

17 **IN directives**

Array of info structures (array of handles)

19 **IN ndirs**

Number of elements in the *info* array (integer)

21 **IN cbfunc**

Callback function to return the credential (**pmix_credential_cbfunc_t** function reference)

23 **IN cbdata**

Data to be passed to the callback function (memory reference)

- 25 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
26 returned in the provided *cbfunc*
- 27 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
28 though the function entry was provided in the server module - the *cbfunc* will not be called
- 29 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
30 and failed - the *cbfunc* will not be called

Required Attributes

If the PMIx library does not itself provide the requested credential, then it is required to pass any attributes provided by the client to the host environment for processing. In addition, it must include the following attributes in the passed *info* array:

PMIX_USERID "pmix.euid" (uint32_t)

Effective user ID of the connecting process.

PMIX_GRPID "pmix.egid" (uint32_t)

Effective group ID of the connecting process.

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_CRED_TYPE "pmix.sec.ctype" (char*)

When passed in **PMIx_Get_credential**, a prioritized, comma-delimited list of desired credential types for use in environments where multiple authentication mechanisms may be available. When returned in a callback function, a string identifier of the credential type.

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Request a credential from the host environment.

17.3.24.1 Credential callback function

Summary

Callback function to return a requested security credential

Format

```
typedef void (*pmix_credential_cbfunc_t) (  
    pmix_status_t status,  
    pmix_byte_object_t *credential,  
    pmix_info_t info[], size_t ninfo,  
    void *cbdata);
```

- IN status**
[pmix_status_t](#) value (handle)
- IN credential**
[pmix_byte_object_t](#) structure containing the security credential (handle)
- IN info**
Array of provided by the system to pass any additional information about the credential - e.g., the identity of the issuing agent. (handle)
- IN ninfo**
Number of elements in *info* (**size_t**)
- IN cbdata**
Object passed in original request (memory reference)

Description

Define a callback function to return a requested security credential. Information provided by the issuing agent can subsequently be used by the application for a variety of purposes. Examples include:

- checking identified authorizations to determine what requests/operations are feasible as a means to steering [workflows](#)
- compare the credential type to that of the local SMS for compatibility

Advice to users

The credential is opaque and therefore understandable only by a service compatible with the issuer. The *info* array is owned by the PMIx library and is not to be released or altered by the receiving party.

17.3.25 pmix_server_validate_cred_fn_t

Summary

Request validation of a credential.

Format

C

```
typedef pmix_status_t (*pmix_server_validate_cred_fn_t) (  
    const pmix_proc_t *proc,  
    const pmix_byte_object_t *cred,  
    const pmix_info_t directives[],  
    size_t ndirs,  
    pmix_validation_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN** `proc`
`pmix_proc_t` structure of requesting process (handle)
- IN** `cred`
Pointer to `pmix_byte_object_t` containing the credential (handle)
- IN** `directives`
Array of info structures (array of handles)
- IN** `ndirs`
Number of elements in the `info` array (integer)
- IN** `cbfunc`
Callback function to return the result (`pmix_validation_cbfunc_t` function reference)
- IN** `cbdata`
Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided `cbfunc`
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned `success` - the `cbfunc` will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the `cbfunc` will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the `cbfunc` will not be called

Required Attributes

If the PMIx library does not itself validate the credential, then it is required to pass any attributes provided by the client to the host environment for processing. In addition, it must include the following attributes in the passed `info` array:

- PMIX_USERID** "pmix.euid" (`uint32_t`)
Effective user ID of the connecting process.
- PMIX_GRPID** "pmix.egid" (`uint32_t`)
Effective group ID of the connecting process.

1
2 Host environments are not required to support any specific attributes.

3
4
5
6
7

Optional Attributes

The following attributes are optional for host environments that support this operation:

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the **PMIX_ERR_TIMEOUT** error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

8
9
10

Description

Request validation of a credential obtained from the host environment via a prior call to the **pmix_server_get_cred_fn_t** module entry.

11

17.3.26 Credential validation callback function

12
13

Summary

Callback function for security credential validation.

14

Format

PMIx v3.0

```
typedef void (*pmix_validation_cfunc_t) (  
    pmix_status_t status,  
    pmix_info_t info[], size_t ninfo,  
    void *cbdata);
```

19

IN status

pmix_status_t value (handle)

20
21

IN info

Array of **pmix_info_t** provided by the system to pass any additional information about the authentication - e.g., the effective userid and group id of the certificate holder, and any related authorizations (handle)

22
23
24
25

IN ninfo

Number of elements in *info* (**size_t**)

26
27

IN cbdata

Object passed in original request (memory reference)

28
29 The returned status shall be one of the following:

- 30
31
32
33
34
- **PMIX_SUCCESS**, indicating that the request was processed and returned *success* (i.e., the credential was both valid and any information it contained was successfully processed). Details of the result will be returned in the *info* array
 - a PMIx error constant indicating either an error in the parsing of the credential or that the request was refused

Description

Define a validation callback function to indicate if a provided credential is valid, and any corresponding information regarding authorizations and other security matters.

Advice to users

The precise contents of the array will depend on the host environment and its associated security system. At the minimum, it is expected (but not required) that the array will contain entries for the `PMIX_USERID` and `PMIX_GRPID` of the client described in the credential. The *info* array is owned by the PMIx library and is not to be released or altered by the receiving party.

17.3.27 pmix_server_iof_fn_t

Summary

Request the specified IO channels be forwarded from the given array of processes.

Format

PMIx v3.0

C

```
typedef pmix_status_t (*pmix_server_iof_fn_t) (  
    const pmix_proc_t procs[],  
    size_t nprocs,  
    const pmix_info_t directives[],  
    size_t ndirs,  
    pmix_iof_channel_t channels,  
    pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

IN `procs`

Array `pmix_proc_t` identifiers whose IO is being requested (handle)

IN `nprocs`

Number of elements in *procs* (`size_t`)

IN `directives`

Array of `pmix_info_t` structures further defining the request (array of handles)

IN `ndirs`

Number of elements in the *info* array (integer)

IN `channels`

Bitmask identifying the channels to be forwarded (`pmix_iof_channel_t`)

IN `cbfunc`

Callback function `pmix_op_cbfunc_t` (function reference)

IN `cbdata`

Data to be passed to the callback function (memory reference)

Returns one of the following:

- `PMIX_SUCCESS`, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API.

- 1 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
2 *success* - the *cbfunc* will not be called
- 3 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
4 though the function entry was provided in the server module - the *cbfunc* will not be called
- 5 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
6 and failed - the *cbfunc* will not be called

Required Attributes

7 The following attributes are required to be included in the passed *info* array:

8 **PMIX_USERID** "pmix.euid" (uint32_t)
9 Effective user ID of the connecting process.

10 **PMIX_GRPID** "pmix.egid" (uint32_t)
11 Effective group ID of the connecting process.

13 Host environments that provide this module entry point are required to support the following attributes:

14 **PMIX_IOF_CACHE_SIZE** "pmix.iof.csize" (uint32_t)
15 The requested size of the PMIx server cache in bytes for each specified channel. By default, the server
16 is allowed (but not required) to drop all bytes received beyond the max size.

17 **PMIX_IOF_DROP_OLDEST** "pmix.iof.old" (bool)
18 In an overflow situation, the PMIx server is to drop the oldest bytes to make room in the cache.

19 **PMIX_IOF_DROP_NEWEST** "pmix.iof.new" (bool)
20 In an overflow situation, the PMIx server is to drop any new bytes received until room becomes
21 available in the cache (default).

Optional Attributes

22 The following attributes may be supported by a host environment.

23 **PMIX_IOF_BUFFERING_SIZE** "pmix.iof.bsize" (uint32_t)
24 Requests that IO on the specified channel(s) be aggregated in the PMIx tool library until the specified
25 number of bytes is collected to avoid being called every time a block of IO arrives. The PMIx tool
26 library will execute the callback and reset the collection counter whenever the specified number of
27 bytes becomes available. Any remaining buffered data will be *flushed* to the callback upon a call to
28 deregister the respective channel.

29 **PMIX_IOF_BUFFERING_TIME** "pmix.iof.btime" (uint32_t)
30 Max time in seconds to buffer IO before delivering it. Used in conjunction with buffering size, this
31 prevents IO from being held indefinitely while waiting for another payload to arrive.

Description

Request the specified IO channels be forwarded from the given array of processes. An error shall be returned in the callback function if the requested service from any of the requested processes cannot be provided.

Advice to PMIx library implementers

The forwarding of stdin is a *push* process - processes cannot request that it be *pulled* from some other source. Requests including the `PMIX_FWD_STDIN_CHANNEL` channel will return a `PMIX_ERR_NOT_SUPPORTED` error.

17.3.27.1 IOF delivery function

Summary

Callback function for delivering forwarded IO to a process.

Format

C

```
typedef void (*pmix_iof_cfunc_t) (  
    size_t iofhdlr, pmix_iof_channel_t channel,  
    pmix_proc_t *source, char *payload,  
    pmix_info_t info[], size_t ninfo);
```

C

IN `iofhdlr`

Registration number of the handler being invoked (`size_t`)

IN `channel`

bitmask identifying the channel the data arrived on (`pmix_iof_channel_t`)

IN `source`

Pointer to a `pmix_proc_t` identifying the namespace/rank of the process that generated the data (`char*`)

IN `payload`

Pointer to character array containing the data.

IN `info`

Array of `pmix_info_t` provided by the source containing metadata about the payload. This could include `PMIX_IOF_COMPLETE` (handle)

IN `ninfo`

Number of elements in `info` (`size_t`)

Description

Define a callback function for delivering forwarded IO to a process. This function will be called whenever data becomes available, or a specified buffering size and/or time has been met.

Advice to users

Multiple strings may be included in a given *payload*, and the *payload* may *not* be `NULL` terminated. The user is responsible for releasing the *payload* memory. The *info* array is owned by the PMIx library and is not to be released or altered by the receiving party.

1 17.3.28 pmix_server_stdin_fn_t

2 Summary

3 Pass standard input data to the host environment for transmission to specified recipients.

4 Format

PMIx v3.0

C

```
5 typedef pmix_status_t (*pmix_server_stdin_fn_t) (  
6     const pmix_proc_t *source,  
7     const pmix_proc_t targets[],  
8     size_t ntargets,  
9     const pmix_info_t directives[],  
10    size_t ndirs,  
11    const pmix_byte_object_t *bo,  
12    pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

13 **IN source**
14 [pmix_proc_t](#) structure of source process (handle)

15 **IN targets**
16 Array of [pmix_proc_t](#) target identifiers (handle)

17 **IN ntargets**
18 Number of elements in the *targets* array (integer)

19 **IN directives**
20 Array of info structures (array of handles)

21 **IN ndirs**
22 Number of elements in the *info* array (integer)

23 **IN bo**
24 Pointer to [pmix_byte_object_t](#) containing the payload (handle)

25 **IN cbfunc**
26 Callback function [pmix_op_cbfunc_t](#) (function reference)

27 **IN cbdata**
28 Data to be passed to the callback function (memory reference)

29 Returns one of the following:

- 30 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
31 returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to
32 returning from the API.
- 33 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
34 *success* - the *cbfunc* will not be called
- 35 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
36 though the function entry was provided in the server module - the *cbfunc* will not be called
- 37 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
38 and failed - the *cbfunc* will not be called

Required Attributes

The following attributes are required to be included in the passed *info* array:

- PMIX_USERID** "pmix.euid" (uint32_t)
Effective user ID of the connecting process.
- PMIX_GRPID** "pmix.egid" (uint32_t)
Effective group ID of the connecting process.

Description

Passes stdin to the host environment for transmission to specified recipients. The host environment is responsible for forwarding the data to all locations that host the specified *targets* and delivering the payload to the PMIx server library connected to those clients.

17.3.29 pmix_server_grp_fn_t

Summary

Request group operations (construct, destruct, etc.) on behalf of a set of processes.

Format

C

```
typedef pmix_status_t (*pmix_server_grp_fn_t) (  
    pmix_group_operation_t op,  
    char grp[],  
    const pmix_proc_t procs[],  
    size_t nprocs,  
    const pmix_info_t directives[],  
    size_t ndirs,  
    pmix_info_cbfunc_t cbfunc,  
    void *cbdata);
```

C

- IN op**
[pmix_group_operation_t](#) value indicating operation the host is requested to perform (integer)
- IN grp**
Character string identifying the group (string)
- IN procs**
Array of [pmix_proc_t](#) identifiers of participants (handle)
- IN nprocs**
Number of elements in the *procs* array (integer)
- IN directives**
Array of info structures (array of handles)
- IN ndirs**
Number of elements in the *info* array (integer)
- IN cbfunc**
Callback function [pmix_info_cbfunc_t](#) (function reference)

1 **IN** `cbdata`

2 Data to be passed to the callback function (memory reference)

3 Returns one of the following:

- 4 • **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be
5 returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to
6 returning from the API.
- 7 • **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned
8 *success* - the *cbfunc* will not be called
- 9 • **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even
10 though the function entry was provided in the server module - the *cbfunc* will not be called
- 11 • a PMIx error constant indicating either an error in the input or that the request was immediately processed
12 and failed - the *cbfunc* will not be called

▼----- Required Attributes -----▼

13 The following attributes are required to be supported by a host environment.

14 **PMIX_LOCAL_COLLECTIVE_STATUS** "`pmix.loc.col.st`" (`pmix_status_t`)

15 Status code for local collective operation being reported to the host by the server library. PMIx servers
16 may aggregate the participation by local client processes in a collective operation - e.g., instead of
17 passing individual client calls to **PMIx_Fence** up to the host environment, the server may pass only a
18 single call to the host when all local participants have executed their **PMIx_Fence** call, thereby
19 reducing the burden placed on the host. However, in cases where the operation locally fails (e.g., if a
20 participating client abnormally terminates prior to calling the operation), the server upcall functions to
21 the host do not include a `pmix_status_t` by which the PMIx server can alert the host to that failure.
22 This attribute resolves that problem by allowing the server to pass the status information regarding the
23 local collective operation.

▲----- Optional Attributes -----▼

24 The following attributes may be supported by a host environment.

25 **PMIX_GROUP_ASSIGN_CONTEXT_ID** "`pmix.grp.actxid`" (`bool`)

26 Requests that the RM assign a new context identifier to the newly created group. The identifier is an
27 unsigned, `size_t` value that the RM guarantees to be unique across the range specified in the request.
28 Thus, the value serves as a means of identifying the group within that range. If no range is specified,
29 then the request defaults to **PMIX_RANGE_SESSION**.

30 **PMIX_GROUP_LOCAL_ONLY** "`pmix.grp.lcl`" (`bool`)

31 Group operation only involves local processes. PMIx implementations are *required* to automatically
32 scan an array of group members for local vs remote processes - if only local processes are detected, the
33 implementation need not execute a global collective for the operation unless a context ID has been
34 requested from the host environment. This can result in significant time savings. This attribute can be
35 used to optimize the operation by indicating whether or not only local processes are represented, thus
36 allowing the implementation to bypass the scan.

37 **PMIX_GROUP_ENDPT_DATA** "`pmix.grp.endpt`" (`pmix_byte_object_t`)

1 Data collected during group construction to ensure communication between group members is
2 supported upon completion of the operation.

3 **PMIX_GROUP_OPTIONAL** "pmix.grp.opt" (bool)

4 Participation is optional - do not return an error if any of the specified processes terminate without
5 having joined. The default is **false**.

6 **PMIX_RANGE** "pmix.range" (pmix_data_range_t)

7 Define constraints on the processes that can access the provided data. Only processes that meet the
8 constraints are allowed to access it.

9 The following attributes may be included in the host's response:

10 **PMIX_GROUP_ID** "pmix.grp.id" (char*)

11 User-provided group identifier - as the group identifier may be used in PMIx operations, the user is
12 required to ensure that the provided ID is unique within the scope of the host environment (e.g., by
13 including some user-specific or application-specific prefix or suffix to the string).

14 **PMIX_GROUP_MEMBERSHIP** "pmix.grp.mbrs" (pmix_data_array_t*)

15 Array **pmix_proc_t** identifiers identifying the members of the specified group.

16 **PMIX_GROUP_CONTEXT_ID** "pmix.grp.ctxid" (size_t)

17 Context identifier assigned to the group by the host RM.

18 **PMIX_GROUP_ENDPT_DATA** "pmix.grp.endpt" (pmix_byte_object_t)

19 Data collected during group construction to ensure communication between group members is
20 supported upon completion of the operation.



21 Description

22 Perform the specified operation across the identified processes, plus any special actions included in the
23 directives. Return the result of any special action requests in the callback function when the operation is
24 completed. Actions may include a request (**PMIX_GROUP_ASSIGN_CONTEXT_ID**) that the host assign a
25 unique numerical (size_t) ID to this group - if given, the **PMIX_RANGE** attribute will specify the range across
26 which the ID must be unique (default to **PMIX_RANGE_SESSION**).

27 17.3.29.1 Group Operation Constants

28 *PMIx v4.0*

29 The **pmix_group_operation_t** structure is a **uint8_t** value for specifying group operations. All
values were originally defined in version 4 of the standard unless otherwise marked.

30 **PMIX_GROUP_CONSTRUCT** 0 Construct a group composed of the specified processes - used by a PMIx
31 server library to direct host operation.

32 **PMIX_GROUP_DESTRUCT** 1 Destruct the specified group - used by a PMIx server library to direct host
33 operation.

34 17.3.30 pmix_server_fabric_fn_t

35 Summary

36 Request fabric-related operations (e.g., information on a fabric) on behalf of a tool or other process.

Format

C

```
typedef pmix_status_t (*pmix_server_fabric_fn_t) (  
    const pmix_proc_t *requestor,  
    pmix_fabric_operation_t op,  
    const pmix_info_t directives[],  
    size_t ndirs,  
    pmix_info_cbfunc_t cbfunc,  
    void *cbdata);
```

C

IN requestor
 pmix_proc_t identifying the requestor (handle)

IN op
 pmix_fabric_operation_t value indicating operation the host is requested to perform (integer)

IN directives
 Array of info structures (array of handles)

IN ndirs
 Number of elements in the *info* array (integer)

IN cbfunc
 Callback function **pmix_info_cbfunc_t** (function reference)

IN cbdata
 Data to be passed to the callback function (memory reference)

Returns one of the following:

- **PMIX_SUCCESS**, indicating that the request is being processed by the host environment - result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API.
- **PMIX_OPERATION_SUCCEEDED**, indicating that the request was immediately processed and returned *success* - the *cbfunc* will not be called
- **PMIX_ERR_NOT_SUPPORTED**, indicating that the host environment does not support the request, even though the function entry was provided in the server module - the *cbfunc* will not be called
- a PMIx error constant indicating either an error in the input or that the request was immediately processed and failed - the *cbfunc* will not be called

Required Attributes

The following directives are required to be supported by all hosts to aid users in identifying the fabric and (if applicable) the device to whom the operation references:

PMIX_FABRIC_VENDOR "pmix.fab.vndr" (**string**)
 Name of the vendor (e.g., Amazon, Mellanox, HPE, Intel) for the specified fabric.

PMIX_FABRIC_IDENTIFIER "pmix.fab.id" (**string**)
 An identifier for the specified fabric (e.g., MgmtEthernet, Slingshot-11, OmniPath-1).

PMIX_FABRIC_PLANE "pmix.fab.plane" (**string**)

1 ID string of a fabric plane (e.g., CIDR for Ethernet). When used as a modifier in a request for
2 information, specifies the plane whose information is to be returned. When used directly as a key in a
3 request, returns a `pmix_data_array_t` of string identifiers for all fabric planes in the overall
4 system.

5 `PMIX_FABRIC_DEVICE_INDEX` "`pmix.fabdev.idx`" (`uint32_t`)

6 Index of the device within an associated communication cost matrix.



7 **Description**

8 Perform the specified operation. Return the result of any requests in the callback function when the operation
9 is completed. Operations may, for example, include a request for fabric information. See `pmix_fabric_t`
10 for a list of expected information to be included in the response. Note that requests for device index are to be
11 returned in the callback function's array of `pmix_info_t` using the `PMIX_FABRIC_DEVICE_INDEX`
12 attribute.

CHAPTER 18

Tools and Debuggers

1 The term *tool* widely refers to programs executed by the user or system administrator on a command line.
2 Tools frequently interact with either the SMS, user applications, or both to perform administrative and support
3 functions. For example, a debugger tool might be used to remotely control the processes of a parallel
4 application, monitoring their behavior on a step-by-step basis. Historically, such tools were custom-written for
5 each specific host environment due to the customized and/or proprietary nature of the environment's interfaces.

6 The advent of PMIx offers the possibility for creating portable tools capable of interacting with multiple RMs
7 without modification. Possible use-cases include:

- 8 • querying the status of scheduling queues and estimated allocation time for various resource options
- 9 • job submission and allocation requests
- 10 • querying job status for executing applications
- 11 • launching, monitoring, and debugging applications

12 Enabling these capabilities requires some extensions to the PMIx Standard (both in terms of APIs and
13 attributes), and utilization of client-side APIs for more tool-oriented purposes.

14 This chapter defines specific APIs related to tools, provides tool developers with an overview of the support
15 provided by PMIx, and serves to guide RM vendors regarding roles and responsibilities of RMs to support
16 tools. As the number of tool-specific APIs and attributes is fairly small, the bulk of the chapter serves to
17 provide a "theory of operation" for tools and debuggers. Description of the APIs themselves is therefore
18 deferred to the Section 18.5 later in the chapter.

18.1 Connection Mechanisms

19 The key to supporting tools lies in providing mechanisms by which a tool can connect to a PMIx server.
20 Application processes are able to connect because their local RM daemon provides them with the necessary
21 contact information upon execution. A command-line tool, however, isn't spawned by an RM daemon, and
22 therefore lacks the information required for rendezvous with a PMIx server.
23

24 Once a tool has started, it initializes PMIx as a tool (via `PMIx_tool_init`) if its access is restricted to
25 PMIx-based informational services such as `PMIx_Query_info`. However, if the tool intends to start jobs,
26 then it must include the `PMIX_LAUNCHER` attribute to inform the library of that intent so that the library can
27 initialize and provide access to the corresponding support.

28 Support for tools requires that the PMIx server be initialized with an appropriate attribute indicating that tool
29 connections are to be allowed. Separate attributes are provided to "fine-tune" this permission by allowing the
30 environment to independently enable (or disable) connections from tools executing on nodes other than the
31 one hosting the server itself. The PMIx server library shall provide an opportunity for the host environment to

1 authenticate and approve each connection request from a specific tool by calling the
 2 `pmix_server_tool_connection_fn_t` "hook" provided in the server module for that purpose.
 3 Servers in environments that do not provide this "hook" shall automatically reject all tool connection requests.
 4 Tools can connect to any local or remote PMIx server provided they are either explicitly given the required
 5 connection information, or are able to discover it via one of several defined rendezvous protocols. Connection
 6 discovery centers around the existence of *rendezvous files* containing the necessary connection information, as
 7 illustrated in Fig. 18.1.

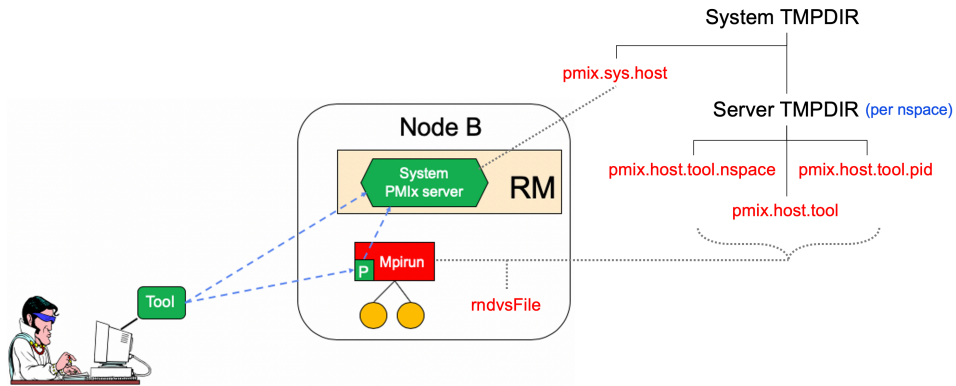


Figure 18.1.: Tool rendezvous files

8 The contents of each rendezvous file are specific to a given PMIx implementation, but should at least contain
 9 the namespace and rank of the server along with its connection URI. Note that tools linked to one PMIx
 10 implementation are therefore unlikely to successfully connect to PMIx server libraries from another
 11 implementation.

12 The top of the directory tree is defined by either the `PMIX_SYSTEM_TMPDIR` attribute (if given) or the
 13 `TMPDIR` environmental variable. PMIx servers that are designated as *system servers* by including the
 14 `PMIX_SERVER_SYSTEM_SUPPORT` attribute when calling `PMIx_server_init` will create a
 15 rendezvous file in this top-level directory. The filename will be of the form `pmix.sys.hostname`, where
 16 *hostname* is the string returned by the `gethostname` system call. Note that only one PMIx server on a node
 17 can be designated as the system server.

18 Non-system PMIx servers will create a set of three rendezvous files in the directory defined by either the
 19 `PMIX_SERVER_TMPDIR` attribute or the `TMPDIR` environmental variable:

- 20 • `pmix.host.tool.namespace` where *host* is the string returned by the `gethostname` system call and *namespace* is the
 21 namespace of the server.
- 22 • `pmix.host.tool.pid` where *host* is the string returned by the `gethostname` system call and *pid* is the PID of
 23 the server.
- 24 • `pmix.host.tool` where *host* is the string returned by the `gethostname` system call. Note that servers which
 25 are not given a namespace-specific `PMIX_SERVER_TMPDIR` attribute may not generate this file due to
 26 conflicts should multiple servers be present on the node.

1 The files are identical and may be implemented as symlinks to a single instance. The individual file names are
2 composed so as to aid the search process should a tool wish to connect to a server identified by its namespace
3 or PID.

4 Servers will additionally provide a rendezvous file in any given location if the path (either absolute or relative)
5 and filename is specified either during `PMIx_server_init` using the
6 `PMIX_LAUNCHER_RENDEZVOUS_FILE` attribute, or by the `PMIX_LAUNCHER_RNDZ_FILE`
7 environmental variable prior to executing the process containing the server. This latter mechanism may be the
8 preferred mechanism for tools such as debuggers that need to fork/exec a launcher (e.g., "mpixec") and then
9 rendezvous with it. This is described in more detail in Section 18.2.2.

10 Rendezvous file ownerships are set to the UID and GID of the server that created them, with permissions set
11 according to the desires of the implementation and/or system administrator policy. All connection attempts are
12 first governed by read access privileges to the target rendezvous file - thus, the combination of permissions,
13 UID, and GID of the rendezvous files act as a first-level of security for tool access.

14 A tool may connect to as many servers at one time as the implementation supports, but is limited to
15 designating only one such connection as its *primary* server. This is done to avoid confusion when the tool calls
16 an API as to which server should service the request. The first server the tool connects to is automatically
17 designated as the *primary* server.

18 Tools are allowed to change their primary server at any time via the `PMIx_tool_set_server` API, and to
19 connect/disconnect from a server as many times as desired. Note that standing requests (e.g., event
20 registrations) with the current primary server may be lost and/or may not be transferred when transitioning to
21 another primary server - PMIx implementors are not required to maintain or transfer state across tool-server
22 connections.

23 Tool process identifiers are assigned by one of the following methods:

- 24 ● If `PMIX_TOOL_NAMESPACE` is given, then the namespace of the tool will be assigned that value.
 - 25 – If `PMIX_TOOL_RANK` is also given, then the rank of the tool will be assigned that value.
 - 26 – If `PMIX_TOOL_RANK` is not given, then the rank will be set to a default value of zero.
- 27 ● If a process ID is not provided and the tool connects to a server, then one will be assigned by the host
28 environment upon connection to that server.
- 29 ● If a process ID is not provided and the tool does not connect to a server (e.g., if
30 `PMIX_TOOL_DO_NOT_CONNECT` is given), then the tool shall self-assign a unique identifier. This is
31 often done using some combination involving hostname and PID.

32 Tool process identifiers remain constant across servers. Thus, it is critical that a system-wide unique
33 namespace be provided if the tool itself sets the identifier, and that host environments provide a system-wide
34 unique identifier in the case where the identifier is set by the server upon connection. The host environment is
35 required to reject any connection request that fails to meet this criterion.

36 For simplicity, the following descriptions will refer to the:

- 37 ● `PMIX_SYSTEM_TMPDIR` as the directory specified by either the `PMIX_SYSTEM_TMPDIR` attribute (if
38 given) or the `TMPDIR` environmental variable.
- 39 ● `PMIX_SERVER_TMPDIR` as the directory specified by either the `PMIX_SERVER_TMPDIR` attribute or
40 the `TMPDIR` environmental variable.

1 The rendezvous methods are automatically employed for the initial tool connection during
2 `PMIx_tool_init` unless the `PMIX_TOOL_DO_NOT_CONNECT` attribute is specified, and on all
3 subsequent calls to `PMIx_tool_attach_to_server`.

4 18.1.1 Rendezvousing with a local server

5 Connection to a local PMIx server is pursued according to the following precedence chain based on attributes
6 contained in the call to the `PMIx_tool_init` or `PMIx_tool_attach_to_server` APIs. Servers to
7 which the tool already holds a connection will be ignored. Except where noted, the PMIx library will return an
8 error if the specified file cannot be found, the caller lacks permissions to read it, or the server specified within
9 the file does not respond to or accept the connection — the library will not proceed to check for other
10 connection options as the user specified a particular one to use.

11 Note that the PMIx implementation may choose to introduce a "delayed connection" protocol between steps in
12 the precedence chain - i.e., the library may cycle several times, checking for creation of the rendezvous file
13 each time after a delay of some period of time, thereby allowing the tool to wait for the server to create the
14 rendezvous file before either returning an error or continuing to the next step in the chain.

- 15 • If `PMIX_TOOL_ATTACHMENT_FILE` is given, then the tool will attempt to read the specified file and
16 connect to the server based on the information contained within it. The format of the attachment file is
17 identical to the rendezvous files described in earlier in this section. An error will be returned if the specified
18 file cannot be found.
- 19 • If `PMIX_SERVER_URI` or `PMIX_TCP_URI` is given, then connection will be attempted to the server at
20 the specified URI. Note that it is an error for both of these attributes to be specified. `PMIX_SERVER_URI`
21 is the preferred method as it is more generalized — `PMIX_TCP_URI` is provided for those cases where the
22 user specifically wants to use a TCP transport for the connection and wants to error out if one isn't available
23 or cannot be used.
- 24 • If `PMIX_SERVER_PIDINFO` was provided, then the tool will search for a rendezvous file created by a
25 PMIx server of the given PID in the `PMIX_SERVER_TMPDIR` directory. An error will be returned if a
26 matching rendezvous file cannot be found.
- 27 • If `PMIX_SERVER_NAMESPACE` is given, then the tool will search for a rendezvous file created by a PMIx
28 server of the given namespace in the `PMIX_SERVER_TMPDIR` directory. An error will be returned if a
29 matching rendezvous file cannot be found.
- 30 • If `PMIX_CONNECT_TO_SYSTEM` is given, then the tool will search for a system-level rendezvous file
31 created by a PMIx server in the `PMIX_SYSTEM_TMPDIR` directory. An error will be returned if a
32 matching rendezvous file cannot be found.
- 33 • If `PMIX_CONNECT_SYSTEM_FIRST` is given, then the tool will look for a system-level rendezvous file
34 created by a PMIx server in the `PMIX_SYSTEM_TMPDIR` directory. If found, then the tool will attempt to
35 connect to it. In this case, no error will be returned if the rendezvous file is not found or connection is
36 refused — the PMIx library will silently continue to the next option.
- 37 • By default, the tool will search the directory tree under the `PMIX_SERVER_TMPDIR` directory for
38 rendezvous files of PMIx servers, attempting to connect to each it finds until one accepts the connection. If
39 no rendezvous files are found, or all contacted servers refuse connection, then the PMIx library will return
40 an error. No "delayed connection" protocols may be utilized at this point.

1 Note that there can be multiple local servers - one from the system plus others from launchers and active jobs.
2 The PMIx tool connection search method is not guaranteed to pick a particular server unless directed to do so.
3 Tools can obtain a list of servers available on their local node using the `PMIx_Query_info` APIs with the
4 `PMIX_QUERY_AVAIL_SERVERS` key.

5 18.1.2 Connecting to a remote server

6 Connecting to remote servers is complicated due to the lack of access to the previously-described rendezvous
7 files. Two methods are required to be supported, both based on the caller having explicit knowledge of either
8 connection information or a path to a local file that contains such information:

- 9 • If `PMIX_TOOL_ATTACHMENT_FILE` is given, then the tool will attempt to read the specified file and
10 connect to the server based on the information contained within it. The format of the attachment file is
11 identical to the rendezvous files described in earlier in this section.
- 12 • If `PMIX_SERVER_URI` or `PMIX_TCP_URI` is given, then connection will be attempted to the server at
13 the specified URI. Note that it is an error for both of these attributes to be specified. `PMIX_SERVER_URI`
14 is the preferred method as it is more generalized — `PMIX_TCP_URI` is provided for those cases where the
15 user specifically wants to use the TCP transport for the connection and wants to error out if it isn't available
16 or cannot be used.

17 Additional methods may be provided by particular PMIx implementations. For example, the tool may use `ssh`
18 to launch a *probe* process onto the remote node so that the probe can search the `PMIX_SYSTEM_TMPDIR`
19 and `PMIX_SERVER_TMPDIR` directories for rendezvous files, relaying the discovered information back to
20 the requesting tool. If sufficient information is found to allow for remote connection, then the tool can use it to
21 establish the connection. Note that this method is not required to be supported - it is provided here as an
22 example and left to the discretion of PMIx implementors.

23 18.1.3 Attaching to running jobs

24 When attaching to a running job, the tool must connect to a PMIx server that is associated with that job - e.g., a
25 server residing in the host environment's local daemon that spawned one or more of the job's processes, or the
26 server residing in the launcher that is overseeing the job. Identifying an appropriate server can sometimes
27 prove challenging, particularly in an environment where multiple job launchers may be in operation, possibly
28 under control of the same user.

29 In cases where the user has only the one job of interest in operation on the local node (e.g., when engaged in an
30 interactive session on the node from which the launcher was executed), the normal rendezvous file discovery
31 method can often be used to successfully connect to the target job, even in the presence of jobs executed by
32 other users. The permissions and security authorizations can, in many cases, reliably ensure that only the one
33 connection can be made. However, this is not guaranteed in all cases.

34 The most common method, therefore, for attaching to a running job is to specify either the PID of the job's
35 launcher or the namespace of the launcher's job (note that the launcher's namespace frequently differs from the
36 namespace of the job it has launched). Unless the application processes themselves act as PMIx servers,
37 connection must be to the servers in the daemons that oversee the application. This is typically either daemons
38 specifically started by the job's launcher process, or daemons belonging to the host environment, that are
39 responsible for starting the application's processes and oversee their execution.

40 Identifying the correct PID or namespace can be accomplished in a variety of ways, including:

- 1 • Using typical OS or host environment tools to obtain a listing of active jobs and perusing those to find the
- 2 target launcher.
- 3 • Using a PMIx-based tool attached to a system-level server to query the active jobs and their command lines,
- 4 thereby identifying the application of interest and its associated launcher.
- 5 • Manually recording the PID of the launcher upon starting the job.
- 6 Once the namespace and/or PID of the target server has been identified, either of the previous methods can be
- 7 used to connect to it.

8 18.1.4 Tool initialization attributes

9 The following attributes are passed to the `PMIx_tool_init` API for use when initializing the PMIx library.

10 **PMIX_TOOL_NAMESPACE** "pmix.tool.namespace" (char*)
 11 Name of the namespace to use for this tool.
 12 **PMIX_TOOL_RANK** "pmix.tool.rank" (uint32_t)
 13 Rank of this tool.
 14 **PMIX_LAUNCHER** "pmix.tool.launcher" (bool)
 15 Tool is a launcher and needs to create rendezvous files.

16 18.1.5 Tool initialization environmental variables

17 The following environmental variables are used during `PMIx_tool_init` and `PMIx_server_init` to
 18 control various rendezvous-related operations when the process is started manually (e.g., on a command line)
 19 or by a fork/exec-like operation.

20 **PMIX_LAUNCHER_RNDZ_URI**

21 The spawned tool is to be connected back to the spawning tool using the given URI so that the
 22 spawning tool can provide directives (e.g., a `PMIx_Spawn` command) to it.

23 **PMIX_LAUNCHER_RNDZ_FILE**

24 If the specified file does not exist, this variable contains the absolute path of the file where the spawned
 25 tool is to store its connection information so that the spawning tool can connect to it. If the file does
 26 exist, it contains the information specifying the server to which the spawned tool is to connect.

27 **PMIX_KEEPLIVE_PIPE**

28 An integer **read**-end of a POSIX pipe that the tool should monitor for closure, thereby indicating that
 29 the parent tool has terminated. Used, for example, when a tool fork/exec's an intermediate launcher
 30 that should self-terminate if the originating tool exits.

31 Note that these environmental variables should be cleared from the environment after use and prior to forking
 32 child processes to avoid potentially unexpected behavior by the child processes.

33 18.1.6 Tool connection attributes

34 These attributes are defined to assist PMIx-enabled tools to connect with a PMIx server by passing them into
 35 either the `PMIx_tool_init` or the `PMIx_tool_attach_to_server` APIs - thus, they are not
 36 typically accessed via the `PMIx_Get` API.

37 **PMIX_SERVER_PIDINFO** "pmix.srvr.pidinfo" (pid_t)

1 PID of the target PMIx server for a tool.

2 **PMIX_CONNECT_TO_SYSTEM** "pmix.cnct.sys" (bool)

3 The requester requires that a connection be made only to a local, system-level PMIx server.

4 **PMIX_CONNECT_SYSTEM_FIRST** "pmix.cnct.sys.first" (bool)

5 Preferentially, look for a system-level PMIx server first.

6 **PMIX_SERVER_URI** "pmix.srvr.uri" (char*)

7 URI of the PMIx server to be contacted.

8 **PMIX_SERVER_HOSTNAME** "pmix.srvr.host" (char*)

9 Host where target PMIx server is located.

10 **PMIX_CONNECT_MAX_RETRIES** "pmix.tool.mretries" (uint32_t)

11 Maximum number of times to try to connect to PMIx server - the default value is implementation

12 specific.

13 **PMIX_CONNECT_RETRY_DELAY** "pmix.tool.retry" (uint32_t)

14 Time in seconds between connection attempts to a PMIx server - the default value is implementation

15 specific.

16 **PMIX_TOOL_DO_NOT_CONNECT** "pmix.tool.nocon" (bool)

17 The tool wants to use internal PMIx support, but does not want to connect to a PMIx server.

18 **PMIX_TOOL_CONNECT_OPTIONAL** "pmix.tool.conopt" (bool)

19 The tool shall connect to a server if available, but otherwise continue to operate unconnected.

20 **PMIX_TOOL_ATTACHMENT_FILE** "pmix.tool.attach" (char*)

21 Pathname of file containing connection information to be used for attaching to a specific server.

22 **PMIX_LAUNCHER_RENDEZVOUS_FILE** "pmix.tool.lncrnd" (char*)

23 Pathname of file where the launcher is to store its connection information so that the spawning tool can

24 connect to it.

25 **PMIX_PRIMARY_SERVER** "pmix.pri.srvr" (bool)

26 The server to which the tool is connecting shall be designated the *primary* server once connection has

27 been accomplished.

28 **PMIX_WAIT_FOR_CONNECTION** "pmix.wait.conn" (bool)

29 Wait until the specified process has connected to the requesting tool or server, or the operation times

30 out (if the **PMIX_TIMEOUT** directive is included in the request).

31 18.2 Launching Applications with Tools

32 Tool-directed launches require that the tool include the **PMIX_LAUNCHER** attribute when calling

33 **PMix_tool_init**. Two launch modes are supported:

- 34 • *Direct launch* where the tool itself is directly responsible for launching all processes, including debugger
- 35 daemons, using either the RM or daemons launched by the tool – i.e., there is no *intermediate launcher* (IL)
- 36 such as *mpiexec*. The case where the tool is self-contained (i.e., uses its own daemons without interacting
- 37 with an external entity such as the RM) lies outside the scope of this Standard; and
- 38 • *Indirect launch* where all processes are started via an IL such as *mpiexec* and the tool itself is not directly
- 39 involved in launching application processes or debugger daemons. Note that the IL may utilize the RM to
- 40 launch processes and/or daemons under the tool's direction.

41 Either of these methods can be executed interactively or by a batch script. Note that not all host environments

42 may support the direct launch method.

1 18.2.1 Direct launch

2 In the direct-launch use-case (Fig. 18.2), the tool itself performs the role of the launcher. Once invoked, the
3 tool connects to an appropriate PMIx server - e.g., a system-level server hosted by the RM. The tool is
4 responsible for assembling the description of the application to be launched (e.g., by parsing its command line)
5 into a spawn request containing an array of `pmix_app_t` applications and `pmix_info_t` job-level
6 information. An allocation of resources may or may not have been made in advance – if not, then the spawn
7 request must include allocation request information.

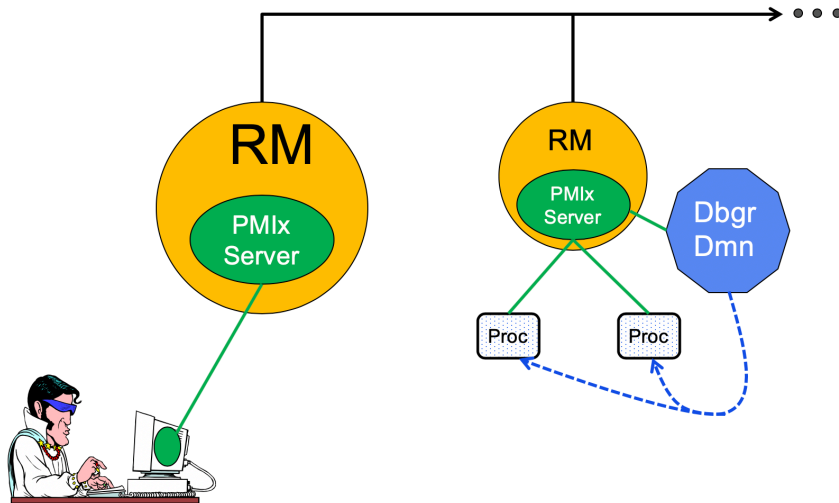


Figure 18.2.: Direct Launch

8 In addition to the attributes described in `PMIx_Spawn`, the tool may optionally wish to include the following
9 tool-specific attributes in the `job_info` argument to that API (the debugger-related attributes are discussed in
10 more detail in Section 18.4):

- 11 • `PMIX_FWD_STDIN` "`pmix.fwd.stdin`" (`pmix_rank_t`)
12 The requester intends to push information from its `stdin` to the indicated process. The local spawn
13 agent should, therefore, ensure that the `stdin` channel to that process remains available. A rank of
14 `PMIX_RANK_WILDCARD` indicates that all processes in the spawned job are potential recipients.
15 The requester will issue a call to `PMIx_IOF_push` to initiate the actual forwarding of information
16 to specified targets - this attribute simply requests that the IL retain the ability to forward the
17 information to the designated targets.
- 18 • `PMIX_FWD_STDOUT` "`pmix.fwd.stdout`" (`bool`)
19 Requests that the ability to forward the `stdout` of the spawned processes be maintained. The
20 requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options
21 for delivery of the forwarded output.
- 22 • `PMIX_FWD_STDERR` "`pmix.fwd.stderr`" (`bool`)

1 Requests that the ability to forward the `stderr` of the spawned processes be maintained. The
2 requester will issue a call to `PMIX_IOF_pull` to specify the callback function and other options
3 for delivery of the forwarded output.

- 4 • `PMIX_FWD_STDDIAG` "`pmix.fwd.stddiag`" (`bool`)
5 Requests that the ability to forward the diagnostic channel (if it exists) of the spawned processes be
6 maintained. The requester will issue a call to `PMIX_IOF_pull` to specify the callback function
7 and other options for delivery of the forwarded output.
- 8 • `PMIX_IOF_CACHE_SIZE` "`pmix.iof.csize`" (`uint32_t`)
9 The requested size of the PMIx server cache in bytes for each specified channel. By default, the
10 server is allowed (but not required) to drop all bytes received beyond the max size.
- 11 • `PMIX_IOF_DROP_OLDEST` "`pmix.iof.old`" (`bool`)
12 In an overflow situation, the PMIx server is to drop the oldest bytes to make room in the cache.
- 13 • `PMIX_IOF_DROP_NEWEST` "`pmix.iof.new`" (`bool`)
14 In an overflow situation, the PMIx server is to drop any new bytes received until room becomes
15 available in the cache (default).
- 16 • `PMIX_IOF_BUFFERING_SIZE` "`pmix.iof.bsize`" (`uint32_t`)
17 Requests that IO on the specified channel(s) be aggregated in the PMIx tool library until the
18 specified number of bytes is collected to avoid being called every time a block of IO arrives. The
19 PMIx tool library will execute the callback and reset the collection counter whenever the specified
20 number of bytes becomes available. Any remaining buffered data will be *flushed* to the callback
21 upon a call to deregister the respective channel.
- 22 • `PMIX_IOF_BUFFERING_TIME` "`pmix.iof.btime`" (`uint32_t`)
23 Max time in seconds to buffer IO before delivering it. Used in conjunction with buffering size, this
24 prevents IO from being held indefinitely while waiting for another payload to arrive.
- 25 • `PMIX_IOF_OUTPUT_RAW` "`pmix.iof.raw`" (`bool`)
26 Do not buffer output to be written as complete lines - output characters as the stream delivers them
- 27 • `PMIX_IOF_TAG_OUTPUT` "`pmix.iof.tag`" (`bool`)
28 Requests that output be prefixed with the `nspac,rank` of the source and a string identifying the
29 channel (`stdout`, `stderr`, etc.).
- 30 • `PMIX_IOF_TIMESTAMP_OUTPUT` "`pmix.iof.ts`" (`bool`)
31 Requests that output be marked with the time at which the data was received by the tool - note that
32 this will differ from the time at which the data was collected from the source.
- 33 • `PMIX_IOF_XML_OUTPUT` "`pmix.iof.xml`" (`bool`)
34 Requests that output be formatted in XML.
- 35 • `PMIX_IOF_RANK_OUTPUT` "`pmix.iof.rank`" (`bool`)
36 Tag output with the rank it came from
- 37 • `PMIX_IOF_OUTPUT_TO_FILE` "`pmix.iof.file`" (`char*`)

1 Direct application output into files of form "<filename>.<nospace>.<rank>.stdout" (for **stdout**) and
2 "<filename>.<nospace>.<rank>.stderr" (for **stderr**). If **PMIX_IOF_MERGE_STDERR_STDOUT**
3 was given, then only the **stdout** file will be created and both streams will be written into it.

- 4 • **PMIX_IOF_OUTPUT_TO_DIRECTORY** "**pmix.iof.dir**" (**char***)
5 Direct application output into files of form "<directory>/<nospace>/rank.<rank>/stdout" (for
6 **stdout**) and "<directory>/<nospace>/rank.<rank>/stderr" (for **stderr**). If
7 **PMIX_IOF_MERGE_STDERR_STDOUT** was given, then only the **stdout** file will be created and
8 both streams will be written into it.
- 9 • **PMIX_IOF_FILE_PATTERN** "**pmix.iof.fpt**" (**bool**)
10 Specified output file is to be treated as a pattern and not automatically annotated by namespace, rank, or
11 other parameters. The pattern can use **%n** for the namespace, and **%r** for the rank wherever those
12 quantities are to be placed. The resulting filename will be appended with ".stdout" for the **stdout**
13 stream and ".stderr" for the **stderr** stream. If **PMIX_IOF_MERGE_STDERR_STDOUT** was
14 given, then only the **stdout** file will be created and both streams will be written into it.
- 15 • **PMIX_IOF_FILE_ONLY** "**pmix.iof.fonly**" (**bool**)
16 Output only into designated files - do not also output a copy to the console's stdout/stderr
- 17 • **PMIX_IOF_MERGE_STDERR_STDOUT** "**pmix.iof.mrg**" (**bool**)
18 Merge stdout and stderr streams from application procs
- 19 • **PMIX_NOHUP** "**pmix.nohup**" (**bool**)
20 Any processes started on behalf of the calling tool (or the specified namespace, if such specification
21 is included in the list of attributes) should continue after the tool disconnects from its server.
- 22 • **PMIX_NOTIFY_JOB_EVENTS** "**pmix.note.jev**" (**bool**)
23 Requests that the launcher generate the **PMIX_EVENT_JOB_START**,
24 **PMIX_LAUNCH_COMPLETE**, and **PMIX_EVENT_JOB_END** events. Each event is to include at
25 least the namespace of the corresponding job and a **PMIX_EVENT_TIMESTAMP** indicating the
26 time the event occurred. Note that the requester must register for these individual events, or capture
27 and process them by registering a default event handler instead of individual handlers and then
28 process the events based on the returned status code. Another common method is to register one
29 event handler for all job-related events, with a separate handler for non-job events - see
30 **PMIx_Register_event_handler** for details.
- 31 • **PMIX_NOTIFY_COMPLETION** "**pmix.notecomp**" (**bool**)
32 Requests that the launcher generate the **PMIX_EVENT_JOB_END** event for normal or abnormal
33 termination of the spawned job. The event shall include the returned status code
34 (**PMIX_JOB_TERM_STATUS**) for the corresponding job; the identity (**PMIX_PROCID**) and exit
35 status (**PMIX_EXIT_CODE**) of the first failed process, if applicable; and a
36 **PMIX_EVENT_TIMESTAMP** indicating the time the termination occurred. Note that the requester
37 must register for the event or capture and process it within a default event handler.
- 38 • **PMIX_LOG_JOB_EVENTS** "**pmix.log.jev**" (**bool**)
39 Requests that the launcher log the **PMIX_EVENT_JOB_START**, **PMIX_LAUNCH_COMPLETE**, and
40 **PMIX_EVENT_JOB_END** events using **PMIx_Log**, subject to the logging attributes of Section
41 13.4.3.
- 42 • **PMIX_LOG_COMPLETION** "**pmix.logcomp**" (**bool**)

1 Requests that the launcher log the `PMIX_EVENT_JOB_END` event for normal or abnormal
2 termination of the spawned job using `PMIx_Log`, subject to the logging attributes of Section
3 13.4.3. The event shall include the returned status code (`PMIX_JOB_TERM_STATUS`) for the
4 corresponding job; the identity (`PMIX_PROCID`) and exit status (`PMIX_EXIT_CODE`) of the first
5 failed process, if applicable; and a `PMIX_EVENT_TIMESTAMP` indicating the time the termination
6 occurred.

- 7 ● `PMIX_DEBUG_STOP_ON_EXEC` "`pmix.dbg.exec`" (bool)
8 Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies
9 only to that application) or in the `job_info` array if it applies to all applications in the given spawn
10 request. Indicates that the application is being spawned under a debugger, and that the local launch
11 agent is to pause the resulting application processes on first instruction for debugger attach. The
12 launcher (RM or IL) is to generate the `PMIX_LAUNCH_COMPLETE` event when all processes are
13 stopped at the exec point.
- 14 ● `PMIX_DEBUG_STOP_IN_INIT` "`pmix.dbg.init`" (bool)
15 Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies
16 only to that application) or in the `job_info` array if it applies to all applications in the given spawn
17 request. Indicates that the specified application process shall notify its PMIx server that it is pausing and then
18 pause during `PMIx_Init` of the spawned processes until either released by debugger modification
19 of an appropriate variable or receipt of the `PMIX_DEBUGGER_RELEASE` event. The launcher (RM
20 or IL) is responsible for generating the `PMIX_READY_FOR_DEBUG` event (stipulating a breakpoint
21 of `pmix-init`) when all processes have reached the pause point.
22
- 23 ● `PMIX_DEBUG_STOP_IN_APP` "`pmix.dbg.notify`" (varies)
24 Direct specified ranks to stop at application-specific point and notify they are ready-to-debug. The
25 attribute's value can be any of three data types:
 - 26 – bool - true indicating all ranks
 - 27 – `pmix_rank_t` - the rank of one proc, or `PMIX_RANK_WILDCARD` for all
 - 28 – a `pmix_data_array_t` if an array of individual processes are specified

29 The resulting application processes are to notify their server (by generating the
30 `PMIX_READY_FOR_DEBUG` event) when they reach some application-determined location - the
31 event shall include the `PMIX_BREAKPOINT` attribute indicating where the application has stopped.
32 The application shall pause at that point until released by debugger modification of an appropriate
33 variable. The launcher (RM or IL) is responsible for generating the `PMIX_READY_FOR_DEBUG`
34 event when all processes have indicated they are at the pause point.

Advice to users

35 The `PMIX_IOF_FILE_ONLY` indicates output is directed to files and no copy is sent back to the application.
36 For example, this can be combined with `PMIX_IOF_OUTPUT_TO_FILE` or
37 `PMIX_IOF_OUTPUT_TO_DIRECTORY` to only output to files.

1 The tool then calls the `PMIx_Spawn` API so that the PMIx library can communicate the spawn request to the
2 server.

3 Upon receipt, the PMIx server library passes the spawn request to its host RM daemon for processing via the
4 `pmix_server_spawn_fn_t` server module function. If this callback was not provided, then the PMIx
5 server library will return the `PMIX_ERR_NOT_SUPPORTED` error status.

6 If an allocation must be made, then the host environment is responsible for communicating the request to its
7 associated scheduler. Once resources are available, the host environment initiates the launch process to start
8 the job. The host environment must parse the spawn request for relevant directives, returning an error if any
9 required directive cannot be supported. Optional directives may be ignored if they cannot be supported.

10 Any error while executing the spawn request must be returned by `PMIx_Spawn` to the requester. Once the
11 spawn request has succeeded in starting the specified processes, the request will return `PMIX_SUCCESS` back
12 to the requester along with the namespace of the started job. Upon termination of the spawned job, the host
13 environment must generate a `PMIX_EVENT_JOB_END` event for normal or abnormal termination if requested
14 to do so. The event shall include:

- 15 • the returned status code (`PMIX_JOB_TERM_STATUS`) for the corresponding job;
- 16 • the identity (`PMIX_PROCID`) and exit status (`PMIX_EXIT_CODE`) of the first failed process, if applicable;
- 17 • a `PMIX_EVENT_TIMESTAMP` indicating the time the termination occurred; plus
- 18 • any other info provided by the host environment.

19 18.2.2 Indirect launch

20 In the indirect launch use-case, the application processes are started via an intermediate launcher (e.g.,
21 `mpiexec`) that is itself started by the tool (see Fig 18.3). Thus, at a high level, this is a two-stage launch
22 procedure to start the application: the tool (henceforth referred to as the *initiator*) starts the IL, which then
23 starts the applications. In practice, additional steps may be involved if, for example, the IL starts its own
24 daemons to shepherd the application processes.

25 A key aspect of this operational mode is the avoidance of any requirement that the initiator parse and/or
26 understand the command line of the IL. Instead, the indirect launch procedure supports either of two methods:
27 one where the initiator assumes responsibility for parsing its command line to obtain the application as well as
28 the IL and its options, and another where the initiator defers the command line parsing to the IL. Both of these
29 methods are described in the following sections.

30 18.2.2.1 Initiator-based command line parsing

31 This method utilizes a first call to the `PMIx_Spawn` API to start the IL itself, and then uses a second call to
32 `PMIx_Spawn` to request that the IL spawn the actual job. The burden of analyzing the initial command line to
33 separately identify the IL's command line from the application itself falls upon the initiator. An example is
34 provided below:

```
35 $ initiator --launcher "mpiexec --verbose" -n 3 ./app <appoptions>
```

1 The initiator spawns the IL using the same procedure for launching an application - it begins by assembling the
 2 description of the IL into a spawn request containing an array of `pmix_app_t` and `pmix_info_t` job-level
 3 information. Note that this step does not include any information regarding the application itself - only the
 4 launcher is included. In addition, the initiator must include the rendezvous URI in the environment so the IL
 5 knows how to connect back to it.

6 An allocation of resources for the IL itself may or may not be required – if it is, then the allocation must be
 7 made in advance or the spawn request must include allocation request information.

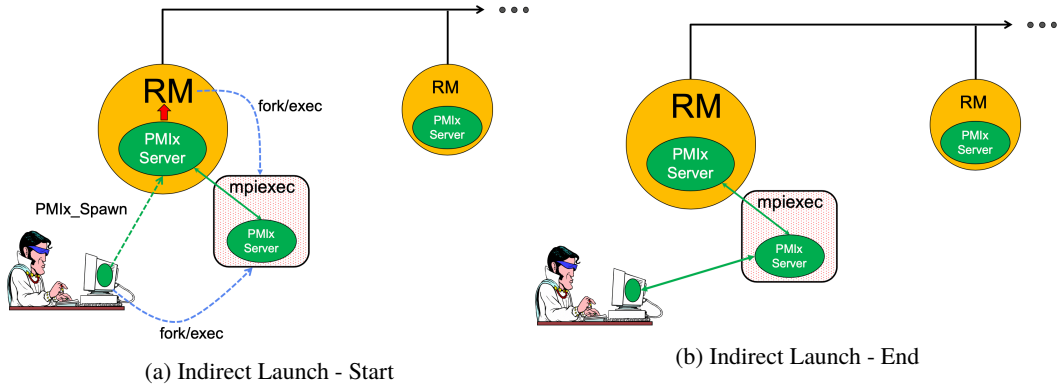


Figure 18.3.: Indirect launch procedure

8 The initiator may optionally wish to include the following tool-specific attributes in the `job_info` argument to
 9 `PMix_Spawn` - note that these attributes refer only to the behavior of the IL itself and not the eventual job to
 10 be launched:

- 11 • `PMIX_FWD_STDIN` "`pmix.fwd.stdin`" (`pmix_rank_t`)
 12 The requester intends to push information from its `stdin` to the indicated process. The local spawn
 13 agent should, therefore, ensure that the `stdin` channel to that process remains available. A rank of
 14 `PMIX_RANK_WILDCARD` indicates that all processes in the spawned job are potential recipients.
 15 The requester will issue a call to `PMix_IOF_push` to initiate the actual forwarding of information
 16 to specified targets - this attribute simply requests that the IL retain the ability to forward the
 17 information to the designated targets.
- 18 • `PMIX_FWD_STDOUT` "`pmix.fwd.stdout`" (`bool`)
 19 Requests that the ability to forward the `stdout` of the spawned processes be maintained. The
 20 requester will issue a call to `PMix_IOF_pull` to specify the callback function and other options
 21 for delivery of the forwarded output.
- 22 • `PMIX_FWD_STDERR` "`pmix.fwd.stderr`" (`bool`)
 23 Requests that the ability to forward the `stderr` of the spawned processes be maintained. The
 24 requester will issue a call to `PMix_IOF_pull` to specify the callback function and other options
 25 for delivery of the forwarded output.
- 26 • `PMIX_FWD_STDDIAG` "`pmix.fwd.stddiag`" (`bool`)

1 Requests that the ability to forward the diagnostic channel (if it exists) of the spawned processes be
2 maintained. The requester will issue a call to `PMIX_IOF_PULL` to specify the callback function
3 and other options for delivery of the forwarded output.

- 4 • `PMIX_IOF_CACHE_SIZE` "pmix.iof.csize" (uint32_t)

5 The requested size of the PMIx server cache in bytes for each specified channel. By default, the
6 server is allowed (but not required) to drop all bytes received beyond the max size.

- 7 • `PMIX_IOF_DROP_OLDEST` "pmix.iof.old" (bool)

8 In an overflow situation, the PMIx server is to drop the oldest bytes to make room in the cache.

- 9 • `PMIX_IOF_DROP_NEWEST` "pmix.iof.new" (bool)

10 In an overflow situation, the PMIx server is to drop any new bytes received until room becomes
11 available in the cache (default).

- 12 • `PMIX_IOF_BUFFERING_SIZE` "pmix.iof.bsize" (uint32_t)

13 Requests that IO on the specified channel(s) be aggregated in the PMIx tool library until the
14 specified number of bytes is collected to avoid being called every time a block of IO arrives. The
15 PMIx tool library will execute the callback and reset the collection counter whenever the specified
16 number of bytes becomes available. Any remaining buffered data will be *flushed* to the callback
17 upon a call to deregister the respective channel.

- 18 • `PMIX_IOF_BUFFERING_TIME` "pmix.iof.btime" (uint32_t)

19 Max time in seconds to buffer IO before delivering it. Used in conjunction with buffering size, this
20 prevents IO from being held indefinitely while waiting for another payload to arrive.

- 21 • `PMIX_IOF_TAG_OUTPUT` "pmix.iof.tag" (bool)

22 Requests that output be prefixed with the nspace,rank of the source and a string identifying the
23 channel (`stdout`, `stderr`, etc.).

- 24 • `PMIX_IOF_TIMESTAMP_OUTPUT` "pmix.iof.ts" (bool)

25 Requests that output be marked with the time at which the data was received by the tool - note that
26 this will differ from the time at which the data was collected from the source.

- 27 • `PMIX_IOF_XML_OUTPUT` "pmix.iof.xml" (bool)

28 Requests that output be formatted in XML.

- 29 • `PMIX_NOHUP` "pmix.nohup" (bool)

30 Any processes started on behalf of the calling tool (or the specified namespace, if such specification
31 is included in the list of attributes) should continue after the tool disconnects from its server.

- 32 • `PMIX_LAUNCHER_DAEMON` "pmix.lnch.dmn" (char*)

33 Path to executable that is to be used as the backend daemon for the launcher. This replaces the
34 launcher's own daemon with the specified executable. Note that the user is therefore responsible for
35 ensuring compatibility of the specified executable and the host launcher.

- 36 • `PMIX_FORKEXEC_AGENT` "pmix.frkex.agnt" (char*)

37 Path to executable that the launcher's backend daemons are to fork/exec in place of the actual
38 application processes. The fork/exec agent shall connect back (as a PMIx tool) to the launcher's

1 daemon to receive its spawn instructions, and is responsible for starting the actual application
2 process it replaced. See Section 18.4.3 for details.

- 3 • **PMIX_EXEC_AGENT** "pmix.exec.agnt" (char*)
4 Path to executable that the launcher's backend daemons are to fork/exec in place of the actual
5 application processes. The launcher's daemon shall pass the full command line of the application on
6 the command line of the exec agent, which shall not connect back to the launcher's daemon. The
7 exec agent is responsible for exec'ing the specified application process in its own place. See Section
8 18.4.3 for details.
- 9 • **PMIX_DEBUG_STOP_IN_INIT** "pmix.dbg.init" (bool)
10 Included in either the **pmix_info_t** array in a **pmix_app_t** description (if the directive applies
11 only to that application) or in the *job_info* array if it applies to all applications in the given spawn
12 request. Indicates that the specified application is being spawned under a debugger. The PMIx client
13 library in each resulting application process shall notify its PMIx server that it is pausing and then
14 pause during **PMIx_Init** of the spawned processes until either released by debugger modification
15 of an appropriate variable or receipt of the **PMIX_DEBUGGER_RELEASE** event. The launcher (RM
16 or IL) is responsible for generating the **PMIX_READY_FOR_DEBUG** event (stipulating a breakpoint
17 of pmix-init) when all processes have reached the pause point. In this context, the initiator is
18 directing the IL to stop in **PMIx_tool_init**. This gives the initiator a chance to connect to the IL
19 and register for events prior to the IL launching the application job.

20 and the following optional variables in the environment of the IL:

- 21 • **PMIX_KEEPLIVE_PIPE** - an integer read-end of a POSIX pipe that the IL should monitor for closure,
22 thereby indicating that the initiator has terminated.

23 The initiator then calls the **PMIx_Spawn** API so that the PMIx library can either communicate the spawn
24 request to a server (if connected to one), or locally spawn the IL itself if not connected to a server and the
25 PMIx implementation includes self-spawn support. **PMIx_Spawn** shall return an error if neither of these
26 conditions is met.

27 When initialized by the IL, the **PMIx_tool_init** function must perform two operations:

- 28 • check for the presence of the **PMIX_KEEPLIVE_PIPE** environmental variable - if provided, then the
29 library shall monitor the pipe for closure, providing a **PMIX_EVENT_JOB_END** event when the pipe closes
30 (thereby indicating the termination of the initiator). The IL should register for this event after completing
31 **PMIx_tool_init** - the initiator's namespace can be obtained via a call to **PMIx_Get** with the
32 **PMIX_PARENT_ID** key. Note that this feature will only be available if the spawned IL is local to the
33 initiator.
- 34 • check for the **PMIX_LAUNCHER_RNDZ_URI** environmental parameter - if found, the library shall connect
35 back to the initiator using the **PMIx_tool_attach_to_server** API, retaining its current server as its
36 primary server.

37 Once the IL completes **PMIx_tool_init**, it must register for the **PMIX_EVENT_JOB_END** termination
38 event and then idle until receiving that event - either directly from the initiator, or from the PMIx library upon
39 detecting closure of the keepalive pipe. The IL idles in the intervening time as it is solely acting as a relay (if
40 connected to a server that is performing the actual application launch) or as a PMIx server responding to
41 spawn requests.

1 Upon return from the `PMIx_Spawn` API, the initiator should set the spawned IL as its primary server using
2 the `PMIx_tool_set_server` API with the namespace returned by `PMIx_Spawn` and any valid rank (a rank
3 of zero would ordinarily be used as only one IL process is typically started). It is advisable to set a connection
4 timeout value when calling this function. The initiator can then proceed to spawn the actual application
5 according to the procedure described in Section 18.2.1.

6 18.2.2.2 IL-based command line parsing

7 In the case where the initiator cannot parse its command line, it must defer that parsing to the IL. A common
8 example is provided below:

```
9 $ initiator mpiexec --verbose -n 3 ./app <appoptions>
```

10 For this situation, the initiator proceeds as above with only one notable exception: instead of calling
11 `PMIx_Spawn` twice (once to start the IL and again to start the actual application), the initiator only calls that
12 API one time:

- 13 • The *app* parameter passed to the spawn request contains only one `pmix_app_t` that contains the entire
14 command line, including both launcher and application(s).
- 15 • The launcher executable must be in the *app.cmd* field and in *app.argv[0]*, with the rest of the command line
16 appended to the *app.argv* array.
- 17 • Any job-level directives for the IL itself (e.g., `PMIX_FORKEXEC_AGENT` or `PMIX_FWD_STDOUT`) are
18 included in the *job_info* parameter of the call to `PMIx_Spawn`.
- 19 • The job-level directives must include both the `PMIX_SPAWN_TOOL` attribute indicating that the initiator is
20 spawning a tool, and the `PMIX_DEBUG_STOP_IN_INIT` attribute directing the IL to stop during the call
21 to `PMIx_tool_init`. The latter directive allows the initiator to connect to the IL prior to launch of the
22 application.
- 23 • The `PMIX_LAUNCHER_RNDZ_URI` and `PMIX_KEEPLIVE_PIPE` environmental variables are
24 provided to the launcher in its environment via the *app.env* field.
- 25 • The IL must use `PMIx_Get` with the `PMIX_LAUNCH_DIRECTIVES` key to obtain any initiator-provided
26 directives (e.g., `PMIX_DEBUG_STOP_IN_INIT` or `PMIX_DEBUG_STOP_ON_EXEC`) aimed at the
27 application(s) it will spawn.

28 Upon return from `PMIx_Spawn`, the initiator must:

- 29 • use the `PMIx_tool_set_server` API to set the spawned IL as its primary server
- 30 • register with that server to receive the `PMIX_LAUNCH_COMPLETE` event. This allows the initiator to know
31 when the IL has completed launch of the application
- 32 • release the IL from its "hold" in `PMIx_tool_init` by issuing the `PMIX_DEBUGGER_RELEASE` event,
33 specifying the IL as the custom range. Upon receipt of the event, the IL is free to parse its command line,
34 apply any provided directives, and execute the application.

35 Upon receipt of the `PMIX_LAUNCH_COMPLETE` event, the initiator should register to receive notification of
36 completion of the returned namespace of the application. Receipt of the `PMIX_EVENT_JOB_END` event
37 provides a signal that the initiator may itself terminate.

18.2.3 Tool spawn-related attributes

Tools are free to utilize the spawn attributes available to applications (see 12.2.4) when constructing a spawn request, but can also utilize the following attributes that are specific to tool-based spawn operations:

PMIX_FWD_STDIN "pmix.fwd.stdin" (pmix_rank_t)

The requester intends to push information from its **stdin** to the indicated process. The local spawn agent should, therefore, ensure that the **stdin** channel to that process remains available. A rank of **PMIX_RANK_WILDCARD** indicates that all processes in the spawned job are potential recipients. The requester will issue a call to **PMIx_IOF_push** to initiate the actual forwarding of information to specified targets - this attribute simply requests that the IL retain the ability to forward the information to the designated targets.

PMIX_FWD_STDOUT "pmix.fwd.stdout" (bool)

Requests that the ability to forward the **stdout** of the spawned processes be maintained. The requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for delivery of the forwarded output.

PMIX_FWD_STDERR "pmix.fwd.stderr" (bool)

Requests that the ability to forward the **stderr** of the spawned processes be maintained. The requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for delivery of the forwarded output.

PMIX_FWD_STDDIAG "pmix.fwd.stddiag" (bool)

Requests that the ability to forward the diagnostic channel (if it exists) of the spawned processes be maintained. The requester will issue a call to **PMIx_IOF_pull** to specify the callback function and other options for delivery of the forwarded output.

PMIX_NOHUP "pmix.nohup" (bool)

Any processes started on behalf of the calling tool (or the specified namespace, if such specification is included in the list of attributes) should continue after the tool disconnects from its server.

PMIX_LAUNCHER_DAEMON "pmix.lnch.dmn" (char*)

Path to executable that is to be used as the backend daemon for the launcher. This replaces the launcher's own daemon with the specified executable. Note that the user is therefore responsible for ensuring compatibility of the specified executable and the host launcher.

PMIX_FORKEXEC_AGENT "pmix.frkex.agnt" (char*)

Path to executable that the launcher's backend daemons are to fork/exec in place of the actual application processes. The fork/exec agent shall connect back (as a PMIx tool) to the launcher's daemon to receive its spawn instructions, and is responsible for starting the actual application process it replaced. See Section 18.4.3 for details.

PMIX_EXEC_AGENT "pmix.exec.agnt" (char*)

Path to executable that the launcher's backend daemons are to fork/exec in place of the actual application processes. The launcher's daemon shall pass the full command line of the application on the command line of the exec agent, which shall not connect back to the launcher's daemon. The exec agent is responsible for exec'ing the specified application process in its own place. See Section 18.4.3 for details.

PMIX_LAUNCH_DIRECTIVES "pmix.lnch.dirs" (pmix_data_array_t*)

Array of **pmix_info_t** containing directives for the launcher - a convenience attribute for retrieving all directives with a single call to **PMIx_Get**.

18.2.4 Tool rendezvous-related events

The following constants refer to events relating to rendezvous of a tool and launcher during spawn of the IL.

1 **PMIX_LAUNCHER_READY** -155 An application launcher (e.g., *mpiexec*) shall generate this event to
2 signal a tool that started it that the launcher is ready to receive directives/commands (e.g.,
3 **PMIx_Spawn**). This is only used when the initiator is able to parse the command line itself, or the
4 launcher is started as a persistent Distributed Virtual Machine (DVM).

5 18.3 IO Forwarding

6 Underlying the operation of many tools is a common need to forward **stdin** from the tool to targeted
7 processes, and to return **stdout/stderr** from those processes to the tool (e.g., for display on the user's
8 console). Historically, each tool developer was responsible for creating their own IO forwarding subsystem.
9 However, the introduction of PMIx as a standard mechanism for interacting between applications and the host
10 environment has made it possible to relieve tool developers of this burden.

11 This section defines functions by which tools can request forwarding of input/output to/from other processes
12 and serves as a design guide to:

- 13 • provide tool developers with an overview of the expected behavior of the PMIx IO forwarding support;
- 14 • guide RM vendors regarding roles and responsibilities expected of the RM to support IO forwarding; and
- 15 • provide insight into the thinking of the PMIx community behind the definition of the PMIx IO forwarding
16 APIs.

17 Note that the forwarding of IO via PMIx requires that both the host environment and the tool support PMIx,
18 but does not impose any similar requirements on the application itself.

19 The responsibility of the host environment in forwarding of IO falls into the following areas:

- 20 • Capturing output from specified processes.
- 21 • Forwarding that output to the host of the PMIx server library that requested it.
- 22 • Delivering that payload to the PMIx server library via the **PMIx_server_IOF_deliver** API for final
23 dispatch to the requesting tool.

24 It is the responsibility of the PMIx library to buffer, format, and deliver the payload to the requesting client.
25 This may require caching of output until a forwarding registration is received, as governed by the
26 corresponding IO forwarding attributes of Section 18.3.5 that are supported by the implementation.

27 18.3.1 Forwarding stdout/stderr

28 At an appropriate point in its operation (usually during startup), a tool will utilize the **PMIx_tool_init**
29 function to connect to a PMIx server. The PMIx server can be hosted by an RM daemon or could be embedded
30 in a library-provided starter program such as *mpiexec* - in terms of IO forwarding, the operations remain the
31 same either way. For purposes of this discussion, we will assume the server is in an RM daemon and that the
32 application processes are directly launched by the RM, as shown in Fig 18.4.

33 Once the tool has connected to the target server, it can request that processes be spawned on its behalf or that
34 output from a specified set of existing processes in a given executing application be forwarded to it. Requests
35 to spawn processes should include the **PMIX_FWD_STDIN**, **PMIX_FWD_STDOUT**, and/or
36 **PMIX_FWD_STDERR** attributes if the tool intends to request that the corresponding streams be forwarded at
37 some point during execution.

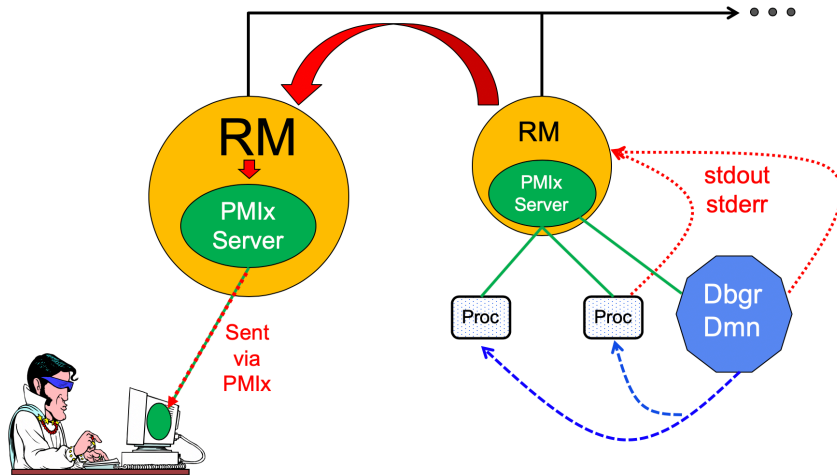


Figure 18.4.: Forwarding stdout/stderr

Note that requests to capture output from existing processes via the `PMIx_IOF_pull` API, and/or to forward input to specified processes via the `PMIx_IOF_push` API, can only succeed if the required attributes to retain that ability were passed when the corresponding job was spawned. The host is required to return an error for all such requests in cases where this condition is not met.

Two modes are supported when requesting that the host forward standard output/error via the `PMIx_IOF_pull` API - these can be controlled by including one of the following attributes in the *info* array passed to that function:

- `PMIX_IOF_COPY` "`pmix.iof.cpy`" (`bool`)
Requests that the host environment deliver a copy of the specified output stream(s) to the tool, letting the stream(s) continue to also be delivered to the default location. This allows the tool to tap into the output stream(s) without redirecting it from its current final destination.
- `PMIX_IOF_REDIRECT` "`pmix.iof.redir`" (`bool`)
Requests that the host environment intercept the specified output stream(s) and deliver it to the requesting tool instead of its current final destination. This might be used, for example, during a debugging procedure to avoid injection of debugger-related output into the application's results file. The original output stream(s) destination is restored upon termination of the tool. This is the default mode of operation.

When requesting to forward `stdout/stderr`, the tool can specify several formatting options to be used on the resulting output stream. These include:

- `PMIX_IOF_TAG_OUTPUT` "`pmix.iof.tag`" (`bool`)
Requests that output be prefixed with the `nspace,rank` of the source and a string identifying the channel (`stdout`, `stderr`, etc.).
- `PMIX_IOF_TIMESTAMP_OUTPUT` "`pmix.iof.ts`" (`bool`)

1 Requests that output be marked with the time at which the data was received by the tool - note that
2 this will differ from the time at which the data was collected from the source.

- 3 • **PMIX_IOF_XML_OUTPUT** "pmix.iof.xml" (bool)
4 Requests that output be formatted in XML.
- 5 • **PMIX_IOF_RANK_OUTPUT** "pmix.iof.rank" (bool)
6 Tag output with the rank it came from
- 7 • **PMIX_IOF_OUTPUT_TO_FILE** "pmix.iof.file" (char*)
8 Direct application output into files of form "<filename>.<nospace>.<rank>.stdout" (for **stdout**) and
9 "<filename>.<nospace>.<rank>.stderr" (for **stderr**). If **PMIX_IOF_MERGE_STDERR_STDOUT**
10 was given, then only the **stdout** file will be created and both streams will be written into it.
- 11 • **PMIX_IOF_OUTPUT_TO_DIRECTORY** "pmix.iof.dir" (char*)
12 Direct application output into files of form "<directory>/<nospace>/rank.<rank>/stdout" (for
13 **stdout**) and "<directory>/<nospace>/rank.<rank>/stderr" (for **stderr**). If
14 **PMIX_IOF_MERGE_STDERR_STDOUT** was given, then only the **stdout** file will be created and
15 both streams will be written into it.
- 16 • **PMIX_IOF_FILE_PATTERN** "pmix.iof.fpt" (bool)
17 Specified output file is to be treated as a pattern and not automatically annotated by nspace, rank, or
18 other parameters. The pattern can use %n for the namespace, and %r for the rank wherever those
19 quantities are to be placed. The resulting filename will be appended with ".stdout" for the **stdout**
20 stream and ".stderr" for the **stderr** stream. If **PMIX_IOF_MERGE_STDERR_STDOUT** was
21 given, then only the **stdout** file will be created and both streams will be written into it.
- 22 • **PMIX_IOF_FILE_ONLY** "pmix.iof.fonly" (bool)
23 Output only into designated files - do not also output a copy to the console's stdout/stderr
- 24 • **PMIX_IOF_MERGE_STDERR_STDOUT** "pmix.iof.mrg" (bool)
25 Merge stdout and stderr streams from application procs

26 The PMIx client in the tool is responsible for formatting the output stream. Note that output from multiple
27 processes will often be interleaved due to variations in arrival time - ordering of output is not guaranteed
28 across processes and/or nodes.

29 18.3.2 Forwarding stdin

30 A tool is not necessarily a child of the RM as it may have been started directly from the command line. Thus,
31 provision must be made for the tool to collect its **stdin** and pass it to the host RM (via the PMIx server) for
32 forwarding. Two methods of support for forwarding of **stdin** are defined:

- 33 • internal collection by the PMIx tool library itself. This is requested via the **PMIX_IOF_PUSH_STDIN**
34 attribute in the **PMIx_IOF_push** call. When this mode is selected, the tool library begins collecting all
35 **stdin** data and internally passing it to the local server for distribution to the specified target processes. All
36 collected data is sent to the same targets until **stdin** is closed, or a subsequent call to **PMIx_IOF_push**
37 is made that includes the **PMIX_IOF_COMPLETE** attribute indicating that forwarding of **stdin** is to be
38 terminated.

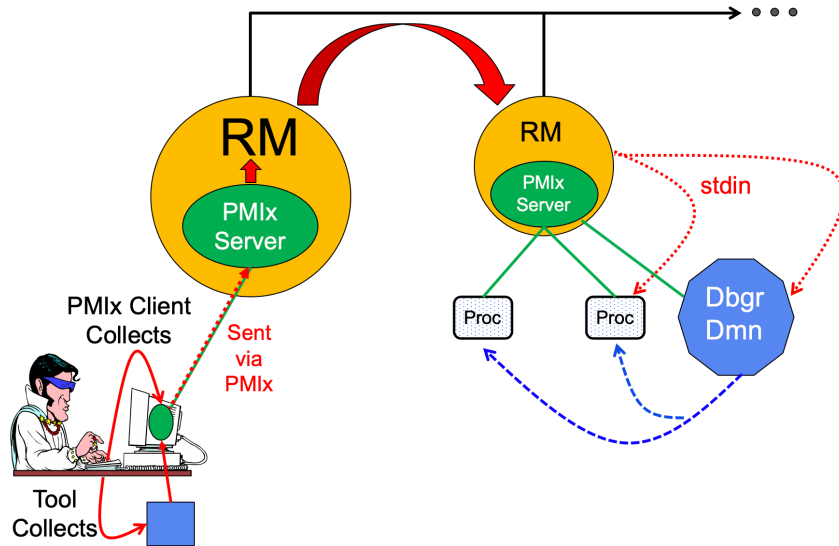


Figure 18.5.: Forwarding stdin

- 1 • external collection directly by the tool. It is assumed that the tool will provide its own code/mechanism for
 2 collecting its **stdin** as the tool developers may choose to insert some filtering and/or editing of the stream
 3 prior to forwarding it. In addition, the tool can directly control the targets for the data on a per-call basis –
 4 i.e., each call to **PMIx_IOF_push** can specify its own set of target recipients for that particular *blob* of
 5 data. Thus, this method provides maximum flexibility, but requires that the tool developer provide their own
 6 code to capture **stdin**.

7 Note that it is the responsibility of the RM to forward data to the host where the target process(es) are
 8 executing, and for the host daemon on that node to deliver the data to the **stdin** of target process(es). The
 9 PMIx server on the remote node is not involved in this process. Systems that do not support forwarding of
 10 **stdin** shall return **PMIX_ERR_NOT_SUPPORTED** in response to a forwarding request.

Advice to users

11 Scalable forwarding of **stdin** represents a significant challenge. Most environments will at least handle a
 12 *send-to-1* model whereby **stdin** is forwarded to a single identified process, and occasionally an additional
 13 *send-to-all* model where **stdin** is forwarded to all processes in the application. Users are advised to check
 14 their host environment for available support as the distribution method lies outside the scope of PMIx.

15 **Stdin** buffering by the RM and/or PMIx library can be problematic. If any targeted recipient is slow reading
 16 data (or decides never to read data), then the data must be buffered in some intermediate daemon or the PMIx
 17 tool library itself. Thus, piping a large amount of data into **stdin** can result in a very large memory footprint
 18 in the system management stack or the tool. Best practices, therefore, typically focus on reading of input files
 19 by application processes as opposed to forwarding of **stdin**.

18.3.3 IO Forwarding Channels

The `pmix_iof_channel_t` structure is a `uint16_t` type that defines a set of bit-mask flags for specifying IO forwarding channels. These can be bitwise OR'd together to reference multiple channels.

```
PMIX_FWD_NO_CHANNELS 0x0000 Forward no channels.
PMIX_FWD_STDIN_CHANNEL 0x0001 Forward stdin.
PMIX_FWD_STDOUT_CHANNEL 0x0002 Forward stdout.
PMIX_FWD_STDErr_CHANNEL 0x0004 Forward stderr.
PMIX_FWD_STDDIAG_CHANNEL 0x0008 Forward stderr, if available.
PMIX_FWD_ALL_CHANNELS 0x00ff Forward all available channels.
```

18.3.4 IO Forwarding constants

```
PMIX_ERR_IOF_FAILURE -172 An IO forwarding operation failed - the affected channel will be
included in the notification.
PMIX_ERR_IOF_COMPLETE -173 IO forwarding of the standard input for this process has completed
- i.e., the stdin file descriptor has closed.
```

18.3.5 IO Forwarding attributes

The following attributes are used to control IO forwarding behavior at the request of tools. Use of the attributes is optional - any option not provided will revert to some implementation-specific value.

```
PMIX_IOF_LOCAL_OUTPUT "pmix.iof.local" (bool) (Provisional)
    Write output streams to local stdout/err
PMIX_IOF_MERGE_STDErr_STDOUT "pmix.iof.mrg" (bool) (Provisional)
    Merge stdout and stderr streams from application procs
PMIX_IOF_CACHE_SIZE "pmix.iof.csize" (uint32_t)
    The requested size of the PMIx server cache in bytes for each specified channel. By default, the server
    is allowed (but not required) to drop all bytes received beyond the max size.
PMIX_IOF_DROP_OLDEST "pmix.iof.old" (bool)
    In an overflow situation, the PMIx server is to drop the oldest bytes to make room in the cache.
PMIX_IOF_DROP_NEWEST "pmix.iof.new" (bool)
    In an overflow situation, the PMIx server is to drop any new bytes received until room becomes
    available in the cache (default).
PMIX_IOF_BUFFERING_SIZE "pmix.iof.bsize" (uint32_t)
    Requests that IO on the specified channel(s) be aggregated in the PMIx tool library until the specified
    number of bytes is collected to avoid being called every time a block of IO arrives. The PMIx tool
    library will execute the callback and reset the collection counter whenever the specified number of
    bytes becomes available. Any remaining buffered data will be flushed to the callback upon a call to
    deregister the respective channel.
PMIX_IOF_BUFFERING_TIME "pmix.iof.btime" (uint32_t)
    Max time in seconds to buffer IO before delivering it. Used in conjunction with buffering size, this
    prevents IO from being held indefinitely while waiting for another payload to arrive.
PMIX_IOF_OUTPUT_RAW "pmix.iof.raw" (bool) (Provisional)
    Do not buffer output to be written as complete lines - output characters as the stream delivers them
```

1 **PMIX_IOF_COMPLETE** "pmix.iof.cmp" (bool)
2 Indicates that the specified IO channel has been closed by the source.

3 **PMIX_IOF_TAG_OUTPUT** "pmix.iof.tag" (bool)
4 Requests that output be prefixed with the nspace,rank of the source and a string identifying the channel
5 (stdout, stderr, etc.).

6 **PMIX_IOF_TIMESTAMP_OUTPUT** "pmix.iof.ts" (bool)
7 Requests that output be marked with the time at which the data was received by the tool - note that this
8 will differ from the time at which the data was collected from the source.

9 **PMIX_IOF_RANK_OUTPUT** "pmix.iof.rank" (bool) *(Provisional)*
10 Tag output with the rank it came from

11 **PMIX_IOF_XML_OUTPUT** "pmix.iof.xml" (bool)
12 Requests that output be formatted in XML.

13 **PMIX_IOF_PUSH_STDIN** "pmix.iof.stdin" (bool)
14 Requests that the PMIx library collect the **stdin** of the requester and forward it to the processes
15 specified in the **PMIX_IOF_push** call. All collected data is sent to the same targets until **stdin** is
16 closed, or a subsequent call to **PMIX_IOF_push** is made that includes the **PMIX_IOF_COMPLETE**
17 attribute indicating that forwarding of **stdin** is to be terminated.

18 **PMIX_IOF_COPY** "pmix.iof.cpy" (bool)
19 Requests that the host environment deliver a copy of the specified output stream(s) to the tool, letting
20 the stream(s) continue to also be delivered to the default location. This allows the tool to tap into the
21 output stream(s) without redirecting it from its current final destination.

22 **PMIX_IOF_REDIRECT** "pmix.iof.redir" (bool)
23 Requests that the host environment intercept the specified output stream(s) and deliver it to the
24 requesting tool instead of its current final destination. This might be used, for example, during a
25 debugging procedure to avoid injection of debugger-related output into the application's results file.
26 The original output stream(s) destination is restored upon termination of the tool.

27 **PMIX_IOF_OUTPUT_TO_FILE** "pmix.iof.file" (char*) *(Provisional)*
28 Direct application output into files of form "<filename>.<namespace>.<rank>.stdout" (for **stdout**) and
29 "<filename>.<namespace>.<rank>.stderr" (for **stderr**). If **PMIX_IOF_MERGE_STDERR_STDOUT** was
30 given, then only the **stdout** file will be created and both streams will be written into it.

31 **PMIX_IOF_OUTPUT_TO_DIRECTORY** "pmix.iof.dir" (char*) *(Provisional)*
32 Direct application output into files of form "<directory>/<namespace>/rank.<rank>/stdout" (for **stdout**)
33 and "<directory>/<namespace>/rank.<rank>/stderr" (for **stderr**). If
34 **PMIX_IOF_MERGE_STDERR_STDOUT** was given, then only the **stdout** file will be created and
35 both streams will be written into it.

36 **PMIX_IOF_FILE_PATTERN** "pmix.iof.fpt" (bool) *(Provisional)*
37 Specified output file is to be treated as a pattern and not automatically annotated by namespace, rank, or
38 other parameters. The pattern can use **%n** for the namespace, and **%r** for the rank wherever those
39 quantities are to be placed. The resulting filename will be appended with ".stdout" for the **stdout**
40 stream and ".stderr" for the **stderr** stream. If **PMIX_IOF_MERGE_STDERR_STDOUT** was given,
41 then only the **stdout** file will be created and both streams will be written into it.

42 **PMIX_IOF_FILE_ONLY** "pmix.iof.only" (bool) *(Provisional)*
43 Output only into designated files - do not also output a copy to the console's stdout/stderr

18.4 Debugger Support

Debuggers are a class of tool that merits special consideration due to their particular requirements for access to job-related information and control over process execution. The primary advantage of using PMIx for these purposes lies in the resulting portability of the debugger as it can be used with any system and/or programming model that supports PMIx. In addition to the general tool support described above, debugger support includes:

- Co-location, co-spawn, and communication wireup of debugger daemons for scalable launch. This includes providing debugger daemons with endpoint connection information across the daemons themselves.
- Identification of the job that is to be debugged. This includes automatically providing debugger daemons with the job-level information for their target job.

Debuggers can also utilize the options in the **PMIx_Spawn** API to exercise a degree of control over spawned jobs for debugging purposes. For example, a debugger can utilize the environmental parameter attributes of Section 12.2.4 to request **LD_PRELOAD** of a memory interceptor library prior to spawning an application process, or interject a custom fork/exec agent to shepherd the application process.

A key element of the debugging process is the ability of the debugger to require that processes *pause* at some well-defined point, thereby providing the debugger with an opportunity to attach and control execution. The actual implementation of the *pause* lies outside the scope of PMIx - it typically requires either the launcher or the application itself to implement the necessary operations. However, PMIx does provide several standard attributes by which the debugger can specify the desired attach point:

- **PMIX_DEBUG_STOP_ON_EXEC** "**pmix.dbg.exec**" (**bool**)
Included in either the **pmix_info_t** array in a **pmix_app_t** description (if the directive applies only to that application) or in the *job_info* array if it applies to all applications in the given spawn request. Indicates that the application is being spawned under a debugger, and that the local launch agent is to pause the resulting application processes on first instruction for debugger attach. The launcher (RM or IL) is to generate the **PMIX_LAUNCH_COMPLETE** event when all processes are stopped at the exec point. Launchers that cannot support this operation shall return an error from the **PMIx_Spawn** API if this behavior is requested.
- **PMIX_DEBUG_STOP_IN_INIT** "**pmix.dbg.init**" (**bool**)
Included in either the **pmix_info_t** array in a **pmix_app_t** description (if the directive applies only to that application) or in the *job_info* array if it applies to all applications in the given spawn request. Indicates that the specified application is being spawned under a debugger. The PMIx client library in each resulting application process shall notify its PMIx server that it is pausing and then pause during **PMIx_Init** of the spawned processes until either released by debugger modification of an appropriate variable or receipt of the **PMIX_DEBUGGER_RELEASE** event. The launcher (RM or IL) is responsible for generating the **PMIX_READY_FOR_DEBUG** event (stipulating a breakpoint of `pmix-init`) when all processes have reached the pause point. PMIx implementations that do not support this operation shall return an error from **PMIx_Init** if this behavior is requested. Launchers that cannot support this operation shall return an error from the **PMIx_Spawn** API if this behavior is requested.
- **PMIX_DEBUG_STOP_IN_APP** "**pmix.dbg.notify**" (**varies**)
Direct specified ranks to stop at application-specific point and notify they are ready-to-debug. The attribute's value can be any of three data types:
 - **bool** - true indicating all ranks

- `pmix_rank_t` - the rank of one proc, or `PMIX_RANK_WILDCARD` for all
- a `pmix_data_array_t` if an array of individual processes are specified

The resulting application processes are to notify their server (by generating the `PMIX_READY_FOR_DEBUG` event) when they reach some application-determined location - the event shall include the `PMIX_BREAKPOINT` attribute indicating where the application has stopped. The application shall pause at that point until released by debugger modification of an appropriate variable. The launcher (RM or IL) is responsible for generating the `PMIX_READY_FOR_DEBUG` event when all processes have indicated they are at the pause point. Launchers that cannot support this operation shall return an error from the `PMIx_Spawn` API if this behavior is requested.

Note that there is no mechanism by which the PMIx library or the launcher can verify that an application will recognize and support the `PMIX_DEBUG_STOP_IN_APP` request. Debuggers utilizing this attachment method must, therefore, be prepared to deal with the case where the application fails to recognize and/or honor the request.

If the PMIx implementation and/or the host environment support it, debuggers can utilize the `PMIx_Query_info` API to determine which features are available via the `PMIX_QUERY_ATTRIBUTE_SUPPORT` attribute.

- `PMIX_DEBUG_STOP_IN_INIT` by checking `PMIX_CLIENT_ATTRIBUTES` for the `PMIx_Init` API.
- `PMIX_DEBUG_STOP_ON_EXEC` by checking `PMIX_HOST_ATTRIBUTES` for the `PMIx_Spawn` API.

The target namespace or process (as given by the debugger in the spawn request) shall be provided to each daemon in its job-level information via the `PMIX_DEBUG_TARGET` attribute. Debugger daemons are responsible for self-determining their specific target process(es), and can then utilize the `PMIx_Query_info` API to obtain information about them (see Fig 18.6) - e.g., to obtain the PIDs of the local processes to which they need to attach. PMIx provides the `pmix_proc_info_t` structure for organizing information about a process' PID, location, and state. Debuggers may request information on a given job at two levels:

- `PMIX_QUERY_PROC_TABLE` "`pmix.qry.ptable`" (`char*`)
Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: `PMIX_NAMESPACE` indicating the namespace whose process table is being queried.
- `PMIX_QUERY_LOCAL_PROC_TABLE` "`pmix.qry.lptable`" (`char*`)
Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in the specified namespace executing on the same node as the requester, ordered by process job rank. REQUIRED QUALIFIER: `PMIX_NAMESPACE` indicating the namespace whose local process table is being queried. OPTIONAL QUALIFIER: `PMIX_HOSTNAME` indicating the host whose local process table is being queried. By default, the query assumes that the host upon which the request was made is to be used.

Note that the information provided in the returned proctable represents a snapshot in time. Any process, regardless of role (tool, client, debugger, etc.) can obtain the proctable of a given namespace so long as it has the system-determined authorizations to do so. The list of namespaces available via a given server can be obtained using the `PMIx_Query_info` API with the `PMIX_QUERY_NAMESPACES` key.

Debugger daemons can be started in two ways - either at the same time the application is spawned, or separately at a later time.

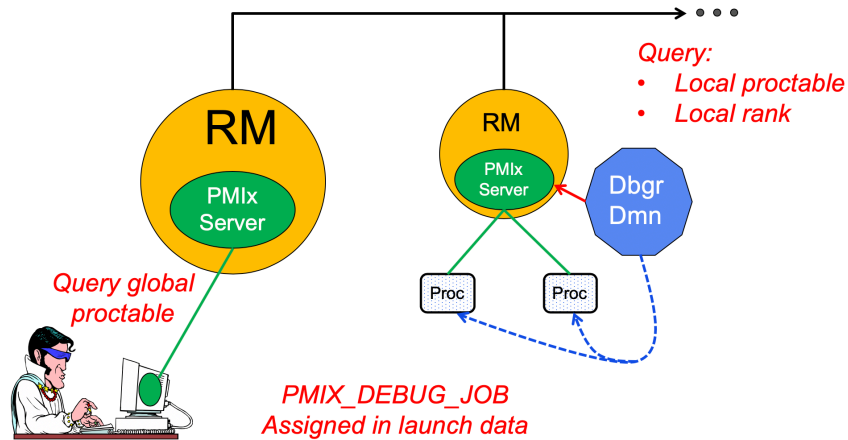


Figure 18.6.: Obtaining proctables

18.4.1 Co-Location of Debugger Daemons

Debugging operations typically require the use of daemons that are located on the same node as the processes they are attempting to debug. The debugger can, of course, specify its own mapping method when issuing its spawn request or utilize its own internal launcher to place the daemons. However, when attaching to a running job, PMIx provides debuggers with a simplified method for requesting that the launcher associated with the job *co-locate* the required daemons. Debuggers can request *co-location* of their daemons by adding the following attributes to the **PMIx_Spawn** used to spawn them:

- **PMIX_DEBUGGER_DAEMONS** - indicating that the launcher is being asked to spawn debugger daemons.
- **PMIX_DEBUG_TARGET** - indicating the job or process that is to be debugged. This allows the launcher to identify the processes to be debugged and their location. Note that the debugger job shall be assigned its own namespace (different from that of the job it is being spawned to debug) and each daemon will be assigned a unique rank within that namespace.
- **PMIX_DEBUG_DAEMONS_PER_PROC** - specifies the number of debugger daemons to be co-located per target process.
- **PMIX_DEBUG_DAEMONS_PER_NODE** - specifies the number of debugger daemons to be co-located per node where at least one target process is executing.

Debugger daemons spawned in this manner shall be provided with the typical PMIx information for their own job plus the target they are to debug via the **PMIX_DEBUG_TARGET** attribute. The debugger daemons spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by referencing their own **PMIX_LOCAL_RANK** in the daemon debugger job versus the corresponding **PMIX_LOCAL_RANK** of the target processes on the node. Note that the debugger will be attaching to the application processes at some arbitrary point in the application's execution unless some method for pausing the application (e.g., by providing a PMIx directive at time of launch, or via a tool using the **PMIx_Job_control** API to direct that the process be paused) has been employed.

Advice to users

Note that the tool calling `PMIx_Spawn` to request the launch of the debugger daemons is *not* included in the resulting job - i.e., the debugger daemons do not inherit the namespace of the tool. Thus, collective operations and notifications that target the debugger daemon job will not include the tool unless the namespace/rank of the tool is explicitly included.

18.4.2 Co-Spawn of Debugger Daemons

In the case where a job is being spawned under the control of a debugger, PMIx provides a shortcut method for spawning the debugger's daemons in parallel with the job. This requires that the debugger be specified as one of the `pmix_app_t` in the same spawn command used to start the job. The debugger application must include at least the `PMIX_DEBUGGER_DAEMONS` attribute identifying itself as a debugger, and may utilize either a mapping option to direct daemon placement, or one of the `PMIX_DEBUG_DAEMONS_PER_PROC` or `PMIX_DEBUG_DAEMONS_PER_NODE` directives.

The launcher must not include information regarding the debugger daemons in the job-level info provided to the rest of the `pmix_app_t`s, nor in any calculated rank values (e.g., `PMIX_NODE_RANK` or `PMIX_LOCAL_RANK`) in those applications. The debugger job is to be assigned its own namespace and each debugger daemon shall receive a unique rank - i.e., the debugger application is to be treated as a completely separate PMIx job that is simply being started in parallel with the user's applications. The launcher is free to implement the launch as a single operation for both the applications and debugger daemons (preferred), or may stage the launches as required. The launcher shall not return from the `PMIx_Spawn` command until all included applications and the debugger daemons have been started.

Attributes that apply to both the debugger daemons and the application processes can be specified in the `job_info` array passed into the `PMIx_Spawn` API. Attributes that either (a) apply solely to the debugger daemons or to one of the applications included in the spawn request, or (b) have values that differ from those provided in the `job_info` array, should be specified in the `info` array in the corresponding `pmix_app_t`. Note that PMIx job *pause* attributes (e.g., `PMIX_DEBUG_STOP_IN_INIT`) do not apply to applications (defined in `pmix_app_t`) where the `PMIX_DEBUGGER_DAEMONS` attribute is set to `true`.

Debugger daemons spawned in this manner shall be provided with the typical PMIx information for their own job plus the target they are to debug via the `PMIX_DEBUG_TARGET` attribute. The debugger daemons spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by referencing their own `PMIX_LOCAL_RANK` in the daemon debugger job versus the corresponding `PMIX_LOCAL_RANK` of the target processes on the node.

Advice to users

Note that the tool calling `PMIx_Spawn` to request the launch of the debugger daemons is *not* included in the resulting job - i.e., the debugger daemons do not inherit the namespace of the tool. Thus, collective operations and notifications that target the debugger daemon job will not include the tool unless the namespace/rank of the tool is explicitly included.

The `PMIx_Spawn` API only supports the return of a single namespace resulting from the spawn request. In the case where the debugger job is co-spawned with the application, the spawn function shall return the namespace of the application and not the debugger job. Tools requiring access to the namespace of the debugger job must query the launcher for the spawned namespaces to find the one belonging to the debugger job.

18.4.3 Debugger Agents

Individual debuggers may, depending upon implementation, require varying degrees of control over each application process when it is started beyond those available via directives to `PMIx_Spawn`. PMIx offers two mechanisms to help provide a means of meeting these needs.

The `PMIX_FORKEXEC_AGENT` attribute allows the debugger to specify an intermediate process (the Fork/Exec Agent (FEA)) for spawning the actual application process (see Fig. 18.7a), thereby interposing the debugger daemon between the application process and the launcher's daemon. Instead of spawning the application process, the launcher will spawn the FEA, which will connect back to the PMIx server as a tool to obtain the spawn description of the application process it is to spawn. The PMIx server in the launcher's daemon shall not register the fork/exec agent as a local client process, nor shall the launcher include the agent in any of the job-level values (e.g., `PMIX_RANK` within the job or `PMIX_LOCAL_RANK` on the node) provided to the application process. The launcher shall treat the collection of FEAs as a debugger job equivalent to the co-spawn use-case described in Section 18.4.2.

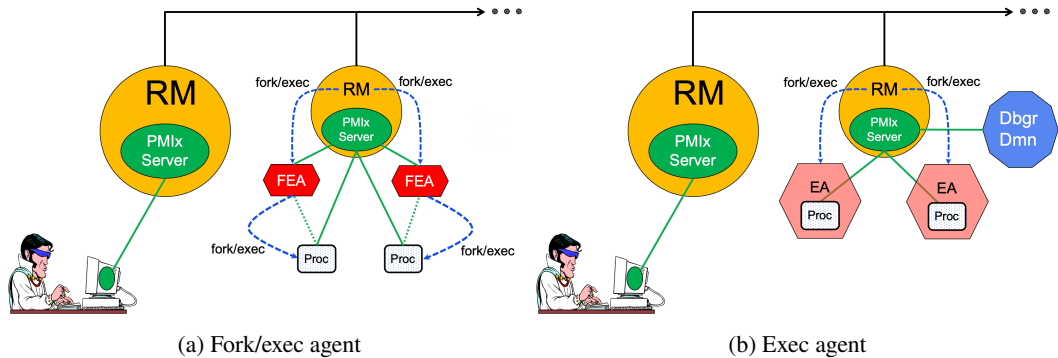


Figure 18.7.: Intermediate agents

In contrast, the `PMIX_EXEC_AGENT` attribute (Fig. 18.7b) allows the debugger to specify an agent that will perform some preparatory actions and then exec the eventual application process to replace itself. In this scenario, the exec agent is provided with the application process' command line as arguments on its command line (e.g., `./agent appargv[0] appargv[1]`) and does not connect back to the host's PMIx server. It is the responsibility of the exec agent to properly separate its own command line arguments (if any) from the application description.

18.4.4 Tracking the job lifecycle

There are a wide range of events a debugger can register to receive, but three are specifically defined for tracking a job's progress:

- `PMIX_EVENT_JOB_START` indicates when the first process in the job has been spawned.
- `PMIX_LAUNCH_COMPLETE` indicates when the last process in the job has been spawned.

- **PMIX_EVENT_JOB_END** indicates that all processes have terminated.

Each event is required to contain at least the namespace of the corresponding job and a **PMIX_EVENT_TIMESTAMP** indicating the time the event occurred. In addition, the **PMIX_EVENT_JOB_END** event shall contain the returned status code (**PMIX_JOB_TERM_STATUS**) for the corresponding job, plus the identity (**PMIX_PROCID**) and exit status (**PMIX_EXIT_CODE**) of the first failed process, if applicable. Generation of these events by the launcher can be requested by including the **PMIX_NOTIFY_JOB_EVENTS** attributes in the spawn request. Note that these events can be logged via the **PMIx_Log** API by including the **PMIX_LOG_JOB_EVENTS** attribute - this can be done either in conjunction with generated events, or in place of them.

Alternatively, if the debugger or tool solely wants to be alerted to job termination, then including the **PMIX_NOTIFY_COMPLETION** attribute in the spawn request would suffice. This attribute directs the launcher to provide just the **PMIX_EVENT_JOB_END** event. Note that this event can be logged via the **PMIx_Log** API by including the **PMIX_LOG_COMPLETION** attribute - this can be done either in conjunction with the generated event, or in place of it.

Advice to users

The PMIx server is required to cache events in order to avoid race conditions - e.g., when a tool is trying to register for the **PMIX_EVENT_JOB_END** event from a very short-lived job. Accordingly, registering for job-related events can result in receiving events relating to jobs other than the one of interest.

Users are therefore advised to specify the job whose events are of interest by including the **PMIX_EVENT_AFFECTED_PROC** or **PMIX_EVENT_AFFECTED_PROCS** attribute in the *info* array passed to the **PMIx_Register_event_handler** API.

18.4.4.1 Job lifecycle events

- PMIX_EVENT_JOB_START** -191 The first process in the job has been spawned - includes **PMIX_EVENT_TIMESTAMP** as well as the **PMIX_JOBID** and/or **PMIX_NAMESPACE** of the job.
- PMIX_LAUNCH_COMPLETE** -174 All processes in the job have been spawned - includes **PMIX_EVENT_TIMESTAMP** as well as the **PMIX_JOBID** and/or **PMIX_NAMESPACE** of the job.
- PMIX_EVENT_JOB_END** -145 All processes in the job have terminated - includes **PMIX_EVENT_TIMESTAMP** when the last process terminated as well as the **PMIX_JOBID** and/or **PMIX_NAMESPACE** of the job.
- PMIX_EVENT_SESSION_START** -192 The allocation has been instantiated and is ready for use - includes **PMIX_EVENT_TIMESTAMP** as well as the **PMIX_SESSION_ID** of the allocation. This event is issued after any system-controlled prologue has completed, but before any user-specified actions are taken.
- PMIX_EVENT_SESSION_END** -193 The allocation has terminated - includes **PMIX_EVENT_TIMESTAMP** as well as the **PMIX_SESSION_ID** of the allocation. This event is issued after any user-specified actions have completed, but before any system-controlled epilogue is performed.

The following events relate to processes within a job:

- PMIX_EVENT_PROC_TERMINATED** -201 The specified process(es) terminated - normal or abnormal termination will be indicated by the **PMIX_PROC_TERM_STATUS** in the *info* array of the notification. Note that a request for individual process events can generate a significant event volume from large-scale jobs.

1 **PMIX_ERR_PROC_TERM_WO_SYNC** -200 Process terminated without calling **PMIx_Finalize**, or
2 was a member of an assemblage formed via **PMIx_Connect** and terminated or called
3 **PMIx_Finalize** without first calling **PMIx_Disconnect** (or its non-blocking form) from that
4 assemblage.

5 The following constants may be included via the **PMIX_JOB_TERM_STATUS** attributed in the *info* array in
6 the **PMIX_EVENT_JOB_END** event notification to provide more detailed information regarding the reason for
7 job abnormal termination:

8 **PMIX_ERR_JOB_CANCELED** -180 The job was canceled by the host environment.

9 **PMIX_ERR_JOB_ABORTED** -182 One or more processes in the job called abort, causing the job to be
10 terminated.

11 **PMIX_ERR_JOB_KILLED_BY_CMD** -183 The job was killed by user command.

12 **PMIX_ERR_JOB_ABORTED_BY_SIG** -184 The job was aborted due to receipt of an error signal
13 (e.g., SIGKILL).

14 **PMIX_ERR_JOB_TERM_WO_SYNC** -185 The job was terminated due to at least one process
15 terminating without calling **PMIx_Finalize**, or was a member of an assemblage formed via
16 **PMIx_Connect** and terminated or called **PMIx_Finalize** without first calling
17 **PMIx_Disconnect** (or its non-blocking form) from that assemblage.

18 **PMIX_ERR_JOB_SENSOR_BOUND_EXCEEDED** -186 The job was terminated due to one or more
19 processes exceeding a specified sensor limit.

20 **PMIX_ERR_JOB_NON_ZERO_TERM** -187 The job was terminated due to one or more processes
21 exiting with a non-zero status.

22 **PMIX_ERR_JOB_ABORTED_BY_SYS_EVENT** -189 The job was aborted due to receipt of a system
23 event.

24 18.4.4.2 Job lifecycle attributes

25 **PMIX_JOB_TERM_STATUS** "pmix.job.term.status" (**pmix_status_t**)

26 Status returned by job upon its termination. The status will be communicated as part of a PMIx event
27 payload provided by the host environment upon termination of a job. Note that generation of the
28 **PMIX_EVENT_JOB_END** event is optional and host environments may choose to provide it only upon
29 request.

30 **PMIX_PROC_STATE_STATUS** "pmix.proc.state" (**pmix_proc_state_t**)

31 State of the specified process as of the last report - may not be the actual current state based on update
32 rate.

33 **PMIX_PROC_TERM_STATUS** "pmix.proc.term.status" (**pmix_status_t**)

34 Status returned by a process upon its termination. The status will be communicated as part of a PMIx
35 event payload provided by the host environment upon termination of a process. Note that generation of
36 the **PMIX_EVENT_PROC_TERMINATED** event is optional and host environments may choose to
37 provide it only upon request.

38 18.4.5 Debugger-related constants

39 The following constants are used in events used to coordinate applications and the debuggers attaching to them.

40 **PMIX_READY_FOR_DEBUG** -58 Event indicating a job (or specified set of processes) is ready for
41 debug - includes identification of the target processes as well as the **PMIX_BREAKPOINT** indicating
42 where the target is waiting

43 **PMIX_DEBUGGER_RELEASE** -3 Release a tool that is paused during **PMIx_tool_init**.

18.4.6 Debugger attributes

Attributes used to assist debuggers - these are values that can either be passed to the `PMIx_Spawn` APIs or accessed by a debugger itself using the `PMIx_Get` API with the `PMIX_RANK_WILDCARD` rank.

PMIX_DEBUG_STOP_ON_EXEC "`pmix.dbg.exec`" (`bool`)

Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies only to that application) or in the `job_info` array if it applies to all applications in the given spawn request. Indicates that the application is being spawned under a debugger, and that the local launch agent is to pause the resulting application processes on first instruction for debugger attach. The launcher (RM or IL) is to generate the `PMIX_LAUNCH_COMPLETE` event when all processes are stopped at the exec point.

PMIX_DEBUG_STOP_IN_INIT "`pmix.dbg.init`" (`bool`)

Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies only to that application) or in the `job_info` array if it applies to all applications in the given spawn request. Indicates that the specified application is being spawned under a debugger. The PMIx client library in each resulting application process shall notify its PMIx server that it is pausing and then pause during `PMIx_Init` of the spawned processes until either released by debugger modification of an appropriate variable or receipt of the `PMIX_DEBUGGER_RELEASE` event. The launcher (RM or IL) is responsible for generating the `PMIX_READY_FOR_DEBUG` event (stipulating a breakpoint of `pmix-init`) when all processes have reached the pause point.

PMIX_DEBUG_STOP_IN_APP "`pmix.dbg.notify`" (`varies`)

Direct specified ranks to stop at application-specific point and notify they are ready-to-debug. The attribute's value can be any of three data types:

- `bool` - true indicating all ranks
- `pmix_rank_t` - the rank of one proc, or `PMIX_RANK_WILDCARD` for all
- a `pmix_data_array_t` if an array of individual processes are specified

The resulting application processes are to notify their server (by generating the `PMIX_READY_FOR_DEBUG` event) when they reach some application-determined location - the event shall include the `PMIX_BREAKPOINT` attribute indicating where the application has stopped. The application shall pause at that point until released by debugger modification of an appropriate variable. The launcher (RM or IL) is responsible for generating the `PMIX_READY_FOR_DEBUG` event when all processes have indicated they are at the pause point.

PMIX_BREAKPOINT "`pmix.brkpnt`" (`char*`)

String ID of the breakpoint where the process(es) is(are) waiting.

PMIX_DEBUG_TARGET "`pmix.dbg.tgt`" (`pmix_proc_t*`)

Identifier of process(es) to be debugged - a rank of `PMIX_RANK_WILDCARD` indicates that all processes in the specified namespace are to be included.

PMIX_DEBUGGER_DAEMONS "`pmix.debugger`" (`bool`)

Included in the `pmix_info_t` array of a `pmix_app_t`, this attribute declares that the application consists of debugger daemons and shall be governed accordingly. If used as the sole `pmix_app_t` in a `PMIx_Spawn` request, then the `PMIX_DEBUG_TARGET` attribute must also be provided (in either the `job_info` or in the `info` array of the `pmix_app_t`) to identify the namespace to be debugged so that the launcher can determine where to place the spawned daemons. If neither `PMIX_DEBUG_DAEMONS_PER_PROC` nor `PMIX_DEBUG_DAEMONS_PER_NODE` is specified, then the launcher shall default to a placement policy of one daemon per process in the target job.

PMIX_COSPAWN_APP "`pmix.cospawn`" (`bool`)

Designated application is to be spawned as a disconnected job - i.e., the launcher shall not include the application in any of the job-level values (e.g., `PMIX_RANK` within the job) provided to any other

1 application process generated by the same spawn request. Typically used to cospawn debugger
2 daemons alongside an application.

3 **PMIX_DEBUG_DAEMONS_PER_PROC** "pmix.dbg.dpproc" (uint16_t)

4 Number of debugger daemons to be spawned per application process. The launcher is to pass the
5 identifier of the namespace to be debugged by including the **PMIX_DEBUG_TARGET** attribute in the
6 daemon's job-level information. The debugger daemons spawned on a given node are responsible for
7 self-determining their specific target process(es) - e.g., by referencing their own **PMIX_LOCAL_RANK**
8 in the daemon debugger job versus the corresponding **PMIX_LOCAL_RANK** of the target processes on
9 the node.

10 **PMIX_DEBUG_DAEMONS_PER_NODE** "pmix.dbg.dpnd" (uint16_t)

11 Number of debugger daemons to be spawned on each node where the target job is executing. The
12 launcher is to pass the identifier of the namespace to be debugged by including the
13 **PMIX_DEBUG_TARGET** attribute in the daemon's job-level information. The debugger daemons
14 spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by
15 referencing their own **PMIX_LOCAL_RANK** in the daemon debugger job versus the corresponding
16 **PMIX_LOCAL_RANK** of the target processes on the node.

17 **PMIX_QUERY_PROC_TABLE** "pmix.qry.ptable" (char*)

18 Returns a (**pmix_data_array_t**) array of **pmix_proc_info_t**, one entry for each process in
19 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: **PMIX_NAMESPACE**
20 indicating the namespace whose process table is being queried.

21 **PMIX_QUERY_LOCAL_PROC_TABLE** "pmix.qry.lptable" (char*)

22 Returns a (**pmix_data_array_t**) array of **pmix_proc_info_t**, one entry for each process in
23 the specified namespace executing on the same node as the requester, ordered by process job rank.
24 REQUIRED QUALIFIER: **PMIX_NAMESPACE** indicating the namespace whose local process table is
25 being queried. OPTIONAL QUALIFIER: **PMIX_HOSTNAME** indicating the host whose local process
26 table is being queried. By default, the query assumes that the host upon which the request was made is
27 to be used.

28 18.5 Tool-Specific APIs

29 PMIx-based tools automatically have access to all PMIx client functions. Tools designated as a *launcher* or a
30 *server* will also have access to all PMIx server functions. There are, however, an additional set of functions
31 (described in this section) that are specific to a PMIx tool. Access to those functions require use of the tool
32 initialization routine.

33 18.5.1 PMIx_tool_init

34 Summary

35 Initialize the PMIx library for operating as a tool, optionally connecting to a specified PMIx server.

36 Format

C

```
37 pmix_status_t  
38 PMIx_tool_init(pmix_proc_t *proc,  
39               pmix_info_t info[], size_t ninfo);
```

```

1  INOUT proc
2      pmix_proc_t structure (handle)
3  IN   info
4      Array of pmix_info_t structures (array of handles)
5  IN   ninfo
6      Number of elements in the info array (size_t)
7
Returns PMIX_SUCCESS or a negative value indicating the error.

```

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

```

9  PMIX_TOOL_NAMESPACE "pmix.tool.namespace" (char*)
10     Name of the namespace to use for this tool.
11  PMIX_TOOL_RANK "pmix.tool.rank" (uint32_t)
12     Rank of this tool.
13  PMIX_TOOL_DO_NOT_CONNECT "pmix.tool.nocon" (bool)
14     The tool wants to use internal PMIx support, but does not want to connect to a PMIx server.
15  PMIX_TOOL_ATTACHMENT_FILE "pmix.tool.attach" (char*)
16     Pathname of file containing connection information to be used for attaching to a specific server.
17  PMIX_SERVER_URI "pmix.srvr.uri" (char*)
18     URI of the PMIx server to be contacted.
19  PMIX_TCP_URI "pmix.tcp.uri" (char*)
20     The URI of the PMIx server to connect to, or a file name containing it in the form of file:<name
21     of file containing it>.
22  PMIX_SERVER_PIDINFO "pmix.srvr.pidinfo" (pid_t)
23     PID of the target PMIx server for a tool.
24  PMIX_SERVER_NAMESPACE "pmix.srv.namespace" (char*)
25     Name of the namespace to use for this PMIx server.
26  PMIX_CONNECT_TO_SYSTEM "pmix.cnct.sys" (bool)
27     The requester requires that a connection be made only to a local, system-level PMIx server.
28  PMIX_CONNECT_SYSTEM_FIRST "pmix.cnct.sys.first" (bool)
29     Preferentially, look for a system-level PMIx server first.

```


Optional Attributes

The following attributes are optional for implementers of PMIx libraries:

PMIX_CONNECT_RETRY_DELAY "pmix.tool.retry" (uint32_t)

Time in seconds between connection attempts to a PMIx server - the default value is implementation specific.

PMIX_CONNECT_MAX_RETRIES "pmix.tool.mretries" (uint32_t)

Maximum number of times to try to connect to PMIx server - the default value is implementation specific.

PMIX_SOCKET_MODE "pmix.sockmode" (uint32_t)

POSIX *mode_t* (9 bits valid). If the library supports socket connections, this attribute may be supported for setting the socket mode.

PMIX_TCP_REPORT_URI "pmix.tcp.repuri" (char*)

If provided, directs that the TCP URI be reported and indicates the desired method of reporting: '-' for stdout, '+' for stderr, or filename. If the library supports TCP socket connections, this attribute may be supported for reporting the URI.

PMIX_TCP_IF_INCLUDE "pmix.tcp.ifinclude" (char*)

Comma-delimited list of devices and/or CIDR notation to include when establishing the TCP connection. If the library supports TCP socket connections, this attribute may be supported for specifying the interfaces to be used.

PMIX_TCP_IF_EXCLUDE "pmix.tcp.ifexclude" (char*)

Comma-delimited list of devices and/or CIDR notation to exclude when establishing the TCP connection. If the library supports TCP socket connections, this attribute may be supported for specifying the interfaces that are *not* to be used.

PMIX_TCP_IPV4_PORT "pmix.tcp.ipv4" (int)

The IPv4 port to be used.. If the library supports IPV4 connections, this attribute may be supported for specifying the port to be used.

PMIX_TCP_IPV6_PORT "pmix.tcp.ipv6" (int)

The IPv6 port to be used. If the library supports IPV6 connections, this attribute may be supported for specifying the port to be used.

PMIX_TCP_DISABLE_IPV4 "pmix.tcp.disipv4" (bool)

Set to **true** to disable IPv4 family of addresses. If the library supports IPV4 connections, this attribute may be supported for disabling it.

PMIX_TCP_DISABLE_IPV6 "pmix.tcp.disipv6" (bool)

Set to **true** to disable IPv6 family of addresses. If the library supports IPV6 connections, this attribute may be supported for disabling it.

PMIX_EXTERNAL_PROGRESS "pmix.evext" (bool)

The host shall progress the PMIx library via calls to **PMIx_Progress**

PMIX_EVENT_BASE "pmix.evbase" (void*)

1 Pointer to an `event_base` to use in place of the internal progress thread. All PMIx library events are
2 to be assigned to the provided event base. The event base *must* be compatible with the event library
3 used by the PMIx implementation - e.g., either both the host and PMIx library must use libevent, or
4 both must use libev. Cross-matches are unlikely to work and should be avoided - it is the responsibility
5 of the host to ensure that the PMIx implementation supports (and was built with) the appropriate event
6 library.

7 **PMIX_IOF_LOCAL_OUTPUT** "pmix.iof.local" (bool)
8 Write output streams to local stdout/err



9 Description

10 Initialize the PMIx tool, returning the process identifier assigned to this tool in the provided `pmix_proc_t`
11 struct. The *info* array is used to pass user requests pertaining to the initialization and subsequent operations.
12 Passing a `NULL` value for the array pointer is supported if no directives are desired.

13 If called with the `PMIX_TOOL_DO_NOT_CONNECT` attribute, the PMIx tool library will fully initialize but
14 not attempt to connect to a PMIx server. The tool can connect to a server at a later point in time, if desired, by
15 calling the `PMIx_tool_attach_to_server` function. If provided, the *proc* structure will be set to a
16 zero-length namespace and a rank of `PMIX_RANK_UNDEF` unless the `PMIX_TOOL_NAMESPACE` and
17 `PMIX_TOOL_RANK` attributes are included in the *info* array.

18 In all other cases, the PMIx tool library will automatically attempt to connect to a PMIx server according to
19 the precedence chain described in Section 18.1. If successful, the function will return `PMIX_SUCCESS` and
20 will fill the process structure (if provided) with the assigned namespace and rank of the tool. The server to
21 which the tool connects will be designated its *primary* server. Note that each connection attempt in the above
22 precedence chain will retry (with delay between each retry) a number of times according to the values of the
23 corresponding attributes.

24 Note that the PMIx tool library is referenced counted, and so multiple calls to `PMIx_tool_init` are
25 allowed. If the tool is not connected to any server when this API is called, then the tool will attempt to connect
26 to a server unless the `PMIX_TOOL_DO_NOT_CONNECT` is included in the call to API.

27 18.5.2 PMIx_tool_finalize

28 Summary

29 Finalize the PMIx tool library.

30 Format

PMIx v2.0

31 `pmix_status_t`
32 `PMIx_tool_finalize(void);`

33 Returns `PMIX_SUCCESS` or a negative value indicating the error.

34 Description

35 Finalize the PMIx tool library, closing all existing connections to servers. An error code will be returned if, for
36 some reason, a connection cannot be cleanly terminated — in such cases, the connection is dropped. Upon
37 detecting loss of the connection, the PMIx server shall cleanup all associated records of the tool.

1 18.5.3 PMIx_tool_disconnect

2 Summary

3 Disconnect the PMIx tool from the specified server connection while leaving the tool library initialized.

4 Format

PMIx v4.0

C

5 `pmix_status_t`

6 `PMIx_tool_disconnect(const pmix_proc_t *server);`

C

7 **IN** `server`

8 `pmix_proc_t` structure (handle)

9 Returns `PMIX_SUCCESS` or a negative value indicating the error.

10 Description

11 Close the current connection to the specified server, if one has been made, while leaving the PMIx library
12 initialized. An error code will be returned if, for some reason, the connection cannot be cleanly terminated - in
13 this case, the connection is dropped. In either case, the library will remain initialized. Upon detecting loss of
14 the connection, the PMIx server shall cleanup all associated records of the tool.

15 Note that if the server being disconnected is the current *primary* server, then all operations requiring support
16 from a server will return the `PMIX_ERR_UNREACH` error until the tool either designates an existing
17 connection to be the *primary* server or, if no other connections exist, the tool establishes a connection to a
18 PMIx server.

19 18.5.4 PMIx_tool_attach_to_server

20 Summary

21 Establish a connection to a PMIx server.

22 Format

PMIx v4.0

C

23 `pmix_status_t`

24 `PMIx_tool_attach_to_server(pmix_proc_t *proc,`
25 `pmix_proc_t *server,`
26 `pmix_info_t info[],`
27 `size_t ninfo);`

C

28 **INOUT** `proc`

29 Pointer to `pmix_proc_t` structure (handle)

30 **INOUT** `server`

31 Pointer to `pmix_proc_t` structure (handle)

32 **IN** `info`

33 Array of `pmix_info_t` structures (array of handles)

34 **IN** `ninfo`

35 Number of elements in the `info` array (`size_t`)

36 Returns `PMIX_SUCCESS` or a negative value indicating the error.

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

PMIX_TOOL_ATTACHMENT_FILE "pmix.tool.attach" (char*)

Pathname of file containing connection information to be used for attaching to a specific server.

PMIX_SERVER_URI "pmix.srvr.uri" (char*)

URI of the PMIx server to be contacted.

PMIX_TCP_URI "pmix.tcp.uri" (char*)

The URI of the PMIx server to connect to, or a file name containing it in the form of **file:<name of file containing it>**.

PMIX_SERVER_PIDINFO "pmix.srvr.pidinfo" (pid_t)

PID of the target PMIx server for a tool.

PMIX_SERVER_NAMESPACE "pmix.srv.namespace" (char*)

Name of the namespace to use for this PMIx server.

PMIX_CONNECT_TO_SYSTEM "pmix.cnct.sys" (bool)

The requester requires that a connection be made only to a local, system-level PMIx server.

PMIX_CONNECT_SYSTEM_FIRST "pmix.cnct.sys.first" (bool)

Preferentially, look for a system-level PMIx server first.

PMIX_PRIMARY_SERVER "pmix.pri.srvr" (bool)

The server to which the tool is connecting shall be designated the *primary* server once connection has been accomplished.

Description

Establish a connection to a server. This function can be called at any time by a PMIx tool to create a new connection to a server. If a specific server is given and the tool is already attached to it, then the API shall return **PMIX_SUCCESS** without taking any further action. In all other cases, the tool will attempt to discover a server using the method described in Section 18.1, ignoring all candidates to which it is already connected. The **PMIX_ERR_UNREACH** error shall be returned if no new connection is made.

The process identifier assigned to this tool is returned in the provided *proc* structure. Passing a value of **NULL** for the *proc* parameter is allowed if the user wishes solely to connect to a PMIx server and does not require return of the identifier at that time.

The process identifier of the server to which the tool attached is returned in the *server* structure. Passing a value of **NULL** for the *proc* parameter is allowed if the user wishes solely to connect to a PMIx server and does not require return of the identifier at that time.

Note that the **PMIX_PRIMARY_SERVER** attribute must be included in the *info* array if the server being connected to is to become the primary server, or a call to **PMIx_tool_set_server** must be provided immediately after the call to this function.

Advice to PMIx library implementers

1 When a tool connects to a server that is under a different namespace manager (e.g., host RM) from the prior
2 server, the namespace in the identifier of the tool must remain unique in the new universe. If the namespace of
3 the tool fails to meet this criteria in the new universe, then the new namespace manager is required to return an
4 error and the connection attempt must fail.

Advice to users

5 Some PMIx implementations may not support connecting to a server that is not under the same namespace
6 manager (e.g., host RM) as the server to which the tool is currently connected.

7 18.5.5 PMIx_tool_get_servers

8 Summary

9 Get an array containing the `pmix_proc_t` process identifiers of all servers to which the tool is currently
10 connected.

11 Format

PMIx v4.0

C

12 `pmix_status_t`

13 `PMIx_tool_get_servers(pmix_proc_t *servers[], size_t *nservers);`

14 OUT `servers`

15 Address where the pointer to an array of `pmix_proc_t` structures shall be returned (handle)

16 INOUT `nservers`

17 Address where the number of elements in `servers` shall be returned (handle)

18 Returns `PMIX_SUCCESS` or a negative value indicating the error.

19 Description

20 Return an array containing the `pmix_proc_t` process identifiers of all servers to which the tool is currently
21 connected. The process identifier of the current primary server shall be the first entry in the array, with the
22 remaining entries in order of attachment from earliest to most recent.

23 18.5.6 PMIx_tool_set_server

24 Summary

25 Designate a server as the tool's *primary* server.

Format

```
pmix_status_t
PMIx_tool_set_server(const pmix_proc_t *server                pmix_info_t
info[], size_t ninfo);
```

IN server
pmix_proc_t structure (handle)

IN info
Array of pmix_info_t structures (array of handles)

IN ninfo
Number of elements in the info array (size_t)

Returns PMIX_SUCCESS or a negative value indicating the error.

Required Attributes

The following attributes are required to be supported by all PMIx libraries:

PMIX_WAIT_FOR_CONNECTION "pmix.wait.conn" (bool)

Wait until the specified process has connected to the requesting tool or server, or the operation times out (if the PMIX_TIMEOUT directive is included in the request).

PMIX_TIMEOUT "pmix.timeout" (int)

Time in seconds before the specified operation should time out (zero indicating infinite) and return the PMIX_ERR_TIMEOUT error. Care should be taken to avoid race conditions caused by multiple layers (client, server, and host) simultaneously timing the operation.

Description

Designate the specified server to be the tool's primary server for all subsequent API calls.

18.5.7 PMIx_IOF_pull

Summary

Register to receive output forwarded from a set of remote processes.

Format

C

```
1 pmix_status_t  
2 PMIx_IOF_pull(const pmix_proc_t procs[], size_t nprocs,  
3               const pmix_info_t directives[], size_t ndirs,  
4               pmix_iof_channel_t channel,  
5               pmix_iof_cbfunc_t cbfunc,  
6               pmix_hdlr_reg_cbfunc_t regcbfunc,  
7               void *regcbdata);  
8
```

C

9 **IN** **procs**
10 Array of proc structures identifying desired source processes (array of handles)

11 **IN** **nprocs**
12 Number of elements in the *procs* array (integer)

13 **IN** **directives**
14 Array of **pmix_info_t** structures (array of handles)

15 **IN** **ndirs**
16 Number of elements in the *directives* array (integer)

17 **IN** **channel**
18 Bitmask of IO channels included in the request (**pmix_iof_channel_t**)

19 **IN** **cbfunc**
20 Callback function for delivering relevant output (**pmix_iof_cbfunc_t** function reference)

21 **IN** **regcbfunc**
22 Function to be called when registration is completed (**pmix_hdlr_reg_cbfunc_t** function
23 reference)

24 **IN** **regcbdata**
25 Data to be passed to the *regcbfunc* callback function (memory reference)

26 Returns **PMIX_SUCCESS** or a negative value indicating the error. In the event the function returns an error, the
27 *regcbfunc* will not be called.

Required Attributes

28 The following attributes are required for PMIx libraries that support IO forwarding:

29 **PMIX_IOF_CACHE_SIZE** "pmix.iof.csize" (**uint32_t**)

30 The requested size of the PMIx server cache in bytes for each specified channel. By default, the server
31 is allowed (but not required) to drop all bytes received beyond the max size.

32 **PMIX_IOF_DROP_OLDEST** "pmix.iof.old" (**bool**)

33 In an overflow situation, the PMIx server is to drop the oldest bytes to make room in the cache.

34 **PMIX_IOF_DROP_NEWEST** "pmix.iof.new" (**bool**)

35 In an overflow situation, the PMIx server is to drop any new bytes received until room becomes
36 available in the cache (default).

Optional Attributes

The following attributes are optional for PMIx libraries that support IO forwarding:

PMIX_IOF_BUFFERING_SIZE "pmix.iof.bsize" (uint32_t)

Requests that IO on the specified channel(s) be aggregated in the PMIx tool library until the specified number of bytes is collected to avoid being called every time a block of IO arrives. The PMIx tool library will execute the callback and reset the collection counter whenever the specified number of bytes becomes available. Any remaining buffered data will be *flushed* to the callback upon a call to deregister the respective channel.

PMIX_IOF_BUFFERING_TIME "pmix.iof.btime" (uint32_t)

Max time in seconds to buffer IO before delivering it. Used in conjunction with buffering size, this prevents IO from being held indefinitely while waiting for another payload to arrive.

PMIX_IOF_TAG_OUTPUT "pmix.iof.tag" (bool)

Requests that output be prefixed with the nspace,rank of the source and a string identifying the channel (`stdout`, `stderr`, etc.).

PMIX_IOF_TIMESTAMP_OUTPUT "pmix.iof.ts" (bool)

Requests that output be marked with the time at which the data was received by the tool - note that this will differ from the time at which the data was collected from the source.

PMIX_IOF_XML_OUTPUT "pmix.iof.xml" (bool)

Requests that output be formatted in XML.

Description

Register to receive output forwarded from a set of remote processes.

Advice to users

Providing a **NULL** function pointer for the *cbfunc* parameter will cause output for the indicated channels to be written to their corresponding `stdout/stderr` file descriptors. Use of **PMIX_RANK_WILDCARD** to specify all processes in a given namespace is supported but should be used carefully due to bandwidth and memory footprint considerations.

18.5.8 PMIx_IOF_deregister

Summary

Deregister from output forwarded from a set of remote processes.

Format

C

```
pmix_status_t
PMIx_IOF_deregister(size_t iofhdlr,
                    const pmix_info_t directives[], size_t ndirs,
                    pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

IN iofhdlr

Registration number returned from the `pmix_hdlr_reg_cbfunc_t` callback from the call to `PMIx_IOF_pull` (`size_t`)

IN directives

Array of `pmix_info_t` structures (array of handles)

IN ndirs

Number of elements in the *directives* array (integer)

IN cbfunc

Callback function to be called when deregistration has been completed. (function reference)

IN cbdata

Data to be passed to the *cbfunc* callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when `PMIX_SUCCESS` is returned.

Returns `PMIX_SUCCESS` or one of the following error codes when the condition described occurs:

`PMIX_OPERATION_SUCCEEDED`, indicating that the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Description

Deregister from output forwarded from a set of remote processes.

Advice to PMIx library implementers

Any currently buffered IO should be flushed upon receipt of a deregistration request. All received IO after receipt of the request shall be discarded.

18.5.9 PMIx_IOF_push

Summary

Push data collected locally (typically from `stdin` or a file) to `stdin` of the target recipients.

Format

C

```
pmix_status_t
PMIx_IOF_push(const pmix_proc_t targets[], size_t ntargets,
              pmix_byte_object_t *bo,
              const pmix_info_t directives[], size_t ndirs,
              pmix_op_cbfunc_t cbfunc, void *cbdata);
```

C

IN targets
Array of proc structures identifying desired target processes (array of handles)

IN ntargets
Number of elements in the *targets* array (integer)

IN bo
Pointer to [pmix_byte_object_t](#) containing the payload to be delivered (handle)

IN directives
Array of [pmix_info_t](#) structures (array of handles)

IN ndirs
Number of elements in the *directives* array (integer)

IN directives
Array of [pmix_info_t](#) structures (array of handles)

IN cbfunc
Callback function to be called when operation has been completed. ([pmix_op_cbfunc_t](#) function reference)

IN cbdata
Data to be passed to the *cbfunc* callback function (memory reference)

A successful return indicates that the request is being processed and the result will be returned in the provided *cbfunc*. Note that the library must not invoke the callback function prior to returning from the API. The callback function, *cbfunc*, is only called when **PMIX_SUCCESS** is returned.

Returns **PMIX_SUCCESS** or one of the following error codes when the condition described occurs:

PMIX_OPERATION_SUCCEEDED, indicating that the request was immediately processed successfully - the *cbfunc* will *not* be called.

If none of the above return codes are appropriate, then an implementation must return either a general PMIx error code or an implementation defined error code as described in Section 3.1.1.

Required Attributes

The following attributes are required for PMIx libraries that support IO forwarding:

PMIX_IOF_CACHE_SIZE "pmix.iof.csize" (uint32_t)

The requested size of the PMIx server cache in bytes for each specified channel. By default, the server is allowed (but not required) to drop all bytes received beyond the max size.

PMIX_IOF_DROP_OLDEST "pmix.iof.old" (bool)

In an overflow situation, the PMIx server is to drop the oldest bytes to make room in the cache.

PMIX_IOF_DROP_NEWEST "pmix.iof.new" (bool)

1 In an overflow situation, the PMIx server is to drop any new bytes received until room becomes
2 available in the cache (default).

Optional Attributes

3 The following attributes are optional for PMIx libraries that support IO forwarding:

4 **PMIX_IOF_BUFFERING_SIZE** "pmix.iof.bsize" (uint32_t)

5 Requests that IO on the specified channel(s) be aggregated in the PMIx tool library until the specified
6 number of bytes is collected to avoid being called every time a block of IO arrives. The PMIx tool
7 library will execute the callback and reset the collection counter whenever the specified number of
8 bytes becomes available. Any remaining buffered data will be *flushed* to the callback upon a call to
9 deregister the respective channel.

10 **PMIX_IOF_BUFFERING_TIME** "pmix.iof.btime" (uint32_t)

11 Max time in seconds to buffer IO before delivering it. Used in conjunction with buffering size, this
12 prevents IO from being held indefinitely while waiting for another payload to arrive.

13 **PMIX_IOF_PUSH_STDIN** "pmix.iof.stdin" (bool)

14 Requests that the PMIx library collect the **stdin** of the requester and forward it to the processes
15 specified in the **PMIX_IOF_push** call. All collected data is sent to the same targets until **stdin** is
16 closed, or a subsequent call to **PMIX_IOF_push** is made that includes the **PMIX_IOF_COMPLETE**
17 attribute indicating that forwarding of **stdin** is to be terminated.

Description

18 Called either to:

- 20 ● push data collected by the caller themselves (typically from **stdin** or a file) to **stdin** of the target
21 recipients;
- 22 ● request that the PMIx library automatically collect and push the **stdin** of the caller to the target recipients;
23 or
- 24 ● indicate that automatic collection and transmittal of **stdin** is to stop

Advice to users

25 Execution of the *cbfunc* callback function serves as notice that the PMIx library no longer requires the caller to
26 maintain the *bo* data object - it does *not* indicate delivery of the payload to the targets. Use of
27 **PMIX_RANK_WILDCARD** to specify all processes in a given namespace is supported but should be used
28 carefully due to bandwidth and memory footprint considerations.

CHAPTER 19

Storage Support Definitions *(Provisional)*

1 Distributed and parallel computing systems are increasingly embracing storage hierarchies to meet the diverse
2 data management needs of applications and other systems software in a cost-effective manner. These
3 hierarchies provide access to a number of distinct storage layers, with each potentially composed of different
4 storage hardware (e.g., HDD, SSD, tape, PMEM), deployed at different locations (e.g., on-node, on-switch,
5 on-site, WAN), and designed using different storage paradigms (e.g., file-based, object-based). Each of these
6 systems offers unique performance and usage characteristics that storage system users should carefully
7 consider to ensure the most efficient use of storage resources.

8 PMIx enables users to better understand storage hierarchies by defining attributes that formalize storage
9 system characteristics, state, and other parameters. These attributes can be queried by applications, I/O
10 libraries and middleware, and workflow systems to discover available storage resources and to inform on which
11 resources are most suitable for different I/O workload requirements.

19.1 Storage support constants *(Provisional)*

13 The `pmix_storage_medium_t` *(Provisional)* is a `uint64_t` type that defines a set of bit-mask flags for
14 specifying different types of storage mediums. These can be bitwise OR'd together to accommodate storage
15 systems that mix storage medium types.

16	<code>PMIX_STORAGE_MEDIUM_UNKNOWN</code> <i>(Provisional)</i>	<code>0x0000000000000001</code>	The storage medium
17			type is unknown.
18	<code>PMIX_STORAGE_MEDIUM_TAPE</code> <i>(Provisional)</i>	<code>0x0000000000000002</code>	The storage system uses
19			tape media.
20	<code>PMIX_STORAGE_MEDIUM_HDD</code> <i>(Provisional)</i>	<code>0x0000000000000004</code>	The storage system uses
21			HDDs with traditional SAS, SATA interfaces.
22	<code>PMIX_STORAGE_MEDIUM_SSD</code> <i>(Provisional)</i>	<code>0x0000000000000008</code>	The storage system uses
23			SSDs with traditional SAS, SATA interfaces.
24	<code>PMIX_STORAGE_MEDIUM_NVME</code> <i>(Provisional)</i>	<code>0x0000000000000010</code>	The storage system uses
25			SSDs with NVMe interface.
26	<code>PMIX_STORAGE_MEDIUM_PMEM</code> <i>(Provisional)</i>	<code>0x0000000000000020</code>	The storage system uses
27			persistent memory.
28	<code>PMIX_STORAGE_MEDIUM_RAM</code> <i>(Provisional)</i>	<code>0x0000000000000040</code>	The storage system is
29			volatile (e.g., tmpfs).

Advice to PMIx library implementers

PMIx implementations should maintain the same ordering for bit-mask values for `pmix_storage_medium_t` struct as provided in this standard, since these constants are ordered to provide semantic information that may be of use to PMIx users. Namely, `pmix_storage_medium_t` constants are ordered in terms of increasing medium bandwidth.

It is further recommended that implementations should try to allocate empty bits in the mask so that they can be extended to account for new constant definitions corresponding to new storage mediums.

The `pmix_storage_accessibility_t` (*Provisional*) is a `uint64_t` type that defines a set of bit-mask flags for specifying different levels of storage accessibility (i.e., from where a storage system may be accessed). These can be bitwise OR'd together to accommodate storage systems that are accessible in multiple ways.

PMIX_STORAGE_ACCESSIBILITY_NODE (*Provisional*) `0x0000000000000001` The storage system resources are accessible within the same node.

PMIX_STORAGE_ACCESSIBILITY_SESSION (*Provisional*) `0x0000000000000002` The storage system resources are accessible within the same session.

PMIX_STORAGE_ACCESSIBILITY_JOB (*Provisional*) `0x0000000000000004` The storage system resources are accessible within the same job.

PMIX_STORAGE_ACCESSIBILITY_RACK (*Provisional*) `0x0000000000000008` The storage system resources are accessible within the same rack.

PMIX_STORAGE_ACCESSIBILITY_CLUSTER (*Provisional*) `0x0000000000000010` The storage system resources are accessible within the same cluster.

PMIX_STORAGE_ACCESSIBILITY_REMOTE (*Provisional*) `0x0000000000000020` The storage system resources are remote.

The `pmix_storage_persistence_t` (*Provisional*) is a `uint64_t` type that defines a set of bit-mask flags for specifying different levels of persistence for a particular storage system.

PMIX_STORAGE_PERSISTENCE_TEMPORARY (*Provisional*) `0x0000000000000001` Data on the storage system is persisted only temporarily (i.e., it does not survive across sessions or node reboots).

PMIX_STORAGE_PERSISTENCE_NODE (*Provisional*) `0x0000000000000002` Data on the storage system is persisted on the node.

PMIX_STORAGE_PERSISTENCE_SESSION (*Provisional*) `0x0000000000000004` Data on the storage system is persisted for the duration of the session.

PMIX_STORAGE_PERSISTENCE_JOB (*Provisional*) `0x0000000000000008` Data on the storage system is persisted for the duration of the job.

PMIX_STORAGE_PERSISTENCE_SCRATCH (*Provisional*) `0x0000000000000010` Data on the storage system is persisted according to scratch storage policies (short-term storage, typically persisted for days to weeks).

PMIX_STORAGE_PERSISTENCE_PROJECT (*Provisional*) `0x0000000000000020` Data on the storage system is persisted according to project storage policies (long-term storage, typically persisted for the duration of a project).

PMIX_STORAGE_PERSISTENCE_ARCHIVE (*Provisional*) `0x0000000000000040` Data on the storage system is persisted according to archive storage policies (long-term storage, typically persisted indefinitely).

The `pmix_storage_access_type_t` (*Provisional*) is a `uint16_t` type that defines a set of bit-mask flags for specifying different storage system access types.

<code>PMIX_STORAGE_ACCESS_RD</code> (<i>Provisional</i>)	<code>0x0001</code>	Provide information on storage system read operations.
<code>PMIX_STORAGE_ACCESS_WR</code> (<i>Provisional</i>)	<code>0x0002</code>	Provide information on storage system write operations.
<code>PMIX_STORAGE_ACCESS_RDWR</code> (<i>Provisional</i>)	<code>0x0003</code>	Provide information on storage system read and write operations.

19.2 Storage support attributes (*Provisional*)

The following attributes may be returned in response to queries (e.g., `PMIx_Get` or `PMIx_Query_info`) made by processes or tools.

<code>PMIX_STORAGE_ID</code> " <code>pmix.strg.id</code> " (<code>char*</code>) (<i>Provisional</i>)	An identifier for the storage system (e.g., <code>lustre-fs1</code> , <code>daos-oss1</code> , <code>home-fs</code>)
<code>PMIX_STORAGE_PATH</code> " <code>pmix.strg.path</code> " (<code>char*</code>) (<i>Provisional</i>)	Mount point path for the storage system (valid only for file-based storage systems)
<code>PMIX_STORAGE_TYPE</code> " <code>pmix.strg.type</code> " (<code>char*</code>) (<i>Provisional</i>)	Type of storage system (i.e., <code>lustre</code> , <code>gpfs</code> , <code>daos</code> , <code>ext4</code>)
<code>PMIX_STORAGE_VERSION</code> " <code>pmix.strg.ver</code> " (<code>char*</code>) (<i>Provisional</i>)	Version string for the storage system
<code>PMIX_STORAGE_MEDIUM</code> " <code>pmix.strg.medium</code> " (<code>pmix_storage_medium_t</code>) (<i>Provisional</i>)	Types of storage mediums utilized by the storage system (e.g., SSDs, HDDs, tape)
<code>PMIX_STORAGE_ACCESSIBILITY</code> " <code>pmix.strg.access</code> " (<code>pmix_storage_accessibility_t</code>) (<i>Provisional</i>)	Accessibility level of the storage system (e.g., within same node, within same session)
<code>PMIX_STORAGE_PERSISTENCE</code> " <code>pmix.strg.persist</code> " (<code>pmix_storage_persistence_t</code>) (<i>Provisional</i>)	Persistence level of the storage system (e.g., scratch storage or archive storage)
<code>PMIX_QUERY_STORAGE_LIST</code> " <code>pmix.strg.list</code> " (<code>char*</code>) (<i>Provisional</i>)	Comma-delimited list of storage identifiers (i.e., <code>PMIX_STORAGE_ID</code> types) for available storage systems
<code>PMIX_STORAGE_CAPACITY_LIMIT</code> " <code>pmix.strg.caplim</code> " (<code>double</code>) (<i>Provisional</i>)	Overall limit on capacity (in bytes) for the storage system
<code>PMIX_STORAGE_CAPACITY_USED</code> " <code>pmix.strg.capuse</code> " (<code>double</code>) (<i>Provisional</i>)	Overall used capacity (in bytes) for the storage system
<code>PMIX_STORAGE_OBJECT_LIMIT</code> " <code>pmix.strg.objlim</code> " (<code>uint64_t</code>) (<i>Provisional</i>)	Overall limit on number of objects (e.g., inodes) for the storage system
<code>PMIX_STORAGE_OBJECTS_USED</code> " <code>pmix.strg.objuse</code> " (<code>uint64_t</code>) (<i>Provisional</i>)	Overall used number of objects (e.g., inodes) for the storage system
<code>PMIX_STORAGE_MINIMAL_XFER_SIZE</code> " <code>pmix.strg.minxfer</code> " (<code>double</code>) (<i>Provisional</i>)	Minimal transfer size (in bytes) for the storage system - this is the storage system's atomic unit of transfer (e.g., block size)
<code>PMIX_STORAGE_SUGGESTED_XFER_SIZE</code> " <code>pmix.strg.sxfer</code> " (<code>double</code>) (<i>Provisional</i>)	

1 Suggested transfer size (in bytes) for the storage system

2 **PMIX_STORAGE_BW_MAX** "pmix.strg.bwmax" (double) (Provisional)

3 Maximum bandwidth (in bytes/sec) for storage system - provided as the theoretical maximum or the
4 maximum observed bandwidth value

5 **PMIX_STORAGE_BW_CUR** "pmix.strg.bwcur" (double) (Provisional)

6 Observed bandwidth (in bytes/sec) for storage system - provided as a recently observed bandwidth
7 value, with the exact measurement interval depending on the storage system and/or PMIx library
8 implementation

9 **PMIX_STORAGE_IOPS_MAX** "pmix.strg.iopsmax" (double) (Provisional)

10 Maximum IOPS (in I/O operations per second) for storage system - provided as the theoretical
11 maximum or the maximum observed IOPS value

12 **PMIX_STORAGE_IOPS_CUR** "pmix.strg.iopscur" (double) (Provisional)

13 Observed IOPS (in I/O operations per second) for storage system - provided as a recently observed
14 IOPS value, with the exact measurement interval depending on the storage system and/or PMIx library
15 implementation

16 **PMIX_STORAGE_ACCESS_TYPE** "pmix.strg.atype" (pmix_storage_access_type_t)
17 (Provisional)

18 Qualifier describing the type of storage access to return information for (e.g., for qualifying

19 **PMIX_STORAGE_BW_CUR**, **PMIX_STORAGE_IOPS_CUR**, or

20 **PMIX_STORAGE_SUGGESTED_XFER_SIZE** attributes)

APPENDIX A

Python Bindings

1 While the PMIx Standard is defined in terms of C-based APIs, there is no intent to limit the use of PMIx to
2 that specific language. Support for other languages is captured in the Standard by describing their equivalent
3 syntax for the PMIx APIs and native forms for the PMIx datatypes. This Appendix specifically deals with
4 Python interfaces, beginning with a review of the PMIx datatypes. Support is restricted to Python 3 and above
5 - i.e., the Python bindings do not support Python 2.

6 Note: the PMIx APIs have been loosely collected into three Python classes based on their PMIx “class” (i.e.,
7 client, server, and tool). All processes have access to a basic set of the APIs, and therefore those have been
8 included in the “client” class. Servers can utilize any of those functions plus a set focused on operations not
9 commonly executed by an application process. Finally, tools can also act as servers but have their own
10 initialization function.

11 A.1 Design Considerations

12 Several issues arose during design of the Python bindings:

13 A.1.1 Error Codes vs Python Exceptions

14 The C programming language reports errors through the return of the corresponding integer status codes.
15 PMIx has defined a range of negative values for this purpose. However, Python has the option of raising
16 *exceptions* that effectively operate as interrupts that can be trapped if the program appropriately tests for them.
17 The PMIx Python bindings opted to follow the C-based standard and return PMIx status codes in lieu of
18 raising exceptions as this method was considered more consistent for those working in both domains.

19 A.1.2 Representation of Structured Data

20 PMIx utilizes a number of C-language structures to efficiently bundle related information. For example, the
21 PMIx process identifier is represented as a struct containing a character array for the namespace and a 32-bit
22 unsigned integer for the process rank. There are several options for translating such objects to Python – e.g.,
23 the PMIx process identifier could be represented as a two-element tuple (nspc, rank) or as a dictionary
24 ‘nspc’: name, ‘rank’: 0. Exploration found no discernible benefit to either representation, nor was any
25 clearly identifiable rationale developed that would lead a user to expect one versus the other for a given PMIx
26 data type. Consistency in the translation (i.e., exclusively using tuple or dictionary) appeared to be the most
27 important criterion. Hence, the decision was made to express all complex datatypes as Python dictionaries.

1 A.2 Datatype Definitions

2 PMIx defines a number of datatypes comprised of fixed-size character arrays, restricted range integers (e.g.,
3 `uint32_t`), and structures. Each datatype is represented by a named unsigned 16-bit integer (`uint16_t`)
4 constant. Users are advised to use the named PMIx constants for indicating datatypes instead of integer values
5 to ensure compatibility with future PMIx versions.

6 With only a few exceptions, the C-based PMIx datatypes defined in Chapter 3 on page 15 directly translate to
7 Python. However, Python lacks the size-specific value definitions of C (e.g., `uint8_t`) and thus some care
8 must be taken to protect against overflow/underflow situations when moving between the languages. Python
9 bindings that accept values including PMIx datatypes shall therefore have the datatype and associated value
10 checked for compatibility with their PMIx-defined equivalents, returning an error if:

- 11 • datatypes not defined by PMIx are encountered
- 12 • provided values fall outside the range of the C-equivalent definition - e.g., if a value identified as
13 `PMIX_UINT8` lies outside the `uint8_t` range

14 Note that explicit labeling of PMIx data type, even when Python itself doesn't care, is often required for the
15 Python bindings to know how to properly interpret and label the provided value when passing it to the PMIx
16 library.

17 Table A.1 lists the correspondence between data types in the two languages.

Table A.1.: C-to-Python Datatype Correspondence

C-Definition	PMIx Name	Python Definition	Notes
<code>bool</code>	<code>PMIX_BOOL</code>	boolean	
<code>byte</code>	<code>PMIX_BYTE</code>	A single element byte array (i.e., a byte array of length one)	
<code>char*</code>	<code>PMIX_STRING</code>	string	
<code>size_t</code>	<code>PMIX_SIZE</code>	integer	
<code>pid_t</code>	<code>PMIX_PID</code>	integer	value shall be limited to the <code>uint32_t</code> range
<code>int, int8_t, int16_t, int32_t, int64_t</code>	<code>PMIX_INT, PMIX_INT8, PMIX_INT16, PMIX_INT32, PMIX_INT64</code>	integer	value shall be limited to its corresponding range
<code>uint, uint8_t, uint16_t, uint32_t, uint64_t</code>	<code>PMIX_UINT, PMIX_UINT8, PMIX_UINT16, PMIX_UINT32, PMIX_UINT64</code>	integer	value shall be limited to its corresponding range
<code>float, double</code>	<code>PMIX_FLOAT, PMIX_DOUBLE</code>	float	value shall be limited to its corresponding range
<code>struct timeval</code>	<code>PMIX_TIMEVAL</code>	{'sec': sec, 'usec': microsec}	each field is an integer value
<code>time_t</code>	<code>PMIX_TIME</code>	integer	limited to positive values
<code>pmix_data_type_t</code>	<code>PMIX_DATA_TYPE</code>	integer	value shall be limited to the <code>uint16_t</code> range
<code>pmix_status_t</code>	<code>PMIX_STATUS</code>	integer	
<code>pmix_key_t</code>	N/A	string	The string's length shall be limited to one less than the size of the <code>pmix_key_t</code> array (to reserve space for the terminating NULL)
<code>pmix_nspace_t</code>	N/A	string	The string's length shall be limited to one less than the size of the <code>pmix_nspace_t</code> array (to reserve space for the terminating NULL)

Table A.1.: C-to-Python Datatype Correspondence

C-Definition	PMIx Name	Python Definition	Notes
<code>pmix_rank_t</code>	<code>PMIX_PROC_RANK</code>	integer	value shall be limited to the <code>uint32_t</code> range excepting the reserved values near <code>UINT32_MAX</code>
<code>pmix_proc_t</code>	<code>PMIX_PROC</code>	{'nspace': nspace, 'rank': rank}	<i>nspace</i> is a Python string and <i>rank</i> is an integer value. The <i>nspace</i> string's length shall be limited to one less than the size of the <code>pmix_nspace_t</code> array (to reserve space for the terminating <code>NULL</code>), and the <i>rank</i> value shall conform to the constraints associated with <code>pmix_rank_t</code>
<code>pmix_byte_object_t</code>	<code>PMIX_BYTE_OBJECT</code>	{'bytes': bytes, 'size': size}	<i>bytes</i> is a Python byte array and <i>size</i> is the integer number of bytes in that array.
<code>pmix_persistence_t</code>	<code>PMIX_PERSISTENCE</code>	integer	value shall be limited to the <code>uint8_t</code> range
<code>pmix_scope_t</code>	<code>PMIX_SCOPE</code>	integer	value shall be limited to the <code>uint8_t</code> range
<code>pmix_data_range_t</code>	<code>PMIX_RANGE</code>	integer	value shall be limited to the <code>uint8_t</code> range
<code>pmix_proc_state_t</code>	<code>PMIX_PROC_STATE</code>	integer	value shall be limited to the <code>uint8_t</code> range
<code>pmix_proc_info_t</code>	<code>PMIX_PROC_INFO</code>	{'proc': {'nspace': nspace, 'rank': rank}, 'hostname': hostname, 'executable': executable, 'pid': pid, 'exitcode': exitcode, 'state': state}	<i>proc</i> is a Python <code>proc</code> dictionary; <i>hostname</i> and <i>executable</i> are Python strings; and <i>pid</i> , <i>exitcode</i> , and <i>state</i> are Python integers

Table A.1.: C-to-Python Datatype Correspondence

C-Definition	PMIx Name	Python Definition	Notes
<code>pmix_data_array_t</code>	<code>PMIX_DATA_ARRAY</code>	{'type': type, 'array': array}	<i>type</i> is the PMIx type of object in the array and <i>array</i> is a Python <i>list</i> containing the individual array elements. Note that <i>array</i> can consist of <i>any</i> PMIx types, including (for example) a Python <code>info</code> object that itself contains an <code>array</code> value
<code>pmix_info_directives_t</code>	<code>PMIX_INFO_DIRECTIVES</code>	list	list of integer values (defined in Section 3.2.10)
<code>pmix_alloc_directive_t</code>	<code>PMIX_ALLOC_DIRECTIVE</code>	integer	value shall be limited to the <code>uint8_t</code> range
<code>pmix_iof_channel_t</code>	<code>PMIX_IOF_CHANNEL</code>	list	list of integer values (defined in Section 18.3.3)
<code>pmix_envvar_t</code>	<code>PMIX_ENVAR</code>	{'envar': envar, 'value': value, 'separator': separator}	<i>envar</i> and <i>value</i> are Python strings, and <i>separator</i> a single-character Python string
<code>pmix_value_t</code>	<code>PMIX_VALUE</code>	{'value': value, 'val_type': type}	<i>type</i> is the PMIx datatype of <i>value</i> , and <i>value</i> is the associated value expressed in the appropriate Python form for the specified datatype
<code>pmix_info_t</code>	<code>PMIX_INFO</code>	{'key': key, 'flags': flags, 'value': value, 'val_type': type}	<i>key</i> is a Python string <code>key</code> ; <i>flags</i> is an <code>info directives</code> value; <i>type</i> is the PMIx datatype of <i>value</i> , and <i>value</i> is the associated value expressed in the appropriate Python form for the specified datatype
<code>pmix_pdata_t</code>	<code>PMIX_PDATA</code>	{'proc': {'nspace': nspace, 'rank': rank}, 'key': key, 'value': value, 'val_type': type}	<i>proc</i> is a Python <code>proc</code> dictionary; <i>key</i> is a Python string <code>key</code> ; <i>type</i> is the PMIx datatype of <i>value</i> ; and <i>value</i> is the associated value expressed in the appropriate Python form for the specified datatype

Table A.1.: C-to-Python Datatype Correspondence

C-Definition	PMIx Name	Python Definition	Notes
<code>pmix_app_t</code>	PMIX_APP	{'cmd': cmd, 'argv': [argv], 'env': [env], 'maxprocs': maxprocs, 'info': [info]}	<i>cmd</i> is a Python string; <i>argv</i> and <i>env</i> are Python <i>lists</i> containing Python strings; <i>maxprocs</i> is an integer; and <i>info</i> is a Python <i>list</i> of info values
<code>pmix_query_t</code>	PMIX_QUERY	{'keys': [keys], 'qualifiers': [info]}	<i>keys</i> is a Python <i>list</i> of Python strings, and <i>qualifiers</i> is a Python <i>list</i> of info values
<code>pmix_regattr_t</code>	PMIX_REGATTR	{'name': name, 'key': key, 'type': type, 'info': [info], 'description': [desc]}	<i>name</i> and <i>string</i> are Python strings; <i>type</i> is the PMIx datatype for the attribute's value; <i>info</i> is a Python <i>list</i> of info values; and <i>description</i> is a list of Python strings describing the attribute
<code>pmix_job_state_t</code>	PMIX_JOB_STATE	integer	value shall be limited to the uint8_t range
<code>pmix_link_state_t</code>	PMIX_LINK_STATE	integer	value shall be limited to the uint8_t range
<code>pmix_cpuset_t</code>	PMIX_PROC_CPuset	{'source': source, 'cpus': bitmap}	<i>source</i> is a string name of the library that created the cpuset; and <i>cpus</i> is a list of string ranges identifying the PUs to which the process is bound (e.g., [1, 3-5, 7])
<code>pmix_locality_t</code>	PMIX_LOCTYPE	list	list of integer values (defined in Section 12.4.2.3) describing the relative locality of the specified local process
<code>pmix_fabric_t</code>	N/A	{'name': name, 'index': idx, 'info': [info]}	<i>name</i> is the string name assigned to the fabric; <i>index</i> is the integer ID assigned to the fabric; <i>info</i> is a list of info describing the fabric
<code>pmix_endpoint_t</code>	PMIX_ENDPOINT	{'uuid': uuid, 'osname': osname, 'endpt': endpt}	<i>uuid</i> is the string system-unique identifier assigned to the device; <i>osname</i> is the operating system name assigned to the device; <i>endpt</i> is a byteobject containing the endpoint information

Table A.1.: C-to-Python Datatype Correspondence

C-Definition	PMIx Name	Python Definition	Notes
<code>pmix_device_distance_t</code>	<code>PMIX_DEVICE_DIST</code>	{'uuid': uuid, 'osname': osname, 'mindist': mindist, 'maxdist': maxdist}	<i>uuid</i> is the string system-unique identifier assigned to the device; <i>osname</i> is the operating system name assigned to the device; and <i>mindist</i> and <i>maxdist</i> are Python integers
<code>pmix_coord_t</code>	<code>PMIX_COORD</code>	{'view': view, 'coord': [coords]}	<i>view</i> is the <code>pmix_coord_view_t</code> of the coordinate; and <i>coord</i> is a list of integer coordinates, one for each dimension of the fabric
<code>pmix_geometry_t</code>	<code>PMIX_GEOMETRY</code>	{'fabric': idx, 'uuid': uuid, 'osname': osname, 'coordinates': [coords]}	<i>fabric</i> is the Python integer index of the fabric; <i>uuid</i> is the string system-unique identifier assigned to the device; <i>osname</i> is the operating system name assigned to the device; and <i>coordinates</i> is a list of <code>coord</code> containing the coordinates for the device across all views
<code>pmix_device_type_t</code>	<code>PMIX_DEVTYPE</code>	list	list of integer values (defined in Section 12.4.8)
<code>pmix_bind_envelope_t</code>	N/A	integer	one of the values defined in Section 12.4.4.1

1 A.2.1 Example

2 Converting a C-based program to its Python equivalent requires translation of the relevant datatypes as well as
3 use of the appropriate API form. An example small program may help illustrate the changes. Consider the
4 following C-based program snippet:

C

```
5 #include <pmix.h>
6 ...
7
8 pmix_info_t info[2];
9
10 PMIx_Info_load(&info[0], PMIX_PROGRAMMING_MODEL, "TEST", PMIX_STRING)
11 PMIx_Info_load(&info[1], PMIX_MODEL_LIBRARY_NAME, "PMIX", PMIX_STRING)
12
13 rc = PMIx_Init(&myproc, info, 2);
14
15 PMIX_INFO_DESTRUCT(&info[0]); // free the copied string
16 PMIX_INFO_DESTRUCT(&info[1]); // free the copied string
```

C

17 Moving to the Python version requires that the `pmix_info_t` be translated to the Python `info` equivalent,
18 and that the returned information be captured in the return parameters as opposed to a pointer parameter in the
19 function call, as shown below:

Python

```
20 import pmix
21 ...
22
23 myclient = PMIxClient()
24 info = [{'key':PMIX_PROGRAMMING_MODEL,
25         'value':'TEST', 'val_type':PMIX_STRING},
26        {'key':PMIX_MODEL_LIBRARY_NAME,
27         'value':'PMIX', 'val_type':PMIX_STRING}]
28 (rc,myproc) = myclient.init(info)
```

Python

29 Note the use of the `PMIX_STRING` identifier to ensure the Python bindings interpret the provided string value
30 as a PMIx "string" and not an array of bytes.

31 A.3 Callback Function Definitions

32 A.3.1 IOF Delivery Function

33 Summary

34 Callback function for delivering forwarded IO to a process

Format

Python

```
def iofcbfunc(iofhdlr:integer, channel:bitarray,  
             source:dict, payload:dict, info:list)
```

Python

IN `iofhdlr`

Registration number of the handler being invoked (integer)

IN `channel`

Python `channel` 16-bit bitarray identifying the channel the data arrived on (bitarray)

IN `source`

Python `proc` identifying the namespace/rank of the process that generated the data (dict)

IN `payload`

Python `byteobject` containing the data (dict)

IN `info`

List of Python `info` provided by the source containing metadata about the payload. This could include `PMIX_IOF_COMPLETE` (list)

Returns: nothing

See `pmix_iof_cbfunc_t` for details

A.3.2 Event Handler

Summary

Callback function for event handlers

Format

Python

```
def evhandler(evhdlr:integer, status:integer,  
             source:dict, info:list, results:list)
```

Python

IN `iofhdlr`

Registration number of the handler being invoked (integer)

IN `status`

Status associated with the operation (integer)

IN `source`

Python `proc` identifying the namespace/rank of the process that generated the event (dict)

IN `info`

List of Python `info` provided by the source containing metadata about the event (list)

IN `results`

List of Python `info` containing the aggregated results of all prior evhandlers (list)

Returns:

- `rc` - Status returned by the event handler's operation (integer)
- `results` - List of Python `info` containing results from this event handler's operation on the event (list)

See `pmix_notification_fn_t` for details

1 A.3.3 Server Module Functions

2 The following definitions represent functions that may be provided to the PMIx server library at time of
3 initialization for servicing of client requests. Module functions that are not provided default to returning "not
4 supported" to the caller.

5 A.3.3.1 Client Connected

6 Summary

7 Notify the host server that a client connected to this server.

8 *PMIx v4.0* Format

Python

9 `def clientconnected2(proc:dict is not None, info:list)`

Python

10 **IN** `proc`

11 Python `proc` identifying the namespace/rank of the process that connected (dict)

12 **IN** `info`

13 list of Python `info` containing information about the process (list)

14 Returns:

- 15 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the connection should be rejected (integer)

16 See [pmix_server_client_connected2_fn_t](#) for details

17 A.3.3.2 Client Finalized

18 Summary

19 Notify the host environment that a client called `PMIx_Finalize`.

20 *PMIx v4.0* Format

Python

21 `def clientfinalized(proc:dict is not None):`

Python

22 **IN** `proc`

23 Python `proc` identifying the namespace/rank of the process that finalized (dict)

24 Returns: nothing

25 See [pmix_server_client_finalized_fn_t](#) for details

26 A.3.3.3 Client Aborted

27 Summary

28 Notify the host environment that a local client called `PMIx_Abort`.

Format

Python

```
def clientaborted(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'caller': Python **proc** identifying the namespace/rank of the process calling abort (dict)
- 'status': PMIx status to be returned on exit (integer)
- 'msg': Optional string message to be printed (string)
- 'targets': Optional list of Python **proc** identifying the namespace/rank of the processes to be aborted (list)

Returns:

- *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)

See [pmix_server_abort_fn_t](#) for details

A.3.3.4 Fence

Summary

At least one client called either **PMIx_Fence** or **PMIx_Fence_nb**

Format

Python

```
def fence(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'procs': List of Python **proc** identifying the namespace/rank of the participating processes (list)
- 'directives': Optional list of Python **info** containing directives controlling the operation (list)
- 'data': Optional Python bytearray of data to be circulated during fence operation (bytearray)

Returns:

- *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- *data* - Python bytearray containing the aggregated data from all participants (bytearray)


See [pmix_server_fence_nb_fn_t](#) for details

A.3.3.5 Direct Modex

Summary

Used by the PMIx server to request its local host contact the PMIx server on the remote node that hosts the specified proc to obtain and return a direct modex blob for that proc.

1 **Format** Python 

2 `def dmodex(args:dict is not None)`
3 Python 

4 **IN** `args`
5 Python dictionary containing:
6

- 'proc': Python **proc** of process whose data is being requested (dict)
- 'directives': Optional list of Python **info** containing directives controlling the operation (list)

7 Returns:
8


- `rc` - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- `data` - Python bytearray containing the data for the specified process (bytearray)

9 See [pmix_server_dmodex_req_fn_t](#) for details

11 A.3.3.6 Publish

12 **Summary**
13 Publish data per the PMIx API specification.

14 *PMIx v4.0* **Format** Python 

15 `def publish(args:dict is not None)`
16 Python 

17 **IN** `args`
18 Python dictionary containing:
19

- 'proc': Python **proc** dictionary of process publishing the data (dict)
- 'directives': List of Python **info** containing data and directives (list)

20 Returns:
21

- `rc` - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)

22 See [pmix_server_publish_fn_t](#) for details

23 A.3.3.7 Lookup

24 **Summary**
25 Lookup published data.

1 **Format** Python

2 `def lookup(args:dict is not None)` Python

- 3 **IN args**
 4 Python dictionary containing:
- 5 • 'proc': Python **proc** of process seeking the data (dict)
 - 6 • 'keys': List of Python strings (list)
 - 7 • 'directives': Optional list of Python **info** containing directives (list)

- 8 Returns:
- 9 • *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
 - 10 • *pdata* - List of **pdata** containing the returned results (list)
- 11 See [pmix_server_lookup_fn_t](#) for details

12 A.3.3.8 Unpublish

13 **Summary**
14 Delete data from the data store.

15 *PMIx v4.0* **Format** Python

16 `def unpublish(args:dict is not None)` Python

- 17 **IN args**
 18 Python dictionary containing:
- 19 • 'proc': Python **proc** of process unpublishing data (dict)
 - 20 • 'keys': List of Python strings (list)
 - 21 • 'directives': Optional list of Python **info** containing directives (list)

- 22 Returns:
- 23 • *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- 24 See [pmix_server_unpublish_fn_t](#) for details

25 A.3.3.9 Spawn

26 **Summary**
27 Spawn a set of applications/processes as per the **PMIx_Spawn** API.

1 **Format** Python

2 `def spawn(args:dict is not None)` Python

- 3 **IN args**
 4 Python dictionary containing:
- 5 • 'proc': Python **proc** of process making the request (dict)
 - 6 • 'jobinfo': Optional list of Python **info** job-level directives and information (list)
 - 7 • 'apps': List of Python **app** describing applications to be spawned (list)

8 Returns:

- 9 • *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- 10 • *nspace* - Python string containing namespace of the spawned job (str)

11 See [pmix_server_spawn_fn_t](#) for details

12 A.3.3.10 Connect

13 **Summary**
 14 Record the specified processes as *connected*.

15 *PMIx v4.0* **Format** Python

16 `def connect(args:dict is not None)` Python

- 17 **IN args**
 18 Python dictionary containing:
- 19 • 'procs': List of Python **proc** identifying the namespace/rank of the participating processes (list)
 - 20 • 'directives': Optional list of Python **info** containing directives controlling the operation (list)

21 Returns:

- 22 • *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)

23 See [pmix_server_connect_fn_t](#) for details

24 A.3.3.11 Disconnect

25 **Summary**
 26 Disconnect a previously connected set of processes.

```

1 Format Python
2 def disconnect (args:dict is not None) Python
3 IN args
4 Python dictionary containing:
5     • 'procs': List of Python proc identifying the namespace/rank of the participating processes (list)
6     • 'directives': Optional list of Python info containing directives controlling the operation (list)
7 Returns:
8     • rc - PMIX_SUCCESS or a PMIx error code indicating the operation failed (integer)
9 See pmix_server_disconnect_fn_t for details

```

A.3.3.12 Register Events

Summary
 Register to receive notifications for the specified events.

```

13 PMIx v4.0 Format Python
14 def register_events (args:dict is not None) Python
15 IN args
16 Python dictionary containing:
17     • 'codes': List of Python integers (list)
18     • 'directives': Optional list of Python info containing directives controlling the operation (list)
19 Returns:
20     • rc - PMIX_SUCCESS or a PMIx error code indicating the operation failed (integer)
21 See pmix_server_register_events_fn_t for details

```

A.3.3.13 Deregister Events

Summary
 Deregister to receive notifications for the specified events.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Format

Python

```
def deregister_events(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'codes': List of Python integers (list)

Returns:

- rc - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)

See [pmix_server_deregister_events_fn_t](#) for details

A.3.3.14 Notify Event

Summary

Notify the specified range of processes of an event.

PMIx v4.0

Format

Python

```
def notify_event(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'code': Python integer **pmix_status_t** (integer)
- 'source': Python **proc** of process that generated the event (dict)
- 'range': Python **range** in which the event is to be reported (integer)
- 'directives': Optional list of Python **info** directives (list)

Returns:

- rc - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)

See [pmix_server_notify_event_fn_t](#) for details

A.3.3.15 Query

Summary

Query information from the resource manager.

1 **Format** Python

2 `def query(args:dict is not None)` Python

3 **IN** `args`
4 Python dictionary containing:
5 • 'source': Python `proc` of requesting process (dict)
6 • 'queries': List of Python `query` directives (list)

7 Returns:
8 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the operation failed (integer)
9 • `info` - List of Python `info` containing the returned results (list)
10 See `pmix_server_query_fn_t` for details

11 A.3.3.16 Tool Connected

12 **Summary**
13 Register that a tool has connected to the server.

14 *PMIx v4.0* **Format** Python

15 `def tool_connected(args:dict is not None)` Python

16 **IN** `args`
17 Python dictionary containing:
18 • 'directives': Optional list of Python `info` info on the connecting tool (list)

19 Returns:
20 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the operation failed (integer)
21 • `proc` - Python `proc` containing the assigned namespace:rank for the tool (dict)
22 See `pmix_server_tool_connection_fn_t` for details

23 A.3.3.17 Log

24 **Summary**
25 Log data on behalf of a client.

1 **Format** Python

2 `def log(args:dict is not None)` Python

- 3 **IN** `args`
 4 Python dictionary containing:
- 5 • 'source': Python `proc` of requesting process (dict)
 - 6 • 'data': Optional list of Python `info` containing data to be logged (list)
 - 7 • 'directives': Optional list of Python `info` containing directives (list)

8 Returns:

- 9 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the operation failed (integer)

10 See `pmix_server_log_fn_t` for details.

11 A.3.3.18 Allocate Resources

12 Summary

13 Request allocation operations on behalf of a client.

14 *PMIx v4.0* **Format** Python

15 `def allocate(args:dict is not None)` Python

- 16 **IN** `args`
 17 Python dictionary containing:
- 18 • 'source': Python `proc` of requesting process (dict)
 - 19 • 'action': Python `allocdir` specifying requested action (integer)
 - 20 • 'directives': Optional list of Python `info` containing directives (list)

21 Returns:

- 22 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the operation failed (integer)
- 23 • `refarginfo` - List of Python `info` containing results of requested operation (list)

24 See `pmix_server_alloc_fn_t` for details.

25 A.3.3.19 Job Control

26 Summary

27 Execute a job control action on behalf of a client.

1 **Format** Python

```
2 def job_control(args:dict is not None)
Python
```

3 **IN args**
4 Python dictionary containing:
5 • 'source': Python **proc** of requesting process (dict)
6 • 'targets': List of Python **proc** specifying target processes (list)
7 • 'directives': Optional list of Python **info** containing directives (list)
8 Returns:
9 • *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
10 See [pmix_server_job_control_fn_t](#) for details.

11 **A.3.3.20 Monitor**

12 **Summary**
13 Request that a client be monitored for activity.

14 *PMIx v4.0* **Format** Python

```
15 def monitor(args:dict is not None)
Python
```

16 **IN args**
17 Python dictionary containing:
18 • 'source': Python **proc** of requesting process (dict)
19 • 'monitor': Python **info** attribute indicating the type of monitor being requested (dict)
20 • 'error': Status code to be used when generating an event notification (integer) alerting that the
21 monitor has been triggered.
22 • 'directives': Optional list of Python **info** containing directives (list)
23 Returns:
24 • *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
25 See [pmix_server_monitor_fn_t](#) for details.

26 **A.3.3.21 Get Credential**

27 **Summary**
28 Request a credential from the host environment.

Format

Python

```
def get_credential(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'source': Python **proc** of requesting process (dict)
- 'directives': Optional list of Python **info** containing directives (list)

Returns:

- *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- *cred* - Python **byteobject** containing returned credential (dict)
- *info* - List of Python **info** containing any additional info about the credential (list)

See [pmix_server_get_cred_fn_t](#) for details.

A.3.3.22 Validate Credential

Summary

Request validation of a credential

Format

Python

```
def validate_credential(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'source': Python **proc** of requesting process (dict)
- 'credential': Python **byteobject** containing credential (dict)
- 'directives': Optional list of Python **info** containing directives (list)

Returns:

- *rc* - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- *info* - List of Python **info** containing any additional info from the credential (list)

See [pmix_server_validate_cred_fn_t](#) for details.

A.3.3.23 IO Forward

Summary

Request the specified IO channels be forwarded from the given array of processes.

PMIx v4.0

1 **Format** Python

2 `def iof_pull(args:dict is not None)` Python

- 3 **IN** `args`
4 Python dictionary containing:
- 5 • 'sources': List of Python `proc` of processes whose IO is being requested (list)
 - 6 • 'channels': Bitmask of Python `channel` identifying IO channels to be forwarded (integer)
 - 7 • 'directives': Optional list of Python `info` containing directives (list)

8 Returns:
9 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the operation failed (integer)
10 See `pmix_server_iof_fn_t` for details.

11 A.3.3.24 IO Push

12 **Summary**
13 Pass standard input data to the host environment for transmission to specified recipients.

14 *PMIx v4.0* **Format** Python

15 `def iof_push(args:dict is not None)` Python

- 16 **IN** `args`
17 Python dictionary containing:
- 18 • 'source': Python `proc` of process whose input is being forwarded (dict)
 - 19 • 'payload': Python `byteobject` containing input bytes (dict)
 - 20 • 'targets': List of `proc` of processes that are to receive the payload (list)
 - 21 • 'directives': Optional list of Python `info` containing directives (list)

22 Returns:
23 • `rc` - `PMIX_SUCCESS` or a PMIx error code indicating the operation failed (integer)
24 See `pmix_server_stdin_fn_t` for details.

25 A.3.3.25 Group Operations

26 **Summary**
27 Request group operations (construct, destruct, etc.) on behalf of a set of processes.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

Format

Python

```
def group(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'op': Operation host is to perform on the specified group (integer)
- 'group': String identifier of target group (str)
- 'procs': List of Python **proc** of participating processes (dict)
- 'directives': Optional list of Python **info** containing directives (list)

Returns:

- rc - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- refarginfo - List of Python **info** containing results of requested operation (list)

See [pmix_server_grp_fn_t](#) for details.

A.3.3.26 Fabric Operations

Summary

Request fabric-related operations (e.g., information on a fabric) on behalf of a tool or other process.

Format

Python

```
def fabric(args:dict is not None)
```

Python

IN args

Python dictionary containing:

- 'source': Python **proc** of requesting process (dict)
- 'index': Identifier of the fabric being operated upon (integer)
- 'op': Operation host is to perform on the specified fabric (integer)
- 'directives': Optional list of Python **info** containing directives (list)

Returns:

- rc - **PMIX_SUCCESS** or a PMIx error code indicating the operation failed (integer)
- refarginfo - List of Python **info** containing results of requested operation (list)

See [pmix_server_fabric_fn_t](#) for details.

1 A.4 PMIxClient

2 The client Python class is by far the richest in terms of APIs as it houses all the APIs that an application might
3 utilize. Due to the datatype translation requirements of the C-Python interface, only the blocking form of each
4 API is supported – providing a Python callback function directly to the C interface underlying the bindings
5 was not a supportable option.

6 A.4.1 Client.init

7 Summary

8 Initialize the PMIx client library after obtaining a new PMIxClient object.

9 Format

PMIx v4.0

Python

```
10 rc, proc = myclient.init(info:list)
```

Python

11 IN info

12 List of Python [info](#) dictionaries (list)

13 Returns:

- 14 • *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- 15 • *proc* - a Python [proc](#) dictionary (dict)

16 See [PMIx_Init](#) for description of all relevant attributes and behaviors.

17 A.4.2 Client.initialized

18 Format

PMIx v4.0

Python

```
19 rc = myclient.initialized()
```

Python

20 Returns:

- 21 • *rc* - a value of **1** (true) will be returned if the PMIx library has been initialized, and **0** (false) otherwise
22 (integer)

23 See [PMIx_Initialized](#) for description of all relevant attributes and behaviors.

24 A.4.3 Client.get_version

25 Format

PMIx v4.0

Python

```
26 vers = myclient.get_version()
```

Python

27 Returns:

- 28 • *vers* - Python string containing the version of the PMIx library (e.g., "3.1.4") (integer)

29 See [PMIx_Get_version](#) for description of all relevant attributes and behaviors.

1 A.4.4 Client.finalize

2 Summary

3 Finalize the PMIx client library.

4 *PMIx v4.0* Format

Python

5 `rc = myclient.finalize(info:list)`

Python

6 **IN** `info`

7 List of Python `info` dictionaries (list)

8 Returns:

- 9 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

10 See `PMIx_Finalize` for description of all relevant attributes and behaviors.

11 A.4.5 Client.abort

12 Summary

13 Request that the provided list of processes be aborted.

14 *PMIx v4.0* Format

Python

15 `rc = myclient.abort(status:integer, msg:str, targets:list)`

Python

16 **IN** `status`

17 PMIx status to be returned on exit (integer)

18 **IN** `msg`

19 String message to be printed (string)

20 **IN** `targets`

21 List of Python `proc` dictionaries (list)

22 Returns:

- 23 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

24 See `PMIx_Abort` for description of all relevant attributes and behaviors.

25 A.4.6 Client.store_internal

26 Summary

27 Store some data locally for retrieval by other areas of the process

Format

Python

```
rc = myclient.store_internal(proc:dict, key:str, value:dict)
```

Python

IN **proc**

Python **proc** dictionary of the process being referenced (dict)

IN **key**

String key of the data (string)

IN **value**

Python **value** dictionary (dict)

Returns:

- *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

See **PMIx_Store_internal** for details.

A.4.7 Client.put

Summary

Push a key/value pair into the client's namespace.

Format

Python

```
rc = myclient.put(scope:integer, key:str, value:dict)
```

Python

IN **scope**

Scope of the data being posted (integer)

IN **key**

String key of the data (string)

IN **value**

Python **value** dictionary (dict)

Returns:

- *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

See **PMIx_Put** for description of all relevant attributes and behaviors.

A.4.8 Client.commit

Summary

Push all previously **PMIxClient.put** values to the local PMIx server.

1 **Format** Python

```
2 rc = myclient.commit()
```

Python

3 Returns:
4 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
5 See **PMIx_Commit** for description of all relevant attributes and behaviors.

6 A.4.9 Client.fence

7 **Summary**
8 Execute a blocking barrier across the processes identified in the specified list.

9 *PMIx v4.0* **Format** Python

```
10 rc = myclient.fence(peers:list, directives:list)
```

Python

11 **IN** **peers**
12 List of Python **proc** dictionaries (list)
13 **IN** **directives**
14 List of Python **info** dictionaries (list)

15 Returns:
16 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
17 See **PMIx_Fence** for description of all relevant attributes and behaviors.

18 A.4.10 Client.get

19 **Summary**
20 Retrieve a key/value pair.

21 *PMIx v4.0* **Format**

Python

```
1 rc, val = myclient.get(proc=dict, key=str, directives=list)
```

Python

2 **IN** `proc`
3 Python `proc` whose data is being requested (dict)
4 **IN** `key`
5 Python string key of the data to be returned (str)
6 **IN** `directives`
7 List of Python `info` dictionaries (list)

8 Returns:

- 9 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)
- 10 • `val` - Python `value` containing the returned data (dict)

11 See `PMIx_Get` for description of all relevant attributes and behaviors.

A.4.11 Client.publish

Summary

14 Publish data for later access via `PMIx_Lookup`.

Format

15 *PMIx v4.0*

Python

```
16 rc = myclient.publish(directives=list)
```

Python

17 **IN** `directives`
18 List of Python `info` dictionaries containing data to be published and directives (list)

19 Returns:

- 20 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

21 See `PMIx_Publish` for description of all relevant attributes and behaviors.

A.4.12 Client.lookup

Summary

24 Lookup information published by this or another process with `PMIx_Publish`.

Format

Python

```
rc, info = myclient.lookup(pdata:list, directives:list)
```

Python

IN `pdata`

List of Python `pdata` dictionaries identifying data to be retrieved (list)

IN `directives`

List of Python `info` dictionaries (list)

Returns:

- `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)
- `info` - Python list of `info` containing the returned data (list)

See [PMIx_Lookup](#) for description of all relevant attributes and behaviors.

A.4.13 Client.unpublish

Summary

Delete data published by this process with [PMIx_Publish](#).

Format

Python

```
rc = myclient.unpublish(keys:list, directives:list)
```

Python

IN `keys`

List of Python string keys identifying data to be deleted (list)

IN `directives`

List of Python `info` dictionaries (list)

Returns:

- `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

See [PMIx_Unpublish](#) for description of all relevant attributes and behaviors.

A.4.14 Client.spawn

Summary

Spawn a new job.

1 **Format** Python

2 `rc, nspace = myclient.spawn(jobinfo:list, apps:list)`

3 **IN** `jobinfo`

4 List of Python `info` dictionaries (list)

5 **IN** `apps`

6 List of Python `app` dictionaries (list)

7 Returns:

8 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

9 • `nspace` - Python `nspace` of the new job (dict)

10 See `PMIx_Spawn` for description of all relevant attributes and behaviors.

11 A.4.15 Client.connect

12 **Summary**

13 Connect namespaces.

14 *PMIx v4.0* **Format** Python

15 `rc = myclient.connect(peers:list, directives:list)`

16 **IN** `peers`

17 List of Python `proc` dictionaries (list)

18 **IN** `directives`

19 List of Python `info` dictionaries (list)

20 Returns:

21 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

22 See `PMIx_Connect` for description of all relevant attributes and behaviors.

23 A.4.16 Client.disconnect

24 **Summary**

25 Disconnect namespaces.

1 **Format** Python

2 `rc = myclient.disconnect(peers:list, directives:list)` Python

3 **IN peers**
 4 List of Python **proc** dictionaries (list)

5 **IN directives**
 6 List of Python **info** dictionaries (list)

7 Returns:

8 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

9 See **PMIx_Disconnect** for description of all relevant attributes and behaviors.

10 A.4.17 Client.resolve_peers

11 **Summary**

12 Return list of processes within the specified **nspace** on the given node.

13 *PMIx v4.0* **Format** Python

14 `rc,procs = myclient.resolve_peers(node:str, nspace:str)` Python

15 **IN node**
 16 Name of node whose processes are being requested (str)

17 **IN nspace**
 18 Python **nnamespace** whose processes are to be returned (str)

19 Returns:

20 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

21 • `procs` - List of Python **proc** dictionaries (list)

22 See **PMIx_Resolve_peers** for description of all relevant attributes and behaviors.

23 A.4.18 Client.resolve_nodes

24 **Summary**

25 Return list of nodes hosting processes within the specified **nnamespace**.

1 **Format** Python

2 `rc, nodes = myclient.resolve_nodes(namespace: str)` Python

3 **IN** `namespace`

4 Python `namespace` (str)

5 Returns:

6 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

7 • `nodes` - List of Python string node names (list)

8 See **PMIx_Resolve_nodes** for description of all relevant attributes and behaviors.

9 **A.4.19 Client.query**

10 **Summary**

11 Query information about the system in general.

12 *PMIx v4.0* **Format** Python

13 `rc, info = myclient.query(queries: list)` Python

14 **IN** `queries`

15 List of Python `query` dictionaries (list)

16 Returns:

17 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

18 • `info` - List of Python `info` containing results of the query (list)

19 See **PMIx_Query_info** for description of all relevant attributes and behaviors.

20 **A.4.20 Client.log**

21 **Summary**

22 Log data to a central data service/store.

1 **Format** Python

2 `rc = myclient.log(data:list, directives:list)`

3 **IN data**

4 List of Python [info](#) (list)

5 **IN directives**

6 Optional list of Python [info](#) (list)

7 Returns:

8 • *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

9 See [PMIx_Log](#) for description of all relevant attributes and behaviors.

10 A.4.21 Client.allocation_request

11 **Summary**

12 Request an allocation operation from the host resource manager.

13 *PMIx v4.0* **Format** Python

14 `rc, info = myclient.allocation_request(request:integer, directives:list)`

15 **IN request**

16 Python [allocdir](#) specifying requested operation (integer)

17 **IN directives**

18 List of Python [info](#) describing request (list)

19 Returns:

20 • *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

21 • *info* - List of Python [info](#) containing results of the request (list)

22 See [PMIx_Allocation_request](#) for description of all relevant attributes and behaviors.

23 A.4.22 Client.job_ctrl

24 **Summary**

25 Request a job control action.

Format

Python

```
rc, info = myclient.job_ctrl(targets:list, directives:list)
```

Python

IN targets

List of Python [proc](#) specifying targets of requested operation (integer)

IN directives

List of Python [info](#) describing operation to be performed (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- *info* - List of Python [info](#) containing results of the request (list)

See [PMIx_Job_control](#) for description of all relevant attributes and behaviors.

A.4.23 Client.monitor

Summary

Request that something be monitored.

Format

Python

```
rc, info = myclient.monitor(monitor:dict, error_code:integer, directives:list)
```

Python

IN monitor

Python [info](#) specifying specifying the type of monitor being requested (dict)

IN error_code

Status code to be used when generating an event notification alerting that the monitor has been triggered (integer)

IN directives

List of Python [info](#) describing request (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- *info* - List of Python [info](#) containing results of the request (list)

See [PMIx_Process_monitor](#) for description of all relevant attributes and behaviors.

A.4.24 Client.get_credential

Summary

Request a credential from the PMIx server/SMS.

1 **Format** Python

2 `rc, cred = myclient.get_credential(directives:list)` Python

3 **IN directives**

4 Optional list of Python **info** describing request (list)

5 Returns:

6 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

7 • *cred* - Python **byteobject** containing returned credential (dict)

8 See **PMIx_Get_credential** for description of all relevant attributes and behaviors.

9 **A.4.25 Client.validate_credential**

10 **Summary**

11 Request validation of a credential by the PMIx server/SMS.

12 *PMIx v4.0* **Format** Python

13 `rc, info = myclient.validate_credential(cred:dict, directives:list)` Python

14 **IN cred**

15 Python **byteobject** containing credential (dict)

16 **IN directives**

17 Optional list of Python **info** describing request (list)

18 Returns:

19 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

20 • *info* - List of Python **info** containing additional results of the request (list)

21 See **PMIx_Validate_credential** for description of all relevant attributes and behaviors.

22 **A.4.26 Client.group_construct**

23 **Summary**

24 Construct a new group composed of the specified processes and identified with the provided group identifier.

Format

Python

```
rc, info = myclient.construct_group(grp:string,  
                                   members:list, directives:list)
```

Python

IN grp

Python string identifier for the group (str)

IN members

List of Python [proc](#) dictionaries identifying group members (list)

IN directives

Optional list of Python [info](#) describing request (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- *info* - List of Python [info](#) containing results of the request (list)

See [PMIx_Group_construct](#) for description of all relevant attributes and behaviors.

A.4.27 Client.group_invite

Summary

Explicitly invite specified processes to join a group.

Format

Python

```
rc, info = myclient.group_invite(grp:string,  
                                 members:list, directives:list)
```

Python

IN grp

Python string identifier for the group (str)

IN members

List of Python [proc](#) dictionaries identifying processes to be invited (list)

IN directives

Optional list of Python [info](#) describing request (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- *info* - List of Python [info](#) containing results of the request (list)

See [PMIx_Group_invite](#) for description of all relevant attributes and behaviors.

A.4.28 Client.group_join

Summary

Respond to an invitation to join a group that is being asynchronously constructed.

Format

Python

```
rc, info = myclient.group_join(grp:string,  
                               leader:dict, opt:integer,  
                               directives:list)
```

Python

IN grp

Python string identifier for the group (str)

IN leader

Python [proc](#) dictionary identifying process leading the group (dict)

IN opt

One of the [pmix_group_opt_t](#) values indicating decline/accept (integer)

IN directives

Optional list of Python [info](#) describing request (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- *info* - List of Python [info](#) containing results of the request (list)

See [PMIx_Group_join](#) for description of all relevant attributes and behaviors.

A.4.29 Client.group_leave

Summary

Leave a PMIx Group.

Format

Python

```
rc = myclient.group_leave(grp:string, directives:list)
```

Python

IN grp

Python string identifier for the group (str)

IN directives

Optional list of Python [info](#) describing request (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

See [PMIx_Group_leave](#) for description of all relevant attributes and behaviors.

A.4.30 Client.group_destruct

Summary

Destruct a PMIx Group.

1 **Format** Python

2 `rc = myclient.group_destruct(grp:string, directives:list)` Python

3 **IN** `grp`
 4 Python string identifier for the group (str)

5 **IN** `directives`
 6 Optional list of Python `info` describing request (list)

7 Returns:

8 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

9 See `PMIx_Group_destruct` for description of all relevant attributes and behaviors.

10 A.4.31 Client.register_event_handler

11 **Summary**
 12 Register an event handler to report events.

13 *PMIx v4.0* **Format** Python

14 `rc, id = myclient.register_event_handler(codes:list,`
 15 `directives:list, cbfunc)` Python

16 **IN** `codes`
 17 List of Python integer status codes that should be reported to this handler (llist)

18 **IN** `directives`
 19 Optional list of Python `info` describing request (list)

20 **IN** `cbfunc`
 21 Python `evhandler` to be called when event is received (func)

22 Returns:

23 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

24 • `id` - PMIx reference identifier for handler (integer)

25 See `PMIx_Register_event_handler` for description of all relevant attributes and behaviors.

26 A.4.32 Client.deregister_event_handler

27 **Summary**
 28 Deregister an event handler.

1 **Format** Python

2 `myclient.deregister_event_handler(id:integer)`

3 **IN id**

4 PMIx reference identifier for handler (integer)

5 Returns: None

6 See [PMIx_Deregister_event_handler](#) for description of all relevant attributes and behaviors.

7 **A.4.33 Client.notify_event**

8 **Summary**

9 Report an event for notification via any registered handler.

10 *PMIx v4.0* **Format** Python

11 `rc = myclient.notify_event(status:integer, source:dict,`

12 `range:integer, directives:list)`

13 **IN status**

14 PMIx status code indicating the event being reported (integer)

15 **IN source**

16 Python [proc](#) of the process that generated the event (dict)

17 **IN range**

18 Python [range](#) in which the event is to be reported (integer)

19 **IN directives**

20 Optional list of Python [info](#) dictionaries describing the event (list)

21 Returns:

22 • `rc` - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

23 See [PMIx_Notify_event](#) for description of all relevant attributes and behaviors.

24 **A.4.34 Client.fabric_register**

25 **Summary**

26 Register for access to fabric-related information, including communication cost matrix.

```
1 Format Python
2 rc, idx, fabricinfo = myclient.fabric_register(directives:list) Python
3 IN directives
4     Optional list of Python info containing directives (list)
5 Returns:
6     • rc - PMIX_SUCCESS or a negative value corresponding to a PMIx error constant (integer)
7     • idx - Index of the registered fabric (integer)
8     • fabricinfo - List of Python info containing fabric info (list)
9 See PMIx_Fabric_register for details.
```

10 A.4.35 Client.fabric_update

11 **Summary**
12 Update fabric-related information, including communication cost matrix.

```
13 PMIx v4.0 Format Python
14 rc, fabricinfo = myclient.fabric_update(idx:integer) Python
```

```
15 IN idx
16     Index of the registered fabric (list)
17 Returns:
18     • rc - PMIX_SUCCESS or a negative value corresponding to a PMIx error constant (integer)
19     • fabricinfo - List of Python info containing updated fabric info (list)
20 See PMIx_Fabric_update for details.
```

21 A.4.36 Client.fabric_deregister

22 **Summary**
23 Deregister fabric.

1 **Format** Python

2 `rc = myclient.fabric_deregister(idx:integer)`
Python

3 **IN** `idx`
4 Index of the registered fabric (list)

5 Returns:

- 6 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

7 See [PMIx_Fabric_deregister](#) for details.

8 **A.4.37 Client.load_topology**

9 **Summary**

10 Load the local hardware topology into the PMIx library.

11 *PMIx v4.0* **Format** Python

12 `rc = myclient.load_topology()`
Python

13 Returns:

- 14 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

15 See [PMIx_Load_topology](#) for details - note that the topology loaded into the PMIx library may be utilized
16 by PMIx and other libraries, but is not directly accessible by Python.

17 **A.4.38 Client.get_relative_locality**

18 **Summary**

19 Get the relative locality of two local processes.

20 *PMIx v4.0* **Format** Python

21 `rc, locality = myclient.get_relative_locality(loc1:str, loc2:str)`
Python

22 **IN** `loc1`
23 Locality string of a process (str)

24 **IN** `loc2`
25 Locality string of a process (str)

26 Returns:

- 27 • `rc` - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

- 28 • `locality` - **locality** list containing the relative locality of the two processes (list)

29 See [PMIx_Get_relative_locality](#) for details.

1 A.4.39 Client.get_cpuset

2 Summary

3 Get the PU binding bitmap of the current process.

4 *PMIx v4.0* Format

Python

```
5 rc, cpuset = myclient.get_cpuset(ref: integer)
```

Python

6 IN ref

7 **bindenv** binding envelope to be used (integer)

8 Returns:

- 9 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 10 • *cpuset* - **cpuset** containing the source and bitmap of the cpuset (dict)

11 See **PMIx_Get_cpuset** for details.

12 A.4.40 Client.parse_cpuset_string

13 Summary

14 Parse the PU binding bitmap from its string representation.

15 *PMIx v4.0* Format

Python

```
16 rc, cpuset = myclient.parse_cpuset_string(cpuset: string)
```

Python

17 IN cpuset

18 String returned by **PMIxServer.generate_cpuset_string** (string)

19 Returns:

- 20 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 21 • *cpuset* - **cpuset** containing the source and bitmap of the cpuset (dict)

22 See **PMIx_Parse_cpuset_string** for details.

23 A.4.41 Client.compute_distances

24 Summary

25 Compute distances from specified process location to local devices.

Format

Python

```
rc, distances = myclient.compute_distances(cpuset:dict, info:list)
```

Python

IN `cpuset`

`cpuset` describing the location of the process (dict)

IN `info`

List of `info` dictionaries describing the devices whose distance is to be computed (list)

Returns:

- `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)
- `distances` - List of `devdist` structures containing the distances from the caller to the specified devices (list)

See `PMIx_Compute_distances` for details. Note that distances can only be computed against the local topology.

A.4.42 Client.error_string

Summary

Pretty-print string representation of `pmix_status_t`.

Format

Python

```
rep = myclient.error_string(status:integer)
```

Python

IN `status`

PMIx status code (integer)

Returns:

- `rep` - String representation of the provided status code (str)

See `PMIx_Error_string` for further details.

A.4.43 Client.proc_state_string

Summary

Pretty-print string representation of `pmix_proc_state_t`.

```
1 Format Python
2 rep = myclient.proc_state_string(state:integer) Python
3 IN state
4     PMIx process state code (integer)
5 Returns:
6 • rep - String representation of the provided process state (str)
7 See PMIx\_Proc\_state\_string for further details.
```

8 A.4.44 Client.scope_string

9 **Summary**
10 Pretty-print string representation of [pmix_scope_t](#).

```
11 PMIx v4.0 Format Python
12 rep = myclient.scope_string(scope:integer) Python
13 IN scope
14     PMIx scope value (integer)
15 Returns:
16 • rep - String representation of the provided scope (str)
17 See PMIx\_Scope\_string for further details
```

18 A.4.45 Client.persistence_string

19 **Summary**
20 Pretty-print string representation of [pmix_persistence_t](#).

```
21 PMIx v4.0 Format Python
22 rep = myclient.persistence_string(persistence:integer) Python
23 IN persistence
24     PMIx persistence value (integer)
25 Returns:
26 • rep - String representation of the provided persistence (str)
27 See PMIx\_Persistence\_string for further details.
```

1 A.4.46 Client.data_range_string

2 Summary

3 Pretty-print string representation of [pmix_data_range_t](#).

4 *PMIx v4.0* Format

Python

```
5 rep = myclient.data_range_string(range:integer)
```

Python

6 IN range

7 PMIx data range value (integer)

8 Returns:

- 9 • *rep* - String representation of the provided data range (str)

10 See [PMIx_Data_range_string](#) for further details.

11 A.4.47 Client.info_directives_string

12 Summary

13 Pretty-print string representation of [pmix_info_directives_t](#).

14 *PMIx v4.0* Format

Python

```
15 rep = myclient.info_directives_string(directives:bitarray)
```

Python

16 IN directives

17 PMIx [info directives](#) value (bitarray)

18 Returns:

- 19 • *rep* - String representation of the provided info directives (str)

20 See [PMIx_Info_directives_string](#) for further details.

21 A.4.48 Client.data_type_string

22 Summary

23 Pretty-print string representation of [pmix_data_type_t](#).

1 **Format** Python

2 `rep = myclient.data_type_string(dtype:integer)` Python

3 **IN dtype**

4 PMIx datatype value (integer)

5 Returns:

6 • *rep* - String representation of the provided datatype (str)

7 See [PMIx_Data_type_string](#) for further details.

8 A.4.49 Client.alloc_directive_string

9 **Summary**

10 Pretty-print string representation of [pmix_alloc_directive_t](#).

11 *PMIx v4.0* **Format** Python

12 `rep = myclient.alloc_directive_string(adir:integer)` Python

13 **IN adir**

14 PMIx allocation directive value (integer)

15 Returns:

16 • *rep* - String representation of the provided allocation directive (str)

17 See [PMIx_Alloc_directive_string](#) for further details.

18 A.4.50 Client.iof_channel_string

19 **Summary**

20 Pretty-print string representation of [pmix_iof_channel_t](#).

21 *PMIx v4.0* **Format** Python

22 `rep = myclient.iof_channel_string(channel:bitarray)` Python

23 **IN channel**

24 PMIx IOF [channel](#) value (bitarray)

25 Returns:

26 • *rep* - String representation of the provided IOF channel (str)

27 See [PMIx_IOF_channel_string](#) for further details.

1 A.4.51 Client.job_state_string

2 Summary

3 Pretty-print string representation of [pmix_job_state_t](#).

4 *PMIx v4.0* Format

Python

```
5 rep = myclient.job_state_string(state:integer)
```

Python

6 IN state

7 PMIx job state value (integer)

8 Returns:

- 9 • *rep* - String representation of the provided job state (str)

10 See [PMIx_Job_state_string](#) for further details.

11 A.4.52 Client.get_attribute_string

12 Summary

13 Pretty-print string representation of a PMIx attribute.

14 *PMIx v4.0* Format

Python

```
15 rep = myclient.get_attribute_string(attribute:str)
```

Python

16 IN attribute

17 PMIx attribute name (string)

18 Returns:

- 19 • *rep* - String representation of the provided attribute (str)

20 See [PMIx_Get_attribute_string](#) for further details.

21 A.4.53 Client.get_attribute_name

22 Summary

23 Pretty-print name of a PMIx attribute corresponding to the provided string.

```

1 Format Python
2 rep = myclient.get_attribute_name(attribute:str) Python
3 IN attributestring
4     Attribute string (string)
5 Returns:
6 • rep - Attribute name corresponding to the provided string (str)
7 See PMIx\_Get\_attribute\_name for further details.

```

8 A.4.54 Client.link_state_string

9 **Summary**
10 Pretty-print string representation of [pmix_link_state_t](#).

```

11 PMIx v4.0 Format Python
12 rep = myclient.link_state_string(state:integer) Python
13 IN state
14     PMIx link state value (integer)
15 Returns:
16 • rep - String representation of the provided link state (str)
17 See PMIx\_Link\_state\_string for further details.

```

18 A.4.55 Client.device_type_string

19 **Summary**
20 Pretty-print string representation of [pmix_device_type_t](#).

```

21 PMIx v4.0 Format Python
22 rep = myclient.device_type_string(type:bitarray) Python
23 IN type
24     PMIx device type value (bitarray)
25 Returns:
26 • rep - String representation of the provided device type (str)
27 See PMIx\_Device\_type\_string for further details.

```

1 A.4.56 Client.progress

2 Summary

3 Progress the PMIx library.

4 *PMIx v4.0* **Format** Python

5 `myclient.progress()` Python

6 See [PMIx_Progress](#) for further details.

7 A.5 PMIxServer

8 The server Python class inherits the Python "client" class as its parent. Thus, it includes all client functions in
9 addition to the ones defined in this section.

10 A.5.1 Server.init

11 Summary

12 Initialize the PMIx server library after obtaining a new PMIxServer object.

13 *PMIx v4.0* **Format** Python

14 `rc = myserver.init(directives:list, map:dict)` Python

15 **IN directives**

16 List of Python [info](#) dictionaries (list)

17 **IN map**

18 Python dictionary key-function pairs that map [server module](#) callback functions to provided
19 implementations (see [pmix_server_module_t](#)) (dict)

20 Returns:

- 21 • *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

22 See [PMIx_server_init](#) for description of all relevant attributes and behaviors.

23 A.5.2 Server.finalize

24 Summary

25 Finalize the PMIx server library.

1 **Format** Python

```
2 rc = myserver.finalize()
```

3 Returns:

- 4 • rc - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

5 See [PMIx_server_finalize](#) for details.

6 A.5.3 Server.generate_regex

7 Summary

8 Generate a regular expression representation of the input strings.

9 *PMIx v4.0* **Format** Python

```
10 rc, regex = myserver.generate_regex(input:list)
```

11 **IN input**
12 List of Python strings (e.g., node names) (list)

13 Returns:

- 14 • rc - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 15 • regex - Python **bytearray** containing regular expression representation of the input list (**bytearray**)

16 See [PMIx_generate_regex](#) for details.

17 A.5.4 Server.generate_ppn

18 Summary

19 Generate a regular expression representation of the input strings.

20 *PMIx v4.0* **Format** Python

```
21 rc, regex = myserver.generate_ppn(input:list)
```

22 **IN input**
23 List of Python strings, each string consisting of a comma-delimited list of ranks on each node, with the
24 strings being in the same order as the node names provided to "generate_regex" (list)

25 Returns:

- 26 • rc - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 27 • regex - Python **bytearray** containing regular expression representation of the input list (**bytearray**)

28 See [PMIx_generate_ppn](#) for details.

1 A.5.5 Server.generate_locality_string

2 Summary

3 Generate a PMIx locality string from a given cpuset.

4 *PMIx v4.0* Format

Python

```
5 rc, locality = myserver.generate_locality_string(cpuset:dict)
```

Python

6 **IN** `cset`

7 `cpuset` containing the bitmap of assigned PUs (dict)

8 Returns:

- 9 • `rc` - **PMIx_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 10 • `locality` - String representation of the PMIx locality corresponding to the input bitmap (string)

11 See [PMIx_server_generate_locality_string](#) for details.

12 A.5.6 Server.generate_cpuset_string

13 Summary

14 Generate a PMIx string representation of the provided cpuset.

15 *PMIx v4.0* Format

Python

```
16 rc, cpustr = myserver.generate_cpuset_string(cpuset:dict)
```

Python

17 **IN** `cset`

18 `cpuset` containing the bitmap of assigned PUs (dict)

19 Returns:

- 20 • `rc` - **PMIx_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 21 • `cpustr` - String representation of the input bitmap (string)

22 See [PMIx_server_generate_cpuset_string](#) for details.

23 A.5.7 Server.register_namespace

24 Summary

25 Setup the data about a particular namespace.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Format

Python

```
rc = myserver.register_namespace(namespace: str,  
                                nlocalprocs: integer,  
                                directives: list)
```

Python

- IN namespace**
Python string containing the namespace (str)
- IN nlocalprocs**
Number of local processes (integer)
- IN directives**
List of Python [info](#) dictionaries (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- See [PMIx_server_register_namespace](#) for description of all relevant attributes and behaviors.

A.5.8 Server.deregister_namespace

Summary

Deregister a namespace.

Format

Python

PMIx v4.0

```
myserver.deregister_namespace(namespace: str)
```

Python

- IN namespace**
Python string containing the namespace (str)

Returns: None

See [PMIx_server_deregister_namespace](#) for details.

A.5.9 Server.register_resources

Summary

Register non-namespace related information with the local PMIx library

Format

Python

PMIx v4.0

```
myserver.register_resources(directives: list)
```

Python

- IN directives**
List of Python [info](#) dictionaries (list)

Returns: None

See [PMIx_server_register_resources](#) for details.

1 A.5.10 Server.deregister_resources

2 Summary

3 Remove non-namespace related information from the local PMIx library

4 *PMIx v4.0*

Format

Python

5 `myserver.deregister_resources(directives:list)`

Python

6 **IN** `directives`

7 List of Python [info](#) dictionaries (list)

8 Returns: None

9 See [PMIx_server_deregister_resources](#) for details.

10 A.5.11 Server.register_client

11 Summary

12 Register a client process with the PMIx server library.

13 *PMIx v4.0*

Format

Python

14 `rc = myserver.register_client(proc:dict, uid:integer, gid:integer)`

Python

15 **IN** `proc`

16 Python [proc](#) dictionary identifying the client process (dict)

17 **IN** `uid`

18 Linux uid value for user executing client process (integer)

19 **IN** `gid`

20 Linux gid value for user executing client process (integer)

21 Returns:

- 22
 - `rc` - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

23 See [PMIx_server_register_client](#) for details.

24 A.5.12 Server.deregister_client

25 Summary

26 Deregister a client process and purge all data relating to it.

1 **Format** Python

2 `myserver.deregister_client(proc:dict)` Python

3 **IN** `proc`

4 Python `proc` dictionary identifying the client process (dict)

5 Returns: None

6 See [PMIx_server_deregister_client](#) for details.

7 A.5.13 Server.setup_fork

8 **Summary**

9 Setup the environment of a child process that is to be forked by the host.

10 *PMIx v4.0* **Format** Python

11 `rc = myserver.setup_fork(proc:dict, environ:dict)` Python

12 **IN** `proc`

13 Python `proc` dictionary identifying the client process (dict)

14 **INOUT** `environ`

15 Python dictionary containing the environment to be passed to the client (dict)

16 Returns:

17 • `rc` - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

18 See [PMIx_server_setup_fork](#) for details.

19 A.5.14 Server.dmodex_request

20 **Summary**

21 Function by which the host server can request modex data from the local PMIx server.

22 *PMIx v4.0* **Format** Python

23 `rc, data = myserver.dmodex_request(proc:dict)` Python

24 **IN** `proc`

25 Python `proc` dictionary identifying the process whose data is requested (dict)

26 Returns:

27 • `rc` - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

28 • `data` - Python [byteobject](#) containing the returned data (dict)

29 See [PMIx_server_dmodex_request](#) for details.

1 A.5.15 Server.setup_application

2 Summary

3 Function by which the resource manager can request application-specific setup data prior to launch of a *job*.

4 *PMIx v4.0* Format

Python

```
5 rc,info = myserver.setup_application(namespace:str, directives:list)
```

Python

6 IN namespace

7 Namespace whose setup information is being requested (str)

8 IN directives

9 Python list of *info* directives

10 Returns:

- 11 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
- 12 • *info* - Python list of *info* dictionaries containing the returned data (list)

13 See [PMIx_server_setup_application](#) for details.

14 A.5.16 Server.register_attributes

15 Summary

16 Register host environment attribute support for a function.

17 *PMIx v4.0* Format

Python

```
18 rc = myserver.register_attributes(function:str, attrs:list)
```

Python

19 IN function

20 Name of the function (str)

21 IN attrs

22 Python list of *regattr* describing the supported attributes

23 Returns:

- 24 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)

25 See [PMIx_Register_attributes](#) for details.

26 A.5.17 Server.setup_local_support

27 Summary

28 Function by which the local PMIx server can perform any application-specific operations prior to spawning
29 local clients of a given application.

1
2
3
4
5
6
7
8
9

Format

Python

```
rc = myserver.setup_local_support(namespace: str, info: list)
```

Python

IN namespace

Namespace whose setup information is being requested (str)

IN info

Python list of [info](#) containing the setup data (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

See [PMIx_server_setup_local_support](#) for details.

10 A.5.18 Server.iof_deliver

11 Summary

12 Function by which the host environment can pass forwarded IO to the PMIx server library for distribution to
13 its clients.

14 Format

PMIx v4.0

Python

```
rc = myserver.iof_deliver(source: dict, channel: integer,  
                           data: dict, directives: list)
```

Python

17 IN source

18 Python [proc](#) dictionary identifying the process who generated the data (dict)

19 IN channel

20 Python [channel](#) bitmask identifying IO channel of the provided data (integer)

21 IN data

22 Python [byteobject](#) containing the data (dict)

23 IN directives

24 Python list of [info](#) containing directives (list)

25 Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

27 See [PMIx_server_IOF_deliver](#) for details.

28 A.5.19 Server.collect_inventory

29 Summary

30 Collect inventory of resources on a node.

1 **Format** Python

```
2 rc, info = myserver.collect_inventory(directives:list)
```

Python

3 **IN directives**
4 Optional Python list of **info** containing directives (list)

5 Returns:
6 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
7 • *info* - Python list of **info** containing the returned data (list)
8 See [PMIx_server_collect_inventory](#) for details.

9 A.5.20 Server.deliver_inventory

10 **Summary**
11 Pass collected inventory to the PMIx server library for storage.

12 *PMIx v4.0* **Format** Python

```
13 rc = myserver.deliver_inventory(info:list, directives:list)
```

Python

14 **IN info**
15 - Python list of **info** dictionaries containing the inventory data (list)

16 **IN directives**
17 Python list of **info** dictionaries containing directives (list)

18 Returns:
19 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
20 See [PMIx_server_deliver_inventory](#) for details.

21 A.5.21 Server.define_process_set

22 **Summary**
23 Add members to a PMIx process set.

```

1 Format Python
2 rc = myserver.define_process_set(members:list, name:str) Python
3 IN members
4 - List of Python proc dictionaries identifying the processes to be added to the process set (list)
5 IN name
6 - Name of the process set (str)
7 Returns:
8 • rc - PMIX_SUCCESS or a negative value corresponding to a PMIx error constant (integer)
9 See PMIx_server_define_process_set for details.

```

A.5.22 Server.delete_process_set

Summary
Delete a PMIx process set.

```

13 PMIx v4.0 Format Python
14 rc = myserver.delete_process_set(name:str) Python
15 IN name
16 - Name of the process set (str)
17 Returns:
18 • rc - PMIX_SUCCESS or a negative value corresponding to a PMIx error constant (integer)
19 See PMIx_server_delete_process_set for details.

```

A.5.23 Server.register_resources

Summary
Register non-namespace related information with the local PMIx server library.

```

23 PMIx v4.0 Format Python
24 rc = myserver.register_resources(info:list) Python
25 IN info
26 - List of Python info dictionaries list)
27 Returns:
28 • rc - PMIX_SUCCESS or a negative value corresponding to a PMIx error constant (integer)
29 See PMIx_server_register_resources for details.

```


1 A.5.24 Server.deregister_resources

2 Summary

3 Deregister non-namespace related information with the local PMIx server library.

4 *PMIx v4.0*

Format

Python

```
5 rc = myserver.deregister_resources(info:list)
```

Python

6 **IN** `info`

7 - List of Python `info` dictionaries list

8 Returns:

- 9 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

10 See [PMIx_server_deregister_resources](#) for details.

11 A.6 PMIxTool

12 The tool Python class inherits the Python "server" class as its parent. Thus, it includes all client and server
13 functions in addition to the ones defined in this section.

14 A.6.1 Tool.init

15 Summary

16 Initialize the PMIx tool library after obtaining a new PMIxTool object.

17 *PMIx v4.0*

Format

Python

```
18 rc,proc = mytool.init(info:list)
```

Python

19 **IN** `info`

20 List of Python `info` directives (list)

21 Returns:

- 22 • `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

- 23 • `proc` - a Python `proc` (dict)

24 See [PMIx_tool_init](#) for description of all relevant attributes and behaviors.

25 A.6.2 Tool.finalize

26 Summary

27 Finalize the PMIx tool library, closing the connection to the server.

1 **Format** Python

```
2 rc = mytool.finalize()
```

3 Returns:
4 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
5 See **PMIx_tool_finalize** for description of all relevant attributes and behaviors.

6 A.6.3 Tool.disconnect

7 **Summary**
8 Disconnect the PMIx tool from the specified server connection while leaving the tool library initialized.

9 *PMIx v4.0* **Format** Python

```
10 rc = mytool.disconnect(server:dict)
```

11 **IN server**
12 Process identifier of server from which the tool is to be disconnected (**proc**)

13 Returns:
14 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
15 See **PMIx_tool_disconnect** for details.

16 A.6.4 Tool.attach_to_server

17 **Summary**
18 Establish a connection to a PMIx server.

19 *PMIx v4.0* **Format** Python

```
20 rc,proc,server = mytool.connect_to_server(info:list)
```

21 **IN info**
22 List of Python **info** dictionaries (list)

23 Returns:
24 • *rc* - **PMIX_SUCCESS** or a negative value corresponding to a PMIx error constant (integer)
25 • *proc* - a Python **proc** containing the tool's identifier (dict)
26 • *server* - a Python **proc** containing the identifier of the server to which the tool attached (dict)
27 See **PMIx_tool_attach_to_server** for details.

1 A.6.5 Tool.get_servers

2 Summary

3 Get a list containing the [proc](#) process identifiers of all servers to which the tool is currently connected.

4 *PMIx v4.0* Format

Python

```
5 rc, servers = mytool.get_servers()
```

Python

6 Returns:

- 7 • *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- 8 • *servers* - a list of Python [proc](#) containing the identifiers of the servers to which the tool is currently attached (dict)

10 See [PMIx_tool_get_servers](#) for details.

11 A.6.6 Tool.set_server

12 Summary

13 Designate a server as the tool's *primary* server.

14 *PMIx v4.0* Format

Python

```
15 rc = mytool.set_server(proc:dict, info:list)
```

Python

16 **IN** *proc*

17 Python [proc](#) containing the identifier of the servers to which the tool is to attach (list)

18 **IN** *info*

19 List of Python [info](#) dictionaries (list)

20 Returns:

- 21 • *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

22 See [PMIx_tool_set_server](#) for details.

23 A.6.7 Tool.iof_pull

24 Summary

25 Register to receive output forwarded from a remote process.

Format

Python

```
rc, id = mytool.iof_pull(sources:list, channel:integer,  
                        directives:list, cbfunc)
```

Python

IN sources

List of Python [proc](#) dictionaries of processes whose IO is being requested (list)

IN channel

Python [channel](#) bitmask identifying IO channels to be forwarded (integer)

IN directives

List of Python [info](#) dictionaries describing request (list)

IN cbfunc

Python [iofcbfunc](#) to receive IO payloads (func)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)
- *id* - PMIx reference identifier for request (integer)

See [PMIx_IOF_pull](#) for description of all relevant attributes and behaviors.

A.6.8 Tool.iof_deregister

Summary

Deregister from output forwarded from a remote process.

Format

Python

```
rc = mytool.iof_deregister(id:integer, directives:list)
```

Python

IN id

PMIx reference identifier returned by pull request (list)

IN directives

List of Python [info](#) dictionaries describing request (list)

Returns:

- *rc* - [PMIX_SUCCESS](#) or a negative value corresponding to a PMIx error constant (integer)

See [PMIx_IOF_deregister](#) for description of all relevant attributes and behaviors.

A.6.9 Tool.iof_push

Summary

Push data collected locally (typically from stdin) to stdin of target recipients.

Format

Python

```
rc = mytool.iof_push(targets:list, data:dict, directives:list)
```

Python

IN sources

List of Python `proc` of target processes (list)

IN data

Python `byteobject` containing data to be delivered (dict)

IN directives

Optional list of Python `info` describing request (list)

Returns:

- `rc` - `PMIX_SUCCESS` or a negative value corresponding to a PMIx error constant (integer)

See `PMIx_IOF_push` for description of all relevant attributes and behaviors.

A.7 Example Usage

The following examples are provided to illustrate the use of the Python bindings.

A.7.1 Python Client

The following example contains a client program that illustrates a fairly common usage pattern. The program instantiates and initializes the `PMIxClient` class, posts some data that is to be shared across all processes in the job, executes a “fence” that circulates the data, and then retrieves a value posted by one of its peers. Note that the example has been formatted to fit the document layout.

Python

```
from pmix import *

def main():
    # Instantiate a client object
    myclient = PMIxClient()
    print("Testing PMIx ", myclient.get_version())

    # Initialize the PMIx client library, declaring the programming model
    # as "TEST" and the library name as "PMIX", just for the example
    info = ['key':PMIX_PROGRAMMING_MODEL,
            'value':'TEST', 'val_type':PMIX_STRING,
            'key':PMIX_MODEL_LIBRARY_NAME,
            'value':'PMIX', 'val_type':PMIX_STRING]
    rc,myname = myclient.init(info)
    if PMIX_SUCCESS != rc:
        print("FAILED TO INIT WITH ERROR", myclient.error_string(rc))
        exit(1)
```

```

1      # try posting a value
2      rc = myclient.put(PMIX_GLOBAL, "mykey",
3                       'value':1, 'val_type':PMIX_INT32)
4      if PMIX_SUCCESS != rc:
5          print("PMIx_Put FAILED WITH ERROR", myclient.error_string(rc))
6          # cleanly finalize
7          myclient.finalize()
8          exit(1)
9
10     # commit it
11     rc = myclient.commit()
12     if PMIX_SUCCESS != rc:
13         print("PMIx_Commit FAILED WITH ERROR",
14              myclient.error_string(rc))
15         # cleanly finalize
16         myclient.finalize()
17         exit(1)
18
19     # execute fence across all processes in my job
20     procs = []
21     info = []
22     rc = myclient.fence(procs, info)
23     if PMIX_SUCCESS != rc:
24         print("PMIx_Fence FAILED WITH ERROR", myclient.error_string(rc))
25         # cleanly finalize
26         myclient.finalize()
27         exit(1)
28
29     # Get a value from a peer
30     if 0 != myname['rank']:
31         info = []
32         rc, get_val = myclient.get('namespace':"testnspace", 'rank': 0,
33                                   "mykey", info)
34         if PMIX_SUCCESS != rc:
35             print("PMIx_Commit FAILED WITH ERROR",
36                  myclient.error_string(rc))
37             # cleanly finalize
38             myclient.finalize()
39             exit(1)
40         print("Get value returned: ", get_val)
41
42     # test a fence that should return not_supported because
43     # we pass a required attribute that the server is known
44     # not to support
45     procs = []
46     info = ['key': 'ARBIT', 'flags': PMIX_INFO_REQD,
47            'value':10, 'val_type':PMIX_INT]

```

```

1     rc = myclient.fence(procs, info)
2     if PMIX_SUCCESS == rc:
3         print("PMIx_Fence SUCCEEDED BUT SHOULD HAVE FAILED")
4         # cleanly finalize
5         myclient.finalize()
6         exit(1)
7
8     # Publish something
9     info = ['key': 'ARBITRARY', 'value':10, 'val_type':PMIX_INT]
10    rc = myclient.publish(info)
11    if PMIX_SUCCESS != rc:
12        print("PMIx_Publish FAILED WITH ERROR",
13              myclient.error_string(rc))
14        # cleanly finalize
15        myclient.finalize()
16        exit(1)
17
18    # finalize
19    info = []
20    myclient.finalize(info)
21    print("Client finalize complete")
22
23    # Python main program entry point
24    if __name__ == '__main__':
25        main()

```

Python

26 A.7.2 Python Server

27 The following example contains a minimum-level server host program that instantiates and initializes the
28 PMIxServer class. The program illustrates passing several server module functions to the bindings and
29 includes code to setup and spawn a simple client application, waiting until the spawned client terminates
30 before finalizing and exiting itself. Note that the example has been formatted to fit the document layout.

```

31 from pmix import *
32 import signal, time
33 import os
34 import select
35 import subprocess
36
37 def clientconnected(proc:tuple is not None):
38     print("CLIENT CONNECTED", proc)
39     return PMIX_OPERATION_SUCCEEDED
40
41 def clientfinalized(proc:tuple is not None):
42     print("CLIENT FINALIZED", proc)

```

Python

```

1     return PMIX_OPERATION_SUCCEEDED
2
3 def clientfence(procs:list, directives:list, data:bytearray):
4     # check directives
5     if directives is not None:
6         for d in directives:
7             # these are each an info dict
8             if "pmix" not in d['key']:
9                 # we do not support such directives - see if
10                # it is required
11                try:
12                    if d['flags'] & PMIX_INFO_REQD:
13                        # return an error
14                        return PMIX_ERR_NOT_SUPPORTED
15                except:
16                    #it can be ignored
17                    pass
18            return PMIX_OPERATION_SUCCEEDED
19
20 def main():
21     try:
22         myserver = PMIXServer()
23     except:
24         print("FAILED TO CREATE SERVER")
25         exit(1)
26     print("Testing server version ", myserver.get_version())
27
28     args = ['key':PMIX_SERVER_SCHEDULER,
29            'value':'T', 'val_type':PMIX_BOOL]
30     map = 'clientconnected': clientconnected,
31          'clientfinalized': clientfinalized,
32          'fencenb': clientfence
33     my_result = myserver.init(args, map)
34
35     # get our environment as a base
36     env = os.environ.copy()
37
38     # register an nspace for the client app
39     (rc, regex) = myserver.generate_regex("test000,test001,test002")
40     (rc, ppn) = myserver.generate_ppn("0")
41     kvals = ['key':PMIX_NODE_MAP,
42            'value':regex, 'val_type':PMIX_STRING,
43            'key':PMIX_PROC_MAP,
44            'value':ppn, 'val_type':PMIX_STRING,
45            'key':PMIX_UNIV_SIZE,
46            'value':1, 'val_type':PMIX_UINT32,
47            'key':PMIX_JOB_SIZE,

```



```

1         'value':1, 'val_type':PMIX_UINT32]
2     rc = foo.register_nspace("testnspace", 1, kval)
3     print("RegNspace ", rc)
4
5     # register a client
6     uid = os.getuid()
7     gid = os.getgid()
8     rc = myserver.register_client('nspace':"testnspace", 'rank':0,
9                                 uid, gid)
10
11     print("RegClient ", rc)
12     # setup the fork
13     rc = myserver.setup_fork('nspace':"testnspace", 'rank':0, env)
14     print("SetupFrk", rc)
15
16     # setup the client argv
17     args = ["/client.py"]
18     # open a subprocess with stdout and stderr
19     # as distinct pipes so we can capture their
20     # output as the process runs
21     p = subprocess.Popen(args, env=env,
22                          stdout=subprocess.PIPE, stderr=subprocess.PIPE)
23     # define storage to catch the output
24     stdout = []
25     stderr = []
26     # loop until the pipes close
27     while True:
28         reads = [p.stdout.fileno(), p.stderr.fileno()]
29         ret = select.select(reads, [], [])
30
31         stdout_done = True
32         stderr_done = True
33
34         for fd in ret[0]:
35             # if the data
36             if fd == p.stdout.fileno():
37                 read = p.stdout.readline()
38                 if read:
39                     read = read.decode('utf-8').rstrip()
40                     print('stdout: ' + read)
41                     stdout_done = False
42             elif fd == p.stderr.fileno():
43                 read = p.stderr.readline()
44                 if read:
45                     read = read.decode('utf-8').rstrip()
46                     print('stderr: ' + read)
47                     stderr_done = False

```

```
1         if stdout_done and stderr_done:
2             break
3         print("FINALIZING")
4         myserver.finalize()
5
6
7     if __name__ == '__main__':
8         main()
```



APPENDIX B

Use-Cases

1 The PMIx standard provides many generic interfaces that can be composed into higher-level use cases in a
2 variety of ways. While the specific interfaces and attributes are standardized, the use cases themselves are not
3 (and should not) be standardized. Common use cases are included here as examples of how PMIx's generic
4 interfaces *might* be composed together for a higher-level purpose. The use cases are intended for both PMIx
5 interface users and library implementors. Whereby a better understanding of the general usage model within
6 the community can help users picking up PMIx for the first and help implementors optimize their
7 implementation for the common cases.

8 Each use case is structured to provide background information about the high-level use case as well as specific
9 details about how the PMIx interfaces are used within the use case. Some use cases even provide code snippets.
10 These code snippets are apart of larger code examples located within the standard's source code repository, and
11 each complete code example is fully compilable and runnable. The related interfaces and attributes collected
12 at the bottom of each use case are mainly for convenience and link to the full standardized definitions.

13 B.1 Business Card Exchange for Process-to- 14 Process Wire-up

15 B.1.1 Use Case Summary

16 Multi-process communication libraries, such as MPI, need to establish communication channels between a set
17 of those processes. In this scenario, each process needs to share connectivity information (a.k.a. Business
18 Cards) with all other processes before communication channels can be established. This connectivity
19 information may take the form of one or more unique strings that allow a different process to establish a
20 communication channel with the originator. The runtime environment must provide a mechanism for the
21 efficient exchange of this connectivity information. Additional information about the current state of the job
22 (e.g., number of processes globally and locally) and of how the process was started (e.g., process binding) is
23 also helpful.

24 Note: The Instant-On wire-up mechanism is a separate, related use case.

25 B.1.2 Use Case Details

26 Each process provides their business card to PMIx via one or more `PMIx_Put` operations to store the tuple of
27 `{UID, key, value}`. The `UID` is the unique name for this process in the PMIx universe (i.e.,
28 `namespace` and `rank`). The `key` is a unique key that other processes can reference generically (note that
29 since the `UID` is also associated with the `key` there is no need to make the `key` uniquely named per process).
30 The `value` is the string representation of the connectivity information.

1 Some business card information is meant for remote processes (e.g., TCP or InfiniBand addresses) while
2 others are meant only for local processes (e.g., shared memory information). As such a **scope** should be
3 associated with the **PMIx_Put** operation to differentiate this intention.

4 The **PMIx_Put** operations may be cached local to the process. Once all **PMIx_Put** operations have been
5 called each process should call **PMIx_Commit** to push those values to the local PMIx server. Note that in a
6 multi-library configuration each library may **PMIx_Put** then **PMIx_Commit** values - so there may be
7 multiple **PMIx_Commit** calls before a Business Card Exchange is activated.

8 After calling **PMIx_Commit** a process can activate the Business Card Exchange collective operation by
9 calling **PMIx_Fence**. The **PMIx_Fence** operation is collective over the set of processes specified in the
10 argument set. That allows for the collective to span a subset of a namespace or multiple namespaces. After the
11 completion of the **PMIx_Fence** operation, the data stored by other processes via **PMIx_Put** is available to
12 the local process through a call to **PMIx_Get** which returns the key/value pairs necessary to establish the
13 connection(s) with the other processes.

14 The **PMIx_Fence** operation has a "Synchronize Only" mode that works as a barrier operation. This is
15 helpful if the communication library requires a synchronization before leaving initialization or starting
16 finalization, for example.

17 The **PMIx_Fence** operation has a "Sparse" mode in addition to a "Full" mode for the data exchange. The
18 "Full" mode will fully exchange all Business Card information with all other processes. This is helpful for
19 tightly communicating applications. The "Sparse" mode will dynamically pull the connectivity information
20 on-demand from inside of **PMIx_Get** (if it is not already available locally). This is helpful for sparsely
21 communicating applications. Since which mode is best for an application cannot be inferred by the PMIx
22 library the caller must specify which mode works best for their application. The **PMIx_Fence** operation has
23 an option for the end user to specify which mode they desire for this operation.

24 Additional information about the current state of the job (e.g., number of processes globally and locally) and of
25 how the process was started (e.g., process binding) is also helpful. This "job level" information is available
26 immediately after **PMIx_Init** without the need for any explicit synchronization.

27 The number of processes globally in the namespace and this process's rank within that namespace is important
28 to know before establishing the Business Card information to best allocate resources.

29 The number of processes local to the node and this process's local rank is important to know before
30 establishing the Business Card information to help the caller determine the scope of the put operation. For
31 example, to designate a leader to set up a shared memory segment of the proper size before putting that
32 information into the locally scoped Business Card information.

33 The number of processes local to a remote node is also helpful to know before establishing the Business Card
34 information. This information is useful to pre-establish local resources before that remote node starts to
35 initiate a connection or to determine the number of connections that need to be advertised in the Business Card
36 when it is sent out.

37 Note that some of the job level information may change over the course of the job in a dynamic application.

38 **Related Interfaces**

39 **PMIx_Put**

PMIx v1.0

```

1      pmix_status_t
2      PMIx_Put (pmix_scope_t scope,
3                const pmix_key_t key,
4                pmix_value_t *val);
5
6      PMIx_Get
7
8      pmix_status_t
9      PMIx_Get (const pmix_proc_t *proc, const pmix_key_t key,
10               const pmix_info_t info[], size_t ninfo,
11               pmix_value_t **val);
12
13     PMIx_Commit
14
15     pmix_status_t PMIx_Commit (void);
16
17     PMIx_Fence
18
19     pmix_status_t
20     PMIx_Fence (const pmix_proc_t procs[], size_t nprocs,
21                 const pmix_info_t info[], size_t ninfo);
22
23     PMIx_Init
24
25     pmix_status_t
26     PMIx_Init (pmix_proc_t *proc,
27                pmix_info_t info[], size_t ninfo)

```

1 Related Attributes

2 The following job level information is useful to have before establishing Business Card information:

3 **PMIX_NODE_LIST** "pmix.nlist" (char*)

4 Comma-delimited list of nodes currently hosting processes in the specified realm. Defaults to the *job*
5 realm.

6 **PMIX_NUM_NODES** "pmix.num.nodes" (uint32_t)

7 Number of nodes currently hosting processes in the specified realm. Defaults to the *job* realm.

8 **PMIX_NODEID** "pmix.nodeid" (uint32_t)

9 Node identifier expressed as the node's index (beginning at zero) in an array of nodes within the active
10 session. The value must be unique and directly correlate to the **PMIX_HOSTNAME** of the node - i.e.,
11 users can interchangeably reference the same location using either the **PMIX_HOSTNAME** or
12 corresponding **PMIX_NODEID**.

13 **PMIX_JOB_SIZE** "pmix.job.size" (uint32_t)

14 Total number of processes in the specified job across all contained applications. Note that this value
15 can be different from **PMIX_MAX_PROCS**. For example, users may choose to subdivide an allocation
16 (running several jobs in parallel within it), and dynamic programming models may support adding and
17 removing processes from a running *job* on-the-fly. In the latter case, PMIx events may be used to notify
18 processes within the job that the job size has changed.

19 **PMIX_PROC_MAP** "pmix.pmap" (char*)

20 Regular expression describing processes on each node in the specified realm - see 17.2.3.2 for an
21 explanation of its generation. Defaults to the *job* realm.

22 **PMIX_LOCAL_PEERS** "pmix.lpeers" (char*)

23 Comma-delimited list of ranks that are executing on the local node within the specified namespace -
24 shortcut for **PMIx_Resolve_peers** for the local node.

25 **PMIX_LOCAL_SIZE** "pmix.local.size" (uint32_t)

26 Number of processes in the specified job or application on the caller's node. Defaults to job unless the
27 **PMIX_APP_INFO** and the **PMIX_APPNUM** qualifiers are given.

28 For each process this information is also useful (note that any one process may want to access this list of
29 information about any other process in the system):

30 **PMIX_RANK** "pmix.rank" (pmix_rank_t)

31 Process rank within the job, starting from zero.

32 **PMIX_LOCAL_RANK** "pmix.lrank" (uint16_t)

33 Rank of the specified process on its node - refers to the numerical location (starting from zero) of the
34 process on its node when counting only those processes from the same job that share the node, ordered
35 by their overall rank within that job.

36 **PMIX_GLOBAL_RANK** "pmix.grank" (pmix_rank_t)

37 Rank of the specified process spanning across all jobs in this session, starting with zero. Note that no
38 ordering of the jobs is implied when computing this value. As jobs can start and end at random times,
39 this is defined as a continually growing number - i.e., it is not dynamically adjusted as individual jobs
40 and processes are started or terminated.

1 **PMIX_LOCALITY_STRING** "pmix.locstr" (char*)

2 String describing a process's bound location - referenced using the process's rank. The string is
3 prefixed by the implementation that created it (e.g., "hwloc") followed by a colon. The remainder of the
4 string represents the corresponding locality as expressed by the underlying implementation. The entire
5 string must be passed to **PMIx_Get_relative_locality** for processing. Note that hosts are
6 only required to provide locality strings for local client processes - thus, a call to **PMIx_Get** for the
7 locality string of a process that returns **PMIX_ERR_NOT_FOUND** indicates that the process is not
8 executing on the same node.

9 **PMIX_HOSTNAME** "pmix.hname" (char*)

10 Name of the host, as returned by the **gethostname** utility or its equivalent.

11 There are other keys that are helpful to have before a synchronization point. This is not meant to be a
12 comprehensive list.

13 B.2 Debugging

14 B.2.1 Terminology

15 B.2.1.1 Tools vs Debuggers

16 A *tool* is a process designed to monitor, record, analyze, or control the execution of another process. Typically
17 used for the purposes of profiling and debugging. A *first-party tool* runs within the address space of the
18 application process while a *third-party tool* run within its own process. A *debugger* is a third-party tool that
19 inspects and controls an application process's execution using system-level debug APIs (e.g., **ptrace**).

20 B.2.1.2 Parallel Launching Methods

21 A *starter* program is a program responsible for launching a parallel runtime, such as MPI. PMIx supports two
22 primary methods for launching parallel applications under tools and debuggers: indirect and direct. In the
23 indirect launching method (Section 18.2.2, the tool is attached to the starter. In the direct launching method
24 (Section 18.2.1, the tool takes the place of the starter. PMIx also supports attaching to already running
25 programs via the *Process Acquisition* interfaces (Section B.2.1.4).

26 B.2.1.3 Process Synchronization

27 Process Synchronization is a technique tools use to start the processes of a parallel application such that the
28 tools can still attach to the process early in its lifetime. Said another way, the tool must be able to start the
29 application processes without them "running away" from the tool. In the case of MPI (Version 3.1 [4] or the
30 MPI World Process in future versions), this means stopping the applications processes before they return from
31 **MPI_Init** or **MPI_Init_thread**.

32 B.2.1.4 Process Acquisition

33 Process Acquisition is a technique tools use to locate all of the processes, local and remote, of a given parallel
34 application. This typically boils down to collecting the following information for every process in the parallel
35 application: the hostname or IP of the machine running the process, the executable name, and the process ID.

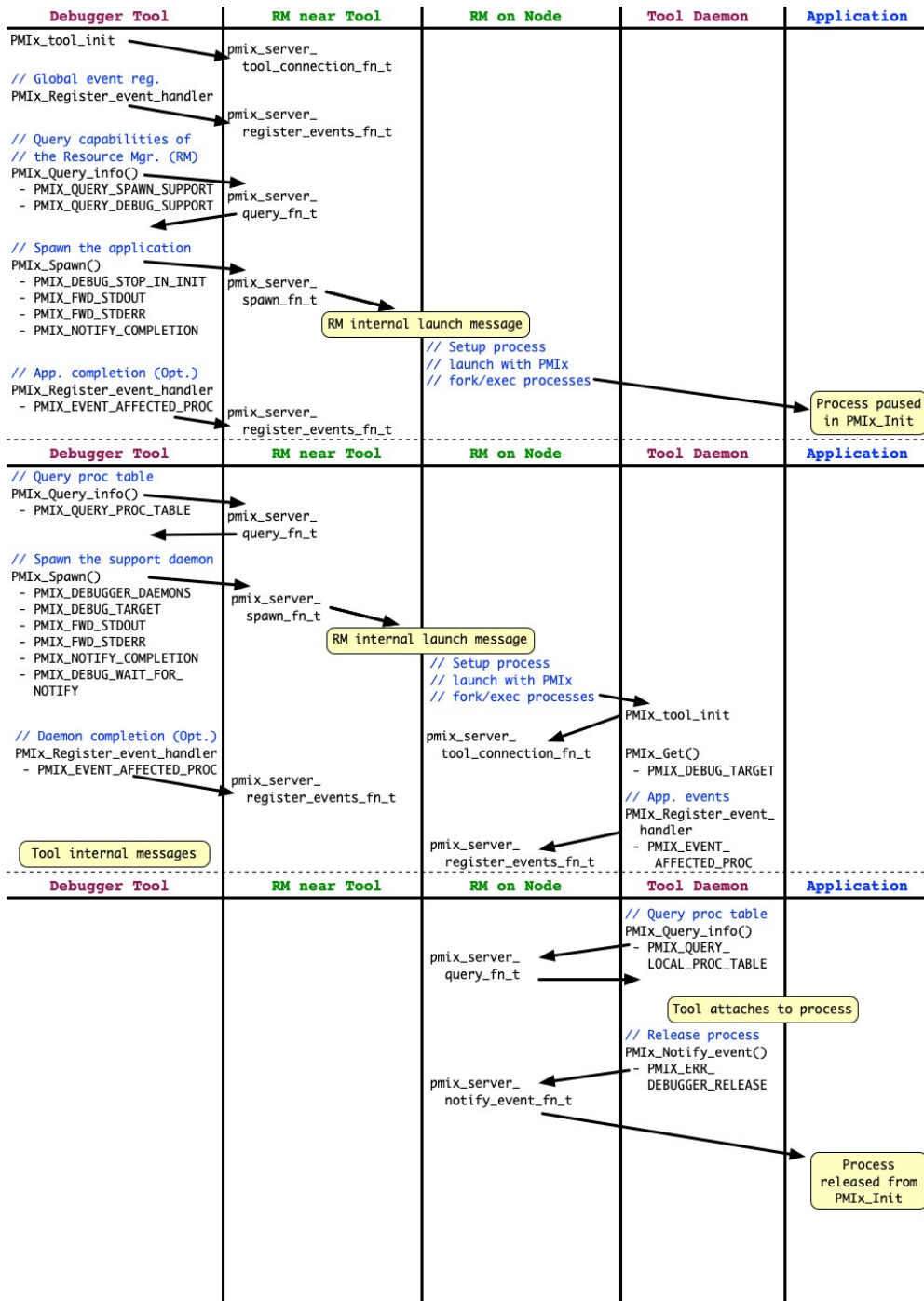


Figure B.1.: Interaction diagram showing an example of the Direct Launch mechanism

```

1  pmix_status_t
2  PMIx_Spawn(const pmix_info_t job_info[], size_t ninfo,
3             const pmix_app_t apps[], size_t napps,
4             char nspace[])

```

5 *PMIx v1.0* **PMIx_Get**

```

6  pmix_status_t
7  PMIx_Get(const pmix_proc_t *proc, const pmix_key_t key,
8           const pmix_info_t info[], size_t ninfo,
9           pmix_value_t **val);

```

10 *PMIx v2.0* **PMIx_Notify_event**

```

11 pmix_status_t
12 PMIx_Notify_event(pmix_status_t status,
13                 const pmix_proc_t *source,
14                 pmix_data_range_t range,
15                 pmix_info_t info[], size_t ninfo,
16                 pmix_op_cbfunc_t cbfunc, void *cbdata);

```

17 Related Attributes

18 **PMIX_QUERY_SPAWN_SUPPORT** "pmix.qry.spawn" (bool)
 19 Return a comma-delimited list of supported spawn attributes. NO QUALIFIERS.

20 **PMIX_QUERY_DEBUG_SUPPORT** "pmix.qry.debug" (bool)
 21 Return a comma-delimited list of supported debug attributes. NO QUALIFIERS.

22 **PMIX_DEBUG_STOP_IN_INIT** "pmix.dbg.init" (bool)
 23 Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies
 24 only to that application) or in the `job_info` array if it applies to all applications in the given spawn
 25 request. Indicates that the specified application is being spawned under a debugger. The PMIx client
 26 library in each resulting application process shall notify its PMIx server that it is pausing and then
 27 pause during **PMIx_Init** of the spawned processes until either released by debugger modification of
 28 an appropriate variable or receipt of the **PMIX_DEBUGGER_RELEASE** event. The launcher (RM or
 29 IL) is responsible for generating the **PMIX_READY_FOR_DEBUG** event (stipulating a breakpoint of
 30 `pmix-init`) when all processes have reached the pause point.

31 **PMIX_DEBUG_STOP_ON_EXEC** "pmix.dbg.exec" (bool)

1 Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies
2 only to that application) or in the `job_info` array if it applies to all applications in the given spawn
3 request. Indicates that the application is being spawned under a debugger, and that the local launch
4 agent is to pause the resulting application processes on first instruction for debugger attach. The
5 launcher (RM or IL) is to generate the `PMIX_LAUNCH_COMPLETE` event when all processes are
6 stopped at the exec point.

7 **PMIX_DEBUG_DAEMONS_PER_PROC** "pmix.dbg.dpproc" (uint16_t)

8 Number of debugger daemons to be spawned per application process. The launcher is to pass the
9 identifier of the namespace to be debugged by including the `PMIX_DEBUG_TARGET` attribute in the
10 daemon's job-level information. The debugger daemons spawned on a given node are responsible for
11 self-determining their specific target process(es) - e.g., by referencing their own `PMIX_LOCAL_RANK`
12 in the daemon debugger job versus the corresponding `PMIX_LOCAL_RANK` of the target processes on
13 the node.

14 **PMIX_DEBUG_DAEMONS_PER_NODE** "pmix.dbg.dpnd" (uint16_t)

15 Number of debugger daemons to be spawned on each node where the target job is executing. The
16 launcher is to pass the identifier of the namespace to be debugged by including the
17 `PMIX_DEBUG_TARGET` attribute in the daemon's job-level information. The debugger daemons
18 spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by
19 referencing their own `PMIX_LOCAL_RANK` in the daemon debugger job versus the corresponding
20 `PMIX_LOCAL_RANK` of the target processes on the node.

21 **PMIX_COSPAWN_APP** "pmix.cospawn" (bool)

22 Designated application is to be spawned as a disconnected job - i.e., the launcher shall not include the
23 application in any of the job-level values (e.g., `PMIX_RANK` within the job) provided to any other
24 application process generated by the same spawn request. Typically used to cospawn debugger
25 daemons alongside an application.

26 **PMIX_MAPBY** "pmix.mapby" (char*)

27 Process mapping policy - when accessed using `PMIx_Get`, use the `PMIX_RANK_WILDCARD` value
28 for the rank to discover the mapping policy used for the provided namespace. Supported values are
29 launcher specific.

30 **PMIX_FWD_STDOUT** "pmix.fwd.stdout" (bool)

31 Requests that the ability to forward the `stdout` of the spawned processes be maintained. The
32 requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options for
33 delivery of the forwarded output.

34 **PMIX_FWD_STDERR** "pmix.fwd.stderr" (bool)

35 Requests that the ability to forward the `stderr` of the spawned processes be maintained. The
36 requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options for
37 delivery of the forwarded output.

38 **PMIX_NOTIFY_COMPLETION** "pmix.notecomp" (bool)

39 Requests that the launcher generate the `PMIX_EVENT_JOB_END` event for normal or abnormal
40 termination of the spawned job. The event shall include the returned status code
41 (`PMIX_JOB_TERM_STATUS`) for the corresponding job; the identity (`PMIX_PROCID`) and exit
42 status (`PMIX_EXIT_CODE`) of the first failed process, if applicable; and a

`PMIX_EVENT_TIMESTAMP` indicating the time the termination occurred. Note that the requester must register for the event or capture and process it within a default event handler.

`PMIX_SETUP_APP_ENVARS` "pmix.setup.env" (bool)

Harvest and include relevant environmental variables.

`PMIX_EVENT_AFFECTED_PROC` "pmix.evproc" (pmix_proc_t)

The single process that was affected.

`PMIX_DEBUGGER_DAEMONS` "pmix.debugger" (bool)

Included in the `pmix_info_t` array of a `pmix_app_t`, this attribute declares that the application consists of debugger daemons and shall be governed accordingly. If used as the sole `pmix_app_t` in a `PMIx_Spawn` request, then the `PMIX_DEBUG_TARGET` attribute must also be provided (in either the `job_info` or in the `info` array of the `pmix_app_t`) to identify the namespace to be debugged so that the launcher can determine where to place the spawned daemons. If neither

`PMIX_DEBUG_DAEMONS_PER_PROC` nor `PMIX_DEBUG_DAEMONS_PER_NODE` is specified, then the launcher shall default to a placement policy of one daemon per process in the target job.

`PMIX_DEBUG_TARGET` "pmix.dbg.tgt" (pmix_proc_t*)

Identifier of process(es) to be debugged - a rank of `PMIX_RANK_WILDCARD` indicates that all processes in the specified namespace are to be included.

`PMIX_DEBUG_STOP_IN_APP` "pmix.dbg.notify" (varies)

Direct specified ranks to stop at application-specific point and notify they are ready-to-debug. The attribute's value can be any of three data types:

- bool - true indicating all ranks
- `pmix_rank_t` - the rank of one proc, or `PMIX_RANK_WILDCARD` for all
- a `pmix_data_array_t` if an array of individual processes are specified

The resulting application processes are to notify their server (by generating the `PMIX_READY_FOR_DEBUG` event) when they reach some application-determined location - the event shall include the `PMIX_BREAKPOINT` attribute indicating where the application has stopped. The application shall pause at that point until released by debugger modification of an appropriate variable. The launcher (RM or IL) is responsible for generating the `PMIX_READY_FOR_DEBUG` event when all processes have indicated they are at the pause point.

`PMIX_QUERY_LOCAL_PROC_TABLE` "pmix.qry.lptable" (char*)

Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in the specified namespace executing on the same node as the requester, ordered by process job rank. REQUIRED QUALIFIER: `PMIX_NAMESPACE` indicating the namespace whose local process table is being queried. OPTIONAL QUALIFIER: `PMIX_HOSTNAME` indicating the host whose local process table is being queried. By default, the query assumes that the host upon which the request was made is to be used.

Related Constants

`PMIX_DEBUG_WAITING_FOR_NOTIFY`

`PMIX_DEBUGGER_RELEASE`

1 B.2.2.2 Indirect-Launch Debugger Tool

2 Executing a program under a tool using an intermediate launcher such as **mpirexec** can also be made possible.
3 This requires some degree of coordination between the tool and the launcher. Ultimately, it is the launcher that
4 is going to launch the application, and the tool must somehow inform the launcher (and the application) that
5 this is being done in a debug session so that the application knows to “block” until the tool attaches to it.

6 In this operational mode, the user invokes a tool (typically on a non-compute, or “head”, node) that in turn
7 uses **mpirexec** to launch their application – a typical command line might look like the following: **dbgx**
8 **-dbgoption mpirexec -n 32 ./myapp**.

9 Related Interfaces

10 *PMIx v2.0*

PMIx_tool_init

▼ C ▼

```
11 pmix_status_t  
12 PMIx_tool_init(pmix_proc_t *proc,  
13               pmix_info_t info[], size_t ninfo);
```

▲ C ▲

14 *PMIx v2.0*

PMIx_Register_event_handler

▼ C ▼

```
15 pmix_status_t  
16 PMIx_Register_event_handler(pmix_status_t codes[], size_t ncodes,  
17                             pmix_info_t info[], size_t ninfo,  
18                             pmix_notification_fn_t evhdlr,  
19                             pmix_hdlr_reg_cbfunc_t cbfunc,  
20                             void *cbdata);
```

▲ C ▲

21 *PMIx v1.0*

PMIx_Spawn

▼ C ▼

```
22 pmix_status_t  
23 PMIx_Spawn(const pmix_info_t job_info[], size_t ninfo,  
24            const pmix_app_t apps[], size_t napps,  
25            char nspace[])
```

▲ C ▲

26 *PMIx v2.0*

PMIx_Notify_event

▼ C ▼

```
27 pmix_status_t  
28 PMIx_Notify_event(pmix_status_t status,  
29                 const pmix_proc_t *source,  
30                 pmix_data_range_t range,  
31                 pmix_info_t info[], size_t ninfo,  
32                 pmix_op_cbfunc_t cbfunc, void *cbdata);
```

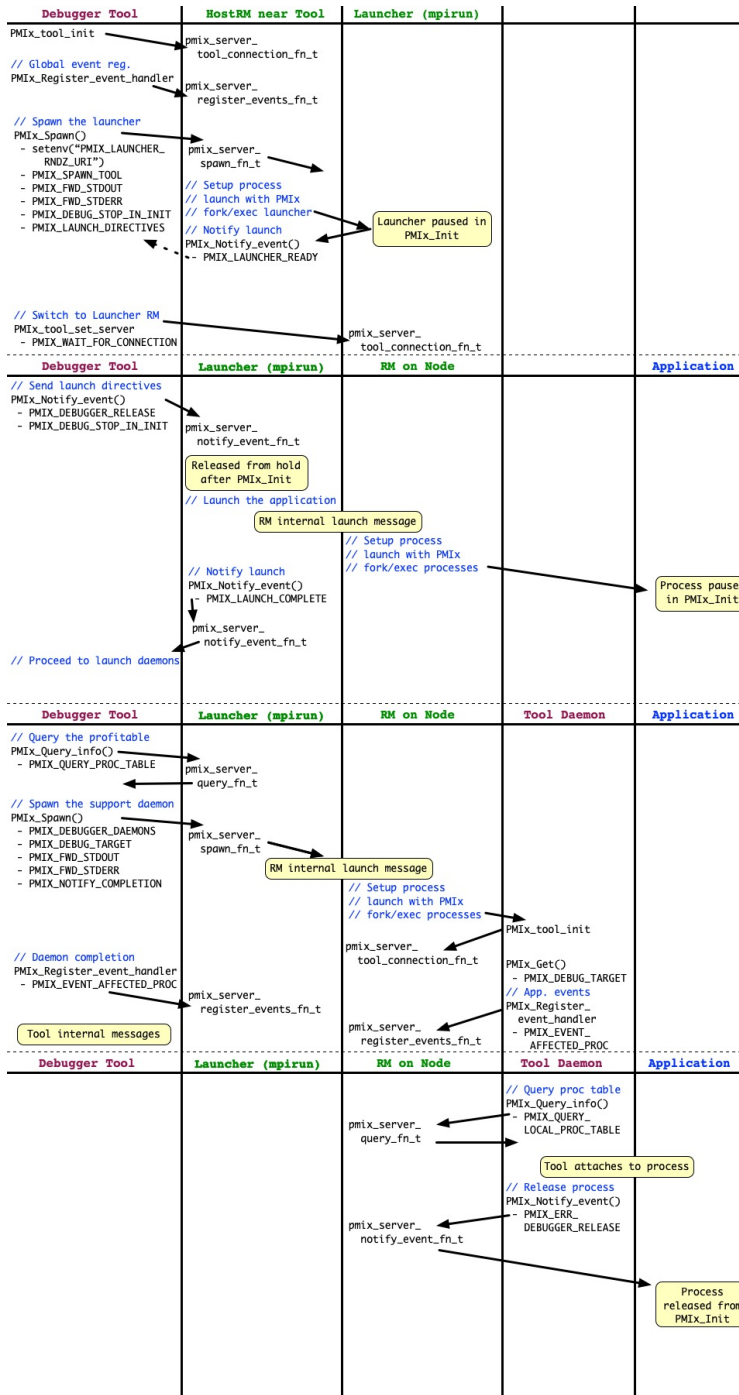


Figure B.2.: Interaction diagram showing an example of the Indirect Launch mechanism

1 *PMIx v4.0*

PMIx_tool_attach_to_server

2

`pmix_status_t`

3

`PMIx_tool_attach_to_server(pmix_proc_t *proc,`

4

`pmix_proc_t *server,`

5

`pmix_info_t info[],`

6

`size_t ninfo);`

7 *PMIx v4.0*

PMIx_Query_info

8

`pmix_status_t`

9

`PMIx_Query_info(pmix_query_t queries[], size_t nqueries,`

10

`pmix_info_t *info[], size_t *ninfo);`

11 *PMIx v1.0*

PMIx_Get

12

`pmix_status_t`

13

`PMIx_Get(const pmix_proc_t *proc, const pmix_key_t key,`

14

`const pmix_info_t info[], size_t ninfo,`

15

`pmix_value_t **val);`

16

Related Attributes

17

PMIX_LAUNCH_DIRECTIVES "`pmix.lnch.dirs`" (`pmix_data_array_t*`)

18

Array of `pmix_info_t` containing directives for the launcher - a convenience attribute for retrieving all directives with a single call to `PMIx_Get`.

19

20

PMIX_SPAWN_TOOL "`pmix.spwn.tool`" (`bool`)

21

Indicate that the job being spawned is a tool.

22

PMIX_COSPAWN_APP "`pmix.cospawn`" (`bool`)

23

Designated application is to be spawned as a disconnected job - i.e., the launcher shall not include the application in any of the job-level values (e.g., `PMIX_RANK` within the job) provided to any other application process generated by the same spawn request. Typically used to cospawn debugger daemons alongside an application.

24

25

26

27

PMIX_FWD_STDOUT "`pmix.fwd.stdout`" (`bool`)

28

Requests that the ability to forward the `stdout` of the spawned processes be maintained. The requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options for delivery of the forwarded output.

29

30

31

PMIX_FWD_STDERR "`pmix.fwd.stderr`" (`bool`)

1 Requests that the ability to forward the `stderr` of the spawned processes be maintained. The
2 requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options for
3 delivery of the forwarded output.

4 **PMIX_SETUP_APP_ENVARS** "pmix.setup.env" (bool)

5 Harvest and include relevant environmental variables.

6 **PMIX_DEBUG_STOP_IN_INIT** "pmix.dbg.init" (bool)

7 Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies
8 only to that application) or in the `job_info` array if it applies to all applications in the given spawn
9 request. Indicates that the specified application is being spawned under a debugger. The PMIx client
10 library in each resulting application process shall notify its PMIx server that it is pausing and then
11 pause during `PMIx_Init` of the spawned processes until either released by debugger modification of
12 an appropriate variable or receipt of the `PMIX_DEBUGGER_RELEASE` event. The launcher (RM or
13 IL) is responsible for generating the `PMIX_READY_FOR_DEBUG` event (stipulating a breakpoint of
14 `pmix-init`) when all processes have reached the pause point.

15 **PMIX_DEBUG_STOP_ON_EXEC** "pmix.dbg.exec" (bool)

16 Included in either the `pmix_info_t` array in a `pmix_app_t` description (if the directive applies
17 only to that application) or in the `job_info` array if it applies to all applications in the given spawn
18 request. Indicates that the application is being spawned under a debugger, and that the local launch
19 agent is to pause the resulting application processes on first instruction for debugger attach. The
20 launcher (RM or IL) is to generate the `PMIX_LAUNCH_COMPLETE` event when all processes are
21 stopped at the exec point.

22 **PMIX_DEBUG_DAEMONS_PER_PROC** "pmix.dbg.dpproc" (uint16_t)

23 Number of debugger daemons to be spawned per application process. The launcher is to pass the
24 identifier of the namespace to be debugged by including the `PMIX_DEBUG_TARGET` attribute in the
25 daemon's job-level information. The debugger daemons spawned on a given node are responsible for
26 self-determining their specific target process(es) - e.g., by referencing their own `PMIX_LOCAL_RANK`
27 in the daemon debugger job versus the corresponding `PMIX_LOCAL_RANK` of the target processes on
28 the node.

29 **PMIX_DEBUG_DAEMONS_PER_NODE** "pmix.dbg.dpnd" (uint16_t)

30 Number of debugger daemons to be spawned on each node where the target job is executing. The
31 launcher is to pass the identifier of the namespace to be debugged by including the
32 `PMIX_DEBUG_TARGET` attribute in the daemon's job-level information. The debugger daemons
33 spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by
34 referencing their own `PMIX_LOCAL_RANK` in the daemon debugger job versus the corresponding
35 `PMIX_LOCAL_RANK` of the target processes on the node.

36 **PMIX_MAPBY** "pmix.mapby" (char*)

37 Process mapping policy - when accessed using `PMIx_Get`, use the `PMIX_RANK_WILDCARD` value
38 for the rank to discover the mapping policy used for the provided namespace. Supported values are
39 launcher specific.

40 **PMIX_QUERY_PROC_TABLE** "pmix.qry.phtable" (char*)

41 Returns a (`pmix_data_array_t`) array of `pmix_proc_info_t`, one entry for each process in
42 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: `PMIX_NAMESPACE`
43 indicating the namespace whose process table is being queried.

1 **PMIX_QUERY_LOCAL_PROC_TABLE** "pmix.qry.lptable" (char*)
2 Returns a ([pmix_data_array_t](#)) array of [pmix_proc_info_t](#), one entry for each process in
3 the specified namespace executing on the same node as the requester, ordered by process job rank.
4 REQUIRED QUALIFIER: [PMIX_NAMESPACE](#) indicating the namespace whose local process table is
5 being queried. OPTIONAL QUALIFIER: [PMIX_HOSTNAME](#) indicating the host whose local process
6 table is being queried. By default, the query assumes that the host upon which the request was made is
7 to be used.

8 **PMIX_DEBUGGER_DAEMONS** "pmix.debugger" (bool)
9 Included in the [pmix_info_t](#) array of a [pmix_app_t](#), this attribute declares that the application
10 consists of debugger daemons and shall be governed accordingly. If used as the sole [pmix_app_t](#) in
11 a [PMIx_Spawn](#) request, then the [PMIX_DEBUG_TARGET](#) attribute must also be provided (in either
12 the *job_info* or in the *info* array of the [pmix_app_t](#)) to identify the namespace to be debugged so that
13 the launcher can determine where to place the spawned daemons. If neither
14 [PMIX_DEBUG_DAEMONS_PER_PROC](#) nor [PMIX_DEBUG_DAEMONS_PER_NODE](#) is specified, then
15 the launcher shall default to a placement policy of one daemon per process in the target job.

16 **PMIX_NOTIFY_COMPLETION** "pmix.notecomp" (bool)
17 Requests that the launcher generate the [PMIX_EVENT_JOB_END](#) event for normal or abnormal
18 termination of the spawned job. The event shall include the returned status code
19 ([PMIX_JOB_TERM_STATUS](#)) for the corresponding job; the identity ([PMIX_PROCID](#)) and exit
20 status ([PMIX_EXIT_CODE](#)) of the first failed process, if applicable; and a
21 [PMIX_EVENT_TIMESTAMP](#) indicating the time the termination occurred. Note that the requester
22 must register for the event or capture and process it within a default event handler.

23 **PMIX_DEBUG_TARGET** "pmix.dbg.tgt" ([pmix_proc_t*](#))
24 Identifier of process(es) to be debugged - a rank of [PMIX_RANK_WILDCARD](#) indicates that all
25 processes in the specified namespace are to be included.

26 **PMIX_WAIT_FOR_CONNECTION** "pmix.wait.conn" (bool)
27 Wait until the specified process has connected to the requesting tool or server, or the operation times
28 out (if the [PMIX_TIMEOUT](#) directive is included in the request).

29 Related Constants

30 [PMIX_LAUNCHER_READY](#)

31 [PMIX_LAUNCH_COMPLETE](#)

32 [PMIX_DEBUG_WAITING_FOR_NOTIFY](#)

33 [PMIX_DEBUGGER_RELEASE](#)

34 [PMIX_LAUNCHER_RNDZ_URI](#)

35 B.2.2.3 Attaching to a Running Job

36 PMIx supports attaching to an already running parallel job in two ways. In the first way, the main process of a
37 tool calls [PMIx_Query_info](#) with the [PMIX_QUERY_PROC_TABLE](#) attribute. This returns an array of
38 structs containing the information required for [process acquisition](#). This includes remote hostnames,
39 executable names, and process IDs. In the second way, every tool daemon calls [PMIx_Query_info](#) with the
40 [PMIX_QUERY_LOCAL_PROC_TABLE](#) attribute. This returns a similar array of structs but only for processes
41 on the same node.

42 An example of this use-case may look like the following: `mpirexec -n 32 ./myApp && dbgr`
43 `attach $!`

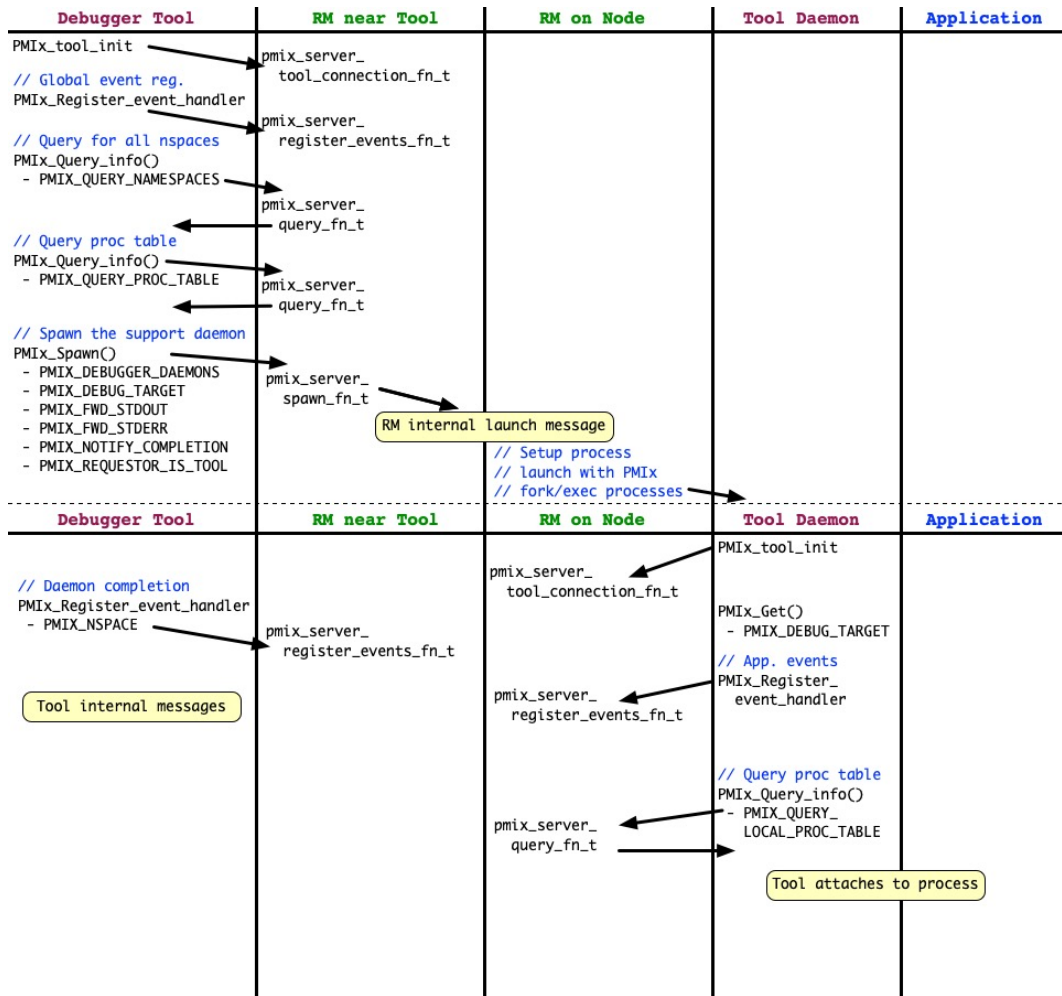


Figure B.3.: Interaction diagram showing an example of the attaching to a running job

1 *PMIx v2.0*

PMIx_tool_init

▼  ▼

```
2 pmix_status_t
3 PMIx_tool_init(pmix_proc_t *proc,
4               pmix_info_t info[], size_t ninfo);
```

▲  ▲

5 *PMIx v2.0*

PMIx_Register_event_handler

▼  ▼

```
6 pmix_status_t
7 PMIx_Register_event_handler(pmix_status_t codes[], size_t ncodes,
8                             pmix_info_t info[], size_t ninfo,
9                             pmix_notification_fn_t evhdlr,
10                            pmix_hdlr_reg_cbfunc_t cbfunc,
11                            void *cbdata);
```

▲  ▲

12 *PMIx v4.0*

PMIx_Query_info

▼  ▼

```
13 pmix_status_t
14 PMIx_Query_info(pmix_query_t queries[], size_t nqueries,
15                pmix_info_t *info[], size_t *ninfo);
```

▲  ▲

16 *PMIx v1.0*

PMIx_Spawn

▼  ▼

```
17 pmix_status_t
18 PMIx_Spawn(const pmix_info_t job_info[], size_t ninfo,
19            const pmix_app_t apps[], size_t napps,
20            char nspace[])
```

▲  ▲

Related Attributes

22 **PMIX_QUERY_PROC_TABLE** "pmix.qry.phtable" (char*)

23 Returns a (**pmix_data_array_t**) array of **pmix_proc_info_t**, one entry for each process in
24 the specified namespace, ordered by process job rank. REQUIRED QUALIFIER: **PMIX_NAMESPACE**
25 indicating the namespace whose process table is being queried.

26 **PMIX_DEBUGGER_DAEMONS** "pmix.debugger" (bool)

1 Included in the `pmix_info_t` array of a `pmix_app_t`, this attribute declares that the application
2 consists of debugger daemons and shall be governed accordingly. If used as the sole `pmix_app_t` in
3 a `PMIx_Spawn` request, then the `PMIX_DEBUG_TARGET` attribute must also be provided (in either
4 the `job_info` or in the `info` array of the `pmix_app_t`) to identify the namespace to be debugged so that
5 the launcher can determine where to place the spawned daemons. If neither
6 `PMIX_DEBUG_DAEMONS_PER_PROC` nor `PMIX_DEBUG_DAEMONS_PER_NODE` is specified, then
7 the launcher shall default to a placement policy of one daemon per process in the target job.

8 **PMIX_DEBUG_TARGET** `"pmix.dbg.tgt"` (`pmix_proc_t*`)

9 Identifier of process(es) to be debugged - a rank of `PMIX_RANK_WILDCARD` indicates that all
10 processes in the specified namespace are to be included.

11 **PMIX_DEBUG_DAEMONS_PER_PROC** `"pmix.dbg.dpproc"` (`uint16_t`)

12 Number of debugger daemons to be spawned per application process. The launcher is to pass the
13 identifier of the namespace to be debugged by including the `PMIX_DEBUG_TARGET` attribute in the
14 daemon's job-level information. The debugger daemons spawned on a given node are responsible for
15 self-determining their specific target process(es) - e.g., by referencing their own `PMIX_LOCAL_RANK`
16 in the daemon debugger job versus the corresponding `PMIX_LOCAL_RANK` of the target processes on
17 the node.

18 **PMIX_DEBUG_DAEMONS_PER_NODE** `"pmix.dbg.dpnd"` (`uint16_t`)

19 Number of debugger daemons to be spawned on each node where the target job is executing. The
20 launcher is to pass the identifier of the namespace to be debugged by including the
21 `PMIX_DEBUG_TARGET` attribute in the daemon's job-level information. The debugger daemons
22 spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by
23 referencing their own `PMIX_LOCAL_RANK` in the daemon debugger job versus the corresponding
24 `PMIX_LOCAL_RANK` of the target processes on the node.

25 **PMIX_MAPBY** `"pmix.mapby"` (`char*`)

26 Process mapping policy - when accessed using `PMIx_Get`, use the `PMIX_RANK_WILDCARD` value
27 for the rank to discover the mapping policy used for the provided namespace. Supported values are
28 launcher specific.

29 **PMIX_FWD_STDOUT** `"pmix.fwd.stdout"` (`bool`)

30 Requests that the ability to forward the `stdout` of the spawned processes be maintained. The
31 requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options for
32 delivery of the forwarded output.

33 **PMIX_FWD_STDERR** `"pmix.fwd.stderr"` (`bool`)

34 Requests that the ability to forward the `stderr` of the spawned processes be maintained. The
35 requester will issue a call to `PMIx_IOF_pull` to specify the callback function and other options for
36 delivery of the forwarded output.

37 **PMIX_NOTIFY_COMPLETION** `"pmix.notecomp"` (`bool`)

38 Requests that the launcher generate the `PMIX_EVENT_JOB_END` event for normal or abnormal
39 termination of the spawned job. The event shall include the returned status code
40 (`PMIX_JOB_TERM_STATUS`) for the corresponding job; the identity (`PMIX_PROCID`) and exit
41 status (`PMIX_EXIT_CODE`) of the first failed process, if applicable; and a
42 `PMIX_EVENT_TIMESTAMP` indicating the time the termination occurred. Note that the requester
43 must register for the event or capture and process it within a default event handler.


```

1 pmix_status_t
2 PMIx_Notify_event (pmix_status_t status,
3                   const pmix_proc_t *source,
4                   pmix_data_range_t range,
5                   pmix_info_t info[], size_t ninfo,
6                   pmix_op_cbfunc_t cbfunc, void *cbdata);

```

7 *PMIx v1.0* **PMIx_server_init**

```

8 pmix_status_t
9 PMIx_server_init (pmix_server_module_t *module,
10                  pmix_info_t info[], size_t ninfo);

```

11 **B.2.2.5 Environmental Parameter Directives for Applications and Launchers**

12
13 It is sometimes desirable or required that standard environmental variables (e.g., **PATH**,
14 **LD_LIBRARY_PATH**, **LD_PRELOAD**) be modified prior to executing an application binary or a starter such
15 as **mpiexec** - this is particularly true when tools/debuggers are used to start the application.

16 **Related Interfaces**

17 *PMIx v1.0* **PMIx_Spawn**

```

18 pmix_status_t
19 PMIx_Spawn (const pmix_info_t job_info[], size_t ninfo,
20             const pmix_app_t apps[], size_t napps,
21             char nspace[])

```

22 **Related Structs**

23 **pmix_envar_t**

Related Attributes

PMIX_SET_ENVAR "pmix.envar.set" (pmix_envar_t*)

Set the envar to the given value, overwriting any pre-existing one

PMIX_ADD_ENVAR "pmix.envar.add" (pmix_envar_t*)

Add the environment variable, but do not overwrite any pre-existing one

PMIX_UNSET_ENVAR "pmix.envar.unset" (char*)

Unset the environment variable specified in the string.

PMIX_PREPEND_ENVAR "pmix.envar.prepnd" (pmix_envar_t*)

Prepend the given value to the specified environmental value using the given separator character, creating the variable if it doesn't already exist

PMIX_APPEND_ENVAR "pmix.envar.appnd" (pmix_envar_t*)

Append the given value to the specified environmental value using the given separator character, creating the variable if it doesn't already exist

Resource managers and launchers must scan for relevant directives, modifying environmental parameters as directed. Directives are to be processed in the order in which they were given, starting with job-level directives (applied to each app) followed by app-level directives.

B.3 Hybrid Applications

B.3.1 Use Case Summary

Hybrid applications (i.e., applications that utilize more than one programming model or runtime system, such as an application using MPI that also uses OpenMP or UPS) are growing in popularity, especially as processors with increasingly large numbers of cores and/or hardware threads proliferate. Unfortunately, the various corresponding runtime systems currently operate under the assumption that they alone control execution. This leads to conflicts in hybrid applications. Deadlock of parallel applications can occur when one runtime system prevents the other from making progress due to lack of coordination between them [3]. Sub-optimal performance can also occur due to uncoordinated division of hardware resources between the runtime systems implementing the different programming models or systems [5, 6]. This use-case offers potential solutions to this problem by providing a pathway for parallel runtime systems to coordinate their actions.

B.3.2 Use Case Details

B.3.2.1 Identifying Active Parallel Runtime Systems

The current state-of-the-practice for concurrently used runtime systems in a single application to detect one another is via set environment variables. For example, some OpenMP implementations look for environment variables to indicate that an MPI library is active. Unfortunately, this technique is not completely reliable as environment variables change over time and with new software versions, and this detection is implementation specific. Also, the fact that an environment variable is present doesn't guarantee that a particular runtime system is in active use since Resource Managers routinely set environment variables "just in case" the application needs them. PMIx provides a reliable mechanism by which each library can determine that another runtime library is in operation.

1 When initializing PMIx, runtime libraries implementing a parallel programming model can register
2 themselves, including their name, the library version, the version of the API they implement, and the threading
3 model. This information is then cached locally and can then be read asynchronously by other runtime systems
4 using PMIx’s Event Notification system.

5 This initialization mechanism also allows runtime libraries to share knowledge of each other’s resources and
6 intended resource utilization. For example, if an OpenMP implementation knows which hardware threads an
7 MPI library is using it could potentially avoid core and cache contention.

8 Code Example

```
1 pmix_proc_t myproc;  
2 pmix_info_t *info;  
3 volatile bool wearedone = false;  
4  
5 PMIX_INFO_CREATE(info, 4);  
6 PMIX_INFO_LOAD(&info[0], PMIX_PROGRAMMING_MODEL, "MPI", PMIX_STRING);  
7 PMIX_INFO_LOAD(&info[1], PMIX_MODEL_LIBRARY_NAME, "FoMPI",  
8 ↪ PMIX_STRING);  
9 PMIX_INFO_LOAD(&info[2], PMIX_MODEL_LIBRARY_VERSION, "1.0.0",  
10 ↪ PMIX_STRING);  
11 PMIX_INFO_LOAD(&info[3], PMIX_THREADING_MODEL, "posix", PMIX_STRING);  
12 pmix_status_t rc = PMIx_Init(&myproc, info, 4);  
13 PMIX_INFO_FREE(info, 4);
```

9

10 Related Interfaces

11 *PMIx v1.2* [PMIx_Init](#)

```
▼ C ▼  
  
12 pmix_status_t  
13 PMIx_Init(pmix_proc_t *proc,  
14 ↪ pmix_info_t info[], size_t ninfo)  
▲ C ▲
```

15 Related Attributes

- 16 [PMIX_PROGRAMMING_MODEL](#) "pmix.pgm.model" (char*)
17 Programming model being initialized (e.g., "MPI" or "OpenMP").
- 18 [PMIX_MODEL_LIBRARY_NAME](#) "pmix.mdl.name" (char*)
19 Programming model implementation ID (e.g., "OpenMPI" or "MPICH").
- 20 [PMIX_MODEL_LIBRARY_VERSION](#) "pmix.mdl.vrs" (char*)
21 Programming model version string (e.g., "2.1.1").
- 22 [PMIX_THREADING_MODEL](#) "pmix.threads" (char*)
23 Threading model used (e.g., "pthreads").
- 24 [PMIX_MODEL_NUM_THREADS](#) "pmix.mdl.nthrds" (uint64_t)

1 Number of active threads being used by the model.

2 **PMIX_MODEL_NUM_CPUS** "pmix.mdl.ncpu" (uint64_t)

3 Number of cpus being used by the model.

4 **PMIX_MODEL_CPU_TYPE** "pmix.mdl.cputype" (char*)

5 Granularity - "hwthread", "core", etc.

6 **PMIX_MODEL_PHASE_NAME** "pmix.mdl.phase" (char*)

7 User-assigned name for a phase in the application execution (e.g., "cfd reduction").

8 **PMIX_MODEL_PHASE_TYPE** "pmix.mdl.ptype" (char*)

9 Type of phase being executed (e.g., "matrix multiply").

10 **PMIX_MODEL_AFFINITY_POLICY** "pmix.mdl.tap" (char*)

11 Thread affinity policy - e.g.: "master" (thread co-located with master thread), "close" (thread located

12 on cpu close to master thread), "spread" (threads load-balanced across available cpus).

13 B.3.2.2 Coordinating at Runtime

14 The PMIx Event Notification system provides a mechanism by which the resource manager can communicate

15 system events to applications, thus providing applications with an opportunity to generate an appropriate

16 response. Hybrid applications can leverage these events for cross-library coordination.

17 Runtime libraries can access the information provided by other runtime libraries during their initialization

18 using the event notification system. In this case, runtime libraries should register a callback for the

19 **PMIX_MODEL_DECLARED** event.

20 Applications, runtime libraries, and resource managers can also use the PMIx event notification system to

21 communicate dynamic information, such as entering a new application phase (**PMIX_MODEL_PHASE_NAME**)

22 or a change in resources used (**PMIX_MODEL_RESOURCES**). This dynamic information can be broadcast

23 using the **PMIx_Notify_event** function. Runtime libraries can register callback functions to run when

24 these events occur using **PMIx_Register_event_handler**.

25 Code Example

26 Registering a callback to run when another runtime library initializes:

```

1  static void model_declared_cb(size_t evhdlr_registration_id,
2                                pmix_status_t status, const pmix_proc_t
3                                ↪ *source,
4                                pmix_info_t info[], size_t ninfo,
5                                pmix_info_t results[], size_t nresults,
6                                pmix_event_notification_cbfnc_fn_t
7                                ↪ cbfunc,
8                                void *cbdata) {
9      printf("Entered %s\n", __FUNCTION__);
10     int n;
11     for (n = 0; n < ninfo; n++) {
12         if (PMIX_CHECK_KEY(&info[n], PMIX_PROGRAMMING_MODEL) &&
13             strcmp(info[n].value.data.string, "MPI") == 0) {
14             /* ignore our own declaration */

```

```

13     break;
14 } else {
15     /* actions to perform when another model registers */
16 }
17 }
18 if (NULL != cbfunc) {
19     /* tell the event handler that we are only a partial step */
20     cbfunc(PMIX_EVENT_PARTIAL_ACTION_TAKEN, NULL, 0, NULL, NULL,
↪     cbdata);
21 }
22 }
23
24 pmix_status_t code = PMIX_MODEL_DECLARED;
25 rc = PMIx_Register_event_handler(&code, 1, NULL, 0, model_declared_cb,
↪     NULL, NULL);

```

1

2

Notifying an event:

```

1 PMIX_INFO_CREATE(info, 1);
2 PMIX_INFO_LOAD(&info[0], PMIX_EVENT_NON_DEFAULT, NULL, PMIX_BOOL);
3 rc = PMIx_Notify_event(PMIX_OPENMP_PARALLEL_ENTERED, &myproc,
↪     PMIX_RANGE_PROC_LOCAL, info, 1, notify_complete, (void*)&wearedone);

```

3

4

Related Interfaces

5 *PMIx v2.0*

[PMIx_Notify_event](#)



6

`pmix_status_t`

7

`PMIx_Notify_event(pmix_status_t status,`

8

`const pmix_proc_t *source,`

9

`pmix_data_range_t range,`

10

`pmix_info_t info[], size_t ninfo,`

11

`pmix_op_cbfunc_t cbfunc, void *cbdata);`



12 *PMIx v2.0*

[PMIx_Register_event_handler](#)

```

1  pmix_status_t
2  PMIx_Register_event_handler(pmix_status_t codes[], size_t ncodes,
3                             pmix_info_t info[], size_t ninfo,
4                             pmix_notification_fn_t evhdlr,
5                             pmix_hdlr_reg_cbfunc_t cbfunc,
6                             void *cbdata);

```

7 *PMIx v2.0*

pmix_event_notification_cbfunc_fn_t

```

8  typedef void (*pmix_event_notification_cbfunc_fn_t)
9          (pmix_status_t status,
10         pmix_info_t *results, size_t nresults,
11         pmix_op_cbfunc_t cbfunc, void *thiscbdata,
12         void *notification_cbdata);

```

Related Constants

- [PMIX_MODEL_DECLARED](#)
- [PMIX_MODEL_RESOURCES](#)
- [PMIX_OPENMP_PARALLEL_ENTERED](#)
- [PMIX_OPENMP_PARALLEL_EXITED](#)
- [PMIX_EVENT_ACTION_COMPLETE](#)

B.3.2.3 Coordinating at Runtime with Multiple Event Handlers

Coordinating with a threading library such as an OpenMP runtime library creates the need for separate event handlers for threads of the same process. For example in an MPI+OpenMP hybrid application, the MPI main thread and the OpenMP primary thread may both want to be notified anytime an OpenMP thread starts executing in a parallel region. This requires support for multiple threads to potentially register different event handlers against the same status code.

Multiple event handlers registered against the same event are processed in a chain-like manner based on the order in which they were registered, as modified by any directives. Registrations against specific event codes are processed first, followed by registrations against multiple event codes and then any default registrations. At each point in the chain, an event handler is called by the PMIx progress thread and given a function to call when that handler has completed its operation. The handler callback notifies PMIx that the handler is done, returning a status code to indicate the result of its work. The results are appended to the array of prior results, with the returned values combined into an array within a single `pmix_info_t` as follows:

- `array[0]`: the event handler name provided at registration (may be an empty field if a string name was not given) will be in the key, with the `pmix_status_t` value returned by the handler
- `array[*]`: the array of results returned by the handler, if any.

The current PMIx standard does not actually specify a default ordering for event handlers as they are being registered. However, it does include an inherent ordering for invocation. Specifically, PMIx stipulates that handlers be called in the following categorical order:

- 1 • single status event handlers - handlers that were registered against a single specific status.
- 2 • multi status event handlers - those registered against more than one specific status.
- 3 • default event handlers - those registered against no specific status.

Code Example

From the OpenMP primary thread:

```

1 static void parallel_region_OMP_cb(size_t evhdlr_registration_id,
2     pmix_status_t status,
3     const pmix_proc_t *source,
4     pmix_info_t info[], size_t ninfo,
5     pmix_info_t results[], size_t
↪ nresults,
6     pmix_event_notification_cbfunc_fn_t
↪ cbfunc,
7     void *cbdata) {
8     printf("Entered %s\n", __FUNCTION__);
9     /* do what we need OpenMP to do on entering a parallel region */
10    if (NULL != cbfunc) {
11        /* tell the event handler that we are only a partial step */
12        cbfunc(PMIX_EVENT_PARTIAL_ACTION_TAKEN, NULL, 0, NULL, NULL,
↪ cbdata);
13    }
14 }
15
16 bool is_true = true;
17 pmix_status_t code = PMIX_OPENMP_PARALLEL_ENTERED;
18 PMIX_INFO_CREATE(info, 2);
19 PMIX_INFO_LOAD(&info[0], PMIX_EVENT_HDLR_NAME, "OpenMP-Primary",
↪ PMIX_STRING);
20 PMIX_INFO_LOAD(&info[1], PMIX_EVENT_HDLR_FIRST, &is_true, PMIX_BOOL);
21 rc = PMIX_Register_event_handler(&code, 1, info, 2,
↪ parallel_region_OMP_cb, NULL, NULL);
22 if (rc < 0)
23     fprintf(stderr, "%s: Failed to register event handler for OpenMP
↪ region entrance\n", __FUNCTION__);
24 PMIX_INFO_FREE(info, 2);

```

- 6
- 7 From the MPI process:

```

1 static void parallel_region_MPI_cb(size_t evhdlr_registration_id,
2     pmix_status_t status,
3     const pmix_proc_t *source,
4     pmix_info_t info[], size_t ninfo,

```

```

5                               pmix_info_t results[], size_t
↳ nresults,
6                               pmix_event_notification_cbfunc_fn_t
↳ cbfunc,
7                               void *cbdata) {
8     printf("Entered %s\n", __FUNCTION__);
9     /* do what we need the MPI library to do on entering a parallel region
↳ */
10    if (NULL != cbfunc) {
11        /* tell the event handler that we are the last step */
12        cbfunc(PMIX_EVENT_ACTION_COMPLETE, NULL, 0, NULL, NULL, cbdata);
13    }
14 }
15
16 pmix_status_t code = PMIX_OPENMP_PARALLEL_ENTERED;
17 PMIX_INFO_CREATE(info, 2);
18 PMIX_INFO_LOAD(&info[0], PMIX_EVENT_HDLR_NAME, "MPI-Thread",
↳ PMIX_STRING);
19 PMIX_INFO_LOAD(&info[1], PMIX_EVENT_HDLR_AFTER, "OpenMP-Primary",
↳ PMIX_STRING);
20 rc = PMIx_Register_event_handler(&code, 1, info, 2,
↳ parallel_region_MPI_cb, NULL, NULL);
21 if (rc < 0)
22     fprintf(stderr, "%s: Failed to register event handler for OpenMP
↳ region entrance\n", __FUNCTION__);
23 PMIX_INFO_FREE(info, 2);

```

1

2 **Related Interfaces**
3 *PMIx v2.0* [PMIx_Register_event_handler](#)



```

4 pmix_status_t
5 PMIx_Register_event_handler(pmix_status_t codes[], size_t ncodes,
6                             pmix_info_t info[], size_t ninfo,
7                             pmix_notification_fn_t evhdlr,
8                             pmix_hdlr_reg_cbfunc_t cbfunc,
9                             void *cbdata);

```



10 *PMIx v2.0* [pmix_event_notification_cbfunc_fn_t](#)

```

1 typedef void (*pmix_event_notification_cbfunc_fn_t)
2     (pmix_status_t status,
3      pmix_info_t *results, size_t nresults,
4      pmix_op_cbfunc_t cbfunc, void *thiscbdata,
5      void *notification_cbdata);

```

Related Attributes

PMIX_EVENT_HDLR_NAME "pmix.evname" (char*)

String name identifying this handler.

PMIX_EVENT_HDLR_FIRST "pmix.evfirst" (bool)

Invoke this event handler before any other handlers.

PMIX_EVENT_HDLR_LAST "pmix.evlast" (bool)

Invoke this event handler after all other handlers have been called.

PMIX_EVENT_HDLR_FIRST_IN_CATEGORY "pmix.evfirstcat" (bool)

Invoke this event handler before any other handlers in this category.

PMIX_EVENT_HDLR_LAST_IN_CATEGORY "pmix.evlastcat" (bool)

Invoke this event handler after all other handlers in this category have been called.

PMIX_EVENT_HDLR_BEFORE "pmix.evbefore" (char*)

Put this event handler immediately before the one specified in the (char*) value.

PMIX_EVENT_HDLR_AFTER "pmix.evafter" (char*)

Put this event handler immediately after the one specified in the (char*) value.

PMIX_EVENT_HDLR_APPEND "pmix.evappend" (bool)

Append this handler to the precedence list within its category.

Related Constants

PMIX_EVENT_NO_ACTION_TAKEN

PMIX_EVENT_PARTIAL_ACTION_TAKEN

PMIX_EVENT_ACTION_DEFERRED

B.4 MPI Sessions

B.4.1 Use Case Summary

MPI Sessions addresses a number of the limitations of the current MPI programming model. Among the immediate problems MPI Sessions is intended to address are the following:

- MPI cannot be initialized within an MPI process from different application components without a priori knowledge or coordination,

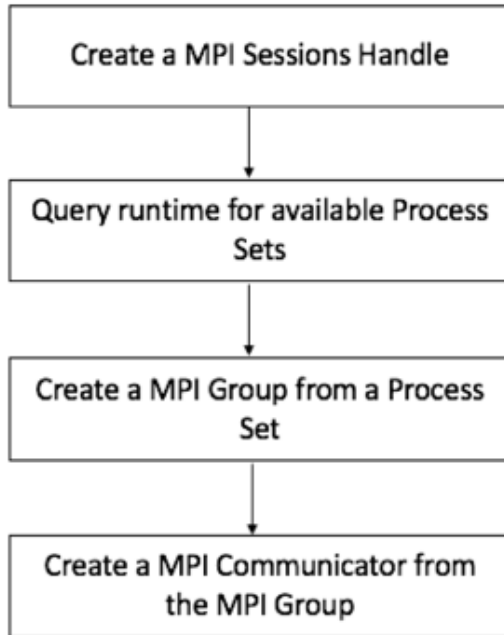


Figure B.4.: MPI Communicator from MPI Session Handle using PMIx

- 1 • MPI cannot be initialized more than once, and MPI cannot be reinitialized after MPI finalize has been
- 2 called.
- 3 • With MPI Sessions, an application no longer needs to explicitly call **MPI_Init** to make use of MPI, but
- 4 rather can use a Session to only initialize MPI resources for specific communication needs.
- 5 • Unless the MPI process explicitly calls **MPI_Init**, there is also no explicit **MPI_COMM_WORLD**
- 6 communicator. Sessions can be created and destroyed multiple times in an MPI process.

7 B.4.2 Use Case Details

8 A PMIx Process Set (PSET) is a user-provided or host environment assigned label associated with a given set
 9 of application processes. Processes can belong to multiple process sets at a time. Definition of a PMIx process
 10 set typically occurs at time of application execution - e.g., on a command line: **prun -n 4 -pset ocean**
 11 **myoceanapp : -n 3 -pset ice myiceapp**

12 PMIx PSETs are used for query functions (**MPI_SESSION_GET_NUM_PSETS**,
 13 **MPI_SESSION_GET_NTH_PSET**) and to create **MPI_GROUP** from a process set name.

14 In OpenMPI's MPI Sessions prototype, PMIx groups are used during creation of **MPI_COMM** from an
 15 **MPI_GROUP**. The PMIx group constructor returns a 64-bit PMIx Group Context Identifier (PGCID) that is
 16 guaranteed to be unique for the duration of an allocation (in the case of a batch managed environment). This

1 B.5 Cross-Version Compatibility

2 B.5.1 Use Case Summary

3 The PMIx interface serves as a conduit between clients (e.g., MPI libraries), tools (e.g., debuggers), and
4 servers (e.g., RMs). As such, it is probable that a process operating in one of these roles (e.g., as a client or
5 tool) is running a different version of the same PMIx implementation than the process with which it is
6 communicating that is operating in a different role (e.g., as a server). For processes running in containers
7 cross-version compatibility is especially important because the container image and the system software levels
8 will naturally evolve and drift apart. As such, there is a need for PMIx implementations to provide
9 cross-version compatibility.

10 The responsibility for providing cross-version compatibility is a feature of a specific PMIx implementation and
11 not necessarily of the PMIx standard. The PMIx standard must strive to enable, and never limit, both the
12 cross-version compatibility in any given PMIx implementation, and the ability for a PMIx consumer to adapt
13 to cross-version differences in capabilities.

14 This use case is focused on cross-version compatibility between different versions of the same PMIx
15 implementation and not between different PMIx implementations.

16 Cross-version compatibility responsibilities are not restricted to PMIx, but a general issue for any library that
17 coordinates across multiple processes. This includes, but not limited to, client/server libraries, and libraries
18 with a user-space and kernel-space component (e.g., high-performance interconnect libraries).

19 B.5.2 Use Case Details

20 There are three scenarios that a PMIx implementation and a PMIx consumer must consider. These scenarios
21 use a PMIx Server and a PMIx Client for clarity, though the scenarios also apply to PMIx Tools.

- 22 1. **PMIx Server version matches PMIx Client version:** No cross-version considerations are necessary since
23 they are running the same version.
- 24 2. **PMIx Server version is older than PMIx Client version:** The implementation must negotiate capabilities
25 during the initial handshake.
26 This scenario is common if the (possibly containerized) PMIx client application is being run on an
27 established system that does not update as frequently as the application requires. Thus the PMIx Server in
28 the RM is locked to an older version of that PMIx implementation.
- 29 3. **PMIx Server version is newer than PMIx Client version:** The implementation must negotiate
30 capabilities during the initial handshake.
31 This scenario is common if the (possibly containerized) PMIx client application is being run after a system
32 software upgrade on the system. Thus the PMIx Server in the RM has been upgraded to a newer version of
33 that PMIx implementation and the client is still linked against the older version.

34 When the two PMIx-enabled processes first connect to each other they need to first check the version of the
35 library that they are each running. This handshake often occurs during initialization (though it could occur on
36 a per-operation basis depending on the specific PMIx implementation), for example during the following
37 operations:

- 38 • PMIx Clients: [PMIx_Init](#)


```
1 typedef pmix_status_t (*pmix_server_client_connected2_fn_t) (  
2     const pmix_proc_t *proc,  
3     void* server_object,  
4     pmix_info_t info[], size_t ninfo,  
5     pmix_op_cbfunc_t cbfunc,  
6     void *cbdata)
```

C

7 *PMIx v2.0*

pmix_server_tool_connection_fn_t

```
8 typedef void (*pmix_server_tool_connection_fn_t) (  
9     pmix_info_t info[], size_t ninfo,  
10    pmix_tool_connection_cbfunc_t cbfunc,  
11    void *cbdata);
```

C

12 *PMIx v4.0*

PMIx_Query_info

```
13 pmix_status_t  
14 PMIx_Query_info(pmix_query_t queries[], size_t nqueries,  
15                pmix_info_t *info[], size_t *ninfo);
```

C

Related Constants

- PMIX_SUCCESS**
- PMIX_ERR_NOT_SUPPORTED**

APPENDIX C

Revision History

1 C.1 Version 1.0: June 12, 2015

2 The PMIx version 1.0 *ad hoc* standard was defined in a set of header files as part of the v1.0.0 release of the
3 OpenPMIx library prior to the creation of the formal PMIx 2.0 standard. Below are a summary listing of the
4 interfaces defined in the 1.0 headers.

5 • Client APIs

- 6 – `PMIx_Init`, `PMIx_Initialized`, `PMIx_Abort`, `PMIx_Finalize`
- 7 – `PMIx_Put`, `PMIx_Commit`,
- 8 – `PMIx_Fence`, `PMIx_Fence_nb`
- 9 – `PMIx_Get`, `PMIx_Get_nb`
- 10 – `PMIx_Publish`, `PMIx_Publish_nb`
- 11 – `PMIx_Lookup`, `PMIx_Lookup_nb`
- 12 – `PMIx_Unpublish`, `PMIx_Unpublish_nb`
- 13 – `PMIx_Spawn`, `PMIx_Spawn_nb`
- 14 – `PMIx_Connect`, `PMIx_Connect_nb`
- 15 – `PMIx_Disconnect`, `PMIx_Disconnect_nb`
- 16 – `PMIx_Resolve_nodes`, `PMIx_Resolve_peers`

17 • Server APIs

- 18 – `PMIx_server_init`, `PMIx_server_finalize`
- 19 – `PMIx_generate_regex`, `PMIx_generate_ppn`
- 20 – `PMIx_server_register_nspace`, `PMIx_server_deregister_nspace`
- 21 – `PMIx_server_register_client`, `PMIx_server_deregister_client`
- 22 – `PMIx_server_setup_fork`, `PMIx_server_dmodex_request`

23 • Common APIs

- 24 – `PMIx_Get_version`, `PMIx_Store_internal`, `PMIx_Error_string`
- 25 – `PMIx_Register_errhandler`, `PMIx_Deregister_errhandler`, `PMIx_Notify_error`

26 The `PMIx_Init` API was subsequently modified in the v1.1.0 release of that library.

1 C.2 Version 2.0: Sept. 2018

2 The following APIs were introduced in v2.0 of the PMIx Standard:

- 3 • Client APIs
 - 4 – `PMIx_Query_info_nb`, `PMIx_Log_nb`
 - 5 – `PMIx_Allocation_request_nb`, `PMIx_Job_control_nb`,
 - 6 `PMIx_Process_monitor_nb`, `PMIx_Heartbeat`
- 7 • Server APIs
 - 8 – `PMIx_server_setup_application`, `PMIx_server_setup_local_support`
- 9 • Tool APIs
 - 10 – `PMIx_tool_init`, `PMIx_tool_finalize`
- 11 • Common APIs
 - 12 – `PMIx_Register_event_handler`, `PMIx_Deregister_event_handler`
 - 13 – `PMIx_Notify_event`
 - 14 – `PMIx_Proc_state_string`, `PMIx_Scope_string`
 - 15 – `PMIx_Persistence_string`, `PMIx_Data_range_string`
 - 16 – `PMIx_Info_directives_string`, `PMIx_Data_type_string`
 - 17 – `PMIx_Alloc_directive_string`
 - 18 – `PMIx_Data_pack`, `PMIx_Data_unpack`, `PMIx_Data_copy`
 - 19 – `PMIx_Data_print`, `PMIx_Data_copy_payload`

20 C.2.1 Removed/Modified APIs

21 The `PMIx_Init` API was modified in v2.0 of the standard from its *ad hoc* v1.0 signature to include passing
22 of a `pmix_info_t` array for flexibility and “future-proofing” of the API. In addition, the
23 `PMIx_Notify_error`, `PMIx_Register_errhandler`, and `PMIx_Deregister_errhandler`
24 APIs were replaced. This pre-dated official adoption of PMIx as a Standard.

25 C.2.2 Deprecated constants

26 The following constants were deprecated in v2.0:

27 `PMIX_MODEX`
28 `PMIX_INFO_ARRAY`

1 C.2.3 Deprecated attributes

2 The following attributes were deprecated in v2.0:

- 3 **PMIX_ERROR_NAME** "pmix.errname" (pmix_status_t)
4 Specific error to be notified
- 5 **PMIX_ERROR_GROUP_COMM** "pmix.errgroup.comm" (bool)
6 Set true to get comm errors notification
- 7 **PMIX_ERROR_GROUP_ABORT** "pmix.errgroup.abort" (bool)
8 Set true to get abort errors notification
- 9 **PMIX_ERROR_GROUP_MIGRATE** "pmix.errgroup.migrate" (bool)
10 Set true to get migrate errors notification
- 11 **PMIX_ERROR_GROUP_RESOURCE** "pmix.errgroup.resource" (bool)
12 Set true to get resource errors notification
- 13 **PMIX_ERROR_GROUP_SPAWN** "pmix.errgroup.spawn" (bool)
14 Set true to get spawn errors notification
- 15 **PMIX_ERROR_GROUP_NODE** "pmix.errgroup.node" (bool)
16 Set true to get node status notification
- 17 **PMIX_ERROR_GROUP_LOCAL** "pmix.errgroup.local" (bool)
18 Set true to get local errors notification
- 19 **PMIX_ERROR_GROUP_GENERAL** "pmix.errgroup.gen" (bool)
20 Set true to get notified of generic errors
- 21 **PMIX_ERROR_HANDLER_ID** "pmix.errhandler.id" (int)
22 Errhandler reference id of notification being reported

23 C.3 Version 2.1: Dec. 2018

24 The v2.1 update includes clarifications and corrections from the v2.0 document, plus addition of examples:

- 25 • Clarify description of **PMIx_Connect** and **PMIx_Disconnect** APIs.
- 26 • Explain that values for the **PMIX_COLLECTIVE_ALGO** are environment-dependent
- 27 • Identify the namespace/rank values required for retrieving attribute-associated information using the
- 28 **PMIx_Get** API
- 29 • Provide definitions for *session*, *job*, *application*, and other terms used throughout the document
- 30 • Clarify definitions of **PMIX_UNIV_SIZE** versus **PMIX_JOB_SIZE**
- 31 • Clarify server module function return values
- 32 • Provide examples of the use of **PMIx_Get** for retrieval of information
- 33 • Clarify the use of **PMIx_Get** versus **PMIx_Query_info_nb**
- 34 • Clarify return values for non-blocking APIs and emphasize that callback functions must not be invoked
- 35 prior to return from the API
- 36 • Provide detailed example for construction of the **PMIx_server_register_nspace** input information
- 37 array
- 38 • Define information levels (e.g., *session* vs *job*) and associated attributes for both storing and retrieving
- 39 values
- 40 • Clarify roles of PMIx server library and host environment for collective operations
- 41 • Clarify definition of **PMIX_UNIV_SIZE**

1 C.4 Version 2.2: Jan 2019

2 The v2.2 update includes the following clarifications and corrections from the v2.1 document:

- 3 • Direct modex upcall function (`pmix_server_dmodex_req_fn_t`) cannot complete atomically as the
- 4 API cannot return the requested information except via the provided callback function
- 5 • Add missing `pmix_data_array_t` definition and support macros
- 6 • Add a rule divider between implementer and host environment required attributes for clarity
- 7 • Add `PMIX_QUERY_QUALIFIERS_CREATE` macro to simplify creation of `pmix_query_t` qualifiers
- 8 • Add `PMIX_APP_INFO_CREATE` macro to simplify creation of `pmix_app_t` directives
- 9 • Add flag and `PMIX_INFO_IS_END` macro for marking and detecting the end of a `pmix_info_t` array
- 10 • Clarify the allowed hierarchical nesting of the `PMIX_SESSION_INFO_ARRAY`,
- 11 `PMIX_JOB_INFO_ARRAY`, and associated attributes

12 C.5 Version 3.0: Dec. 2018

13 The following APIs were introduced in v3.0 of the PMIx Standard:

- 14 • Client APIs
 - 15 – `PMIx_Log`, `PMIx_Job_control`
 - 16 – `PMIx_Allocation_request`, `PMIx_Process_monitor`
 - 17 – `PMIx_Get_credential`, `PMIx_Validate_credential`
- 18 • Server APIs
 - 19 – `PMIx_server_IOF_deliver`
 - 20 – `PMIx_server_collect_inventory`, `PMIx_server_deliver_inventory`
- 21 • Tool APIs
 - 22 – `PMIx_IOF_pull`, `PMIx_IOF_push`, `PMIx_IOF_deregister`
 - 23 – `PMIx_tool_connect_to_server`
- 24 • Common APIs
 - 25 – `PMIx_IOF_channel_string`

26 The document added a chapter on security credentials, a new section for IO forwarding to the Process
27 Management chapter, and a few blocking forms of previously-existing non-blocking APIs. Attributes
28 supporting the new APIs were introduced, as well as additional attributes for a few existing functions.

29 C.5.1 Removed constants

30 The following constants were removed in v3.0:

- 31 `PMIX_MODEX`
- 32 `PMIX_INFO_ARRAY`

1 C.5.2 Deprecated attributes

2 The following attributes were deprecated in v3.0:

3 **PMIX_COLLECTIVE_ALGO_REQD** "pmix.calreqd" (bool)
4 If **true**, indicates that the requested choice of algorithm is mandatory.

5 C.5.3 Removed attributes

6 The following attributes were removed in v3.0:

7 **PMIX_ERROR_NAME** "pmix.errname" (pmix_status_t)
8 Specific error to be notified
9 **PMIX_ERROR_GROUP_COMM** "pmix.errgroup.comm" (bool)
10 Set true to get comm errors notification
11 **PMIX_ERROR_GROUP_ABORT** "pmix.errgroup.abort" (bool)
12 Set true to get abort errors notification
13 **PMIX_ERROR_GROUP_MIGRATE** "pmix.errgroup.migrate" (bool)
14 Set true to get migrate errors notification
15 **PMIX_ERROR_GROUP_RESOURCE** "pmix.errgroup.resource" (bool)
16 Set true to get resource errors notification
17 **PMIX_ERROR_GROUP_SPAWN** "pmix.errgroup.spawn" (bool)
18 Set true to get spawn errors notification
19 **PMIX_ERROR_GROUP_NODE** "pmix.errgroup.node" (bool)
20 Set true to get node status notification
21 **PMIX_ERROR_GROUP_LOCAL** "pmix.errgroup.local" (bool)
22 Set true to get local errors notification
23 **PMIX_ERROR_GROUP_GENERAL** "pmix.errgroup.gen" (bool)
24 Set true to get notified of generic errors
25 **PMIX_ERROR_HANDLER_ID** "pmix.errhandler.id" (int)
26 Errhandler reference id of notification being reported

27 C.6 Version 3.1: Jan. 2019

28 The v3.1 update includes clarifications and corrections from the v3.0 document:

- 29
- 30 • Direct modex upcall function (**pmix_server_dmodex_req_fn_t**) cannot complete atomically as the
31 API cannot return the requested information except via the provided callback function
 - 32 • Fix typo in name of **PMIX_FWD_STDDIAG** attribute
 - 33 • Correctly identify the information retrieval and storage attributes as “new” to v3 of the standard
 - 34 • Add missing **pmix_data_array_t** definition and support macros
 - 35 • Add a rule divider between implementer and host environment required attributes for clarity
 - 36 • Add **PMIX_QUERY_QUALIFIERS_CREATE** macro to simplify creation of **pmix_query_t** qualifiers
 - 37 • Add **PMIX_APP_INFO_CREATE** macro to simplify creation of **pmix_app_t** directives
 - 38 • Add new attributes to specify the level of information being requested where ambiguity may exist (see 6.1)
 - 39 • Add new attributes to assemble information by its level for storage where ambiguity may exist (see 17.2.3.1)
 - Add flag and **PMIX_INFO_IS_END** macro for marking and detecting the end of a **pmix_info_t** array

- Clarify that `PMIX_NUM_SLOTS` is duplicative of (a) `PMIX_UNIV_SIZE` when used at the *session* level and (b) `PMIX_MAX_PROCS` when used at the *job* and *application* levels, but leave it in for backward compatibility.
- Clarify difference between `PMIX_JOB_SIZE` and `PMIX_MAX_PROCS`
- Clarify that `PMIx_server_setup_application` must be called per-*job* instead of per-*application* as the name implies. Unfortunately, this is a historical artifact. Note that both `PMIX_NODE_MAP` and `PMIX_PROC_MAP` must be included as input in the *info* array provided to that function. Further descriptive explanation of the “instant on” procedure will be provided in the next version of the PMIx Standard.
- Clarify how the PMIx server expects data passed to the host by `pmix_server_fence_nb_fn_t` should be aggregated across nodes, and provide a code snippet example

C.7 Version 3.2: Oct. 2020

The v3.2 update includes clarifications and corrections from the v3.1 document:

- Correct an error in the `PMIx_Allocation_request` function signature, and clarify the allocation ID attributes
- Rename the `PMIX_ALLOC_ID` attribute to `PMIX_ALLOC_REQ_ID` to clarify that this is a string the user provides as a means to identify their request to query status
- Add a new `PMIX_ALLOC_ID` attribute that contains the identifier (provided by the host environment) for the resulting allocation which can later be used to reference the allocated resources in, for example, a call to `PMIx_Spawn`
- Update the `PMIx_generate_regex` and `PMIx_generate_ppn` descriptions to clarify that the output from these generator functions may not be a NULL-terminated string, but instead could be a byte array of arbitrary binary content.
- Add a new `PMIX_REGEX` constant that represents a regular expression data type.

C.7.1 Deprecated constants

The following constants were deprecated in v3.2:

<code>PMIX_ERR_DATA_VALUE_NOT_FOUND</code>	Data value not found
<code>PMIX_ERR_HANDSHAKE_FAILED</code>	Connection handshake failed
<code>PMIX_ERR_IN_ERRNO</code>	Error defined in <code>errno</code>
<code>PMIX_ERR_INVALID_ARG</code>	Invalid argument
<code>PMIX_ERR_INVALID_ARGS</code>	Invalid arguments
<code>PMIX_ERR_INVALID_KEY</code>	Invalid key
<code>PMIX_ERR_INVALID_KEY_LENGTH</code>	Invalid key length
<code>PMIX_ERR_INVALID_KEYVALP</code>	Invalid key/value pair
<code>PMIX_ERR_INVALID_LENGTH</code>	Invalid argument length
<code>PMIX_ERR_INVALID_NAMESPACE</code>	Invalid namespace
<code>PMIX_ERR_INVALID_NUM_ARGS</code>	Invalid number of arguments
<code>PMIX_ERR_INVALID_NUM_PARSED</code>	Invalid number parsed
<code>PMIX_ERR_INVALID_SIZE</code>	Invalid size
<code>PMIX_ERR_INVALID_VAL</code>	Invalid value

1	PMIX_ERR_INVALID_VAL_LENGTH	Invalid value length
2	PMIX_ERR_NOT_IMPLEMENTED	Not implemented
3	PMIX_ERR_PACK_MISMATCH	Pack mismatch
4	PMIX_ERR_PROC_ENTRY_NOT_FOUND	Process not found
5	PMIX_ERR_PROC_REQUESTED_ABORT	Process is already requested to abort
6	PMIX_ERR_READY_FOR_HANDSHAKE	Ready for handshake
7	PMIX_ERR_SERVER_FAILED_REQUEST	Failed to connect to the server
8	PMIX_ERR_SERVER_NOT_AVAIL	Server is not available
9	PMIX_ERR_SILENT	Silent error
10	PMIX_GDS_ACTION_COMPLETE	The Global Data Storage (GDS) action has completed
11	PMIX_NOTIFY_ALLOC_COMPLETE	Notify that a requested allocation operation is complete - the result
12		of the request will be included in the <i>info</i> array

13 C.7.2 Deprecated attributes

14 The following attributes were deprecated in v3.2:

15	PMIX_ARCH	" pmix.arch " (uint32_t)
16		Architecture flag.
17	PMIX_COLLECTIVE_ALGO	" pmix.calgo " (char*)
18		Comma-delimited list of algorithms to use for the collective operation. PMIx does not impose any
19		requirements on a host environment's collective algorithms. Thus, the acceptable values for this
20		attribute will be environment-dependent - users are encouraged to check their host environment for
21		supported values.
22	PMIX_DSTPATH	" pmix.dstpath " (char*)
23		Path to shared memory data storage (dstore) files. Deprecated from Standard as being implementation
24		specific.
25	PMIX_HWLOC_HOLE_KIND	" pmix.hwlocholek " (char*)
26		Kind of VM "hole" HWLOC should use for shared memory
27	PMIX_HWLOC_SHARE_TOPO	" pmix.hwlocsh " (bool)
28		Share the HWLOC topology via shared memory
29	PMIX_HWLOC_SHMEM_ADDR	" pmix.hwlocaddr " (size_t)
30		Address of the HWLOC shared memory segment.
31	PMIX_HWLOC_SHMEM_FILE	" pmix.hwlocfile " (char*)
32		Path to the HWLOC shared memory file.
33	PMIX_HWLOC_SHMEM_SIZE	" pmix.hwlocsize " (size_t)
34		Size of the HWLOC shared memory segment.
35	PMIX_HWLOC_XML_V1	" pmix.hwlocxml1 " (char*)
36		XML representation of local topology using HWLOC's v1.x format.
37	PMIX_HWLOC_XML_V2	" pmix.hwlocxml2 " (char*)
38		XML representation of local topology using HWLOC's v2.x format.
39	PMIX_LOCAL_TOPO	" pmix.ltopo " (char*)
40		XML representation of local node topology.
41	PMIX_MAPPER	" pmix.mapper " (char*)

1 Mapping mechanism to use for placing spawned processes - when accessed using [PMIx_Get](#), use the
2 [PMIX_RANK_WILDCARD](#) value for the rank to discover the mapping mechanism used for the
3 provided namespace.

4 [PMIX_MAP_BLOB](#) "pmix.mblob" ([pmix_byte_object_t](#))

5 Packed blob of process location.

6 [PMIX_NON_PMI](#) "pmix.nonpmi" ([bool](#))

7 Spawned processes will not call [PMIx_Init](#).

8 [PMIX_PROC_BLOB](#) "pmix.pblob" ([pmix_byte_object_t](#))

9 Packed blob of process data.

10 [PMIX_PROC_URI](#) "pmix.puri" ([char*](#))

11 URI containing contact information for the specified process.

12 [PMIX_TOPOLOGY_FILE](#) "pmix.topo.file" ([char*](#))

13 Full path to file containing XML topology description

14 [PMIX_TOPOLOGY_SIGNATURE](#) "pmix.toposig" ([char*](#))

15 Topology signature string.

16 [PMIX_TOPOLOGY_XML](#) "pmix.topo.xml" ([char*](#))

17 XML-based description of topology

18 C.8 Version 4.0: Dec. 2020

19 NOTE: The PMIx Standard document has undergone significant reorganization in an effort to become more
20 user-friendly. Highlights include:

- 21 • Moving all added, deprecated, and removed items to this revision log section to make them more visible
- 22 • Co-locating constants and attribute definitions with the primary API that uses them - citations and
23 hyperlinks are retained elsewhere
- 24 • Splitting the Key-Value Management chapter into separate chapters on the use of reserved keys,
25 non-reserved keys, and non-process-related key-value data exchange
- 26 • Creating a new chapter on synchronization and data access methods
- 27 • Removing references to specific implementations of PMIx and to implementation-specific features and/or
28 behaviors

29 In addition to the reorganization, the following changes were introduced in v4.0 of the PMIx Standard:

- 30 • Clarified that the [PMIx_Fence_nb](#) operation can immediately return [PMIX_OPERATION_SUCCEEDED](#)
31 in lieu of passing the request to a PMIx server if only the calling process is involved in the operation
- 32 • Added the [PMIx_Register_attributes](#) API by which a host environment can register the attributes
33 it supports for each server-to-host operation
- 34 • Added the ability to query supported attributes from the PMIx tool, client and server libraries, as well as the
35 host environment via the new [pmix_regattr_t](#) structure. Both human-readable and machine-parsable
36 output is supported. New attributes to support this operation include:
 - 37 – [PMIX_CLIENT_ATTRIBUTES](#), [PMIX_SERVER_ATTRIBUTES](#), [PMIX_TOOL_ATTRIBUTES](#), and
38 [PMIX_HOST_ATTRIBUTES](#) to identify which library supports the attribute; and
 - 39 – [PMIX_MAX_VALUE](#), [PMIX_MIN_VALUE](#), and [PMIX_ENUM_VALUE](#) to provide machine-parsable
40 description of accepted values
- 41 • Add [PMIX_APP_WILDCARD](#) to reference all applications within a given job

- 1 • Fix signature of blocking APIs `PMIx_Allocation_request`, `PMIx_Job_control`,
- 2 `PMIx_Process_monitor`, `PMIx_Get_credential`, and `PMIx_Validate_credential` to
- 3 allow return of results
- 4 • Update description to provide an option for blocking behavior of the
- 5 `PMIx_Register_event_handler`, `PMIx_Deregister_event_handler`,
- 6 `PMIx_Notify_event`, `PMIx_IOF_pull`, `PMIx_IOF_deregister`, and `PMIx_IOF_push` APIs.
- 7 The need for blocking forms of these functions was not initially anticipated but has emerged over time. For
- 8 these functions, the return value is sufficient to provide the caller with information otherwise returned via
- 9 callback. Thus, use of a **NULL** value as the callback function parameter was deemed a minimal disruption
- 10 method for providing the desired capability
- 11 • Added a chapter on fabric support that includes new APIs, datatypes, and attributes
- 12 • Added a chapter on process sets and groups that includes new APIs and attributes
- 13 • Added APIs and a new datatypes to support generation and parsing of PMIx locality and cpuset strings
- 14 • Added a new chapter on tools that provides deeper explanation on their operation and collecting all
- 15 tool-relevant definitions into one location. Also introduced two new APIs and removed restriction that
- 16 limited tools to being connected to only one server at a time.
- 17 • Extended behavior of `PMIx_server_init` to scalably expose the topology description to the local
- 18 clients. This includes creating any required shared memory backing stores and/or XML representations,
- 19 plus ensuring that all necessary key-value pairs for clients to access the description are included in the
- 20 job-level information provided to each client.
- 21 • Added a new API by which the host can manually progress the PMIx library in lieu of the library's own
- 22 progress thread. s

23 The above changes included introduction of the following APIs and data types:

- 24 • Client APIs
 - 25 – `PMIx_Group_construct`, `PMIx_Group_construct_nb`
 - 26 – `PMIx_Group_destruct`, `PMIx_Group_destruct_nb`
 - 27 – `PMIx_Group_invite`, `PMIx_Group_invite_nb`
 - 28 – `PMIx_Group_join`, `PMIx_Group_join_nb`
 - 29 – `PMIx_Group_leave`, `PMIx_Group_leave_nb`
 - 30 – `PMIx_Get_relative_locality`, `PMIx_Load_topology`
 - 31 – `PMIx_Parse_cpuset_string`, `PMIx_Get_cpuset`
 - 32 – `PMIx_Link_state_string`, `PMIx_Job_state_string`
 - 33 – `PMIx_Device_type_string`
 - 34 – `PMIx_Fabric_register`, `PMIx_Fabric_register_nb`
 - 35 – `PMIx_Fabric_update`, `PMIx_Fabric_update_nb`
 - 36 – `PMIx_Fabric_deregister`, `PMIx_Fabric_deregister_nb`
 - 37 – `PMIx_Compute_distances`, `PMIx_Compute_distances_nb`
 - 38 – `PMIx_Get_attribute_string`, `PMIx_Get_attribute_name`
 - 39 – `PMIx_Progress`
- 40 • Server APIs
 - 41 – `PMIx_server_generate_locality_string`
 - 42 – `PMIx_Register_attributes`

```
1      - PMIx_server_define_process_set, PMIx_server_delete_process_set
2      - pmix_server_grp_fn_t, pmix_server_fabric_fn_t
3      - pmix_server_client_connected2_fn_t
4      - PMIx_server_generate_cpuset_string
5      - PMIx_server_register_resources, PMIx_server_deregister_resources
6
7      • Tool APIs
8
9      - PMIx_tool_disconnect
10     - PMIx_tool_set_server
11     - PMIx_tool_attach_to_server
12     - PMIx_tool_get_servers
13
14     • Data types
15
16     - pmix_regattr_t
17     - pmix_cpuset_t
18     - pmix_topology_t
19     - pmix_locality_t
20     - pmix_bind_envelope_t
21     - pmix_group_opt_t
22     - pmix_group_operation_t
23     - pmix_fabric_t
24     - pmix_device_distance_t
25     - pmix_coord_t
26     - pmix_coord_view_t
27     - pmix_geometry_t
28     - pmix_link_state_t
29     - pmix_job_state_t
30     - pmix_device_type_t
31
32     • Callback functions
33
34     - pmix_device_dist_cbfunc_t
```

29 C.8.1 Added Constants

30 General error constants

```
31 PMIX_ERR_EXISTS_OUTSIDE_SCOPE
32 PMIX_ERR_PARAM_VALUE_NOT_SUPPORTED
33 PMIX_ERR_EMPTY
34
```

1 **Data type constants**
2 PMIX_COORD
3 PMIX_REGATTR
4 PMIX_REGEX
5 PMIX_JOB_STATE
6 PMIX_LINK_STATE
7 PMIX_PROC_CPUSSET
8 PMIX_GEOMETRY
9 PMIX_DEVICE_DIST
10 PMIX_ENDPOINT
11 PMIX_TOPO
12 PMIX_DEVTYPE
13 PMIX_LOCTYPE
14 PMIX_DATA_TYPE_MAX
15 PMIX_COMPRESSED_BYTE_OBJECT
16

17 **Info directives**
18 PMIX_INFO_REQD_PROCESSED
19

20 **Server constants**
21 PMIX_ERR_REPEAT_ATTR_REGISTRATION
22

23 **Job-Mgmt constants**
24 PMIX_ERR_CONFLICTING_CLEANUP_DIRECTIVES
25

26 **Publish constants**
27 PMIX_ERR_DUPLICATE_KEY
28

29 **Tool constants**
30 PMIX_LAUNCHER_READY
31 PMIX_ERR_IOF_FAILURE
32 PMIX_ERR_IOF_COMPLETE
33 PMIX_EVENT_JOB_START
34 PMIX_LAUNCH_COMPLETE
35 PMIX_EVENT_JOB_END
36 PMIX_EVENT_SESSION_START
37 PMIX_EVENT_SESSION_END
38 PMIX_ERR_PROC_TERM_WO_SYNC
39 PMIX_ERR_JOB_CANCELED
40 PMIX_ERR_JOB_ABORTED
41 PMIX_ERR_JOB_KILLED_BY_CMD
42 PMIX_ERR_JOB_ABORTED_BY_SIG
43 PMIX_ERR_JOB_TERM_WO_SYNC

1 PMIX_ERR_JOB_SENSOR_BOUND_EXCEEDED
2 PMIX_ERR_JOB_NON_ZERO_TERM
3 PMIX_ERR_JOB_ABORTED_BY_SYS_EVENT
4 PMIX_DEBUG_WAITING_FOR_NOTIFY
5 PMIX_DEBUGGER_RELEASE
6

Fabric constants

7
8 PMIX_FABRIC_UPDATE_PENDING
9 PMIX_FABRIC_UPDATED
10 PMIX_FABRIC_UPDATE_ENDPOINTS
11 PMIX_COORD_VIEW_UNDEF
12 PMIX_COORD_LOGICAL_VIEW
13 PMIX_COORD_PHYSICAL_VIEW
14 PMIX_LINK_STATE_UNKNOWN
15 PMIX_LINK_DOWN
16 PMIX_LINK_UP
17 PMIX_FABRIC_REQUEST_INFO
18 PMIX_FABRIC_UPDATE_INFO
19

Sets-Groups constants

20
21 PMIX_PROCESS_SET_DEFINE
22 PMIX_PROCESS_SET_DELETE
23 PMIX_GROUP_INVITED
24 PMIX_GROUP_LEFT
25 PMIX_GROUP_MEMBER_FAILED
26 PMIX_GROUP_INVITE_ACCEPTED
27 PMIX_GROUP_INVITE_DECLINED
28 PMIX_GROUP_INVITE_FAILED
29 PMIX_GROUP_MEMBERSHIP_UPDATE
30 PMIX_GROUP_CONSTRUCT_ABORT
31 PMIX_GROUP_CONSTRUCT_COMPLETE
32 PMIX_GROUP_LEADER_FAILED
33 PMIX_GROUP_LEADER_SELECTED
34 PMIX_GROUP_CONTEXT_ID_ASSIGNED
35

Process-Mgmt constants

36
37 PMIX_ERR_JOB_ALLOC_FAILED
38 PMIX_ERR_JOB_APP_NOT_EXECUTABLE
39 PMIX_ERR_JOB_NO_EXE_SPECIFIED
40 PMIX_ERR_JOB_FAILED_TO_MAP
41 PMIX_ERR_JOB_FAILED_TO_LAUNCH
42 PMIX_LOCALITY_UNKNOWN
43 PMIX_LOCALITY_NONLOCAL
44 PMIX_LOCALITY_SHARE_HWTHREAD
45 PMIX_LOCALITY_SHARE_CORE

1 `PMIX_LOCALITY_SHARE_L1CACHE`
2 `PMIX_LOCALITY_SHARE_L2CACHE`
3 `PMIX_LOCALITY_SHARE_L3CACHE`
4 `PMIX_LOCALITY_SHARE_PACKAGE`
5 `PMIX_LOCALITY_SHARE_NUMA`
6 `PMIX_LOCALITY_SHARE_NODE`
7

8 **Events**

9 `PMIX_EVENT_SYS_BASE`
10 `PMIX_EVENT_NODE_DOWN`
11 `PMIX_EVENT_NODE_OFFLINE`
12 `PMIX_EVENT_SYS_OTHER`
13

14 **C.8.2 Added Attributes**

15 **Sync-Access attributes**

16 `PMIX_COLLECT_GENERATED_JOB_INFO` "pmix.collect.gen" (bool)

17 Collect all job-level information (i.e., reserved keys) that was locally generated by PMIx servers. Some
18 job-level information (e.g., distance between processes and fabric devices) is best determined on a
19 distributed basis as it primarily pertains to local processes. Should remote processes need to access the
20 information, it can either be obtained collectively using the `PMIx_Fence` operation with this
21 directive, or can be retrieved one peer at a time using `PMIx_Get` without first having performed the
22 job-wide collection.

23 `PMIX_ALL_CLONES_PARTICIPATE` "pmix.clone.part" (bool)

24 All *clones* of the calling process must participate in the collective operation.

25 `PMIX_GET_POINTER_VALUES` "pmix.get.pntrs" (bool)

26 Request that any pointers in the returned value point directly to values in the key-value store. The user
27 *must not* release any returned data pointers.

28 `PMIX_GET_STATIC_VALUES` "pmix.get.static" (bool)

29 Request that the data be returned in the provided storage location. The caller is responsible for
30 destructing the `pmix_value_t` using the `PMIX_VALUE_DESTRUCT` macro when done.

31 `PMIX_GET_REFRESH_CACHE` "pmix.get.refresh" (bool)

32 When retrieving data for a remote process, refresh the existing local data cache for the process in case
33 new values have been put and committed by the process since the last refresh. Local process
34 information is assumed to be automatically updated upon posting by the process. A `NULL` key will
35 cause all values associated with the process to be refreshed - otherwise, only the indicated key will be
36 updated. A process rank of `PMIX_RANK_WILDCARD` can be used to update job-related information in
37 dynamic environments. The user is responsible for subsequently updating refreshed values they may
38 have cached in their own local memory.

39 `PMIX_QUERY_RESULTS` "pmix.qry.res" (`pmix_data_array_t`)

1 Contains an array of query results for a given `pmix_query_t` passed to the `PMIx_Query_info`
2 APIs. If qualifiers were included in the query, then the first element of the array shall be the
3 `PMIX_QUERY_QUALIFIERS` attribute containing those qualifiers. Each of the remaining elements
4 of the array is a `pmix_info_t` containing the query key and the corresponding value returned by the
5 query. This attribute is solely for reporting purposes and cannot be used in `PMIx_Get` or other query
6 operations.

7 `PMIX_QUERY_QUALIFIERS` "pmix.qry.qual" (`pmix_data_array_t`)

8 Contains an array of qualifiers that were included in the query that produced the provided results. This
9 attribute is solely for reporting purposes and cannot be used in `PMIx_Get` or other query operations.

10 `PMIX_QUERY_SUPPORTED_KEYS` "pmix.qry.keys" (`char*`)

11 Returns comma-delimited list of keys supported by the query function. NO QUALIFIERS.

12 `PMIX_QUERY_SUPPORTED_QUALIFIERS` "pmix.qry.qual" (`char*`)

13 Return comma-delimited list of qualifiers supported by a query on the provided key, instead of actually
14 performing the query on the key. NO QUALIFIERS.

15 `PMIX_QUERY_NAMESPACE_INFO` "pmix.qry.nsinfo" (`pmix_data_array_t*`)

16 Return an array of active namespace information - each element will itself contain an array including
17 the namespace plus the command line of the application executing within it. OPTIONAL
18 QUALIFIERS: `PMIX_NAMESPACE` of specific namespace whose info is being requested.

19 `PMIX_QUERY_ATTRIBUTE_SUPPORT` "pmix.qry.attrs" (`bool`)

20 Query list of supported attributes for specified APIs. REQUIRED QUALIFIERS: one or more of
21 `PMIX_CLIENT_FUNCTIONS`, `PMIX_SERVER_FUNCTIONS`, `PMIX_TOOL_FUNCTIONS`, and
22 `PMIX_HOST_FUNCTIONS`.

23 `PMIX_QUERY_AVAIL_SERVERS` "pmix.qry.asrvrs" (`pmix_data_array_t*`)

24 Return an array of `pmix_info_t`, each element itself containing a `PMIX_SERVER_INFO_ARRAY`
25 entry holding all available data for a server on this node to which the caller might be able to connect.

26 `PMIX_SERVER_INFO_ARRAY` "pmix.srv.arr" (`pmix_data_array_t`)

27 Array of `pmix_info_t` about a given server, starting with its `PMIX_NAMESPACE` and including at least
28 one of the rendezvous-required pieces of information.

29 `PMIX_CLIENT_FUNCTIONS` "pmix.client.fns" (`bool`)

30 Request a list of functions supported by the PMIx client library.

31 `PMIX_CLIENT_ATTRIBUTES` "pmix.client.attrs" (`bool`)

32 Request attributes supported by the PMIx client library.

33 `PMIX_SERVER_FUNCTIONS` "pmix.srvr.fns" (`bool`)

34 Request a list of functions supported by the PMIx server library.

35 `PMIX_SERVER_ATTRIBUTES` "pmix.srvr.attrs" (`bool`)

36 Request attributes supported by the PMIx server library.

37 `PMIX_HOST_FUNCTIONS` "pmix.srvr.fns" (`bool`)

38 Request a list of functions supported by the host environment.

39 `PMIX_HOST_ATTRIBUTES` "pmix.host.attrs" (`bool`)

1 Request attributes supported by the host environment.

2 **PMIX_TOOL_FUNCTIONS** "pmix.tool.fns" (bool)

3 Request a list of functions supported by the PMIx tool library.

4 **PMIX_TOOL_ATTRIBUTES** "pmix.setup.env" (bool)

5 Request attributes supported by the PMIx tool library functions.

6 Server attributes

7 **PMIX_TOPOLOGY2** "pmix.topo2" (pmix_topology_t)

8 Provide a pointer to an implementation-specific description of the local node topology.

9 **PMIX_SERVER_SHARE_TOPOLOGY** "pmix.srvr.share" (bool)

10 The PMIx server is to share its copy of the local node topology (whether given to it or self-discovered)
11 with any clients.

12 **PMIX_SERVER_SESSION_SUPPORT** "pmix.srvr.sess" (bool)

13 The host RM wants to declare itself as being the local session server for PMIx connection requests.

14 **PMIX_SERVER_START_TIME** "pmix.srvr.strtime" (char*)

15 Time when the server started - i.e., when the server created it's rendezvous file (given in ctime string
16 format).

17 **PMIX_SERVER_SCHEDULER** "pmix.srv.sched" (bool)

18 Server is supporting system scheduler and desires access to appropriate WLM-supporting features.
19 Indicates that the library is to be initialized for scheduler support.

20 **PMIX_JOB_INFO_ARRAY** "pmix.job.arr" (pmix_data_array_t)

21 Provide an array of **pmix_info_t** containing job-realm information. The **PMIX_SESSION_ID**
22 attribute of the *session* containing the *job* is required to be included in the array whenever the PMIx
23 server library may host multiple sessions (e.g., when executing with a host RM daemon). As
24 information is registered one job (aka namespace) at a time via the

25 **PMIx_server_register_namespace** API, there is no requirement that the array contain either the
26 **PMIX_NAMESPACE** or **PMIX_JOBID** attributes when used in that context (though either or both of them
27 may be included). At least one of the job identifiers must be provided in all other contexts where the
28 job being referenced is ambiguous.

29 **PMIX_APP_INFO_ARRAY** "pmix.app.arr" (pmix_data_array_t)

30 Provide an array of **pmix_info_t** containing application-realm information. The **PMIX_NAMESPACE**
31 or **PMIX_JOBID** attributes of the *job* containing the application, plus its **PMIX_APPNUM** attribute,
32 must be included in the array when the array is *not* included as part of a call to
33 **PMIx_server_register_namespace** - i.e., when the job containing the application is ambiguous.
34 The job identification is otherwise optional.

35 **PMIX_PROC_INFO_ARRAY** "pmix.pdata" (pmix_data_array_t)

36 Provide an array of **pmix_info_t** containing process-realm information. The **PMIX_RANK** and
37 **PMIX_NAMESPACE** attributes, or the **PMIX_PROCID** attribute, are required to be included in the array
38 when the array is not included as part of a call to **PMIx_server_register_namespace** - i.e., when
39 the job containing the process is ambiguous. All three may be included if desired. When the array is
40 included in some broader structure that identifies the job, then only the **PMIX_RANK** or the
41 **PMIX_PROCID** attribute must be included (the others are optional).

1 **PMIX_NODE_INFO_ARRAY** "pmix.node.arr" (**pmix_data_array_t**)
2 Provide an array of **pmix_info_t** containing node-realm information. At a minimum, either the
3 **PMIX_NODEID** or **PMIX_HOSTNAME** attribute is required to be included in the array, though both
4 may be included.

5 **PMIX_MAX_VALUE** "pmix.descr.maxval" (**varies**)
6 Used in **pmix_regattr_t** to describe the maximum valid value for the associated attribute.

7 **PMIX_MIN_VALUE** "pmix.descr.minval" (**varies**)
8 Used in **pmix_regattr_t** to describe the minimum valid value for the associated attribute.

9 **PMIX_ENUM_VALUE** "pmix.descr.enum" (**char***)
10 Used in **pmix_regattr_t** to describe accepted values for the associated attribute. Numerical values
11 shall be presented in a form convertible to the attribute's declared data type. Named values (i.e., values
12 defined by constant names via a typical C-language enum declaration) must be provided as their
13 numerical equivalent.

14 **PMIX_HOMOGENEOUS_SYSTEM** "pmix.homo" (**bool**)
15 The nodes comprising the session are homogeneous - i.e., they each contain the same number of
16 identical packages, fabric interfaces, GPUs, and other devices.

17 **PMIX_REQUIRED_KEY** "pmix.req.key" (**char***)
18 Identifies a key that must be included in the requested information. If the specified key is not already
19 available, then the PMIx servers are required to delay response to the dmodex request until either the
20 key becomes available or the request times out.

21 Job-Mgmt attributes

22 **PMIX_ALLOC_ID** "pmix.alloc.id" (**char***)
23 A string identifier (provided by the host environment) for the resulting allocation which can later be
24 used to reference the allocated resources in, for example, a call to **PMIx_Spawn**.

25 **PMIX_ALLOC_QUEUE** "pmix.alloc.queue" (**char***)
26 Name of the WLM queue to which the allocation request is to be directed, or the queue being
27 referenced in a query.

28 Publish attributes

29 **PMIX_ACCESS_PERMISSIONS** "pmix.aperms" (**pmix_data_array_t**)
30 Define access permissions for the published data. The value shall contain an array of **pmix_info_t**
31 structs containing the specified permissions.

32 **PMIX_ACCESS_USERIDS** "pmix.auids" (**pmix_data_array_t**)
33 Array of effective UIDs that are allowed to access the published data.

34 **PMIX_ACCESS_GRPIDS** "pmix.agids" (**pmix_data_array_t**)
35 Array of effective GIDs that are allowed to access the published data.

Reserved keys

PMIX_NUM_ALLOCATED_NODES "pmix.num.anodes" (uint32_t)

Number of nodes in the specified realm regardless of whether or not they currently host processes. Defaults to the *job* realm.

PMIX_NUM_NODES "pmix.num.nodes" (uint32_t)

Number of nodes currently hosting processes in the specified realm. Defaults to the *job* realm.

PMIX_CMD_LINE "pmix.cmd.line" (char*)

Command line used to execute the specified job (e.g., "mpirun -n 2 -map-by foo ./myapp : -n 4 ./myapp2"). If the job was created by a call to **PMIx_Spawn**, the string is an in-order concatenation of the values of **PMIX_APP_ARGV** for each application in the job using the character ':' as a separator.

PMIX_APP_ARGV "pmix.app.argv" (char*)

Consolidated argv passed to the spawn command for the given application (e.g., "./myapp arg1 arg2 arg3").

PMIX_PACKAGE_RANK "pmix.pkgrank" (uint16_t)

Rank of the specified process on the *package* where this process resides - refers to the numerical location (starting from zero) of the process on its package when counting only those processes from the same job that share the package, ordered by their overall rank within that job. Note that processes that are not bound to PUs within a single specific package cannot have a package rank.

PMIX_REINCARNATION "pmix.reinc" (uint32_t)

Number of times this process has been re-instantiated - i.e, a value of zero indicates that the process has never been restarted.

PMIX_HOSTNAME_ALIASES "pmix.alias" (char*)

Comma-delimited list of names by which the target node is known.

PMIX_HOSTNAME_KEEP_FQDN "pmix.fqdn" (bool)

FQDNs are being retained by the PMIx library.

PMIX_CPUSSET_BITMAP "pmix.bitmap" (pmix_cpuset_t*)

Bitmap applied to the process upon launch.

PMIX_EXTERNAL_PROGRESS "pmix.evext" (bool)

The host shall progress the PMIx library via calls to **PMIx_Progress**

PMIX_NODE_MAP_RAW "pmix.nmap.raw" (char*)

Comma-delimited list of nodes containing procs within the specified realm. Defaults to the *job* realm.

PMIX_PROC_MAP_RAW "pmix.pmap.raw" (char*)

Semi-colon delimited list of strings, each string containing a comma-delimited list of ranks on the corresponding node within the specified realm. Defaults to the *job* realm.

Tool attributes

PMIX_TOOL_CONNECT_OPTIONAL "pmix.tool.conopt" (bool)

The tool shall connect to a server if available, but otherwise continue to operate unconnected.

PMIX_TOOL_ATTACHMENT_FILE "pmix.tool.attach" (char*)

Pathname of file containing connection information to be used for attaching to a specific server.

1 **PMIX_LAUNCHER_RENDEZVOUS_FILE** "pmix.tool.lncrnd" (char*)
2 Pathname of file where the launcher is to store its connection information so that the spawning tool can
3 connect to it.

4 **PMIX_PRIMARY_SERVER** "pmix.pri.srvr" (bool)
5 The server to which the tool is connecting shall be designated the *primary* server once connection has
6 been accomplished.

7 **PMIX_NOHUP** "pmix.nohup" (bool)
8 Any processes started on behalf of the calling tool (or the specified namespace, if such specification is
9 included in the list of attributes) should continue after the tool disconnects from its server.

10 **PMIX_LAUNCHER_DAEMON** "pmix.lnch.dmn" (char*)
11 Path to executable that is to be used as the backend daemon for the launcher. This replaces the
12 launcher's own daemon with the specified executable. Note that the user is therefore responsible for
13 ensuring compatibility of the specified executable and the host launcher.

14 **PMIX_FORKEXEC_AGENT** "pmix.frkex.agnt" (char*)
15 Path to executable that the launcher's backend daemons are to fork/exec in place of the actual
16 application processes. The fork/exec agent shall connect back (as a PMIx tool) to the launcher's
17 daemon to receive its spawn instructions, and is responsible for starting the actual application process it
18 replaced. See Section 18.4.3 for details.

19 **PMIX_EXEC_AGENT** "pmix.exec.agnt" (char*)
20 Path to executable that the launcher's backend daemons are to fork/exec in place of the actual
21 application processes. The launcher's daemon shall pass the full command line of the application on
22 the command line of the exec agent, which shall not connect back to the launcher's daemon. The exec
23 agent is responsible for exec'ing the specified application process in its own place. See Section 18.4.3
24 for details.

25 **PMIX_IOF_PUSH_STDIN** "pmix.iof.stdin" (bool)
26 Requests that the PMIx library collect the **stdin** of the requester and forward it to the processes
27 specified in the **PMIx_IOF_push** call. All collected data is sent to the same targets until **stdin** is
28 closed, or a subsequent call to **PMIx_IOF_push** is made that includes the **PMIX_IOF_COMPLETE**
29 attribute indicating that forwarding of **stdin** is to be terminated.

30 **PMIX_IOF_COPY** "pmix.iof.cpy" (bool)
31 Requests that the host environment deliver a copy of the specified output stream(s) to the tool, letting
32 the stream(s) continue to also be delivered to the default location. This allows the tool to tap into the
33 output stream(s) without redirecting it from its current final destination.

34 **PMIX_IOF_REDIRECT** "pmix.iof.redir" (bool)
35 Requests that the host environment intercept the specified output stream(s) and deliver it to the
36 requesting tool instead of its current final destination. This might be used, for example, during a
37 debugging procedure to avoid injection of debugger-related output into the application's results file.
38 The original output stream(s) destination is restored upon termination of the tool.

39 **PMIX_DEBUG_TARGET** "pmix.dbg.tgt" (pmix_proc_t*)
40 Identifier of process(es) to be debugged - a rank of **PMIX_RANK_WILDCARD** indicates that all
41 processes in the specified namespace are to be included.

1 **PMIX_DEBUG_DAEMONS_PER_PROC** "pmix.dbg.dpproc" (uint16_t)
2 Number of debugger daemons to be spawned per application process. The launcher is to pass the
3 identifier of the namespace to be debugged by including the **PMIX_DEBUG_TARGET** attribute in the
4 daemon's job-level information. The debugger daemons spawned on a given node are responsible for
5 self-determining their specific target process(es) - e.g., by referencing their own **PMIX_LOCAL_RANK**
6 in the daemon debugger job versus the corresponding **PMIX_LOCAL_RANK** of the target processes on
7 the node.

8 **PMIX_DEBUG_DAEMONS_PER_NODE** "pmix.dbg.dpnd" (uint16_t)
9 Number of debugger daemons to be spawned on each node where the target job is executing. The
10 launcher is to pass the identifier of the namespace to be debugged by including the
11 **PMIX_DEBUG_TARGET** attribute in the daemon's job-level information. The debugger daemons
12 spawned on a given node are responsible for self-determining their specific target process(es) - e.g., by
13 referencing their own **PMIX_LOCAL_RANK** in the daemon debugger job versus the corresponding
14 **PMIX_LOCAL_RANK** of the target processes on the node.

15 **PMIX_WAIT_FOR_CONNECTION** "pmix.wait.conn" (bool)
16 Wait until the specified process has connected to the requesting tool or server, or the operation times
17 out (if the **PMIX_TIMEOUT** directive is included in the request).

18 **PMIX_LAUNCH_DIRECTIVES** "pmix.lnch.dirs" (pmix_data_array_t*)
19 Array of **pmix_info_t** containing directives for the launcher - a convenience attribute for retrieving
20 all directives with a single call to **PMIx_Get**.

21 **Fabric attributes**

22 **PMIX_SERVER_SCHEDULER** "pmix.srv.sched" (bool)
23 Server is supporting system scheduler and desires access to appropriate WLM-supporting features.
24 Indicates that the library is to be initialized for scheduler support.

25 **PMIX_FABRIC_COST_MATRIX** "pmix.fab.cm" (pointer)
26 Pointer to a two-dimensional square array of point-to-point relative communication costs expressed as
27 **uint16_t** values.

28 **PMIX_FABRIC_GROUPS** "pmix.fab.grps" (string)
29 A string delineating the group membership of nodes in the overall system, where each fabric group
30 consists of the group number followed by a colon and a comma-delimited list of nodes in that group,
31 with the groups delimited by semi-colons (e.g., **0:node000,node002,node004,node006;**
32 **1:node001,node003,node005,node007**)

33 **PMIX_FABRIC_VENDOR** "pmix.fab.vndr" (string)
34 Name of the vendor (e.g., Amazon, Mellanox, HPE, Intel) for the specified fabric.

35 **PMIX_FABRIC_IDENTIFIER** "pmix.fab.id" (string)
36 An identifier for the specified fabric (e.g., MgmtEthernet, Slingshot-11, OmniPath-1).

37 **PMIX_FABRIC_INDEX** "pmix.fab.idx" (size_t)
38 The index of the fabric as returned in **pmix_fabric_t**.

39 **PMIX_FABRIC_NUM_DEVICES** "pmix.fab.nverts" (size_t)
40 Total number of fabric devices in the overall system - corresponds to the number of rows or columns in
41 the cost matrix.

1 **PMIX_FABRIC_COORDINATES** "pmix.fab.coords" (pmix_data_array_t)
2 Array of pmix_geometry_t fabric coordinates for devices on the specified node. The array will
3 contain the coordinates of all devices on the node, including values for all supported coordinate views.
4 The information for devices on the local node shall be provided if the node is not specified in the
5 request.

6 **PMIX_FABRIC_DIMS** "pmix.fab.dims" (uint32_t)
7 Number of dimensions in the specified fabric plane/view. If no plane is specified in a request, then the
8 dimensions of all planes in the overall system will be returned as a pmix_data_array_t
9 containing an array of uint32_t values. Default is to provide dimensions in *logical* view.

10 **PMIX_FABRIC_ENDPT** "pmix.fab.endpt" (pmix_data_array_t)
11 Fabric endpoints for a specified process. As multiple endpoints may be assigned to a given process
12 (e.g., in the case where multiple devices are associated with a package to which the process is bound),
13 the returned values will be provided in a pmix_data_array_t of pmix_endpoint_t elements.
14

15 **PMIX_FABRIC_SHAPE** "pmix.fab.shape" (pmix_data_array_t*)
16 The size of each dimension in the specified fabric plane/view, returned in a pmix_data_array_t
17 containing an array of uint32_t values. The size is defined as the number of elements present in
18 that dimension - e.g., the number of devices in one dimension of a physical view of a fabric plane. If no
19 plane is specified, then the shape of each plane in the overall system will be returned in a
20 pmix_data_array_t array where each element is itself a two-element array containing the
21 PMIX_FABRIC_PLANE followed by that plane's fabric shape. Default is to provide the shape in
22 *logical* view.

23 **PMIX_FABRIC_SHAPE_STRING** "pmix.fab.shapestr" (string)
24 Network shape expressed as a string (e.g., "10x12x2"). If no plane is specified, then the shape of
25 each plane in the overall system will be returned in a pmix_data_array_t array where each
26 element is itself a two-element array containing the PMIX_FABRIC_PLANE followed by that plane's
27 fabric shape string. Default is to provide the shape in *logical* view.

28 **PMIX_SWITCH_PEERS** "pmix.speers" (pmix_data_array_t)
29 Peer ranks that share the same switch as the process specified in the call to PMIx_Get. Returns a
30 pmix_data_array_t array of pmix_info_t results, each element containing the
31 PMIX_SWITCH_PEERS key with a three-element pmix_data_array_t array of pmix_info_t
32 containing the PMIX_DEVICE_ID of the local fabric device, the PMIX_FABRIC_SWITCH
33 identifying the switch to which it is connected, and a comma-delimited string of peer ranks sharing the
34 switch to which that device is connected.

35 **PMIX_FABRIC_PLANE** "pmix.fab.plane" (string)
36 ID string of a fabric plane (e.g., CIDR for Ethernet). When used as a modifier in a request for
37 information, specifies the plane whose information is to be returned. When used directly as a key in a
38 request, returns a pmix_data_array_t of string identifiers for all fabric planes in the overall
39 system.

40 **PMIX_FABRIC_SWITCH** "pmix.fab.switch" (string)
41 ID string of a fabric switch. When used as a modifier in a request for information, specifies the switch
42 whose information is to be returned. When used directly as a key in a request, returns a
43 pmix_data_array_t of string identifiers for all fabric switches in the overall system.

1 **PMIX_FABRIC_DEVICE** "pmix.fabdev" (**pmix_data_array_t**)
2 An array of **pmix_info_t** describing a particular fabric device using one or more of the attributes
3 defined below. The first element in the array shall be the **PMIX_DEVICE_ID** of the device.

4 **PMIX_FABRIC_DEVICE_INDEX** "pmix.fabdev.idx" (**uint32_t**)
5 Index of the device within an associated communication cost matrix.

6 **PMIX_FABRIC_DEVICE_NAME** "pmix.fabdev.nm" (**string**)
7 The operating system name associated with the device. This may be a logical fabric interface name
8 (e.g. "eth0" or "eno1") or an absolute filename.

9 **PMIX_FABRIC_DEVICE_VENDOR** "pmix.fabdev.vndr" (**string**)
10 Indicates the name of the vendor that distributes the device.

11 **PMIX_FABRIC_DEVICE_BUS_TYPE** "pmix.fabdev.btyp" (**string**)
12 The type of bus to which the device is attached (e.g., "PCI", "GEN-Z").

13 **PMIX_FABRIC_DEVICE_VENDORID** "pmix.fabdev.vendid" (**string**)
14 This is a vendor-provided identifier for the device or product.

15 **PMIX_FABRIC_DEVICE_DRIVER** "pmix.fabdev.driver" (**string**)
16 The name of the driver associated with the device.

17 **PMIX_FABRIC_DEVICE_FIRMWARE** "pmix.fabdev.fmwr" (**string**)
18 The device's firmware version.

19 **PMIX_FABRIC_DEVICE_ADDRESS** "pmix.fabdev.addr" (**string**)
20 The primary link-level address associated with the device, such as a MAC address. If multiple
21 addresses are available, only one will be reported.

22 **PMIX_FABRIC_DEVICE_COORDINATES** "pmix.fab.coord" (**pmix_geometry_t**)
23 The **pmix_geometry_t** fabric coordinates for the device, including values for all supported
24 coordinate views.

25 **PMIX_FABRIC_DEVICE_MTU** "pmix.fabdev.mtu" (**size_t**)
26 The maximum transfer unit of link level frames or packets, in bytes.

27 **PMIX_FABRIC_DEVICE_SPEED** "pmix.fabdev.speed" (**size_t**)
28 The active link data rate, given in bits per second.

29 **PMIX_FABRIC_DEVICE_STATE** "pmix.fabdev.state" (**pmix_link_state_t**)
30 The last available physical port state for the specified device. Possible values are
31 **PMIX_LINK_STATE_UNKNOWN**, **PMIX_LINK_DOWN**, and **PMIX_LINK_UP**, to indicate if the port
32 state is unknown or not applicable (unknown), inactive (down), or active (up).

33 **PMIX_FABRIC_DEVICE_TYPE** "pmix.fabdev.type" (**string**)
34 Specifies the type of fabric interface currently active on the device, such as Ethernet or InfiniBand.

35 **PMIX_FABRIC_DEVICE_PCI_DEVID** "pmix.fabdev.pcidevid" (**string**)
36 A node-level unique identifier for a PCI device. Provided only if the device is located on a PCI bus.
37 The identifier is constructed as a four-part tuple delimited by colons comprised of the PCI 16-bit
38 domain, 8-bit bus, 8-bit device, and 8-bit function IDs, each expressed in zero-extended hexadecimal
39 form. Thus, an example identifier might be "abc1:0f:23:01". The combination of node identifier

([PMIX_HOSTNAME](#) or [PMIX_NODEID](#)) and [PMIX_FABRIC_DEVICE_PCI_DEVID](#) shall be unique within the overall system.

Device attributes

[PMIX_DEVICE_DISTANCES](#) "pmix.dev.dist" ([pmix_data_array_t](#))

Return an array of [pmix_device_distance_t](#) containing the minimum and maximum distances of the given process location to all devices of the specified type on the local node.

[PMIX_DEVICE_TYPE](#) "pmix.dev.type" ([pmix_device_type_t](#))

Bitmask specifying the type(s) of device(s) whose information is being requested. Only used as a directive/qualifier.

[PMIX_DEVICE_ID](#) "pmix.dev.id" ([string](#))

System-wide UUID or node-local OS name of a particular device.

Sets-Groups attributes

[PMIX_QUERY_NUM_PSETS](#) "pmix.qry.psetnum" ([size_t](#))

Return the number of process sets defined in the specified range (defaults to [PMIX_RANGE_SESSION](#)).

[PMIX_QUERY_PSET_NAMES](#) "pmix.qry.psets" ([pmix_data_array_t*](#))

Return a [pmix_data_array_t](#) containing an array of strings of the process set names defined in the specified range (defaults to [PMIX_RANGE_SESSION](#)).

[PMIX_QUERY_PSET_MEMBERSHIP](#) "pmix.qry.pmems" ([pmix_data_array_t*](#))

Return an array of [pmix_proc_t](#) containing the members of the specified process set.

[PMIX_PSET_NAME](#) "pmix.pset.nm" ([char*](#))

The name of the newly defined process set.

[PMIX_PSET_MEMBERS](#) "pmix.pset.mems" ([pmix_data_array_t*](#))

An array of [pmix_proc_t](#) containing the members of the newly defined process set.

[PMIX_PSET_NAMES](#) "pmix.pset.nms" ([pmix_data_array_t*](#))

Returns an array of [char*](#) string names of the process sets in which the given process is a member.

[PMIX_QUERY_NUM_GROUPS](#) "pmix.qry.pgrpnum" ([size_t](#))

Return the number of process groups defined in the specified range (defaults to session). OPTIONAL QUALIFIERS: [PMIX_RANGE](#).

[PMIX_QUERY_GROUP_NAMES](#) "pmix.qry.pgrp" ([pmix_data_array_t*](#))

Return a [pmix_data_array_t](#) containing an array of string names of the process groups defined in the specified range (defaults to session). OPTIONAL QUALIFIERS: [PMIX_RANGE](#).

[PMIX_QUERY_GROUP_MEMBERSHIP](#) "pmix.qry.pgrpmems" ([pmix_data_array_t*](#))

Return a [pmix_data_array_t](#) of [pmix_proc_t](#) containing the members of the specified process group. REQUIRED QUALIFIERS: [PMIX_GROUP_ID](#).

[PMIX_GROUP_ID](#) "pmix.grp.id" ([char*](#))

User-provided group identifier - as the group identifier may be used in PMIx operations, the user is required to ensure that the provided ID is unique within the scope of the host environment (e.g., by including some user-specific or application-specific prefix or suffix to the string).

1 **PMIX_GROUP_LEADER** "pmix.grp.ldr" (bool)
 2 This process is the leader of the group.

3 **PMIX_GROUP_OPTIONAL** "pmix.grp.opt" (bool)
 4 Participation is optional - do not return an error if any of the specified processes terminate without
 5 having joined. The default is **false**.

6 **PMIX_GROUP_NOTIFY_TERMINATION** "pmix.grp.notterm" (bool)
 7 Notify remaining members when another member terminates without first leaving the group.

8 **PMIX_GROUP_FT_COLLECTIVE** "pmix.grp.ftcoll" (bool)
 9 Adjust internal tracking on-the-fly for terminated processes during a PMIx group collective operation.

10 **PMIX_GROUP_ASSIGN_CONTEXT_ID** "pmix.grp.actxid" (bool)
 11 Requests that the RM assign a new context identifier to the newly created group. The identifier is an
 12 unsigned, **size_t** value that the RM guarantees to be unique across the range specified in the request.
 13 Thus, the value serves as a means of identifying the group within that range. If no range is specified,
 14 then the request defaults to **PMIX_RANGE_SESSION**.

15 **PMIX_GROUP_LOCAL_ONLY** "pmix.grp.lcl" (bool)
 16 Group operation only involves local processes. PMIx implementations are *required* to automatically
 17 scan an array of group members for local vs remote processes - if only local processes are detected, the
 18 implementation need not execute a global collective for the operation unless a context ID has been
 19 requested from the host environment. This can result in significant time savings. This attribute can be
 20 used to optimize the operation by indicating whether or not only local processes are represented, thus
 21 allowing the implementation to bypass the scan.

22 **PMIX_GROUP_CONTEXT_ID** "pmix.grp.ctxid" (**size_t**)
 23 Context identifier assigned to the group by the host RM.

24 **PMIX_GROUP_ENDPT_DATA** "pmix.grp.endpt" (**pmix_byte_object_t**)
 25 Data collected during group construction to ensure communication between group members is
 26 supported upon completion of the operation.

27 **PMIX_GROUP_NAMES** "pmix.pgrp.nm" (**pmix_data_array_t***)
 28 Returns an array of **char*** string names of the process groups in which the given process is a member.
 29

30 Process Mgmt attributes

31 **PMIX_OUTPUT_TO_DIRECTORY** "pmix.outdir" (**char***)
 32 Direct output into files of form "<directory>/<jobid>/rank.<rank>/stdout[err]" -
 33 can be assigned to the entire job (by including attribute in the *job_info* array) or on a per-application
 34 basis in the *info* array for each **pmix_app_t**.

35 **PMIX_TIMEOUT_STACKTRACES** "pmix.tim.stack" (bool)
 36 Include process stacktraces in timeout report from a job.

37 **PMIX_TIMEOUT_REPORT_STATE** "pmix.tim.state" (bool)
 38 Report process states in timeout report from a job.

39 **PMIX_NOTIFY_JOB_EVENTS** "pmix.note.jev" (bool)

1 Requests that the launcher generate the [PMIX_EVENT_JOB_START](#), [PMIX_LAUNCH_COMPLETE](#),
2 and [PMIX_EVENT_JOB_END](#) events. Each event is to include at least the namespace of the
3 corresponding job and a [PMIX_EVENT_TIMESTAMP](#) indicating the time the event occurred. Note
4 that the requester must register for these individual events, or capture and process them by registering a
5 default event handler instead of individual handlers and then process the events based on the returned
6 status code. Another common method is to register one event handler for all job-related events, with a
7 separate handler for non-job events - see [PMIx_Register_event_handler](#) for details.

8 **PMIX_NOTIFY_PROC_TERMINATION** "pmix.noteproc" (bool)

9 Requests that the launcher generate the [PMIX_EVENT_PROC_TERMINATED](#) event whenever a
10 process either normally or abnormally terminates.

11 **PMIX_NOTIFY_PROC_ABNORMAL_TERMINATION** "pmix.noteabproc" (bool)

12 Requests that the launcher generate the [PMIX_EVENT_PROC_TERMINATED](#) event only when a
13 process abnormally terminates.

14 **PMIX_LOG_PROC_TERMINATION** "pmix.logproc" (bool)

15 Requests that the launcher log the [PMIX_EVENT_PROC_TERMINATED](#) event whenever a process
16 either normally or abnormally terminates.

17 **PMIX_LOG_PROC_ABNORMAL_TERMINATION** "pmix.logabproc" (bool)

18 Requests that the launcher log the [PMIX_EVENT_PROC_TERMINATED](#) event only when a process
19 abnormally terminates.

20 **PMIX_LOG_JOB_EVENTS** "pmix.log.jev" (bool)

21 Requests that the launcher log the [PMIX_EVENT_JOB_START](#), [PMIX_LAUNCH_COMPLETE](#), and
22 [PMIX_EVENT_JOB_END](#) events using [PMIx_Log](#), subject to the logging attributes of Section
23 13.4.3.

24 **PMIX_LOG_COMPLETION** "pmix.logcomp" (bool)

25 Requests that the launcher log the [PMIX_EVENT_JOB_END](#) event for normal or abnormal
26 termination of the spawned job using [PMIx_Log](#), subject to the logging attributes of Section 13.4.3.
27 The event shall include the returned status code ([PMIX_JOB_TERM_STATUS](#)) for the corresponding
28 job; the identity ([PMIX_PROCID](#)) and exit status ([PMIX_EXIT_CODE](#)) of the first failed process, if
29 applicable; and a [PMIX_EVENT_TIMESTAMP](#) indicating the time the termination occurred.

30 **PMIX_FIRST_ENVAR** "pmix.envar.first" (pmix_envar_t*)

31 Ensure the given value appears first in the specified envvar using the separator character, creating the
32 envvar if it doesn't already exist

33 Event attributes

34 **PMIX_EVENT_TIMESTAMP** "pmix.evtstamp" (time_t)

35 System time when the associated event occurred.

36 C.8.3 Added Environmental Variables

37 Tool environmental variables

38 [PMIX_LAUNCHER_RNDZ_URI](#)

39 [PMIX_LAUNCHER_RNDZ_FILE](#)

40 [PMIX_KEEPLIVE_PIPE](#)

41

1 C.8.4 Added Macros

2 `PMIX_CHECK_RESERVED_KEY` `PMIX_INFO_WAS_PROCESSED` `PMIX_INFO_PROCESSED`
3 `PMIX_INFO_LIST_START` `PMIX_INFO_LIST_ADD` `PMIX_INFO_LIST_XFER`
4 `PMIX_INFO_LIST_CONVERT` `PMIX_INFO_LIST_RELEASE`

5 C.8.5 Deprecated APIs

6 `pmix_evhdlr_reg_cbfunc_t` Renamed to `pmix_hdlr_reg_cbfunc_t`

7 The `pmix_server_client_connected_fn_t` server module entry point has been *deprecated* in favor
8 of `pmix_server_client_connected2_fn_t`

9 `PMIx_tool_connect_to_server` Replaced by `PMIx_tool_attach_to_server` to allow return
10 of the process identifier of the server to which the tool has attached.

11 C.8.6 Deprecated constants

12 The following constants were deprecated in v4.0:

13 `PMIX_ERR_DEBUGGER_RELEASE` Renamed to `PMIX_DEBUGGER_RELEASE`

14 `PMIX_ERR_JOB_TERMINATED` Renamed to `PMIX_EVENT_JOB_END`

15 `PMIX_EXISTS` Renamed to `PMIX_ERR_EXISTS`

16 `PMIX_ERR_PROC_ABORTED` Consolidated with `PMIX_EVENT_PROC_TERMINATED`

17 `PMIX_ERR_PROC_ABORTING` Consolidated with `PMIX_EVENT_PROC_TERMINATED`

18 `PMIX_ERR_LOST_CONNECTION_TO_SERVER` Consolidated into `PMIX_ERR_LOST_CONNECTION`

19 `PMIX_ERR_LOST_PEER_CONNECTION` Consolidated into `PMIX_ERR_LOST_CONNECTION`

20 `PMIX_ERR_LOST_CONNECTION_TO_CLIENT` Consolidated into `PMIX_ERR_LOST_CONNECTION`

21 `PMIX_ERR_INVALID_TERMINATION` Renamed to `PMIX_ERR_JOB_TERM_WO_SYNC`

22 `PMIX_PROC_TERMINATED` Renamed to `PMIX_EVENT_PROC_TERMINATED`

23 `PMIX_ERR_NODE_DOWN` Renamed to `PMIX_EVENT_NODE_DOWN`

24 `PMIX_ERR_NODE_OFFLINE` Renamed to `PMIX_EVENT_NODE_OFFLINE`

25 `PMIX_ERR_SYS_OTHER` Renamed to `PMIX_EVENT_SYS_OTHER`

26 `PMIX_CONNECT_REQUESTED` Connection has been requested by a PMIx-based tool - deprecated as not
27 required.

28 `PMIX_PROC_HAS_CONNECTED` A tool or client has connected to the PMIx server - deprecated in favor
29 of the new `pmix_server_client_connected2_fn_t` server module API

30 C.8.7 Removed constants

31 The following constants were removed from the PMIx Standard in v4.0 as they are internal to a particular
32 PMIx implementation.

33 `PMIX_ERR_HANDSHAKE_FAILED` Connection handshake failed

34 `PMIX_ERR_READY_FOR_HANDSHAKE` Ready for handshake

35 `PMIX_ERR_IN_ERRNO` Error defined in `errno`

36 `PMIX_ERR_INVALID_VAL_LENGTH` Invalid value length

1	PMIX_ERR_INVALID_LENGTH	Invalid argument length
2	PMIX_ERR_INVALID_NUM_ARGS	Invalid number of arguments
3	PMIX_ERR_INVALID_ARGS	Invalid arguments
4	PMIX_ERR_INVALID_NUM_PARSED	Invalid number parsed
5	PMIX_ERR_INVALID_KEYVALP	Invalid key/value pair
6	PMIX_ERR_INVALID_SIZE	Invalid size
7	PMIX_ERR_PROC_REQUESTED_ABORT	Process is already requested to abort
8	PMIX_ERR_SERVER_FAILED_REQUEST	Failed to connect to the server
9	PMIX_ERR_PROC_ENTRY_NOT_FOUND	Process not found
10	PMIX_ERR_INVALID_ARG	Invalid argument
11	PMIX_ERR_INVALID_KEY	Invalid key
12	PMIX_ERR_INVALID_KEY_LENGTH	Invalid key length
13	PMIX_ERR_INVALID_VAL	Invalid value
14	PMIX_ERR_INVALID_NAMESPACE	Invalid namespace
15	PMIX_ERR_SERVER_NOT_AVAIL	Server is not available
16	PMIX_ERR_SILENT	Silent error
17	PMIX_ERR_PACK_MISMATCH	Pack mismatch
18	PMIX_ERR_DATA_VALUE_NOT_FOUND	Data value not found
19	PMIX_ERR_NOT_IMPLEMENTED	Not implemented
20	PMIX_GDS_ACTION_COMPLETE	The GDS action has completed
21	PMIX_NOTIFY_ALLOC_COMPLETE	Notify that a requested allocation operation is complete - the result
22		of the request will be included in the <i>info</i> array

23 C.8.8 Deprecated attributes

24 The following attributes were deprecated in v4.0:

25	PMIX_TOPOLOGY	" <code>pmix.topo</code> " (<code>hwloc_topology_t</code>)
26		Renamed to PMIX_TOPOLOGY2 .
27	PMIX_DEBUG_JOB	" <code>pmix.dbg.job</code> " (<code>char*</code>)
28		Renamed to PMIX_DEBUG_TARGET
29	PMIX_RECONNECT_SERVER	" <code>pmix.tool.recon</code> " (<code>bool</code>)
30		Renamed to the PMIX_tool_connect_to_server API
31	PMIX_ALLOC_NETWORK	" <code>pmix.alloc.net</code> " (<code>array</code>)
32		Renamed to PMIX_ALLOC_FABRIC
33	PMIX_ALLOC_NETWORK_ID	" <code>pmix.alloc.netid</code> " (<code>char*</code>)
34		Renamed to PMIX_ALLOC_FABRIC_ID
35	PMIX_ALLOC_NETWORK_QOS	" <code>pmix.alloc.netqos</code> " (<code>char*</code>)
36		Renamed to PMIX_ALLOC_FABRIC_QOS
37	PMIX_ALLOC_NETWORK_TYPE	" <code>pmix.alloc.nettype</code> " (<code>char*</code>)
38		Renamed to PMIX_ALLOC_FABRIC_TYPE
39	PMIX_ALLOC_NETWORK_PLANE	" <code>pmix.alloc.netplane</code> " (<code>char*</code>)
40		Renamed to PMIX_ALLOC_FABRIC_PLANE
41	PMIX_ALLOC_NETWORK_ENDPTS	" <code>pmix.alloc.endpts</code> " (<code>size_t</code>)
42		Renamed to PMIX_ALLOC_FABRIC_ENDPTS

1 **PMIX_ALLOC_NETWORK_ENDPTS_NODE** "pmix.alloc.endpts.nd" (size_t)
 2 Renamed to **PMIX_ALLOC_FABRIC_ENDPTS_NODE**
 3 **PMIX_ALLOC_NETWORK_SEC_KEY** "pmix.alloc.nsec" (pmix_byte_object_t)
 4 Renamed to **PMIX_ALLOC_FABRIC_SEC_KEY**
 5 **PMIX_PROC_DATA** "pmix.pdata" (pmix_data_array_t)
 6 Renamed to **PMIX_PROC_INFO_ARRAY**
 7 **PMIX_LOCALITY** "pmix.loc" (pmix_locality_t)
 8 Relative locality of the specified process to the requester, expressed as a bitmask as per the description
 9 in the **pmix_locality_t** section. This value is unique to the requesting process and thus cannot be
 10 communicated by the server as part of the job-level information. Its use has been replaced by the
 11 **PMIx_Get_relative_locality** function.

12 C.8.9 Removed attributes

13 The following attributes were removed from the PMIx Standard in v4.0 as they are internal to a particular
 14 PMIx implementation. Users are referred to the **PMIx_Load_topology** API for obtaining the local
 15 topology description.

16 **PMIX_LOCAL_TOPO** "pmix.ltopo" (char*)
 17 XML representation of local node topology.
 18 **PMIX_TOPOLOGY_XML** "pmix.topo.xml" (char*)
 19 XML-based description of topology
 20 **PMIX_TOPOLOGY_FILE** "pmix.topo.file" (char*)
 21 Full path to file containing XML topology description
 22 **PMIX_TOPOLOGY_SIGNATURE** "pmix.toposig" (char*)
 23 Topology signature string.
 24 **PMIX_HWLOC_SHMEM_ADDR** "pmix.hwlocaddr" (size_t)
 25 Address of the HWLOC shared memory segment.
 26 **PMIX_HWLOC_SHMEM_SIZE** "pmix.hwlocsize" (size_t)
 27 Size of the HWLOC shared memory segment.
 28 **PMIX_HWLOC_SHMEM_FILE** "pmix.hwlocfile" (char*)
 29 Path to the HWLOC shared memory file.
 30 **PMIX_HWLOC_XML_V1** "pmix.hwlocxml1" (char*)
 31 XML representation of local topology using HWLOC's v1.x format.
 32 **PMIX_HWLOC_XML_V2** "pmix.hwlocxml2" (char*)
 33 XML representation of local topology using HWLOC's v2.x format.
 34 **PMIX_HWLOC_SHARE_TOPO** "pmix.hwlocsh" (bool)
 35 Share the HWLOC topology via shared memory
 36 **PMIX_HWLOC_HOLE_KIND** "pmix.hwlocholek" (char*)
 37 Kind of VM "hole" HWLOC should use for shared memory
 38 **PMIX_DSTPATH** "pmix.dstpath" (char*)
 39 Path to shared memory data storage (dstore) files. Deprecated from Standard as being implementation
 40 specific.
 41 **PMIX_COLLECTIVE_ALGO** "pmix.calgo" (char*)
 42 Comma-delimited list of algorithms to use for the collective operation. PMIx does not impose any
 43 requirements on a host environment's collective algorithms. Thus, the acceptable values for this
 44 attribute will be environment-dependent - users are encouraged to check their host environment for
 45 supported values.

1 **PMIX_COLLECTIVE_ALGO_REQD** "pmix.calreqd" (bool)
2 If **true**, indicates that the requested choice of algorithm is mandatory.
3 **PMIX_PROC_BLOB** "pmix.pblob" (pmix_byte_object_t)
4 Packed blob of process data.
5 **PMIX_MAP_BLOB** "pmix.mblob" (pmix_byte_object_t)
6 Packed blob of process location.
7 **PMIX_MAPPER** "pmix.mapper" (char*)
8 Mapping mechanism to use for placing spawned processes - when accessed using **PMIx_Get**, use the
9 **PMIX_RANK_WILDCARD** value for the rank to discover the mapping mechanism used for the
10 provided namespace.
11 **PMIX_NON_PMI** "pmix.nonpmi" (bool)
12 Spawned processes will not call **PMIx_Init**.
13 **PMIX_PROC_URI** "pmix.puri" (char*)
14 URI containing contact information for the specified process.
15 **PMIX_ARCH** "pmix.arch" (uint32_t)
16 Architecture flag.

17 C.9 Version 4.1: Oct. 2021

18 The v4.1 update includes clarifications and corrections from the v4.0 document:

19 C.9.1 Removed constants

20 The following constants were removed from the PMIx Standard in v4.1 as they are internal to a particular
21 PMIx implementation.

22 **PMIX_BUFFER** Buffer.

- 23 • Remove some stale language in [Chapter 9.1](#).
- 24 • Provisional Items:
 - 25 – Storage Chapter [19](#) on page [431](#)

26 C.9.2 Added Functions (Provisional)

- 27 • [PMIx_Data_load](#)
- 28 • [PMIx_Data_unload](#)
- 29 • [PMIx_Data_compress](#)
- 30 • [PMIx_Data_decompress](#)

31 C.9.3 Added Data Structures (Provisional)

- 32 • [pmix_storage_medium_t](#)
- 33 • [pmix_storage_accessibility_t](#)
- 34 • [pmix_storage_persistence_t](#)
- 35 • [pmix_storage_access_type_t](#)

1 C.9.4 Added Macros (Provisional)

- 2 • `PMIX_NAMESPACE_INVALID`
- 3 • `PMIX_RANK_IS_VALID`
- 4 • `PMIX_PROCID_INVALID`
- 5 • `PMIX_PROCID_XFER`

6 C.9.5 Added Constants (Provisional)

- 7 • `PMIX_PROC_NAMESPACE`

8 Storage constants

- 9 • `PMIX_STORAGE_MEDIUM_UNKNOWN`
- 10 • `PMIX_STORAGE_MEDIUM_TAPE`
- 11 • `PMIX_STORAGE_MEDIUM_HDD`
- 12 • `PMIX_STORAGE_MEDIUM_SSD`
- 13 • `PMIX_STORAGE_MEDIUM_NVME`
- 14 • `PMIX_STORAGE_MEDIUM_PMEM`
- 15 • `PMIX_STORAGE_MEDIUM_RAM`
- 16 • `PMIX_STORAGE_ACCESSIBILITY_NODE`
- 17 • `PMIX_STORAGE_ACCESSIBILITY_SESSION`
- 18 • `PMIX_STORAGE_ACCESSIBILITY_JOB`
- 19 • `PMIX_STORAGE_ACCESSIBILITY_RACK`
- 20 • `PMIX_STORAGE_ACCESSIBILITY_CLUSTER`
- 21 • `PMIX_STORAGE_ACCESSIBILITY_REMOTE`
- 22 • `PMIX_STORAGE_PERSISTENCE_TEMPORARY`
- 23 • `PMIX_STORAGE_PERSISTENCE_NODE`
- 24 • `PMIX_STORAGE_PERSISTENCE_SESSION`
- 25 • `PMIX_STORAGE_PERSISTENCE_JOB`
- 26 • `PMIX_STORAGE_PERSISTENCE_SCRATCH`
- 27 • `PMIX_STORAGE_PERSISTENCE_PROJECT`
- 28 • `PMIX_STORAGE_PERSISTENCE_ARCHIVE`
- 29 • `PMIX_STORAGE_ACCESS_RD`
- 30 • `PMIX_STORAGE_ACCESS_WR`
- 31 • `PMIX_STORAGE_ACCESS_RDWR`

32 C.9.6 Added Attributes (Provisional)

33 Storage attributes

34 `PMIX_STORAGE_ID` "pmix.strg.id" (char*)

35 An identifier for the storage system (e.g., lustre-fs1, daos-oss1, home-fs)

36 `PMIX_STORAGE_PATH` "pmix.strg.path" (char*)

37 Mount point path for the storage system (valid only for file-based storage systems)

38 `PMIX_STORAGE_TYPE` "pmix.strg.type" (char*)

1 Type of storage system (i.e., "lustre", "gpfs", "daos", "ext4")

2 **PMIX_STORAGE_VERSION** "pmix.strg.ver" (char*)

3 Version string for the storage system

4 **PMIX_STORAGE_MEDIUM** "pmix.strg.medium" (pmix_storage_medium_t)

5 Types of storage mediums utilized by the storage system (e.g., SSDs, HDDs, tape)

6 **PMIX_STORAGE_ACCESSIBILITY**

7 **"pmix.strg.access"** (pmix_storage_accessibility_t)

8 Accessibility level of the storage system (e.g., within same node, within same session)

9 **PMIX_STORAGE_PERSISTENCE** "pmix.strg.persist" (pmix_storage_persistence_t)

10 Persistence level of the storage system (e.g., sratch storage or achive storage)

11 **PMIX_QUERY_STORAGE_LIST** "pmix.strg.list" (char*)

12 Comma-delimited list of storage identifiers (i.e., **PMIX_STORAGE_ID** types) for available storage

13 systems

14 **PMIX_STORAGE_CAPACITY_LIMIT** "pmix.strg.caplim" (double)

15 Overall limit on capacity (in bytes) for the storage system

16 **PMIX_STORAGE_CAPACITY_USED** "pmix.strg.capuse" (double)

17 Overall used capacity (in bytes) for the storage system

18 **PMIX_STORAGE_OBJECT_LIMIT** "pmix.strg.objlim" (uint64_t)

19 Overall limit on number of objects (e.g., inodes) for the storage system

20 **PMIX_STORAGE_OBJECTS_USED** "pmix.strg.objuse" (uint64_t)

21 Overall used number of objects (e.g., inodes) for the storage system

22 **PMIX_STORAGE_MINIMAL_XFER_SIZE** "pmix.strg.minxfer" (double)

23 Minimal transfer size (in bytes) for the storage system - this is the storage system's atomic unit of

24 transfer (e.g., block size)

25 **PMIX_STORAGE_SUGGESTED_XFER_SIZE** "pmix.strg.sxfer" (double)

26 Suggested transfer size (in bytes) for the storage system

27 **PMIX_STORAGE_BW_MAX** "pmix.strg.bwmax" (double)

28 Maximum bandwidth (in bytes/sec) for storage system - provided as the theoretical maximum or the

29 maximum observed bandwidth value

30 **PMIX_STORAGE_BW_CUR** "pmix.strg.bwcur" (double)

31 Observed bandwidth (in bytes/sec) for storage system - provided as a recently observed bandwidth

32 value, with the exact measurement interval depending on the storage system and/or PMIx library

33 implementation

34 **PMIX_STORAGE_IOPS_MAX** "pmix.strg.iopsmax" (double)

35 Maximum IOPS (in I/O operations per second) for storage system - provided as the theoretical

36 maximum or the maximum observed IOPS value

37 **PMIX_STORAGE_IOPS_CUR** "pmix.strg.iopscur" (double)

1 Observed IOPS (in I/O operations per second) for storage system - provided as a recently observed
2 IOPS value, with the exact measurement interval depending on the storage system and/or PMIx library
3 implementation

4 `PMIX_STORAGE_ACCESS_TYPE` "pmix.strg.atype" (`pmix_storage_access_type_t`)
5 Qualifier describing the type of storage access to return information for (e.g., for qualifying
6 `PMIX_STORAGE_BW_CUR`, `PMIX_STORAGE_IOPS_CUR`, or
7 `PMIX_STORAGE_SUGGESTED_XFER_SIZE` attributes)

8 C.10 Version 5.0: May 2023

9 The v5.0 update includes the following changes from the v4.1 document:

- 10 • First release prepared using procedures defined in the PMIx Governance v1.7 document¹.
- 11 • Add specific values to constant definitions to ensure consistency across implementations.
- 12 • Add [Use-Cases](#) appendix with descriptions for Business Card Exchange, Debugging, Hybrid Applications,
13 MPI Sessions, and Cross-Version Compatibility.
- 14 • Add guidance on how PMIx defines an Application Binary Interface (ABI).
- 15 • Add ABI query attributes.
- 16 • Clarify three roles of consumers of the PMIx interface (client, server, tool).
- 17 • Clarify when `PMIX_PARENT_ID` attribute is provided.
- 18 • Clarify the value of `PMIX_CMD_LINE` attribute in spawn case.
- 19 • Clarifications to Terms and Conventions chapter and addition of additional term definitions.
- 20 • Re-organize the presentation of data access, synchronization, reserved keys and non-reserved keys.
- 21 • Make presentation of return values consistent across APIs.
- 22 • Attributes supported by PRRTE are no longer color coded. Refer to PRRTE documentation to see what is
23 supported for a particular PRRTE version.
- 24 • NEW markers are removed from item declarations. Refer to [Revision History](#) to see when something was
25 added.

26 C.10.1 Added Functions (Provisional)

- 27 • `PMIx_Data_embed`
- 28 • `PMIx_Value_load`
- 29 • `PMIx_Value_unload`
- 30 • `PMIx_Value_xfer`
- 31 • `PMIx_Info_list_start`
- 32 • `PMIx_Info_list_add`
- 33 • `PMIx_Info_list_xfer`
- 34 • `PMIx_Info_list_convert`
- 35 • `PMIx_Info_list_release`
- 36 • `PMIx_Topology_destruct`

¹<https://github.com/pmix/governance/releases/tag/v1.7>

1 C.10.2 Added Macros (Provisional)

- 2 • `PMIX_APP_STATIC_INIT`
- 3 • `PMIX_BYTE_OBJECT_STATIC_INIT`
- 4 • `PMIX_COORD_STATIC_INIT`
- 5 • `PMIX_CPUSET_STATIC_INIT`
- 6 • `PMIX_DATA_ARRAY_STATIC_INIT`
- 7 • `PMIX_DATA_BUFFER_STATIC_INIT`
- 8 • `PMIX_DEVICE_DIST_STATIC_INIT`
- 9 • `PMIX_ENDPOINT_STATIC_INIT`
- 10 • `PMIX_ENVAR_STATIC_INIT`
- 11 • `PMIX_FABRIC_STATIC_INIT`
- 12 • `PMIX_GEOMETRY_STATIC_INIT`
- 13 • `PMIX_INFO_STATIC_INIT`
- 14 • `PMIX_LOOKUP_STATIC_INIT`
- 15 • `PMIX_PROC_INFO_STATIC_INIT`
- 16 • `PMIX_PROC_STATIC_INIT`
- 17 • `PMIX_QUERY_STATIC_INIT`
- 18 • `PMIX_REGATTR_STATIC_INIT`
- 19 • `PMIX_TOPOLOGY_STATIC_INIT`
- 20 • `PMIX_VALUE_STATIC_INIT`

21 C.10.3 Added Constants (Provisional)

22 Spawn constants

- 23 • `PMIX_ERR_JOB_EXE_NOT_FOUND`
- 24 • `PMIX_ERR_JOB_INSUFFICIENT_RESOURCES`
- 25 • `PMIX_ERR_JOB_SYS_OP_FAILED`
- 26 • `PMIX_ERR_JOB_WDIR_NOT_FOUND`

27 C.10.4 Added Attributes

28 ABI attributes

29 `PMIX_QUERY_STABLE_ABI_VERSION` "`pmix.qry.stabiver`" (`char *`)

30 Query the PMIx Standard Stable ABI version(s) supported by the PMIx library. The version returned
31 will be of the form "`MAJOR.MINOR`". If multiple versions are supported then a comma-separated list
32 of version numbers will be returned. See Section 1.3.3.1 for versioning advice. NO QUALIFIERS.

33 `PMIX_QUERY_PROVISIONAL_ABI_VERSION` "`pmix.qry.prabiver`" (`char *`)

34 Query the PMIx Standard Provisional ABI version(s) supported by the PMIx library. The version
35 returned will be of the form "`MAJOR.MINOR`". If multiple versions are supported then a
36 comma-separated list of version numbers will be returned. See Section 1.3.3.1 for versioning advice.
37 NO QUALIFIERS.

1 C.10.5 Added Attributes (Provisional)

2 Spawn attributes

3 **PMIX_ENVARS_HARVESTED** "pmix.evar.hvstd" (bool)

4 Environmental parameters have been harvested by the spawn requestor - the server does not need to
5 harvest them.

6 **PMIX_JOB_TIMEOUT** "pmix.job.time" (int)

7 Time in seconds before the spawned job should time out and be terminated (0 => infinite), defined as
8 the total runtime of the job (equivalent to the walltime limit of typical batch schedulers).

9 **PMIX_LOCAL_COLLECTIVE_STATUS** "pmix.loc.col.st" (pmix_status_t)

10 Status code for local collective operation being reported to the host by the server library. PMIx servers
11 may aggregate the participation by local client processes in a collective operation - e.g., instead of
12 passing individual client calls to **PMIx_Fence** up to the host environment, the server may pass only a
13 single call to the host when all local participants have executed their **PMIx_Fence** call, thereby
14 reducing the burden placed on the host. However, in cases where the operation locally fails (e.g., if a
15 participating client abnormally terminates prior to calling the operation), the server upcall functions to
16 the host do not include a **pmix_status_t** by which the PMIx server can alert the host to that failure.
17 This attribute resolves that problem by allowing the server to pass the status information regarding the
18 local collective operation.

19 **PMIX_NODE_OVERSUBSCRIBED** "pmix.ndosub" (bool)

20 True if the number of processes from this job on this node exceeds the number of slots allocated to it

21 **PMIX_SINGLETON** "pmix.singleton" (char*)

22 String representation (nspace.rank) of proc ID for the singleton the server was started to support

23 **PMIX_SPAWN_TIMEOUT** "pmix.sp.time" (int)

24 Time in seconds before spawn operation should time out (0 => infinite). Logically equivalent to
25 passing the **PMIX_TIMEOUT** attribute to the **PMIx_Spawn** API, it is provided as a separate attribute
26 to distinguish it from the **PMIX_JOB_TIMEOUT** attribute

27 Tool attributes

28 **PMIX_IOF_FILE_PATTERN** "pmix.iof.fpt" (bool)

29 Specified output file is to be treated as a pattern and not automatically annotated by nspace, rank, or
30 other parameters. The pattern can use %n for the namespace, and %r for the rank wherever those
31 quantities are to be placed. The resulting filename will be appended with ".stdout" for the **stdout**
32 stream and ".stderr" for the **stderr** stream. If **PMIX_IOF_MERGE_STDERR_STDOUT** was given,
33 then only the **stdout** file will be created and both streams will be written into it.

34 **PMIX_IOF_FILE_ONLY** "pmix.iof.only" (bool)

35 Output only into designated files - do not also output a copy to the console's stdout/stderr

36 **PMIX_IOF_LOCAL_OUTPUT** "pmix.iof.local" (bool)

37 Write output streams to local stdout/err

38 **PMIX_IOF_MERGE_STDERR_STDOUT** "pmix.iof.mrg" (bool)

39 Merge stdout and stderr streams from application procs

40 **PMIX_IOF_RANK_OUTPUT** "pmix.iof.rank" (bool)

1 Tag output with the rank it came from

2 **PMIX_IOF_OUTPUT_RAW** "pmix.iof.raw" (bool)

3 Do not buffer output to be written as complete lines - output characters as the stream delivers them

4 **PMIX_IOF_OUTPUT_TO_DIRECTORY** "pmix.iof.dir" (char*)

5 Direct application output into files of form "<directory>/<nospace>/rank.<rank>/stdout" (for **stdout**)
6 and "<directory>/<nospace>/rank.<rank>/stderr" (for **stderr**). If

7 **PMIX_IOF_MERGE_STDERR_STDOUT** was given, then only the **stdout** file will be created and
8 both streams will be written into it.

9 **PMIX_IOF_OUTPUT_TO_FILE** "pmix.iof.file" (char*)

10 Direct application output into files of form "<filename>.<nospace>.<rank>.stdout" (for **stdout**) and
11 "<filename>.<nospace>.<rank>.stderr" (for **stderr**). If **PMIX_IOF_MERGE_STDERR_STDOUT** was
12 given, then only the **stdout** file will be created and both streams will be written into it.

13 **PMIX_BREAKPOINT** "pmix.brkpt" (char*)

14 String ID of the breakpoint where the process(es) is(are) waiting.

15 **PMIX_DEBUG_STOP_IN_APP** "pmix.dbg.notify" (varies)

16 Direct specified ranks to stop at application-specific point and notify they are ready-to-debug. The
17 attribute's value can be any of three data types:

- 18 • bool - true indicating all ranks
- 19 • **pmix_rank_t** - the rank of one proc, or **PMIX_RANK_WILDCARD** for all
- 20 • a **pmix_data_array_t** if an array of individual processes are specified

21 The resulting application processes are to notify their server (by generating the
22 **PMIX_READY_FOR_DEBUG** event) when they reach some application-determined location - the event
23 shall include the **PMIX_BREAKPOINT** attribute indicating where the application has stopped. The
24 application shall pause at that point until released by debugger modification of an appropriate variable.
25 The launcher (RM or IL) is responsible for generating the **PMIX_READY_FOR_DEBUG** event when all
26 processes have indicated they are at the pause point.

27 C.10.6 Deprecated constants

28 The following constants were deprecated in v5.0:

29 **PMIX_DEBUG_WAITING_FOR_NOTIFY** Renamed to **PMIX_READY_FOR_DEBUG**

30 C.10.7 Deprecated attributes

31 The following attributes were deprecated in v5.0:

32 **PMIX_DEBUG_WAIT_FOR_NOTIFY** "pmix.dbg.notify" (bool)
33 Renamed to **PMIX_DEBUG_STOP_IN_APP**

1 C.10.8 Deprecated macros

2 The following macros were deprecated in v5.0:

- 3 ● **PMIX_VALUE_LOAD** Replaced by the **PMIx_Value_load** API
- 4 ● **PMIX_VALUE_UNLOAD** Replaced by the **PMIx_Value_unload** API
- 5 ● **PMIX_VALUE_XFER** Replaced by the **PMIx_Value_xfer** API
- 6 ● **PMIX_INFO_LOAD** Replaced by the **PMIx_Info_load** API
- 7 ● **PMIX_INFO_XFER** Replaced by the **PMIx_Info_xfer** API
- 8 ● **PMIX_INFO_LIST_START** Replaced by the **PMIx_Info_list_start** API
- 9 ● **PMIX_INFO_LIST_ADD** Replaced by the **PMIx_Info_list_add** API
- 10 ● **PMIX_INFO_LIST_XFER** Replaced by the **PMIx_Info_list_xfer** API
- 11 ● **PMIX_INFO_LIST_CONVERT** Replaced by the **PMIx_Info_list_convert** API
- 12 ● **PMIX_INFO_LIST_RELEASE** Replaced by the **PMIx_Info_list_release** API
- 13 ● **PMIX_TOPOLOGY_DESTRUCT** Replaced by the **PMIx_Topology_destruct** API
- 14 ● **PMIX_TOPOLOGY_FREE** Not replaced.

15 C.10.9 Removed Constants

16 The following constants were removed:

- 17 **PMIX_QUERY_PARTIAL_SUCCESS** Some, but not all, of the requested information was returned.
- 18 Replaced by **PMIX_ERR_PARTIAL_SUCCESS**.

APPENDIX D

Acknowledgements

1 This document represents the work of many people who have contributed to the PMIx community. Without
2 the hard work and dedication of these people this document would not have been possible. The sections below
3 list some of the active participants and organizations in the various PMIx standard iterations.

4 **D.1 Version 5.0**

5 The following list includes some of the active participants in the PMIx v5 standardization process.

6 **Release Managers for v5.0**

- 7 • Ken Raffanetti
- 8 • David Solt

9 **ASC Chairs terms during v5.0 preparation**

- 10 • Aurelien Bouteiller (2023-2025)
- 11 • Joshua Hursey (2019-2024)
- 12 • Kathryn Mohror (2020-2023)
- 13 • Ralph Castain (2018-2020)

14 **Working Group Chairs during v5.0 preparation**

- 15 • Isafías A. Comprés (Tools Working Group)
- 16 • Stephen Herbein (Use-Cases Working Group)
- 17 • Shane Snyder (Storage Working Group)
- 18 • David Solt (Implementation Agnostic Document Working Group)
- 19 • Justin Wozniak (Dynamic Workflows Working Group)

20 **ASC Secretaries terms during v5.0 preparation**

- 21 • Norbert Eicker (2023-2025)
- 22 • Thomas Naughton (2020-2024)
- 23 • Aurelien Bouteiller (2021-2023)
- 24 • Stephen Herbein (2019-2021)

Contributors

- 1
- 2 ● Julien Adam
- 3 ● William E. Allcock
- 4 ● Brian Barrett
- 5 ● Albeaus Bayucan
- 6 ● David Bernholdt
- 7 ● Wesley Bland
- 8 ● Swen Boehm
- 9 ● George Bosilca
- 10 ● Aurelien Bouteiller
- 11 ● Suren Byna
- 12 ● Paul Carpenter
- 13 ● John Carrier
- 14 ● Ralph H. Castain
- 15 ● Sourav Chakraborty
- 16 ● Michael Chuvelev
- 17 ● Isaiás A. Comprés
- 18 ● Jai Dayal
- 19 ● John DelSignore
- 20 ● Andreas Dilger
- 21 ● Dmitry Durnov
- 22 ● Norbert Eicker
- 23 ● Bengisu Elis
- 24 ● Noah Evans
- 25 ● Jim Garlick
- 26 ● Mahdieh Ghazimirsaeed
- 27 ● Brice Goglin
- 28 ● Andrew Gontarek
- 29 ● Stephen Herbein
- 30 ● Thomas Hines
- 31 ● Daniel J. Holmes

- 1 • Kaiyuan Hou
- 2 • Dominik Huber
- 3 • Joshua Hursey
- 4 • Julien Jaeger
- 5 • Sid Jana
- 6 • Jithin Jose
- 7 • Michael Karo
- 8 • Quincey Koziol
- 9 • Stephan Krempel
- 10 • Gregory Kurtzer
- 11 • Ignacio Laguna
- 12 • Ti Leggett
- 13 • Karthik Vadambacheri Manian
- 14 • Pat McCarthy
- 15 • Guillaume Mercier
- 16 • Kathryn Mohror
- 17 • Grace Nansamba
- 18 • Thomas Naughton
- 19 • Bogdan Nicolae
- 20 • Guillaume Papauré
- 21 • Trupeshkumar Patel
- 22 • Artem Polyakov
- 23 • Swaroop Pophale
- 24 • Howard Pritchard
- 25 • Nick Radcliffe
- 26 • Ken Raffenetti
- 27 • Bharath Ramesh
- 28 • Michael A Raymond
- 29 • Paul Rich
- 30 • Barry Rountree
- 31 • Anatoliy Rozanov

- 1 • Amit Ruhela
- 2 • Derek Schafer
- 3 • Dirk Schubert
- 4 • Martin Schulz
- 5 • Tom Scogland
- 6 • Nat Shineman
- 7 • Danielle Sikich
- 8 • Shane Snyder
- 9 • David Solt
- 10 • Jeff Squyres
- 11 • Hari Subramoni
- 12 • Shinji Sumimoto
- 13 • Geoffroy Vallee
- 14 • Justin Wozniak
- 15 • Andrew Younge

16 **Institutions**

17 The following institutions supported this effort through time and travel support for the people listed above.

- 18 • Altair
- 19 • Amazon Web Services
- 20 • AMD
- 21 • Argonne National Laboratory
- 22 • Arm, Inc
- 23 • Barcelona Supercomputing Center
- 24 • CEA
- 25 • Cisco
- 26 • The Exascale Computing Project, an initiative of the US Department of Energy
- 27 • Fujitsu
- 28 • HPE Co.
- 29 • IBM, Inc.
- 30 • INRIA
- 31 • Intel Corporation

- 1 • Jülich Supercomputing Center
- 2 • Lawrence Berkeley National Laboratory
- 3 • Lawrence Livermore National Laboratory
- 4 • Los Alamos National Laboratory
- 5 • Meta
- 6 • Microsoft
- 7 • Nanook Consulting
- 8 • National Science Foundation
- 9 • Northwestern University
- 10 • NVIDIA
- 11 • Oak Ridge National Laboratory
- 12 • The Ohio State University
- 13 • ParTec AG
- 14 • Perforce Software, Inc.
- 15 • Sandia National Laboratories
- 16 • Sylabs
- 17 • Tennessee Technological University
- 18 • TU Munich
- 19 • University of Alabama
- 20 • University of Tennessee, Chattanooga
- 21 • University of Tennessee, Knoxville
- 22 • Whamcloud

23 **D.2 Version 4.0**

24 The following list includes some of the active participants in the PMIx v4 standardization process.

- 25 • Ralph H. Castain and Danielle Sikich
- 26 • Joshua Hursey and David Solt
- 27 • Dirk Schubert
- 28 • John DelSignore
- 29 • Aurelien Bouteiller
- 30 • Michael A Raymond

- 1 • Howard Pritchard and Nathan Hjelm
- 2 • Brice Goglin
- 3 • Kathryn Mohror and Stephen Herbein
- 4 • Thomas Naughton and Swaroop Pophale
- 5 • William E. Allcock and Paul Rich
- 6 • Michael Karo
- 7 • Artem Polyakov

8 The following institutions supported this effort through time and travel support for the people listed above.

- 9 • Intel Corporation
- 10 • IBM, Inc.
- 11 • Allinea (ARM)
- 12 • Perforce
- 13 • University of Tennessee, Knoxville
- 14 • The Exascale Computing Project, an initiative of the US Department of Energy
- 15 • National Science Foundation
- 16 • HPE Co.
- 17 • Los Alamos National Laboratory
- 18 • INRIA
- 19 • Lawrence Livermore National Laboratory
- 20 • Oak Ridge National Laboratory
- 21 • Argonne National Laboratory
- 22 • Altair
- 23 • NVIDIA

24 **D.3 Version 3.0**

25 The following list includes some of the active participants in the PMIx v3 standardization process.

- 26 • Ralph H. Castain, Andrew Friedley, Brandon Yates
- 27 • Joshua Hursey and David Solt
- 28 • Aurelien Bouteiller and George Bosilca
- 29 • Dirk Schubert
- 30 • Kevin Harms

- 1 • Artem Polyakov
- 2 The following institutions supported this effort through time and travel support for the people listed above.
- 3 • Intel Corporation
- 4 • IBM, Inc.
- 5 • University of Tennessee, Knoxville
- 6 • The Exascale Computing Project, an initiative of the US Department of Energy
- 7 • National Science Foundation
- 8 • Argonne National Laboratory
- 9 • Allinea (ARM)
- 10 • NVIDIA

11 **D.4 Version 2.0**

12 The following list includes some of the active participants in the PMIX v2 standardization process.

- 13 • Ralph H. Castain, Annapurna Dasari, Christopher A. Holguin, Andrew Friedley, Michael Klemm and Terry
- 14 Wilmarth
- 15 • Joshua Hursey, David Solt, Alexander Eichenberger, Geoff Paulsen, and Sameh Sharkawi
- 16 • Aurelien Bouteiller and George Bosilca
- 17 • Artem Polyakov, Igor Ivanov and Boris Karasev
- 18 • Gilles Gouaillardet
- 19 • Michael A Raymond and Jim Stoffel
- 20 • Dirk Schubert
- 21 • Moe Jette
- 22 • Takahiro Kawashima and Shinji Sumimoto
- 23 • Howard Pritchard
- 24 • David Beer
- 25 • Brice Goglin
- 26 • Geoffroy Vallee, Swen Boehm, Thomas Naughton and David Bernholdt
- 27 • Adam Moody and Martin Schulz
- 28 • Ryan Grant and Stephen Olivier
- 29 • Michael Karo

30 The following institutions supported this effort through time and travel support for the people listed above.

- 1 • Intel Corporation
- 2 • IBM, Inc.
- 3 • University of Tennessee, Knoxville
- 4 • The Exascale Computing Project, an initiative of the US Department of Energy
- 5 • National Science Foundation
- 6 • Mellanox, Inc.
- 7 • Research Organization for Information Science and Technology
- 8 • HPE Co.
- 9 • Allinea (ARM)
- 10 • SchedMD, Inc.
- 11 • Fujitsu Limited
- 12 • Los Alamos National Laboratory
- 13 • Adaptive Solutions, Inc.
- 14 • INRIA
- 15 • Oak Ridge National Laboratory
- 16 • Lawrence Livermore National Laboratory
- 17 • Sandia National Laboratory
- 18 • Altair

19 **D.5 Version 1.0**

20 The following list includes some of the active participants in the PMIx v1 standardization process.

- 21 • Ralph H. Castain, Annapurna Dasari and Christopher A. Holguin
- 22 • Joshua Hursey and David Solt
- 23 • Aurelien Bouteiller and George Bosilca
- 24 • Artem Polyakov, Elena Shipunova, Igor Ivanov, and Joshua Ladd
- 25 • Gilles Gouaillardet
- 26 • Gary Brown
- 27 • Moe Jette

28 The following institutions supported this effort through time and travel support for the people listed above.

- 29 • Intel Corporation
- 30 • IBM, Inc.

- 1 ● University of Tennessee, Knoxville
- 2 ● Mellanox, Inc.
- 3 ● Research Organization for Information Science and Technology
- 4 ● Adaptive Solutions, Inc.
- 5 ● SchedMD, Inc.

Bibliography

- [1] Ralph H. Castain, David Solt, Joshua Hursey, and Aurelien Bouteiller. PMIx: Process management for exascale environments. In *Proceedings of the 24th European MPI Users' Group Meeting*, EuroMPI '17, pages 14:1–14:10, New York, NY, USA, 2017. ACM.
- [2] Balaji P. et al. PMI: A scalable parallel process-management interface for extreme-scale systems. In *Recent Advances in the Message Passing Interface*, EuroMPI '10, pages 31–41, Berlin, Heidelberg, 2010. Springer.
- [3] Khaled Hamidouche, Jian Lin, Mingzhe Li, Jie Zhang, and D.K.Panda. Supporting hybrid MPI+PGAS programming models through unified communication runtime: An MVAPICH2-X approach. In *4th MVAPICH User's Group*, MUG, 2016.
- [4] Message Passing Interface Forum. *MPI: A Message-passing Interface Standard, Version 3.1*. High-Performance Computing Center Stuttgart, University of Stuttgart, June 2015.
- [5] Geoffroy R. Vallée. MOC - MPI Open MP Coordination library. <https://github.com/OMPI-X/MOC>, 2018. [Online; accessed 20-Dec-2019].
- [6] Geoffroy R. Vallée and David Bernhold. Improving support of MPI+OpenMP applications. In *Proceedings of the 25th European MPI Users' Group Meeting*, EuroMPI, 2018.

Index

General terms and other items not induced in the other indices.

application, [10](#), [90](#), [110](#), [295](#), [297](#), [300](#), [537](#), [540](#)
attribute, [11](#)

Build ABI, [7](#), [8](#)
business card exchange, [69](#)

client, [5](#), [10](#), [61](#)
clients, [10](#)
clone, [10](#)
clones, [10](#), [114](#), [117](#), [118](#), [181](#), [183](#), [185](#), [187](#), [208](#), [211](#), [547](#)

data realm, [81](#), [263](#), [264](#)
data realms, [81](#)
device, [11](#)
devices, [11](#)
Direct Modex, [252](#), [310](#)

fabric, [11](#)
fabric device, [11](#)
fabric devices, [11](#)
fabric plane, [11](#), [166](#), [171](#), [201](#), [204](#), [205](#), [268](#), [312](#)
fabric planes, [11](#)
fabrics, [11](#)

host environment, [10](#)

instant on, [11](#), [70](#), [251](#)

job, [9](#), [10](#), [83–86](#), [90](#), [110](#), [287–290](#), [292](#), [295–297](#), [299](#), [300](#), [311–313](#), [488](#), [505](#), [537](#), [540](#), [549](#), [551](#)

key, [11](#)

Linker ABI, [7](#), [8](#)

modex exchange, [69](#)

namespace, [9](#)
node, [10](#), [90](#), [110](#), [166](#), [172](#), [201](#), [204](#), [205](#), [295](#), [312](#)

package, [10](#), [88](#), [293](#), [551](#)

peer, [10](#), [85](#), [292](#)
peers, [10](#)
process, [10](#), [90](#), [110](#), [166](#), [172](#), [201](#), [204](#), [205](#), [295](#), [312](#)
processing unit, [11](#)

rank, [10](#), [301](#)
realm, [81](#)
realms, [81](#)
resource manager, [10](#)
RM, [10](#)

scheduler, [10](#), [266](#)
server, [5](#)
session, [9](#), [82](#), [90](#), [110](#), [287](#), [295](#), [296](#), [537](#), [540](#), [549](#)
singleton, [5](#)

thread, [10](#)
threads, [10](#)
tool, [5](#), [10](#)
tools, [10](#)

workflow, [10](#)
workflows, [10](#), [375](#)

Index of APIs

PMIx_Abort, [28](#), [161](#), [162](#), [332](#), [333](#), [444](#), [458](#), [535](#)
 PMIxClient.abort (Python), [458](#)

PMIx_Alloc_directive_string, [59](#), [479](#), [536](#)
 PMIxClient.alloc_directive_string (Python), [479](#)

PMIx_Allocation_request, [111](#), [199](#), [199](#), [204](#), [466](#), [538](#), [540](#), [543](#)
 PMIxClient.allocation_request (Python), [466](#)

PMIx_Allocation_request_nb, [202](#), [204](#), [206](#), [536](#)

PMIx_Commit, [69](#), [71](#), [73](#), [73](#), [74](#), [114](#), [116](#), [117](#), [310](#), [311](#), [334](#), [338](#), [460](#), [503](#), [504](#), [535](#)
 PMIxClient.commit (Python), [459](#)

PMIx_Compute_distances, [192](#), [194](#), [476](#), [543](#)
 PMIxClient.compute_distances (Python), [475](#)

PMIx_Compute_distances_nb, [193](#), [543](#)

PMIx_Connect, [180](#), [182](#), [184](#), [187](#), [228–230](#), [416](#), [463](#), [535](#), [537](#)
 PMIxClient.connect (Python), [463](#)

PMIx_Connect_nb, [182](#), [182](#), [535](#)

pmix_connection_cbfunc_t, [357](#), [358](#)

pmix_credential_cbfunc_t, [274](#), [373](#), [374](#)

PMIx_Data_compress, [158](#), [159](#), [562](#)

PMIx_Data_copy, [154](#), [536](#)

PMIx_Data_copy_payload, [156](#), [536](#)

PMIx_Data_decompress, [158](#), [562](#)

PMIx_Data_embed, [159](#), [565](#)

PMIx_Data_load, [156](#), [160](#), [562](#)

PMIx_Data_pack, [151](#), [152](#), [285](#), [536](#)

PMIx_Data_print, [155](#), [536](#)

PMIx_Data_range_string, [58](#), [478](#), [536](#)
 PMIxClient.data_range_string (Python), [478](#)

PMIx_Data_type_string, [58](#), [479](#), [536](#)
 PMIxClient.data_type_string (Python), [478](#)

PMIx_Data_unload, [157](#), [562](#)

PMIx_Data_unpack, [153](#), [157](#), [536](#)

PMIx_Deregister_event_handler, [143](#), [472](#), [520](#), [536](#), [543](#)
 PMIxClient.deregister_event_handler (Python), [471](#)

pmix_device_dist_cbfunc_t, [194](#), [194](#), [544](#)

PMIx_Device_type_string, [60](#), [481](#), [543](#)
 PMIxClient.device_type_string (Python), [481](#)

PMIx_Disconnect, [184](#), [185–187](#), [230](#), [416](#), [464](#), [535](#), [537](#)
 PMIxClient.disconnect (Python), [463](#)

PMIx_Disconnect_nb, [186](#), [187](#), [230](#), [535](#)

pmix_dmodex_response_fn_t, [309](#), [310](#)

PMIx_Error_string, [57](#), [476](#), [535](#)

PMIxCient.error_string (Python), 476
 pmix_event_notification_cbfunc_fn_t, 142, **147**, 147, 526, 528
 PMIx_Fabric_deregister, **270**, 271, 474, 543
 PMIxCient.fabric_deregister (Python), 473
 PMIx_Fabric_deregister_nb, **271**, 543
 PMIx_Fabric_register, 260, **267**, 269, 473, 543
 PMIxCient.fabric_register (Python), 472
 PMIx_Fabric_register_nb, **268**, 543
 PMIx_Fabric_update, 268, **269**, 270, 473, 543
 PMIxCient.fabric_update (Python), 473
 PMIx_Fabric_update_nb, **269**, 543
 PMIx_Fence, 4, 70, **113**, 113–118, 182, 185, 227, 235, 239, 252, 282, 310, 333–335, 350, 351, 383, 445, 460,
 503, 504, 535, 547, 567
 PMIxCient.fence (Python), 460
 PMIx_Fence_nb, 56, 113, **115**, 333, 335, 445, 535, 542
 PMIx_Finalize, 28, 64, **66**, 66, 67, 99, 180, 331, 416, 444, 458, 535
 PMIxCient.finalize (Python), 458
 PMIx_generate_ppn, **285**, 483, 535, 540
 PMIxCient.generate_ppn (Python), 483
 PMIx_generate_regex, **284**, 285, 296, 483, 535, 540
 PMIxCient.generate_regex (Python), 483
 PMIx_Get, 3, 11, 31, 64, 65, 68–70, 73, **74**, 75, 77–83, 85–90, 94, 96, 97, 100, 101, 106, 110–114, 116, 118,
 164, 165, 169, 170, 173–175, 189, 191, 192, 204, 212, 218, 224, 226, 228, 232, 235, 252, 254, 263,
 264, 266, 288, 289, 293, 296, 323, 346, 348, 392, 401–403, 417, 433, 461, 503, 504, 506, 509, 510,
 514, 515, 519, 531, 535, 537, 542, 547, 548, 553, 554, 562
 PMIxCient.get (Python), 460
 PMIx_Get_attribute_name, **59**, 481, 543
 PMIxCient.get_attribute_name (Python), 480
 PMIx_Get_attribute_string, **59**, 480, 543
 PMIxCient.get_attribute_string (Python), 480
 PMIx_Get_cpuset, **192**, 475, 543
 PMIxCient.get_cpuset (Python), 475
 PMIx_Get_credential, **273**, 275, 374, 468, 538, 543
 PMIxCient.get_credential (Python), 467
 PMIx_Get_credential_nb, **274**
 PMIx_Get_nb, 56, 69, 70, **77**, 113, 535
 PMIx_Get_relative_locality, **188**, 191, 293, 323, 474, 506, 543, 561
 PMIxCient.get_relative_locality (Python), 474
 PMIx_Get_version, 13, **62**, 457, 535
 PMIxCient.get_version (Python), 457
 PMIx_Group_construct, 227, 228, **232**, 234, 235, 237, 469, 531, 543
 PMIxCient.group_construct (Python), 468
 PMIx_Group_construct_nb, **235**, 237, 543
 PMIx_Group_destruct, 230, **238**, 239, 240, 249, 471, 543
 PMIxCient.group_destruct (Python), 470
 PMIx_Group_destruct_nb, **239**, 240, 543
 PMIx_Group_invite, 229, **240**, 242–244, 469, 543

PMIxCliEnt.group_invite (Python), 469
 PMIx_Group_invite_nb, [243](#), 543
 PMIx_Group_join, [229](#), [242](#), [244](#), [245](#), 246, 248, 470, 543
 PMIxCliEnt.group_join (Python), 469
 PMIx_Group_join_nb, [244](#), [246](#), 248, 543
 PMIx_Group_leave, [230](#), [248](#), 249, 250, 470, 543
 PMIxCliEnt.group_leave (Python), 470
 PMIx_Group_leave_nb, [249](#), 543
 pmix_hdlr_reg_cbfunc_t, [57](#), 138, 426, 428, 559
 pmix_info_cbfunc_t, [55](#), [56](#), 56, 100, 194, 202, 209, 211, 215, 217, 236, 243, 247, 321, 359, 366, 368, 370,
 [371](#), [382](#), [385](#)
 PMIx_Info_directives_string, [58](#), 478, 536
 PMIxCliEnt.info_directives_string (Python), 478
 PMIx_Info_list_add, [39](#), 565, 569
 PMIx_Info_list_convert, [40](#), 565, 569
 PMIx_Info_list_release, [41](#), 565, 569
 PMIx_Info_list_start, [39](#), 40, 41, 565, 569
 PMIx_Info_list_xfer, [40](#), 565, 569
 PMIx_Info_load, [37](#), 569
 PMIx_Info_xfer, [38](#), 296, 569
 PMIx_Init, 10, 61, [62](#), 64, 65, 87, 96, 99, 101, 330, 397, 401, 410, 411, 417, 457, 503, 504, 509, 515, 523,
 [532](#), [533](#), [536](#), 542, 562
 PMIxCliEnt.init (Python), 457
 PMIx_Initialized, [61](#), 457, 535
 PMIxCliEnt.initialized (Python), 457
 pmix_iof_cbfunc_t, [380](#), 426, 443
 iofcbfunc (Python), 442
 PMIx_IOF_channel_string, [59](#), 479, 538
 PMIxCliEnt.iof_channel_string (Python), 479
 PMIx_IOF_deregister, [427](#), 495, 538, 543
 PMIxTool.iof_deregister (Python), 495
 PMIx_IOF_pull, [347](#), [362](#), 394, 395, 399, 400, 403, 405, [425](#), 428, 495, 510, 514, 515, 519, 538, 543
 PMIxTool.iof_pull (Python), 494
 PMIx_IOF_push, [347](#), [362](#), 394, 399, 403, 405–407, 409, [428](#), 430, 496, 538, 543, 552
 PMIxTool.iof_push (Python), 495
 PMIx_Job_control, [199](#), [206](#), 208, 210–212, 370, 412, 467, 538, 543
 PMIxCliEnt.job_ctrl (Python), 466
 PMIx_Job_control_nb, 94, 206, [208](#), 294, 536
 PMIx_Job_state_string, [59](#), 480, 543
 PMIxCliEnt.job_state_string (Python), 480
 PMIx_Link_state_string, [59](#), 481, 543
 PMIxCliEnt.link_state_string (Python), 481
 PMIx_Load_topology, [187](#), 474, 543, 561
 PMIxCliEnt.load_topology (Python), 474
 PMIx_Log, [176](#), [218](#), 221, 224, 396, 397, 415, 466, 538, 558
 PMIxCliEnt.log (Python), 465
 PMIx_Log_nb, [221](#), 224, 536

- PMIx_Lookup, [69](#), [119](#), [123](#), [125](#), [127](#), [461](#), [462](#), [535](#)
 - PMIxClient.lookup (Python), [461](#)
- pmix_lookup_cbfunc_t, [130](#), [130](#), [340](#)
- PMIx_Lookup_nb, [69](#), [125](#), [130](#), [535](#)
- pmix_modex_cbfunc_t, [55](#), [334](#), [336](#), [336](#), [337](#)
- pmix_notification_fn_t, [137](#), [142](#), [142](#), [443](#)
 - evhandler (Python), [443](#)
- PMIx_Notify_event, [144](#), [357](#), [472](#), [509](#), [512](#), [520](#), [524](#), [525](#), [536](#), [543](#)
 - PMIxClient.notify_event (Python), [472](#)
- pmix_op_cbfunc_t, [55](#), [55](#), [121](#), [133](#), [144](#), [145](#), [148](#), [183](#), [186](#), [221](#), [239](#), [249](#), [268](#), [270](#), [271](#), [286](#), [304–308](#), [314](#), [319](#), [320](#), [322](#), [329–332](#), [338](#), [342](#), [349](#), [351](#), [353](#), [355](#), [356](#), [364](#), [378](#), [381](#), [429](#)
- PMIx_Parse_cpuset_string, [191](#), [324](#), [475](#), [543](#)
 - PMIxClient.parse_cpuset_string (Python), [475](#)
- PMIx_Persistence_string, [58](#), [477](#), [536](#)
 - PMIxClient.persistence_string (Python), [477](#)
- PMIx_Proc_state_string, [58](#), [477](#), [536](#)
 - PMIxClient.proc_state_string (Python), [476](#)
- PMIx_Process_monitor, [199](#), [213](#), [217](#), [467](#), [538](#), [543](#)
 - PMIxClient.monitor (Python), [467](#)
- PMIx_Process_monitor_nb, [215](#), [217](#), [218](#), [536](#)
- PMIx_Progress, [63](#), [67](#), [281](#), [284](#), [420](#), [482](#), [543](#), [551](#)
 - PMIxClient.progress (Python), [482](#)
- PMIx_Publish, [69](#), [119](#), [120](#), [122](#), [123](#), [340](#), [461](#), [462](#), [535](#)
 - PMIxClient.publish (Python), [461](#)
- PMIx_Publish_nb, [120](#), [123](#), [535](#)
- PMIx_Put, [31](#), [69](#), [70](#), [71](#), [71–74](#), [110](#), [113–117](#), [180](#), [235](#), [242](#), [310](#), [311](#), [334](#), [338](#), [459](#), [502](#), [503](#), [535](#)
 - PMIxClient.put (Python), [459](#)
- PMIx_Query_info, [11](#), [94](#), [95](#), [99](#), [100](#), [104](#), [106](#), [108](#), [110](#), [226](#), [228](#), [263](#), [264](#), [387](#), [391](#), [411](#), [433](#), [465](#), [507](#), [514](#), [516](#), [518](#), [533](#), [534](#), [548](#)
 - PMIxClient.query (Python), [465](#)
- PMIx_Query_info_nb, [94](#), [100](#), [100](#), [104](#), [106](#), [108](#), [112](#), [180](#), [296](#), [315](#), [520](#), [536](#), [537](#)
- PMIx_Register_attributes, [314](#), [488](#), [542](#), [543](#)
 - PMIxServer.register_attributes (Python), [488](#)
- PMIx_Register_event_handler, [94](#), [137](#), [175](#), [396](#), [415](#), [471](#), [507](#), [512](#), [518](#), [520](#), [524](#), [525](#), [528](#), [536](#), [543](#), [558](#)
 - PMIxClient.register_event_handler (Python), [471](#)
- pmix_release_cbfunc_t, [55](#), [55](#)
- PMIx_Resolve_nodes, [109](#), [465](#), [535](#)
 - PMIxClient.resolve_nodes (Python), [464](#)
- PMIx_Resolve_peers, [85](#), [89](#), [108](#), [291](#), [292](#), [464](#), [505](#), [535](#)
 - PMIxClient.resolve_peers (Python), [464](#)
- PMIx_Scope_string, [58](#), [477](#), [536](#)
 - PMIxClient.scope_string (Python), [477](#)
- pmix_server_abort_fn_t, [332](#), [445](#)
 - clientaborted (Python), [444](#)
- pmix_server_alloc_fn_t, [365](#), [452](#)
 - allocate (Python), [452](#)
- pmix_server_client_connected2_fn_t, [56](#), [272](#), [307](#), [329](#), [329](#), [330](#), [444](#), [533](#), [544](#), [559](#)

- clientconnected2 (Python), [444](#)
- pmix_server_client_finalized_fn_t, [331](#), [331](#), [444](#)
 - clientfinalized (Python), [444](#)
- PMIx_server_collect_inventory, [321](#), [323](#), [490](#), [538](#)
 - PMIxServer.collect_inventory (Python), [489](#)
- pmix_server_connect_fn_t, [180](#), [349](#), [350](#), [352](#), [448](#)
 - connect (Python), [448](#)
- PMIx_server_define_process_set, [226](#), [326](#), [491](#), [544](#)
 - PMIxServer.define_process_set (Python), [490](#)
- PMIx_server_delete_process_set, [226](#), [326](#), [491](#), [544](#)
 - PMIxServer.delete_process_set (Python), [491](#)
- PMIx_server_deliver_inventory, [322](#), [490](#), [538](#)
 - PMIxServer.deliver_inventory (Python), [490](#)
- PMIx_server_deregister_client, [308](#), [487](#), [535](#)
 - PMIxServer.deregister_client (Python), [486](#)
- pmix_server_deregister_events_fn_t, [354](#), [450](#)
 - deregister_events (Python), [449](#)
- PMIx_server_deregister_nspace, [303](#), [308](#), [485](#), [535](#)
 - PMIxServer.deregister_nspace (Python), [485](#)
- PMIx_server_deregister_resources, [305](#), [486](#), [492](#), [544](#)
 - PMIxServer.deregister_resources (Python), [486](#), [492](#)
- pmix_server_disconnect_fn_t, [350](#), [352](#), [449](#)
 - disconnect (Python), [448](#)
- pmix_server_dmodex_req_fn_t, [74](#), [89](#), [336](#), [336](#), [446](#), [538](#), [539](#)
 - dmodex (Python), [445](#)
- PMIx_server_dmodex_request, [309](#), [310](#), [487](#), [535](#)
 - PMIxServer.dmodex_request (Python), [487](#)
- pmix_server_fabric_fn_t, [260](#), [266](#), [384](#), [456](#), [544](#)
 - fabric (Python), [456](#)
- pmix_server_fencenb_fn_t, [333](#), [335](#), [336](#), [445](#), [540](#)
 - fence (Python), [445](#)
- PMIx_server_finalize, [282](#), [483](#), [535](#)
 - PMIxServer.finalize (Python), [482](#)
- PMIx_server_generate_cpuset_string, [191](#), [324](#), [484](#), [544](#)
 - PMIxServer.generate_cpuset_string (Python), [484](#)
- PMIx_server_generate_locality_string, [187](#), [189](#), [323](#), [484](#), [543](#)
 - PMIxServer.generate_locality_string (Python), [484](#)
- pmix_server_get_cred_fn_t, [373](#), [377](#), [454](#)
 - get_credential (Python), [453](#)
- pmix_server_grp_fn_t, [382](#), [456](#), [544](#)
 - group (Python), [455](#)
- PMIx_server_init, [61](#), [279](#), [283](#), [315](#), [327](#), [388](#), [389](#), [392](#), [482](#), [521](#), [533](#), [535](#), [543](#)
 - PMIxServer.init (Python), [482](#)
- PMIx_server_IOF_deliver, [320](#), [404](#), [489](#), [538](#)
 - PMIxServer.iof_deliver (Python), [489](#)
- pmix_server_iof_fn_t, [378](#), [455](#)
 - iof_pull (Python), [454](#)

- pmix_server_job_control_fn_t, [368](#), [453](#)
 - job_control (Python), [452](#)
- pmix_server_listener_fn_t, [357](#)
- pmix_server_log_fn_t, [363](#), [452](#)
 - log (Python), [451](#)
- pmix_server_lookup_fn_t, [340](#), [447](#)
 - lookup (Python), [446](#)
- pmix_server_module_t, [279](#), [282](#), [315](#), [327](#), [327](#), [328](#), [482](#)
- pmix_server_monitor_fn_t, [370](#), [453](#)
 - monitor (Python), [453](#)
- pmix_server_notify_event_fn_t, [143](#), [147](#), [356](#), [357](#), [450](#)
 - notify_event (Python), [450](#)
- pmix_server_publish_fn_t, [338](#), [446](#)
 - publish (Python), [446](#)
- pmix_server_query_fn_t, [358](#), [451](#)
 - query (Python), [450](#)
- PMIx_server_register_client, [272](#), [306](#), [308](#), [330](#), [331](#), [486](#), [535](#)
 - PMIxServer.register_client (Python), [486](#)
- pmix_server_register_events_fn_t, [352](#), [449](#)
 - register_events (Python), [449](#)
- PMIx_server_register_nspace, [13](#), [56](#), [284](#), [285](#), [285](#), [287](#), [295](#), [296](#), [299](#), [305](#), [320](#), [323](#), [324](#), [485](#), [535](#), [537](#), [549](#)
 - PMIxServer.register_nspace (Python), [484](#)
- PMIx_server_register_resources, [288](#), [291](#), [292](#), [305](#), [485](#), [491](#), [544](#)
 - PMIxServer.register_resources (Python), [485](#), [491](#)
- PMIx_server_setup_application, [311](#), [313](#), [314](#), [320](#), [323](#), [488](#), [536](#), [540](#)
 - PMIxServer.setup_application (Python), [488](#)
- PMIx_server_setup_fork, [308](#), [487](#), [535](#)
 - PMIxServer.setup_fork (Python), [487](#)
- PMIx_server_setup_local_support, [319](#), [489](#), [536](#)
 - PMIxServer.setup_local_support (Python), [488](#)
- pmix_server_spawn_fn_t, [179](#), [344](#), [398](#), [448](#)
 - spawn (Python), [447](#)
- pmix_server_stdin_fn_t, [381](#), [455](#)
 - iof_push (Python), [455](#)
- pmix_server_tool_connection_fn_t, [272](#), [361](#), [388](#), [451](#), [533](#), [534](#)
 - tool_connected (Python), [451](#)
- pmix_server_unpublish_fn_t, [342](#), [447](#)
 - unpublish (Python), [447](#)
- pmix_server_validate_cred_fn_t, [375](#), [454](#)
 - validate_credential (Python), [454](#)
- pmix_setup_application_cbfunc_t, [311](#), [313](#)
- PMIx_Spawn, [84](#), [85](#), [87](#), [88](#), [162](#), [162](#), [167](#), [168](#), [172](#), [173](#), [176](#), [202](#), [204](#), [293](#), [294](#), [309](#), [344](#), [345](#), [347](#), [348](#), [366](#), [392](#), [394](#), [398](#), [399](#), [401](#), [402](#), [404](#), [410–414](#), [417](#), [447](#), [463](#), [507](#), [511](#), [512](#), [516](#), [518](#), [519](#), [521](#), [535](#), [540](#), [550](#), [551](#), [567](#)
 - PMIxClient.spawn (Python), [462](#)
- pmix_spawn_cbfunc_t, [168](#), [179](#), [179](#), [344](#)

PMIx_Spawn_nb, [87](#), [168](#), [176](#), [179](#), [345](#), [535](#)
 PMIx_Store_internal, [70](#), [72](#), [459](#), [535](#)
 PMIxClient.store_internal (Python), [458](#)
 PMIx_tool_attach_to_server, [390](#), [392](#), [401](#), [421](#), [422](#), [493](#), [514](#), [533](#), [544](#), [559](#)
 PMIxTool.attach_to_server (Python), [493](#)
 PMIx_tool_connect_to_server, [538](#), [560](#)
 pmix_tool_connection_cbfunc_t, [361](#), [362](#), [363](#)
 PMIx_tool_disconnect, [422](#), [493](#), [544](#)
 PMIxTool.disconnect (Python), [493](#)
 PMIx_tool_finalize, [421](#), [493](#), [536](#)
 PMIxTool.finalize (Python), [492](#)
 PMIx_tool_get_servers, [424](#), [494](#), [544](#)
 PMIxTool.get_servers (Python), [494](#)
 PMIx_tool_init, [10](#), [61](#), [87](#), [96](#), [101](#), [387](#), [390](#), [392](#), [393](#), [401](#), [402](#), [404](#), [416](#), [418](#), [421](#), [492](#), [507](#), [512](#), [516](#),
 [533](#), [536](#)
 PMIxTool.init (Python), [492](#)
 PMIx_tool_set_server, [389](#), [402](#), [423](#), [424](#), [494](#), [544](#)
 PMIxTool.set_server (Python), [494](#)
 PMIx_Topology_destruct, [190](#), [565](#), [569](#)
 PMIx_Unpublish, [131](#), [132](#), [134](#), [462](#), [535](#)
 PMIxClient.unpublish (Python), [462](#)
 PMIx_Unpublish_nb, [132](#), [535](#)
 PMIx_Validate_credential, [275](#), [468](#), [538](#), [543](#)
 PMIxClient.validate_credential (Python), [468](#)
 PMIx_Validate_credential_nb, [277](#)
 pmix_validation_cbfunc_t, [277](#), [376](#), [377](#)
 pmix_value_cbfunc_t, [56](#), [56](#)
 PMIx_Value_load, [34](#), [565](#), [569](#)
 PMIx_Value_unload, [34](#), [565](#), [569](#)
 PMIx_Value_xfer, [35](#), [565](#), [569](#)

 pmix_evhdlr_reg_cbfunc_t
 (Deprecated), [559](#)
 pmix_server_client_connected_fn_t
 (Deprecated), [329](#), [559](#)
 PMIx_tool_connect_to_server
 (Deprecated), [559](#)

Index of Support Macros

PMIX_APP_CONSTRUCT, [177](#)
PMIX_APP_CREATE, [178](#)
PMIX_APP_DESTRUCT, [178](#)
PMIX_APP_FREE, [178](#)
PMIX_APP_INFO_CREATE, [179](#), [538](#), [539](#)
PMIX_APP_RELEASE, [178](#)
PMIX_APP_STATIC_INIT, [177](#), [566](#)
PMIX_ARGV_APPEND, [49](#)
PMIX_ARGV_APPEND_UNIQUE, [50](#)
PMIX_ARGV_COPY, [52](#)
PMIX_ARGV_COUNT, [52](#)
PMIX_ARGV_FREE, [51](#)
PMIX_ARGV_JOIN, [51](#)
PMIX_ARGV_PREPEND, [50](#)
PMIX_ARGV_SPLIT, [51](#)
PMIX_BYTE_OBJECT_CONSTRUCT, [46](#)
PMIX_BYTE_OBJECT_CREATE, [47](#)
PMIX_BYTE_OBJECT_DESTRUCT, [47](#)
PMIX_BYTE_OBJECT_FREE, [47](#)
PMIX_BYTE_OBJECT_LOAD, [47](#)
PMIX_BYTE_OBJECT_STATIC_INIT, [46](#), [566](#)
PMIX_CHECK_KEY, [20](#)
PMIX_CHECK_NAMESPACE, [21](#)
PMIX_CHECK_PROCID, [25](#)
PMIX_CHECK_RANK, [23](#)
PMIX_CHECK_RESERVED_KEY, [20](#), [559](#)
PMIX_COORD_CONSTRUCT, [256](#)
PMIX_COORD_CREATE, [257](#)
PMIX_COORD_DESTRUCT, [257](#)
PMIX_COORD_FREE, [257](#)
PMIX_COORD_STATIC_INIT, [256](#), [566](#)
PMIX_CPUSET_CONSTRUCT, [325](#)
PMIX_CPUSET_CREATE, [325](#)
PMIX_CPUSET_DESTRUCT, [325](#)
PMIX_CPUSET_FREE, [325](#)
PMIX_CPUSET_STATIC_INIT, [325](#), [566](#)
PMIX_DATA_ARRAY_CONSTRUCT, [48](#)
PMIX_DATA_ARRAY_CREATE, [49](#)
PMIX_DATA_ARRAY_DESTRUCT, [48](#)
PMIX_DATA_ARRAY_FREE, [49](#)
PMIX_DATA_ARRAY_STATIC_INIT, [48](#), [566](#)

PMIX_DATA_BUFFER_CONSTRUCT, [150](#), [152](#), [154](#)
PMIX_DATA_BUFFER_CREATE, [150](#), [152](#), [154](#)
PMIX_DATA_BUFFER_DESTRUCT, [150](#)
PMIX_DATA_BUFFER_LOAD, [151](#)
PMIX_DATA_BUFFER_RELEASE, [150](#)
PMIX_DATA_BUFFER_STATIC_INIT, [150](#), [566](#)
PMIX_DATA_BUFFER_UNLOAD, [151](#), [285](#)
PMIX_DEVICE_DIST_CONSTRUCT, [197](#)
PMIX_DEVICE_DIST_CREATE, [197](#)
PMIX_DEVICE_DIST_DESTRUCT, [197](#)
PMIX_DEVICE_DIST_FREE, [198](#)
PMIX_DEVICE_DIST_STATIC_INIT, [197](#), [566](#)
PMIX_ENDPOINT_CONSTRUCT, [255](#)
PMIX_ENDPOINT_CREATE, [255](#)
PMIX_ENDPOINT_DESTRUCT, [255](#)
PMIX_ENDPOINT_FREE, [255](#)
PMIX_ENDPOINT_STATIC_INIT, [254](#), [566](#)
PMIX_ENVAR_CONSTRUCT, [45](#)
PMIX_ENVAR_CREATE, [45](#)
PMIX_ENVAR_DESTRUCT, [15](#), [45](#)
PMIX_ENVAR_FREE, [45](#)
PMIX_ENVAR_LOAD, [46](#)
PMIX_ENVAR_STATIC_INIT, [45](#), [566](#)
PMIX_FABRIC_CONSTRUCT, [263](#)
PMIX_FABRIC_STATIC_INIT, [263](#), [566](#)
PMIX_GEOMETRY_CONSTRUCT, [258](#)
PMIX_GEOMETRY_CREATE, [259](#)
PMIX_GEOMETRY_DESTRUCT, [258](#)
PMIX_GEOMETRY_FREE, [259](#)
PMIX_GEOMETRY_STATIC_INIT, [258](#), [566](#)
PMIx_Heartbeat, [217](#), [536](#)
PMIX_INFO_CONSTRUCT, [37](#)
PMIX_INFO_CREATE, [37](#), [42](#), [44](#)
PMIX_INFO_DESTRUCT, [37](#)
PMIX_INFO_FREE, [37](#)
PMIX_INFO_IS_END, [44](#), [538](#), [539](#)
PMIX_INFO_IS_OPTIONAL, [43](#)
PMIX_INFO_IS_REQUIRED, [42](#), [43](#)
PMIX_INFO_LIST_ADD, [559](#)
PMIX_INFO_LIST_CONVERT, [559](#)
PMIX_INFO_LIST_RELEASE, [559](#)
PMIX_INFO_LIST_START, [559](#)
PMIX_INFO_LIST_XFER, [559](#)
PMIX_INFO_OPTIONAL, [43](#)
PMIX_INFO_PROCESSED, [43](#), [559](#)
PMIX_INFO_REQUIRED, [41](#), [42](#)
PMIX_INFO_STATIC_INIT, [36](#), [566](#)

PMIX_INFO_TRUE, [39](#)
PMIX_INFO_WAS_PROCESSED, [44](#), [559](#)
PMIX_LOAD_KEY, [20](#)
PMIX_LOAD_NAMESPACE, [22](#)
PMIX_LOAD_PROCID, [25](#), [26](#)
PMIX_LOOKUP_STATIC_INIT, [128](#), [566](#)
PMIX_MULTICLUSTER_NAMESPACE_CONSTRUCT, [27](#)
PMIX_MULTICLUSTER_NAMESPACE_PARSE, [27](#)
PMIX_NAMESPACE_INVALID, [22](#), [563](#)
PMIX_PDATA_CONSTRUCT, [128](#)
PMIX_PDATA_CREATE, [128](#)
PMIX_PDATA_DESTRUCT, [128](#)
PMIX_PDATA_FREE, [129](#)
PMIX_PDATA_LOAD, [129](#)
PMIX_PDATA_RELEASE, [128](#)
PMIX_PDATA_XFER, [130](#)
PMIX_PROC_CONSTRUCT, [24](#)
PMIX_PROC_CREATE, [24](#)
PMIX_PROC_DESTRUCT, [24](#)
PMIX_PROC_FREE, [25](#), [109](#)
PMIX_PROC_INFO_CONSTRUCT, [29](#)
PMIX_PROC_INFO_CREATE, [30](#)
PMIX_PROC_INFO_DESTRUCT, [30](#)
PMIX_PROC_INFO_FREE, [30](#)
PMIX_PROC_INFO_RELEASE, [30](#)
PMIX_PROC_INFO_STATIC_INIT, [29](#), [566](#)
PMIX_PROC_LOAD, [25](#)
PMIX_PROC_RELEASE, [25](#)
PMIX_PROC_STATIC_INIT, [24](#), [566](#)
PMIX_PROCID_INVALID, [26](#), [563](#)
PMIX_PROCID_XFER, [27](#), [563](#)
PMIX_QUERY_CONSTRUCT, [107](#)
PMIX_QUERY_CREATE, [107](#)
PMIX_QUERY_DESTRUCT, [107](#)
PMIX_QUERY_FREE, [108](#)
PMIX_QUERY_QUALIFIERS_CREATE, [108](#), [538](#), [539](#)
PMIX_QUERY_RELEASE, [108](#)
PMIX_QUERY_STATIC_INIT, [107](#), [566](#)
PMIX_RANK_IS_VALID, [23](#), [563](#)
PMIX_REGATTR_CONSTRUCT, [317](#)
PMIX_REGATTR_CREATE, [318](#)
PMIX_REGATTR_DESTRUCT, [317](#)
PMIX_REGATTR_FREE, [318](#)
PMIX_REGATTR_LOAD, [318](#)
PMIX_REGATTR_STATIC_INIT, [317](#), [566](#)
PMIX_REGATTR_XFER, [319](#)
PMIX_SETENV, [52](#)

PMIX_SYSTEM_EVENT, [140](#)
PMIX_TOPOLOGY_CONSTRUCT, [190](#)
PMIX_TOPOLOGY_CREATE, [190](#)
PMIX_TOPOLOGY_STATIC_INIT, [189](#), [566](#)
PMIX_VALUE_CONSTRUCT, [33](#)
PMIX_VALUE_CREATE, [33](#)
PMIX_VALUE_DESTRUCT, [33](#), [76](#), [80](#), [547](#)
PMIX_VALUE_FREE, [34](#)
PMIX_VALUE_GET_NUMBER, [36](#)
PMIX_VALUE_RELEASE, [33](#)
PMIX_VALUE_STATIC_INIT, [32](#), [566](#)

PMIX_INFO_LIST_ADD
(Deprecated), [569](#)
PMIX_INFO_LIST_CONVERT
(Deprecated), [569](#)
PMIX_INFO_LIST_RELEASE
(Deprecated), [569](#)
PMIX_INFO_LIST_START
(Deprecated), [569](#)
PMIX_INFO_LIST_XFER
(Deprecated), [569](#)
PMIX_INFO_LOAD
(Deprecated), [569](#)
PMIX_INFO_XFER
(Deprecated), [569](#)
PMIX_TOPOLOGY_DESTRUCT
(Deprecated), [569](#)
PMIX_TOPOLOGY_FREE
(Deprecated), [569](#)
PMIX_VALUE_LOAD
(Deprecated), [569](#)
PMIX_VALUE_UNLOAD
(Deprecated), [569](#)
PMIX_VALUE_XFER
(Deprecated), [569](#)

Index of Data Structures

pmix_alloc_directive_t, 54, 59, 199, 202, [205](#), 205, 365, 439, 479
pmix_app_t, 49, 50, 53, 163–165, 168, 170, 174, [176](#), 176–179, 344, 345, 347, 394, 397, 399, 401, 402, 410, 413, 417, 440, 509–511, 515, 516, 519, 538, 539, 557
pmix_bind_envelope_t, [192](#), 192, 441, 544
pmix_byte_object_t, [46](#), 46, 47, 54, 156, 157, 159, 273, 274, 276, 277, 320, 375, 376, 381, 429, 438
pmix_coord_t, 54, [256](#), 256–258, 441, 544
pmix_coord_view_t, [259](#), 441, 544
pmix_cpuset_t, 54, 193, 194, 323, [324](#), 324, 325, 440, 544
pmix_data_array_t, 31, 40, 41, [48](#), 48, 49, 54, 85, 98, 99, 102–105, 111, 201, 203, 205, 228, 231, 261, 262, 264–267, 287, 288, 290–292, 299–301, 312, 360, 367, 386, 397, 411, 417, 418, 439, 511, 515, 516, 518, 531, 538, 539, 554–556, 568
pmix_data_buffer_t, [149](#), 149–153, 156, 157
pmix_data_range_t, 54, 58, [123](#), 123, 145, 356, 438, 478
pmix_data_type_t, 34, 36, 38, 40, 48, 49, [53](#), 53, 54, 58, 129, 152, 153, 155, 318, 437, 478
pmix_device_distance_t, 54, 193, 195, [196](#), 196–198, 294, 441, 544, 556
pmix_device_type_t, 54, 60, [195](#), 195, 266, 441, 481, 544
pmix_endpoint_t, 54, [254](#), 254, 255, 266, 440, 554
pmix_envar_t, 15, [44](#), 44–46, 54, 439, 521
pmix_fabric_operation_t, [260](#), 260, 385
pmix_fabric_t, 254, [260](#), 260, 263, 264, 267–271, 386, 440, 544, 553
pmix_geometry_t, 54, 253, [257](#), 257–259, 265, 441, 544, 554, 555
pmix_group_operation_t, 382, [384](#), 384, 544
pmix_group_opt_t, 245, 247, [248](#), 248, 470, 544
pmix_info_directives_t, [41](#), 41, 42, 54, 58, 439, 478
pmix_info_t, 4, 5, 11, 20, [36](#), 36–44, 54, 56, 57, 62, 64, 66, 95, 99, 105, 106, 108, 111, 120, 122–125, 142, 145, 148, 179, 193, 194, 199–203, 205, 206, 208, 211–213, 217, 220, 223, 225, 233, 234, 236, 238, 239, 241, 243, 245, 247–249, 260, 262, 265, 266, 273, 274, 276, 277, 279, 282, 286–288, 290, 291, 295, 296, 299–301, 312, 316, 318, 320–322, 327, 329, 347, 356, 361, 362, 365, 367, 368, 370, 371, 377, 378, 380, 386, 394, 397, 399, 401, 403, 410, 417, 419, 422, 425, 426, 428, 429, 439, 442, 509–511, 514–516, 519, 526, 536, 538, 539, 548–550, 553–555
pmix_iof_channel_t, 54, 59, 320, 378, 380, [408](#), 408, 426, 439, 479
pmix_job_state_t, [31](#), 31, 54, 59, 440, 480, 544
pmix_key_t, 11, [19](#), 19, 71, 75, 318, 437
pmix_link_state_t, 54, 59, 253, [260](#), 260, 262, 265, 440, 481, 544, 555
pmix_locality_t, 54, 189, [191](#), 191, 440, 544, 561
pmix_nspace_t, [21](#), 21, 22, 25–27, 54, 179, 437, 438
pmix_pdata_t, 124, 125, [127](#), 127–130, 439
pmix_persistence_t, 54, 58, [123](#), 123, 438, 477
pmix_proc_info_t, [29](#), 29, 30, 54, 96, 98, 101–104, 360, 411, 418, 438, 511, 515, 516, 518
pmix_proc_state_t, [28](#), 28, 54, 58, 438, 476
pmix_proc_t, 22, [23](#), 23–27, 53, 64, 77, 89, 105, 114–116, 129, 139, 141, 142, 145, 146, 152, 153, 161, 162, 228, 231–233, 236, 241, 243, 246, 292, 307–309, 318, 320, 326, 329–333, 337, 338, 340, 342, 344,

349, 351, 356, 358, 363–365, 368, 371, 373, 376, 378, 380–382, 384, 385, 419, 421, 422, 424, 425, 438, 531, 556

pmix_query_t, 54, [94](#), 94, 97, 102, 106–108, 110, 358, 360, 440, 538, 539, 548

pmix_rank_t, [22](#), 22, 23, 25, 26, 54, 397, 411, 417, 438, 511, 568

pmix_regattr_t, 54, 111, 315, [316](#), 316–319, 440, 542, 544, 550

pmix_scope_t, 54, 58, [72](#), 72, 438, 477

pmix_status_t, [17](#), 17, 36, 49, 50, 52, 53, 56, 57, 117, 137, 140, 142, 145, 148, 195, 310, 314, 334, 350, 351, 353, 355, 356, 363, 375, 377, 383, 437, 450, 476, 526, 567

pmix_storage_access_type_t, 55, [433](#), 433, 562

pmix_storage_accessibility_t, 55, [432](#), 432, 562

pmix_storage_medium_t, 54, [431](#), 431, 432, 562

pmix_storage_persistence_t, 55, [432](#), 432, 562

pmix_topology_t, 54, 188, [189](#), 189, 190, 193, 194, 544

pmix_value_t, 7, 11, [31](#), 31–36, 53, 56, 71, 76, 77, 80, 439, 547

Index of Constants

PMIX_ALLOC_DIRECTIVE, [54](#)
PMIX_ALLOC_EXTEND, [205](#)
PMIX_ALLOC_EXTERNAL, [205](#)
PMIX_ALLOC_NEW, [205](#)
PMIX_ALLOC_REQUIRE, [205](#)
PMIX_ALLOC_RELEASE, [205](#)
PMIX_APP, [54](#)
PMIX_APP_WILDCARD, [16](#)
PMIX_BOOL, [53](#)
PMIX_BYTE, [53](#)
PMIX_BYTE_OBJECT, [54](#)
PMIX_COMMAND, [54](#)
PMIX_COMPRESSED_BYTE_OBJECT, [54](#)
PMIX_COMPRESSED_STRING, [54](#)
PMIX_COORD, [54](#)
PMIX_COORD_LOGICAL_VIEW, [259](#)
PMIX_COORD_PHYSICAL_VIEW, [259](#)
PMIX_COORD_VIEW_UNDEF, [259](#)
PMIX_CPUBIND_PROCESS, [192](#)
PMIX_CPUBIND_THREAD, [192](#)
PMIX_DATA_ARRAY, [54](#)
PMIX_DATA_RANGE, [54](#)
PMIX_DATA_TYPE, [54](#)
PMIX_DATA_TYPE_MAX, [55](#)
PMIX_DEBUGGER_RELEASE, [416](#)
PMIX_DEVICE_DIST, [54](#)
PMIX_DEVTYPE, [54](#)
PMIX_DEVTYPE_BLOCK, [195](#)
PMIX_DEVTYPE_COPROC, [196](#)
PMIX_DEVTYPE_DMA, [196](#)
PMIX_DEVTYPE_GPU, [195](#)
PMIX_DEVTYPE_NETWORK, [195](#)
PMIX_DEVTYPE_OPENFABRICS, [196](#)
PMIX_DEVTYPE_UNKNOWN, [195](#)
PMIX_DOUBLE, [53](#)
PMIX_ENDPOINT, [54](#)
PMIX_ENVAR, [54](#)
PMIX_ERR_BAD_PARAM, [18](#)
PMIX_ERR_COMM_FAILURE, [18](#)
PMIX_ERR_CONFLICTING_CLEANUP_DIRECTIVES, [211](#)
PMIX_ERR_DUPLICATE_KEY, [122](#)

PMIX_ERR_EMPTY, [18](#)
PMIX_ERR_EVENT_REGISTRATION, [140](#)
PMIX_ERR_EXISTS, [17](#)
PMIX_ERR_EXISTS_OUTSIDE_SCOPE, [17](#)
PMIX_ERR_INIT, [18](#)
PMIX_ERR_INVALID_CRED, [17](#)
PMIX_ERR_INVALID_OPERATION, [18](#)
PMIX_ERR_IOF_COMPLETE, [408](#)
PMIX_ERR_IOF_FAILURE, [408](#)
PMIX_ERR_JOB_ABORTED, [416](#)
PMIX_ERR_JOB_ABORTED_BY_SIG, [416](#)
PMIX_ERR_JOB_ABORTED_BY_SYS_EVENT, [416](#)
PMIX_ERR_JOB_ALLOC_FAILED, [173](#)
PMIX_ERR_JOB_APP_NOT_EXECUTABLE, [173](#)
PMIX_ERR_JOB_CANCELED, [416](#)
PMIX_ERR_JOB_EXE_NOT_FOUND, [173](#)
PMIX_ERR_JOB_FAILED_TO_LAUNCH, [173](#)
PMIX_ERR_JOB_FAILED_TO_MAP, [173](#)
PMIX_ERR_JOB_INSUFFICIENT_RESOURCES, [173](#)
PMIX_ERR_JOB_KILLED_BY_CMD, [416](#)
PMIX_ERR_JOB_NO_EXE_SPECIFIED, [173](#)
PMIX_ERR_JOB_NON_ZERO_TERM, [416](#)
PMIX_ERR_JOB_SENSOR_BOUND_EXCEEDED, [416](#)
PMIX_ERR_JOB_SYS_OP_FAILED, [173](#)
PMIX_ERR_JOB_TERM_WO_SYNC, [416](#)
PMIX_ERR_JOB_WDIR_NOT_FOUND, [173](#)
PMIX_ERR_LOST_CONNECTION, [18](#)
PMIX_ERR_NO_PERMISSIONS, [18](#)
PMIX_ERR_NOMEM, [18](#)
PMIX_ERR_NOT_FOUND, [18](#)
PMIX_ERR_NOT_SUPPORTED, [18](#)
PMIX_ERR_OUT_OF_RESOURCE, [18](#)
PMIX_ERR_PACK_FAILURE, [18](#)
PMIX_ERR_PARAM_VALUE_NOT_SUPPORTED, [18](#)
PMIX_ERR_PARTIAL_SUCCESS, [19](#)
PMIX_ERR_PROC_CHECKPOINT, [211](#)
PMIX_ERR_PROC_MIGRATE, [211](#)
PMIX_ERR_PROC_RESTART, [211](#)
PMIX_ERR_PROC_TERM_WO_SYNC, [416](#)
PMIX_ERR_REPEAT_ATTR_REGISTRATION, [316](#)
PMIX_ERR_RESOURCE_BUSY, [18](#)
PMIX_ERR_TIMEOUT, [18](#)
PMIX_ERR_TYPE_MISMATCH, [17](#)
PMIX_ERR_UNKNOWN_DATA_TYPE, [17](#)
PMIX_ERR_UNPACK_FAILURE, [18](#)
PMIX_ERR_UNPACK_INADEQUATE_SPACE, [17](#)
PMIX_ERR_UNPACK_READ_PAST_END_OF_BUFFER, [18](#)

PMIX_ERR_UNREACH, [18](#)
PMIX_ERR_WOULD_BLOCK, [17](#)
PMIX_ERROR, [17](#)
PMIX_EVENT_ACTION_COMPLETE, [148](#)
PMIX_EVENT_ACTION_DEFERRED, [148](#)
PMIX_EVENT_JOB_END, [415](#)
PMIX_EVENT_JOB_START, [415](#)
PMIX_EVENT_NO_ACTION_TAKEN, [148](#)
PMIX_EVENT_NODE_DOWN, [140](#)
PMIX_EVENT_NODE_OFFLINE, [140](#)
PMIX_EVENT_PARTIAL_ACTION_TAKEN, [148](#)
PMIX_EVENT_PROC_TERMINATED, [415](#)
PMIX_EVENT_SESSION_END, [415](#)
PMIX_EVENT_SESSION_START, [415](#)
PMIX_EVENT_SYS_BASE, [140](#)
PMIX_EVENT_SYS_OTHER, [140](#)
PMIX_EXTERNAL_ERR_BASE, [19](#)
PMIX_FABRIC_REQUEST_INFO, [260](#)
PMIX_FABRIC_UPDATE_ENDPOINTS, [254](#)
PMIX_FABRIC_UPDATE_INFO, [260](#)
PMIX_FABRIC_UPDATE_PENDING, [254](#)
PMIX_FABRIC_UPDATED, [254](#)
PMIX_FLOAT, [53](#)
PMIX_FWD_ALL_CHANNELS, [408](#)
PMIX_FWD_NO_CHANNELS, [408](#)
PMIX_FWD_STDDIAG_CHANNEL, [408](#)
PMIX_FWD_STDERR_CHANNEL, [408](#)
PMIX_FWD_STDIN_CHANNEL, [408](#)
PMIX_FWD_STDOUT_CHANNEL, [408](#)
PMIX_GEOMETRY, [54](#)
PMIX_GLOBAL, [72](#)
PMIX_GROUP_ACCEPT, [248](#)
PMIX_GROUP_CONSTRUCT, [384](#)
PMIX_GROUP_CONSTRUCT_ABORT, [231](#)
PMIX_GROUP_CONSTRUCT_COMPLETE, [231](#)
PMIX_GROUP_CONTEXT_ID_ASSIGNED, [231](#)
PMIX_GROUP_DECLINE, [248](#)
PMIX_GROUP_DESTRUCT, [384](#)
PMIX_GROUP_INVITE_ACCEPTED, [230](#)
PMIX_GROUP_INVITE_DECLINED, [230](#)
PMIX_GROUP_INVITE_FAILED, [230](#)
PMIX_GROUP_INVITED, [230](#)
PMIX_GROUP_LEADER_FAILED, [231](#)
PMIX_GROUP_LEADER_SELECTED, [231](#)
PMIX_GROUP_LEFT, [230](#)
PMIX_GROUP_MEMBER_FAILED, [230](#)
PMIX_GROUP_MEMBERSHIP_UPDATE, [230](#)

PMIX_INFO, [54](#)
PMIX_INFO_ARRAY_END, [42](#)
PMIX_INFO_DIR_RESERVED, [42](#)
PMIX_INFO_DIRECTIVES, [54](#)
PMIX_INFO_REQD, [42](#)
PMIX_INFO_REQD_PROCESSED, [42](#)
PMIX_INT, [53](#)
PMIX_INT16, [53](#)
PMIX_INT32, [53](#)
PMIX_INT64, [53](#)
PMIX_INT8, [53](#)
PMIX_INTERNAL, [72](#)
PMIX_IOF_CHANNEL, [54](#)
PMIX_JCTRL_CHECKPOINT, [211](#)
PMIX_JCTRL_CHECKPOINT_COMPLETE, [211](#)
PMIX_JCTRL_PREEMPT_ALERT, [211](#)
PMIX_JOB_STATE, [54](#)
PMIX_JOB_STATE_AWAITING_ALLOC, [31](#)
PMIX_JOB_STATE_CONNECTED, [31](#)
PMIX_JOB_STATE_LAUNCH_UNDERWAY, [31](#)
PMIX_JOB_STATE_RUNNING, [31](#)
PMIX_JOB_STATE_SUSPENDED, [31](#)
PMIX_JOB_STATE_TERMINATED, [31](#)
PMIX_JOB_STATE_TERMINATED_WITH_ERROR, [31](#)
PMIX_JOB_STATE_UNDEF, [31](#)
PMIX_JOB_STATE_UNTERMINATED, [31](#)
PMIX_KVAL, [54](#)
PMIX_LAUNCH_COMPLETE, [415](#)
PMIX_LAUNCHER_READY, [404](#)
PMIX_LINK_DOWN, [260](#)
PMIX_LINK_STATE, [54](#)
PMIX_LINK_STATE_UNKNOWN, [260](#)
PMIX_LINK_UP, [260](#)
PMIX_LOCAL, [72](#)
PMIX_LOCALITY_NONLOCAL, [191](#)
PMIX_LOCALITY_SHARE_CORE, [191](#)
PMIX_LOCALITY_SHARE_HWTHREAD, [191](#)
PMIX_LOCALITY_SHARE_L1CACHE, [191](#)
PMIX_LOCALITY_SHARE_L2CACHE, [191](#)
PMIX_LOCALITY_SHARE_L3CACHE, [191](#)
PMIX_LOCALITY_SHARE_NODE, [191](#)
PMIX_LOCALITY_SHARE_NUMA, [191](#)
PMIX_LOCALITY_SHARE_PACKAGE, [191](#)
PMIX_LOCALITY_UNKNOWN, [191](#)
PMIX_LOCTYPE, [54](#)
PMIX_MAX_KEYLEN, [16](#)
PMIX_MAX_NSLEN, [16](#)

PMIX_MODEL_DECLARED, [65](#)
PMIX_MODEL_RESOURCES, [65](#)
PMIX_MONITOR_FILE_ALERT, [217](#)
PMIX_MONITOR_HEARTBEAT_ALERT, [217](#)
PMIX_OPENMP_PARALLEL_ENTERED, [65](#)
PMIX_OPENMP_PARALLEL_EXITED, [65](#)
PMIX_OPERATION_IN_PROGRESS, [18](#)
PMIX_OPERATION_SUCCEEDED, [18](#)
PMIX_PDATA, [54](#)
PMIX_PERSIST, [54](#)
PMIX_PERSIST_APP, [123](#)
PMIX_PERSIST_FIRST_READ, [123](#)
PMIX_PERSIST_INDEF, [123](#)
PMIX_PERSIST_INVALID, [123](#)
PMIX_PERSIST_PROC, [123](#)
PMIX_PERSIST_SESSION, [123](#)
PMIX_PID, [53](#)
PMIX_POINTER, [54](#)
PMIX_PROC, [53](#)
PMIX_PROC_CPUSET, [54](#)
PMIX_PROC_INFO, [54](#)
PMIX_PROC_NSSPACE, [54](#)
PMIX_PROC_RANK, [54](#)
PMIX_PROC_STATE, [54](#)
PMIX_PROC_STATE_ABORTED, [28](#)
PMIX_PROC_STATE_ABORTED_BY_SIG, [28](#)
PMIX_PROC_STATE_CALLED_ABORT, [28](#)
PMIX_PROC_STATE_CANNOT_RESTART, [28](#)
PMIX_PROC_STATE_COMM_FAILED, [28](#)
PMIX_PROC_STATE_CONNECTED, [28](#)
PMIX_PROC_STATE_ERROR, [28](#)
PMIX_PROC_STATE_FAILED_TO_LAUNCH, [28](#)
PMIX_PROC_STATE_FAILED_TO_START, [28](#)
PMIX_PROC_STATE_HEARTBEAT_FAILED, [28](#)
PMIX_PROC_STATE_KILLED_BY_CMD, [28](#)
PMIX_PROC_STATE_LAUNCH_UNDERWAY, [28](#)
PMIX_PROC_STATE_MIGRATING, [28](#)
PMIX_PROC_STATE_PREPPED, [28](#)
PMIX_PROC_STATE_RESTART, [28](#)
PMIX_PROC_STATE_RUNNING, [28](#)
PMIX_PROC_STATE_SENSOR_BOUND_EXCEEDED, [28](#)
PMIX_PROC_STATE_TERM_NON_ZERO, [28](#)
PMIX_PROC_STATE_TERM_WO_SYNC, [28](#)
PMIX_PROC_STATE_TERMINATE, [28](#)
PMIX_PROC_STATE_TERMINATED, [28](#)
PMIX_PROC_STATE_UNDEF, [28](#)
PMIX_PROC_STATE_UNTERMINATED, [28](#)

PMIX_PROCESS_SET_DEFINE, [227](#)
PMIX_PROCESS_SET_DELETE, [227](#)
PMIX_QUERY, [54](#)
PMIX_RANGE_CUSTOM, [123](#)
PMIX_RANGE_GLOBAL, [123](#)
PMIX_RANGE_INVALID, [123](#)
PMIX_RANGE_LOCAL, [123](#)
PMIX_RANGE_NAMESPACE, [123](#)
PMIX_RANGE_PROC_LOCAL, [123](#)
PMIX_RANGE_RM, [123](#)
PMIX_RANGE_SESSION, [123](#)
PMIX_RANGE_UNDEF, [123](#)
PMIX_RANK_INVALID, [23](#)
PMIX_RANK_LOCAL_NODE, [22](#)
PMIX_RANK_LOCAL_PEERS, [23](#)
PMIX_RANK_UNDEF, [22](#)
PMIX_RANK_VALID, [23](#)
PMIX_RANK_WILDCARD, [22](#)
PMIX_READY_FOR_DEBUG, [416](#)
PMIX_REGATTR, [54](#)
PMIX_REGEX, [54](#)
PMIX_REMOTE, [72](#)
PMIX_SCOPE, [54](#)
PMIX_SCOPE_UNDEF, [72](#)
PMIX_SIZE, [53](#)
PMIX_STATUS, [53](#)
PMIX_STOR_ACCESS, [55](#)
PMIX_STOR_ACCESS_TYPE, [55](#)
PMIX_STOR_MEDIUM, [54](#)
PMIX_STOR_PERSIST, [55](#)
PMIX_STORAGE_ACCESS_RD, [433](#)
PMIX_STORAGE_ACCESS_RDWR, [433](#)
PMIX_STORAGE_ACCESS_WR, [433](#)
PMIX_STORAGE_ACCESSIBILITY_CLUSTER, [432](#)
PMIX_STORAGE_ACCESSIBILITY_JOB, [432](#)
PMIX_STORAGE_ACCESSIBILITY_NODE, [432](#)
PMIX_STORAGE_ACCESSIBILITY_RACK, [432](#)
PMIX_STORAGE_ACCESSIBILITY_REMOTE, [432](#)
PMIX_STORAGE_ACCESSIBILITY_SESSION, [432](#)
PMIX_STORAGE_MEDIUM_HDD, [431](#)
PMIX_STORAGE_MEDIUM_NVME, [431](#)
PMIX_STORAGE_MEDIUM_PMEM, [431](#)
PMIX_STORAGE_MEDIUM_RAM, [431](#)
PMIX_STORAGE_MEDIUM_SSD, [431](#)
PMIX_STORAGE_MEDIUM_TAPE, [431](#)
PMIX_STORAGE_MEDIUM_UNKNOWN, [431](#)
PMIX_STORAGE_PERSISTENCE_ARCHIVE, [432](#)

PMIX_STORAGE_PERSISTENCE_JOB, [432](#)
PMIX_STORAGE_PERSISTENCE_NODE, [432](#)
PMIX_STORAGE_PERSISTENCE_PROJECT, [432](#)
PMIX_STORAGE_PERSISTENCE_SCRATCH, [432](#)
PMIX_STORAGE_PERSISTENCE_SESSION, [432](#)
PMIX_STORAGE_PERSISTENCE_TEMPORARY, [432](#)
PMIX_STRING, [53](#)
PMIX_SUCCESS, [17](#)
PMIX_TIME, [53](#)
PMIX_TIMEVAL, [53](#)
PMIX_TOPO, [54](#)
PMIX_UINT, [53](#)
PMIX_UINT16, [53](#)
PMIX_UINT32, [53](#)
PMIX_UINT64, [53](#)
PMIX_UINT8, [53](#)
PMIX_UNDEF, [53](#)
PMIX_VALUE, [53](#)

PMIX_BUFFER

Removed, [562](#)

PMIX_CONNECT_REQUESTED

Deprecated, [559](#)

PMIX_DEBUG_WAITING_FOR_NOTIFY

Deprecated, [568](#)

PMIX_ERR_DATA_VALUE_NOT_FOUND

Deprecated, [540](#)

Removed, [560](#)

PMIX_ERR_DEBUGGER_RELEASE

Deprecated, [559](#)

PMIX_ERR_HANDSHAKE_FAILED

Deprecated, [540](#)

Removed, [559](#)

PMIX_ERR_IN_ERRNO

Deprecated, [540](#)

Removed, [559](#)

PMIX_ERR_INVALID_ARG

Deprecated, [540](#)

Removed, [560](#)

PMIX_ERR_INVALID_ARGS

Deprecated, [540](#)

Removed, [560](#)

PMIX_ERR_INVALID_KEY

Deprecated, [540](#)

Removed, [560](#)

PMIX_ERR_INVALID_KEY_LENGTH

Deprecated, [540](#)

Removed, [560](#)
PMIX_ERR_INVALID_KEYVALP
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_LENGTH
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_NAMESPACE
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_NUM_ARGS
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_NUM_PARSED
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_SIZE
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_TERMINATION
Deprecated, [559](#)
PMIX_ERR_INVALID_VAL
Deprecated, [540](#)
Removed, [560](#)
PMIX_ERR_INVALID_VAL_LENGTH
Deprecated, [541](#)
Removed, [559](#)
PMIX_ERR_JOB_TERMINATED
Deprecated, [559](#)
PMIX_ERR_LOST_CONNECTION_TO_CLIENT
Deprecated, [559](#)
PMIX_ERR_LOST_CONNECTION_TO_SERVER
Deprecated, [559](#)
PMIX_ERR_LOST_PEER_CONNECTION
Deprecated, [559](#)
PMIX_ERR_NODE_DOWN
Deprecated, [559](#)
PMIX_ERR_NODE_OFFLINE
Deprecated, [559](#)
PMIX_ERR_NOT_IMPLEMENTED
Deprecated, [541](#)
Removed, [560](#)
PMIX_ERR_PACK_MISMATCH
Deprecated, [541](#)
Removed, [560](#)
PMIX_ERR_PROC_ABORTED
Deprecated, [559](#)

PMIX_ERR_PROC_ABORTING
Deprecated, [559](#)

PMIX_ERR_PROC_ENTRY_NOT_FOUND
Deprecated, [541](#)
Removed, [560](#)

PMIX_ERR_PROC_REQUESTED_ABORT
Deprecated, [541](#)
Removed, [560](#)

PMIX_ERR_READY_FOR_HANDSHAKE
Deprecated, [541](#)
Removed, [559](#)

PMIX_ERR_SERVER_FAILED_REQUEST
Deprecated, [541](#)
Removed, [560](#)

PMIX_ERR_SERVER_NOT_AVAIL
Deprecated, [541](#)
Removed, [560](#)

PMIX_ERR_SILENT
Deprecated, [541](#)
Removed, [560](#)

PMIX_ERR_SYS_OTHER
Deprecated, [559](#)

PMIX_EXISTS
Deprecated, [559](#)

PMIX_GDS_ACTION_COMPLETE
Deprecated, [541](#)
Removed, [560](#)

PMIX_INFO_ARRAY
Deprecated, [536](#)

PMIX_MODEX
Deprecated, [536](#)

PMIX_NOTIFY_ALLOC_COMPLETE
Deprecated, [541](#)
Removed, [560](#)

PMIX_PROC_HAS_CONNECTED
Deprecated, [559](#)

PMIX_PROC_TERMINATED
Deprecated, [559](#)

PMIX_QUERY_PARTIAL_SUCCESS
Removed, [569](#)

Index of Environmental Variables

PMIX_KEEPALIVE_PIPE, [392](#), 401, 402, 558

PMIX_LAUNCHER_RNDZ_FILE, 389, [392](#), 558

PMIX_LAUNCHER_RNDZ_URI, [392](#), 401, 402, 516, 558

Index of Attributes

PMIX_ACCESS_GRPIDS, [122](#), [550](#)
PMIX_ACCESS_PERMISSIONS, [120](#), [122](#), [122](#), [550](#)
PMIX_ACCESS_USERIDS, [122](#), [550](#)
PMIX_ADD_ENVAR, [165](#), [171](#), [176](#), [522](#)
PMIX_ADD_HOST, [164](#), [169](#), [173](#), [346](#)
PMIX_ADD_HOSTFILE, [164](#), [169](#), [173](#), [346](#)
PMIX_ALL_CLONES_PARTICIPATE, [114](#), [117](#), [118](#), [181](#), [183](#), [185](#), [187](#), [547](#)
PMIX_ALLOC_BANDWIDTH, [166](#), [171](#), [201](#), [203](#), [205](#), [205](#), [312](#), [367](#)
PMIX_ALLOC_CPU_LIST, [166](#), [171](#), [200](#), [203](#), [204](#), [367](#)
PMIX_ALLOC_FABRIC, [200](#), [203](#), [204](#), [311](#), [367](#), [560](#)
PMIX_ALLOC_FABRIC_ENDPTS, [166](#), [172](#), [200](#), [201](#), [203](#), [204](#), [205](#), [205](#), [312](#), [367](#), [560](#)
PMIX_ALLOC_FABRIC_ENDPTS_NODE, [166](#), [172](#), [201](#), [204](#), [205](#), [312](#), [561](#)
PMIX_ALLOC_FABRIC_ID, [200](#), [203](#), [205](#), [205](#), [312](#), [367](#), [560](#)
PMIX_ALLOC_FABRIC_PLANE, [166](#), [171](#), [201](#), [203](#), [204](#), [205](#), [205](#), [312](#), [367](#), [560](#)
PMIX_ALLOC_FABRIC_QOS, [166](#), [171](#), [201](#), [203](#), [205](#), [205](#), [312](#), [367](#), [560](#)
PMIX_ALLOC_FABRIC_SEC_KEY, [201](#), [203](#), [204](#), [205](#), [205](#), [312](#), [367](#), [561](#)
PMIX_ALLOC_FABRIC_TYPE, [166](#), [171](#), [200](#), [201](#), [203](#), [204](#), [205](#), [205](#), [312](#), [367](#), [560](#)
PMIX_ALLOC_ID, [202](#), [204](#), [366](#), [540](#), [550](#)
PMIX_ALLOC_MEM_SIZE, [166](#), [171](#), [200](#), [203](#), [204](#), [367](#)
PMIX_ALLOC_NODE_LIST, [166](#), [171](#), [200](#), [203](#), [204](#), [367](#)
PMIX_ALLOC_NUM_CPU_LIST, [166](#), [171](#), [200](#), [203](#), [204](#), [367](#)
PMIX_ALLOC_NUM_CPUS, [166](#), [171](#), [200](#), [203](#), [204](#), [366](#)
PMIX_ALLOC_NUM_NODES, [166](#), [171](#), [200](#), [203](#), [204](#), [366](#)
PMIX_ALLOC_QUEUE, [98](#), [102](#), [104](#), [166](#), [171](#), [204](#), [359](#), [550](#)
PMIX_ALLOC_REQ_ID, [200](#), [202](#), [204](#), [540](#)
PMIX_ALLOC_TIME, [166](#), [171](#), [200](#), [203](#), [205](#), [366](#)
PMIX_ALLOCATED_NODELIST, [83](#), [288](#)
PMIX_ANL_MAP, [83](#), [84](#), [289](#)
PMIX_APP_ARGV, [85](#), [86](#), [290](#), [551](#)
PMIX_APP_INFO, [76](#), [78](#), [81](#), [85](#), [86](#), [96](#), [101](#), [290](#), [291](#), [505](#)
PMIX_APP_INFO_ARRAY, [287](#), [290](#), [295](#), [295](#), [300](#), [549](#)
PMIX_APP_MAP_REGEX, [86](#), [290](#)
PMIX_APP_MAP_TYPE, [86](#), [290](#)
PMIX_APP_RANK, [87](#), [292](#)
PMIX_APP_SIZE, [85](#), [290](#), [300](#)
PMIX_APPEND_ENVAR, [166](#), [171](#), [176](#), [522](#)
PMIX_APPLDR, [85](#), [290](#), [300](#)
PMIX_APPNUM, [76](#), [78](#), [81](#), [82](#), [85](#), [87](#), [96](#), [101](#), [287](#), [290–292](#), [295](#), [300](#), [505](#), [549](#)
PMIX_ATTR_UNDEF, [5](#)
PMIX_AVAIL_PHYS_MEMORY, [89](#), [105](#), [292](#)
PMIX_BINDTO, [164](#), [169](#), [174](#), [289](#), [346](#)

PMIX_BREAKPOINT, 397, 411, 416, **417**, 417, 511, 568
 PMIX_CLEANUP_EMPTY, 207, 210, **213**
 PMIX_CLEANUP_IGNORE, 207, 210, **213**
 PMIX_CLEANUP_LEAVE_TOPDIR, 207, 210, **213**
 PMIX_CLEANUP_RECURSIVE, 207, 210, **212**
 PMIX_CLIENT_ATTRIBUTES, 97, 102, **106**, 111, 411, 542, 548
 PMIX_CLIENT_AVG_MEMORY, 99, 103, **105**
 PMIX_CLIENT_FUNCTIONS, 97, 102, 105, **106**, 111, 548
 PMIX_CLUSTER_ID, **82**, 288
 PMIX_CMD_LINE, **85**, 551, 565
 PMIX_COLLECT_DATA, 70, 113, 114, 116, **117**, 334
 PMIX_COLLECT_GENERATED_JOB_INFO, 114, 116, 117, **118**, 252, 334, 547
 PMIX_COLLECTIVE_ALGO, 537
 PMIX_CONNECT_MAX_RETRIES, **393**, 420
 PMIX_CONNECT_RETRY_DELAY, **393**, 420
 PMIX_CONNECT_SYSTEM_FIRST, 390, **393**, 419, 423
 PMIX_CONNECT_TO_SYSTEM, 390, **393**, 419, 423
 PMIX_COSPAWN_APP, 166, 172, **417**, 510, 514
 PMIX_CPU_LIST, 165, 170, **175**, 348
 PMIX_CPUS_PER_PROC, 165, 170, **174**, 347
 PMIX_CPUSET, **88**, 192, 293, 324
 PMIX_CPUSET_BITMAP, **88**, 294, 551
 PMIX_CRED_TYPE, **275**, 374
 PMIX_CREDENTIAL, **88**, 361
 PMIX_CRYPTO_KEY, **275**, 290
 PMIX_DAEMON_MEMORY, 99, 103, **105**
 PMIX_DATA_SCOPE, 75, 78, **80**
 PMIX_DEBUG_DAEMONS_PER_NODE, 347, 412, 413, 417, **418**, 510, 511, 515, 516, 519, 553
 PMIX_DEBUG_DAEMONS_PER_PROC, 347, 412, 413, 417, **418**, 510, 511, 515, 516, 519, 553
 PMIX_DEBUG_STOP_IN_APP, 397, 410, 411, **417**, 511, 568
 PMIX_DEBUG_STOP_IN_INIT, 397, 401, 402, 410, 411, 413, **417**, 509, 515
 PMIX_DEBUG_STOP_ON_EXEC, 397, 402, 410, 411, **417**, 509, 515
 PMIX_DEBUG_TARGET, 347, 411–413, **417**, 417, 418, 510, 511, 515, 516, 519, 552, 553, 560
 PMIX_DEBUGGER_DAEMONS, 347, 412, 413, **417**, 507, 511, 516, 518
 PMIX_DEVICE_DISTANCES, **198**, 266, 294, 556
 PMIX_DEVICE_ID, **198**, 252, 262, 265, 266, 306, 554–556
 PMIX_DEVICE_TYPE, **198**, 556
 PMIX_DISPLAY_MAP, 164, 169, **173**, 346
 PMIX_EMBED_BARRIER, 66, **67**
 PMIX_ENUM_VALUE, 316, **317**, 542, 550
 PMIX_ENVARS_HARVESTED, 167, 172, **176**, 567
 PMIX_EVENT_ACTION_TIMEOUT, **141**, 146
 PMIX_EVENT_AFFECTED_PROC, 139, **141**, 146, 415, 511
 PMIX_EVENT_AFFECTED_PROCS, 139, **141**, 146, 415
 PMIX_EVENT_BASE, **63**, 281, 420
 PMIX_EVENT_CUSTOM_RANGE, 139, **141**, 146
 PMIX_EVENT_DO_NOT_CACHE, **141**, 146

PMIX_EVENT_HDLR_AFTER, 138, [140](#), 529
 PMIX_EVENT_HDLR_APPEND, 139, [140](#), 529
 PMIX_EVENT_HDLR_BEFORE, 138, [140](#), 529
 PMIX_EVENT_HDLR_FIRST, 138, [140](#), 529
 PMIX_EVENT_HDLR_FIRST_IN_CATEGORY, 138, [140](#), 529
 PMIX_EVENT_HDLR_LAST, 138, [140](#), 529
 PMIX_EVENT_HDLR_LAST_IN_CATEGORY, 138, [140](#), 529
 PMIX_EVENT_HDLR_NAME, 138, [140](#), 529
 PMIX_EVENT_HDLR_PREPEND, 138, [140](#)
 PMIX_EVENT_NON_DEFAULT, [141](#), 145
 PMIX_EVENT_PROXY, [141](#), 146
 PMIX_EVENT_RETURN_OBJECT, 139, [141](#)
 PMIX_EVENT_SILENT_TERMINATION, 167, 172, [176](#)
 PMIX_EVENT_TERMINATE_JOB, [141](#), 146
 PMIX_EVENT_TERMINATE_NODE, [141](#), 146
 PMIX_EVENT_TERMINATE_PROC, [141](#), 146
 PMIX_EVENT_TERMINATE_SESSION, [141](#), 146
 PMIX_EVENT_TEXT_MESSAGE, [141](#), 146
 PMIX_EVENT_TIMESTAMP, [141](#), 175, 176, 396–398, 415, 511, 516, 519, 558
 PMIX_EXEC_AGENT, 401, [403](#), 414, 552
 PMIX_EXIT_CODE, [87](#), 175, 176, 396–398, 415, 510, 516, 519, 558
 PMIX_EXTERNAL_PROGRESS, 63, 281, [284](#), 420, 551
 PMIX_FABRIC_COORDINATES, [265](#), 554
 PMIX_FABRIC_COST_MATRIX, 261, [263](#), 553
 PMIX_FABRIC_DEVICE, 252, 262, [265](#), 265, 555
 PMIX_FABRIC_DEVICE_ADDRESS, 253, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_BUS_TYPE, 252, 263, [265](#), 555
 PMIX_FABRIC_DEVICE_COORDINATES, 253, [265](#), 555
 PMIX_FABRIC_DEVICE_DRIVER, 253, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_FIRMWARE, 253, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_INDEX, 253, [265](#), 386, 555
 PMIX_FABRIC_DEVICE_MTU, 253, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_NAME, 252, 262, [265](#), 306, 555
 PMIX_FABRIC_DEVICE_PCI_DEVID, 253, 263, [266](#), 266, 555, 556
 PMIX_FABRIC_DEVICE_SPEED, 253, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_STATE, 253, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_TYPE, 253, 263, [266](#), 555
 PMIX_FABRIC_DEVICE_VENDOR, 252, 262, [265](#), 555
 PMIX_FABRIC_DEVICE_VENDORID, 253, [265](#), 555
 PMIX_FABRIC_DEVICES, 252, [265](#)
 PMIX_FABRIC_DIMS, 261, [264](#), 554
 PMIX_FABRIC_ENDPT, [266](#), 554
 PMIX_FABRIC_GROUPS, 261, [264](#), 553
 PMIX_FABRIC_IDENTIFIER, 261, [264](#), 267, 385, 553
 PMIX_FABRIC_INDEX, 260, [264](#), 264, 553
 PMIX_FABRIC_NUM_DEVICES, 261, [264](#), 553
 PMIX_FABRIC_PLANE, 261, 262, [264](#), 264, 265, 267, 268, 385, 554

PMIX_FABRIC_SHAPE, 261, [264](#), 554
 PMIX_FABRIC_SHAPE_STRING, 262, [264](#), 554
 PMIX_FABRIC_SWITCH, [264](#), 266, 554
 PMIX_FABRIC_VENDOR, 261, [264](#), 267, 385, 553
 PMIX_FIRST_ENVAR, 166, 171, [176](#), 558
 PMIX_FORKEXEC_AGENT, 400, 402, [403](#), 414, 552
 PMIX_FWD_STDDIAG, 395, 399, [403](#), 539
 PMIX_FWD_STDERR, 347, 362, 394, 399, [403](#), 404, 510, 514, 519
 PMIX_FWD_STDIN, 347, 362, 394, 399, [403](#), 404
 PMIX_FWD_STDOUT, 347, 362, 394, 399, 402, [403](#), 404, 510, 514, 519
 PMIX_GET_POINTER_VALUES, 76, 77, 79, [80](#), 547
 PMIX_GET_REFRESH_CACHE, 73, 76, 79, [80](#), 547
 PMIX_GET_STATIC_VALUES, 75–77, 79, [80](#), 547
 PMIX_GLOBAL_RANK, 84, [87](#), 289, 292, 505
 PMIX_GROUP_ASSIGN_CONTEXT_ID, [232](#), 233, 237, 241, 244, 383, 384, 557
 PMIX_GROUP_CONTEXT_ID, [232](#), 384, 557
 PMIX_GROUP_ENDPT_DATA, [232](#), 383, 384, 557
 PMIX_GROUP_FT_COLLECTIVE, [231](#), 233, 237, 241, 244, 557
 PMIX_GROUP_ID, [231](#), 231, 384, 531, 556
 PMIX_GROUP_LEADER, [231](#), 233, 234, 237, 243, 246, 557
 PMIX_GROUP_LOCAL_ONLY, [232](#), 233, 237, 383, 557
 PMIX_GROUP_MEMBERSHIP, [231](#), 234, 384
 PMIX_GROUP_NAMES, [232](#), 557
 PMIX_GROUP_NOTIFY_TERMINATION, [231](#), 233, 234, 237, 239, 241, 244, 557
 PMIX_GROUP_OPTIONAL, [231](#), 233, 235, 237, 241, 244, 384, 557
 PMIX_GRPID, 119, 121, 124, 126, 132, 133, 200, 202, 207, 209, 214, 216, 219, 222, 273, 274, 276, 278, 339–345, 353, 359, 361, [363](#), 364, 366, 369, 372, 374, 376, 378, 379, 382
 PMIX_HOMOGENEOUS_SYSTEM, 282, [284](#), 550
 PMIX_HOST, 163, 169, [173](#), 346
 PMIX_HOST_ATTRIBUTES, 97, 102, [106](#), 111, 112, 411, 542, 548
 PMIX_HOST_FUNCTIONS, 97, 102, 105, [106](#), 111, 548
 PMIX_HOSTFILE, 163, 169, [173](#), 346
 PMIX_HOSTNAME, 76, 79, 82, 85, [88](#), 88, 89, 96, 98, 99, 101, 103–105, 252, 253, 262, 263, 266, 287, 291, 293, 295, 306, 360, 411, 418, 505, 506, 511, 516, 550, 556
 PMIX_HOSTNAME_ALIASES, [88](#), 291, 551
 PMIX_HOSTNAME_KEEP_FQDN, [82](#), 289, 551
 PMIX_IMMEDIATE, 73, 75, 78, [80](#), 89
 PMIX_INDEX_ARGV, 165, 170, [174](#), 347
 PMIX_IOF_BUFFERING_SIZE, 379, 395, 400, [408](#), 427, 430
 PMIX_IOF_BUFFERING_TIME, 379, 395, 400, [408](#), 427, 430
 PMIX_IOF_CACHE_SIZE, 379, 395, 400, [408](#), 426, 429
 PMIX_IOF_COMPLETE, 380, 406, [409](#), 409, 430, 443, 552
 PMIX_IOF_COPY, 405, [409](#), 552
 PMIX_IOF_DROP_NEWEST, 379, 395, 400, [408](#), 426, 429
 PMIX_IOF_DROP_OLDEST, 379, 395, 400, [408](#), 426, 429
 PMIX_IOF_FILE_ONLY, 396, 397, 406, [409](#), 567
 PMIX_IOF_FILE_PATTERN, 396, 406, [409](#), 567

PMIX_IOF_LOCAL_OUTPUT, 282, [408](#), 421, 567
 PMIX_IOF_MERGE_STDERR_STDOUT, 396, 406, [408](#), 409, 567, 568
 PMIX_IOF_OUTPUT_RAW, 395, [408](#), 568
 PMIX_IOF_OUTPUT_TO_DIRECTORY, 396, 397, 406, [409](#), 568
 PMIX_IOF_OUTPUT_TO_FILE, 395, 397, 406, [409](#), 568
 PMIX_IOF_PUSH_STDIN, 406, [409](#), 430, 552
 PMIX_IOF_RANK_OUTPUT, 395, 406, [409](#), 567
 PMIX_IOF_REDIRECT, 405, [409](#), 552
 PMIX_IOF_TAG_OUTPUT, 395, 400, 405, [409](#), 427
 PMIX_IOF_TIMESTAMP_OUTPUT, 395, 400, 405, [409](#), 427
 PMIX_IOF_XML_OUTPUT, 395, 400, 406, [409](#), 427
 PMIX_JOB_CONTINUOUS, 165, 170, [175](#), 348
 PMIX_JOB_CTRL_CANCEL, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_CHECKPOINT, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_CHECKPOINT_EVENT, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_CHECKPOINT_METHOD, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_CHECKPOINT_SIGNAL, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_CHECKPOINT_TIMEOUT, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_ID, 207–210, [212](#), 212, 369, 370
 PMIX_JOB_CTRL_KILL, 207, 210, [212](#), 369
 PMIX_JOB_CTRL_PAUSE, 207, 209, [212](#), 369
 PMIX_JOB_CTRL_PREEMPTIBLE, 208, 211, [212](#), 370
 PMIX_JOB_CTRL_PROVISION, 208, 211, [212](#), 370
 PMIX_JOB_CTRL_PROVISION_IMAGE, 208, 211, [212](#), 370
 PMIX_JOB_CTRL_RESTART, 208, 210, [212](#), 370
 PMIX_JOB_CTRL_RESUME, 207, 209, [212](#), 369
 PMIX_JOB_CTRL_SIGNAL, 207, 210, [212](#), 369
 PMIX_JOB_CTRL_TERMINATE, 207, 210, [212](#), 369
 PMIX_JOB_INFO, 76, 78, [81](#), 83, 96, 101
 PMIX_JOB_INFO_ARRAY, 286, 288, [295](#), 295, 299, 538, 549
 PMIX_JOB_NUM_APPS, [85](#), 289, 299
 PMIX_JOB_RECOVERABLE, 165, 170, [175](#), 348
 PMIX_JOB_SIZE, [85](#), 288, 299, 505, 537, 540
 PMIX_JOB_TERM_STATUS, 175, 176, 396–398, 415, [416](#), 416, 510, 516, 519, 558
 PMIX_JOB_TIMEOUT, 167, 172, [176](#), 176, 348, 567
 PMIX_JOBID, [84](#), 287, 288, 295, 299, 415, 549
 PMIX_LAUNCH_DIRECTIVES, 402, [403](#), 514, 553
 PMIX_LAUNCHER, 387, [392](#), 393
 PMIX_LAUNCHER_DAEMON, 400, [403](#), 552
 PMIX_LAUNCHER_RENDEZVOUS_FILE, 389, [393](#), 552
 PMIX_LOCAL_COLLECTIVE_STATUS, [117](#), 334, 350, 351, 383, 567
 PMIX_LOCAL_CPUSSETS, [85](#), 292, 303
 PMIX_LOCAL_PEERS, [85](#), 85, 291, 292, 301, 505
 PMIX_LOCAL_PROCS, [89](#), 292
 PMIX_LOCAL_RANK, [87](#), 292, 412–414, 418, 505, 510, 515, 519, 553
 PMIX_LOCAL_SIZE, [85](#), 291, 505
 PMIX_LOCALITY_STRING, 189, [191](#), 293, 323, 506

PMIX_LOCALLDR, [85](#), [291](#)
PMIX_LOG_COMPLETION, [176](#), [396](#), [415](#), [558](#)
PMIX_LOG_EMAIL, [220](#), [223](#), [225](#), [365](#)
PMIX_LOG_EMAIL_ADDR, [220](#), [223](#), [225](#), [365](#)
PMIX_LOG_EMAIL_MSG, [220](#), [223](#), [225](#), [365](#)
PMIX_LOG_EMAIL_SENDER_ADDR, [220](#), [223](#), [225](#)
PMIX_LOG_EMAIL_SERVER, [220](#), [223](#), [225](#)
PMIX_LOG_EMAIL_SRVR_PORT, [220](#), [223](#), [225](#)
PMIX_LOG_EMAIL_SUBJECT, [220](#), [223](#), [225](#), [365](#)
PMIX_LOG_GENERATE_TIMESTAMP, [220](#), [223](#), [224](#)
PMIX_LOG_GLOBAL_DATASTORE, [220](#), [223](#), [225](#)
PMIX_LOG_GLOBAL_SYSLOG, [219](#), [222](#), [224](#)
PMIX_LOG_JOB_EVENTS, [176](#), [396](#), [415](#), [558](#)
PMIX_LOG_JOB_RECORD, [220](#), [223](#), [225](#)
PMIX_LOG_LOCAL_SYSLOG, [219](#), [222](#), [224](#)
PMIX_LOG_MSG, [225](#), [365](#)
PMIX_LOG_ONCE, [219](#), [222](#), [224](#)
PMIX_LOG_PROC_ABNORMAL_TERMINATION, [176](#), [558](#)
PMIX_LOG_PROC_TERMINATION, [175](#), [558](#)
PMIX_LOG_SOURCE, [220](#), [223](#), [224](#)
PMIX_LOG_STDERR, [219](#), [222](#), [224](#), [364](#)
PMIX_LOG_STDOUT, [219](#), [222](#), [224](#), [364](#)
PMIX_LOG_SYSLOG, [219](#), [222](#), [224](#), [364](#)
PMIX_LOG_SYSLOG_PRI, [219](#), [222](#), [224](#)
PMIX_LOG_TAG_OUTPUT, [220](#), [223](#), [224](#)
PMIX_LOG_TIMESTAMP, [220](#), [223](#), [224](#)
PMIX_LOG_TIMESTAMP_OUTPUT, [220](#), [223](#), [224](#)
PMIX_LOG_XML_OUTPUT, [220](#), [223](#), [224](#)
PMIX_MAPBY, [164](#), [169](#), [173](#), [174](#), [289](#), [346](#), [510](#), [515](#), [519](#)
PMIX_MAX_PROCS, [82](#), [83](#), [83–86](#), [287–290](#), [292](#), [316](#), [505](#), [540](#)
PMIX_MAX_RESTARTS, [165](#), [170](#), [175](#), [348](#)
PMIX_MAX_VALUE, [316](#), [317](#), [542](#), [550](#)
PMIX_MERGE_STDERR_STDOUT, [165](#), [170](#), [174](#), [347](#)
PMIX_MIN_VALUE, [316](#), [317](#), [542](#), [550](#)
PMIX_MODEL_AFFINITY_POLICY, [64](#), [66](#), [524](#)
PMIX_MODEL_CPU_TYPE, [64](#), [66](#), [524](#)
PMIX_MODEL_LIBRARY_NAME, [63](#), [65](#), [291](#), [313](#), [523](#)
PMIX_MODEL_LIBRARY_VERSION, [64](#), [65](#), [291](#), [313](#), [523](#)
PMIX_MODEL_NUM_CPUS, [64](#), [66](#), [524](#)
PMIX_MODEL_NUM_THREADS, [64](#), [66](#), [523](#)
PMIX_MODEL_PHASE_NAME, [66](#), [141](#), [524](#)
PMIX_MODEL_PHASE_TYPE, [66](#), [141](#), [524](#)
PMIX_MONITOR_APP_CONTROL, [214](#), [216](#), [218](#), [372](#)
PMIX_MONITOR_CANCEL, [214](#), [216](#), [218](#), [372](#)
PMIX_MONITOR_FILE, [214–216](#), [218](#), [372](#)
PMIX_MONITOR_FILE_ACCESS, [214](#), [216](#), [218](#), [372](#)
PMIX_MONITOR_FILE_CHECK_TIME, [214](#), [217](#), [218](#), [372](#)

PMIX_MONITOR_FILE_DROPS, 214, 217, **218**, 373
 PMIX_MONITOR_FILE_MODIFY, 214, 216, **218**, 372
 PMIX_MONITOR_FILE_SIZE, 214, 216, **218**, 372
 PMIX_MONITOR_HEARTBEAT, 214, 216, **218**, 372
 PMIX_MONITOR_HEARTBEAT_DROPS, 214, 216, **218**, 372
 PMIX_MONITOR_HEARTBEAT_TIME, 214, 216, **218**, 372
 PMIX_MONITOR_ID, 214, 216, **218**, 372
 PMIX_NO_OVERSUBSCRIBE, 165, 170, **175**, 348
 PMIX_NO_PROCS_ON_HEAD, 165, 170, **175**, 348
 PMIX_NODE_INFO, 76, 78, **81**, 88, 96, 101, 292
 PMIX_NODE_INFO_ARRAY, 287, 291, **295**, 295, 300, 301, 305, 306, 550
 PMIX_NODE_LIST, **83**, 84, 86, 505
 PMIX_NODE_MAP, **83**, 84, 86, 289, 299, 300, 312, 313, 540
 PMIX_NODE_MAP_RAW, **83**, 551
 PMIX_NODE_OVERSUBSCRIBED, **89**, 291, 567
 PMIX_NODE_RANK, **87**, 293, 413
 PMIX_NODE_SIZE, **89**, 291
 PMIX_NODEID, 76, 79, 82, 85, **88**, 88, 89, 96, 99, 101, 103, 105, 252, 253, 263, 266, 287, 291, 293, 295,
 306, 505, 550, 556
 PMIX_NOHUP, 396, 400, **403**, 552
 PMIX_NOTIFY_COMPLETION, **175**, 396, 415, 510, 516, 519
 PMIX_NOTIFY_JOB_EVENTS, **175**, 396, 415, 557
 PMIX_NOTIFY_PROC_ABNORMAL_TERMINATION, **175**, 558
 PMIX_NOTIFY_PROC_TERMINATION, **175**, 558
 PMIX_NPROC_OFFSET, **84**, 289
 PMIX_NSDIR, **85**, 88, 292, 293
 PMIX_NSPACE, 82, **87**, 87, 96–98, 101–106, 287, 288, 295, 299, 359, 360, 411, 415, 418, 511, 515, 516,
 518, 548, 549
 PMIX_NUM_ALLOCATED_NODES, **83**, 551
 PMIX_NUM_NODES, 69, 81, **83**, 84, 86, 299, 505, 551
 PMIX_NUM_SLOTS, **83**, 84, 86
 PMIX_OPTIONAL, 73, 75, 78, **80**
 PMIX_OUTPUT_TO_DIRECTORY, **174**, 557
 PMIX_OUTPUT_TO_FILE, 165, 170, **174**, 347
 PMIX_PACKAGE_RANK, **88**, 293, 551
 PMIX_PARENT_ID, **87**, 345, 401, 565
 PMIX_PERSISTENCE, 120, 121, **122**, 339, 438
 PMIX_PERSONALITY, 164, 169, **173**, 346
 PMIX_PPR, 164, 169, **173**, 346
 PMIX_PREFIX, 163, 169, **173**, 346
 PMIX_PRELOAD_BIN, 164, 169, **174**, 346
 PMIX_PRELOAD_FILES, 164, 169, **174**, 346
 PMIX_PREPEND_ENVAR, 165, 171, **176**, 522
 PMIX_PRIMARY_SERVER, **393**, 423, 552
 PMIX_PROC_INFO, **81**, 86, 96, 101
 PMIX_PROC_INFO_ARRAY, 287, 292, **295**, 301, 549, 561
 PMIX_PROC_MAP, **83**, 83, 84, 86, 289, 299, 312, 313, 505, 540

PMIX_PROC_MAP_RAW, [83](#), [551](#)
 PMIX_PROC_PID, [88](#), [99](#), [103](#)
 PMIX_PROC_STATE_STATUS, [99](#), [103](#), [416](#)
 PMIX_PROC_TERM_STATUS, [415](#), [416](#)
 PMIX_PROCDIR, [88](#), [293](#)
 PMIX_PROCID, [87](#), [87](#), [96–98](#), [101–103](#), [105](#), [175](#), [176](#), [287](#), [295](#), [360](#), [396–398](#), [415](#), [510](#), [516](#), [519](#), [549](#),
 [558](#)
 PMIX_PROGRAMMING_MODEL, [63](#), [65](#), [291](#), [313](#), [523](#)
 PMIX_PSET_MEMBERS, [227](#), [228](#), [556](#)
 PMIX_PSET_NAME, [227](#), [228](#), [556](#)
 PMIX_PSET_NAMES, [226](#), [228](#), [290](#), [531](#), [556](#)
 PMIX_QUERY_ALLOC_STATUS, [98](#), [103](#), [106](#), [360](#)
 PMIX_QUERY_ATTRIBUTE_SUPPORT, [97](#), [102](#), [105](#), [110](#), [411](#), [548](#)
 PMIX_QUERY_AUTHORIZATIONS, [99](#), [103](#), [104](#)
 PMIX_QUERY_AVAIL_SERVERS, [105](#), [391](#), [548](#)
 PMIX_QUERY_DEBUG_SUPPORT, [98](#), [103](#), [104](#), [360](#), [509](#)
 PMIX_QUERY_GROUP_MEMBERSHIP, [231](#), [531](#), [556](#)
 PMIX_QUERY_GROUP_NAMES, [231](#), [531](#), [556](#)
 PMIX_QUERY_JOB_STATUS, [97](#), [102](#), [104](#), [359](#)
 PMIX_QUERY_LOCAL_ONLY, [106](#), [360](#)
 PMIX_QUERY_LOCAL_PROC_TABLE, [98](#), [102](#), [104](#), [360](#), [411](#), [418](#), [511](#), [516](#)
 PMIX_QUERY_MEMORY_USAGE, [98](#), [103](#), [105](#), [360](#)
 PMIX_QUERY_NAMESPACE_INFO, [104](#), [548](#)
 PMIX_QUERY_NAMESPACES, [97](#), [102](#), [104](#), [359](#), [411](#), [520](#)
 PMIX_QUERY_NUM_GROUPS, [231](#), [531](#), [556](#)
 PMIX_QUERY_NUM_PSETS, [105](#), [228](#), [556](#)
 PMIX_QUERY_PROC_TABLE, [98](#), [102](#), [104](#), [359](#), [411](#), [418](#), [515](#), [516](#), [518](#)
 PMIX_QUERY_PROVISIONAL_ABI_VERSION, [99](#), [105](#), [566](#)
 PMIX_QUERY_PSET_MEMBERSHIP, [105](#), [228](#), [556](#)
 PMIX_QUERY_PSET_NAMES, [105](#), [228](#), [556](#)
 PMIX_QUERY_QUALIFIERS, [99](#), [106](#), [106](#), [548](#)
 PMIX_QUERY_QUEUE_LIST, [98](#), [102](#), [104](#), [359](#)
 PMIX_QUERY_QUEUE_STATUS, [98](#), [102](#), [104](#), [359](#)
 PMIX_QUERY_REFRESH_CACHE, [96](#), [100](#), [101](#), [106](#), [110](#)
 PMIX_QUERY_REPORT_AVG, [98](#), [103](#), [106](#), [360](#)
 PMIX_QUERY_REPORT_MINMAX, [98](#), [103](#), [106](#), [360](#)
 PMIX_QUERY_RESULTS, [99](#), [106](#), [547](#)
 PMIX_QUERY_SPAWN_SUPPORT, [98](#), [103](#), [104](#), [360](#), [509](#)
 PMIX_QUERY_STABLE_ABI_VERSION, [99](#), [105](#), [566](#)
 PMIX_QUERY_STORAGE_LIST, [433](#), [564](#)
 PMIX_QUERY_SUPPORTED_KEYS, [104](#), [548](#)
 PMIX_QUERY_SUPPORTED_QUALIFIERS, [104](#), [548](#)
 PMIX_RANGE, [120](#), [121](#), [122](#), [124](#), [127](#), [132](#), [133](#), [139](#), [215](#), [231](#), [339](#), [341](#), [343](#), [357](#), [384](#), [438](#), [531](#), [556](#)
 PMIX_RANK, [87](#), [97](#), [98](#), [101–103](#), [105](#), [167](#), [172](#), [287](#), [292](#), [295](#), [360](#), [414](#), [417](#), [505](#), [510](#), [514](#), [549](#)
 PMIX_RANKBY, [164](#), [169](#), [174](#), [289](#), [346](#)
 PMIX_REGISTER_CLEANUP, [207](#), [210](#), [212](#)
 PMIX_REGISTER_CLEANUP_DIR, [207](#), [210](#), [212](#)

PMIX_REGISTER_NODATA, 286, [295](#)
 PMIX_REINCARNATION, [88](#), 293, 551
 PMIX_REPORT_BINDINGS, 165, 170, [175](#), 348
 PMIX_REQUESTOR_IS_CLIENT, 345, [349](#)
 PMIX_REQUESTOR_IS_TOOL, 345, [349](#), 520
 PMIX_REQUIRED_KEY, 337, [338](#), 550
 PMIX_RM_NAME, [82](#), 288
 PMIX_RM_VERSION, [82](#), 288
 PMIX_SEND_HEARTBEAT, 214, 217, [218](#)
 PMIX_SERVER_ATTRIBUTES, 97, 102, [106](#), 111, 542, 548
 PMIX_SERVER_ENABLE_MONITORING, 282, [283](#)
 PMIX_SERVER_FUNCTIONS, 97, 102, 105, [106](#), 111, 548
 PMIX_SERVER_GATEWAY, 280, [283](#)
 PMIX_SERVER_HOSTNAME, 288, [393](#)
 PMIX_SERVER_INFO_ARRAY, 105, [106](#), 548
 PMIX_SERVER_NAMESPACE, 280, [283](#), 288, 390, 419, 423
 PMIX_SERVER_PIDINFO, 390, [392](#), 419, 423
 PMIX_SERVER_RANK, 280, [283](#), 288
 PMIX_SERVER_REMOTE_CONNECTIONS, 281, [283](#)
 PMIX_SERVER_SCHEDULER, 263, 266, 280, [284](#), 549, 553
 PMIX_SERVER_SESSION_SUPPORT, 280, [283](#), 549
 PMIX_SERVER_SHARE_TOPOLOGY, 281, [283](#), 549
 PMIX_SERVER_START_TIME, [283](#), 549
 PMIX_SERVER_SYSTEM_SUPPORT, 280, [283](#), 388
 PMIX_SERVER_TMPDIR, 280, 282, [283](#), 388, 389
 PMIX_SERVER_TOOL_SUPPORT, 272, 280, 282, [283](#)
 PMIX_SERVER_URI, 98, 103, 390, 391, [393](#), 419, 423
 PMIX_SESSION_ID, 82, 83, [87](#), 90, 286–288, 295, 299, 415, 549
 PMIX_SESSION_INFO, 75, 78, [81](#), 83, 90, 96, 101, 287, 289, 312
 PMIX_SESSION_INFO_ARRAY, 286, 287, [295](#), 295, 299, 538
 PMIX_SET_ENVAR, 165, 170, [176](#), 522
 PMIX_SET_SESSION_CWD, 163, 168, [174](#), 345
 PMIX_SETUP_APP_ALL, 311, [314](#)
 PMIX_SETUP_APP_ENVARS, 311, [314](#), 511, 515
 PMIX_SETUP_APP_NONENVAR, 311, [314](#)
 PMIX_SINGLE_LISTENER, 62, 280, [283](#)
 PMIX_SINGLETON, 282, [284](#), 567
 PMIX_SOCKET_MODE, 62, 280, [283](#), 420
 PMIX_SPAWN_TIMEOUT, 167, 172, [176](#), 348, 567
 PMIX_SPAWN_TOOL, 167, 172, [175](#), 402, 514
 PMIX_SPAWNED, [88](#), 293, 345
 PMIX_STDIN_TGT, 164, 169, [174](#), 346
 PMIX_STORAGE_ACCESS_TYPE, [434](#), 565
 PMIX_STORAGE_ACCESSIBILITY, [433](#), 564
 PMIX_STORAGE_BW_CUR, [434](#), 434, 564, 565
 PMIX_STORAGE_BW_MAX, [434](#), 564
 PMIX_STORAGE_CAPACITY_LIMIT, [433](#), 564

PMIX_STORAGE_CAPACITY_USED, [433](#), 564
 PMIX_STORAGE_ID, [433](#), 433, 563, 564
 PMIX_STORAGE_IOPS_CUR, [434](#), 434, 564, 565
 PMIX_STORAGE_IOPS_MAX, [434](#), 564
 PMIX_STORAGE_MEDIUM, [433](#), 564
 PMIX_STORAGE_MINIMAL_XFER_SIZE, [433](#), 564
 PMIX_STORAGE_OBJECT_LIMIT, [433](#), 564
 PMIX_STORAGE_OBJECTS_USED, [433](#), 564
 PMIX_STORAGE_PATH, [433](#), 563
 PMIX_STORAGE_PERSISTENCE, [433](#), 564
 PMIX_STORAGE_SUGGESTED_XFER_SIZE, [433](#), 434, 564, 565
 PMIX_STORAGE_TYPE, [433](#), 563
 PMIX_STORAGE_VERSION, [433](#), 564
 PMIX_SWITCH_PEERS, [266](#), 266, 554
 PMIX_SYSTEM_TMPDIR, 280, [283](#), 388, 389
 PMIX_TAG_OUTPUT, 164, 170, [174](#), 347
 PMIX_TCP_DISABLE_IPV4, 63, [65](#), 281, 420
 PMIX_TCP_DISABLE_IPV6, 63, [65](#), 281, 420
 PMIX_TCP_IF_EXCLUDE, 63, [65](#), 281, 420
 PMIX_TCP_IF_INCLUDE, 63, [65](#), 281, 420
 PMIX_TCP_IPV4_PORT, 63, [65](#), 281, 420
 PMIX_TCP_IPV6_PORT, 63, [65](#), 281, 420
 PMIX_TCP_REPORT_URI, 63, [65](#), 280, 420
 PMIX_TCP_URI, [65](#), 390, 391, 419, 423
 PMIX_TDIR_RMCLEAN, [82](#), 289
 PMIX_THREADING_MODEL, 64, [65](#), 523
 PMIX_TIME_REMAINING, 94, 98, 103, [105](#), 360
 PMIX_TIMEOUT, 4, 74, 76, 79, [80](#), 89, 114, 117, 120, 121, 124, 127, 132, 133, 167, 172, 176, 181, 183, 185,
 187, 230, 234, 235, 237, 238, 240, 242, 244, 245, 248, 273, 275, 276, 278, 334, 337, 339, 341, 343,
 348, 350, 352, 374, 377, 393, 425, 516, 553, 567
 PMIX_TIMEOUT_REPORT_STATE, [175](#), 557
 PMIX_TIMEOUT_STACKTRACES, [175](#), 557
 PMIX_TIMESTAMP_OUTPUT, 164, 170, [174](#), 347
 PMIX_TMPDIR, [82](#), 82, 85, 289, 292
 PMIX_TOOL_ATTACHMENT_FILE, 390, 391, [393](#), 419, 423, 551
 PMIX_TOOL_ATTRIBUTES, 97, 102, [106](#), 111, 542, 549
 PMIX_TOOL_CONNECT_OPTIONAL, [393](#), 551
 PMIX_TOOL_DO_NOT_CONNECT, 389, 390, [393](#), 419, 421
 PMIX_TOOL_FUNCTIONS, 97, 102, 105, [106](#), 111, 548, 549
 PMIX_TOOL_NAMESPACE, 361, 389, [392](#), 419, 421
 PMIX_TOOL_RANK, 361, 389, [392](#), 419, 421
 PMIX_TOPOLOGY2, 281, [283](#), 549, 560
 PMIX_UNIV_SIZE, 11, [82](#), 287, 299, 537, 540
 PMIX_UNSET_ENVAR, 165, 170, [176](#), 522
 PMIX_USERID, 119, 121, 124, 126, 132, 133, 200, 202, 207, 209, 214, 216, 219, 222, 273, 274, 276, 277,
 339–345, 353, 359, 361, [363](#), 364, 366, 369, 372, 374, 376, 378, 379, 382
 PMIX_USOCK_DISABLE, 62, 280, [283](#)

PMIX_VERSION_INFO, 362, [363](#)
PMIX_WAIT, 79, [80](#), 124, 125, 127, 341
PMIX_WAIT_FOR_CONNECTION, [393](#), 425, 516, 553
PMIX_WDIR, 163, 168, [173](#), 290, 345

PMIX_ALLOC_NETWORK

Deprecated, [560](#)

PMIX_ALLOC_NETWORK_ENDPTS

Deprecated, [560](#)

PMIX_ALLOC_NETWORK_ENDPTS_NODE

Deprecated, [561](#)

PMIX_ALLOC_NETWORK_ID

Deprecated, [560](#)

PMIX_ALLOC_NETWORK_PLANE

Deprecated, [560](#)

PMIX_ALLOC_NETWORK_QOS

Deprecated, [560](#)

PMIX_ALLOC_NETWORK_SEC_KEY

Deprecated, [561](#)

PMIX_ALLOC_NETWORK_TYPE

Deprecated, [560](#)

PMIX_ARCH

Deprecated, [541](#)

Removed, [562](#)

PMIX_COLLECTIVE_ALGO

Deprecated, [541](#)

Removed, [561](#)

PMIX_COLLECTIVE_ALGO_REQD

Deprecated, [539](#)

Removed, [562](#)

PMIX_DEBUG_JOB

Deprecated, [560](#)

PMIX_DEBUG_WAIT_FOR_NOTIFY

Deprecated, [568](#)

PMIX_DSTPATH

Deprecated, [541](#)

Removed, [561](#)

PMIX_ERROR_GROUP_ABORT

Deprecated, [537](#)

Removed, [539](#)

PMIX_ERROR_GROUP_COMM

Deprecated, [537](#)

Removed, [539](#)

PMIX_ERROR_GROUP_GENERAL

Deprecated, [537](#)

Removed, [539](#)

PMIX_ERROR_GROUP_LOCAL

Deprecated, [537](#)
Removed, [539](#)
PMIX_ERROR_GROUP_MIGRATE
Deprecated, [537](#)
Removed, [539](#)
PMIX_ERROR_GROUP_NODE
Deprecated, [537](#)
Removed, [539](#)
PMIX_ERROR_GROUP_RESOURCE
Deprecated, [537](#)
Removed, [539](#)
PMIX_ERROR_GROUP_SPAWN
Deprecated, [537](#)
Removed, [539](#)
PMIX_ERROR_HANDLER_ID
Deprecated, [537](#)
Removed, [539](#)
PMIX_ERROR_NAME
Deprecated, [537](#)
Removed, [539](#)
PMIX_HWLOC_HOLE_KIND
Deprecated, [541](#)
Removed, [561](#)
PMIX_HWLOC_SHARE_TOPO
Deprecated, [541](#)
Removed, [561](#)
PMIX_HWLOC_SHMEM_ADDR
Deprecated, [541](#)
Removed, [561](#)
PMIX_HWLOC_SHMEM_FILE
Deprecated, [541](#)
Removed, [561](#)
PMIX_HWLOC_SHMEM_SIZE
Deprecated, [541](#)
Removed, [561](#)
PMIX_HWLOC_XML_V1
Deprecated, [541](#)
Removed, [561](#)
PMIX_HWLOC_XML_V2
Deprecated, [541](#)
Removed, [561](#)
PMIX_LOCAL_TOPO
Deprecated, [541](#)
Removed, [561](#)
PMIX_LOCALITY
Deprecated, [561](#)
PMIX_MAP_BLOB

Deprecated, [542](#)
Removed, [562](#)
PMIX_MAPPER
Deprecated, [541](#)
Removed, [562](#)
PMIX_NON_PMI
Deprecated, [542](#)
Removed, [562](#)
PMIX_PROC_BLOB
Deprecated, [542](#)
Removed, [562](#)
PMIX_PROC_DATA
Deprecated, [561](#)
PMIX_PROC_URI
Deprecated, [542](#)
Removed, [562](#)
PMIX_RECONNECT_SERVER
Deprecated, [560](#)
PMIX_TOPOLOGY
Deprecated, [560](#)
PMIX_TOPOLOGY_FILE
Deprecated, [542](#)
Removed, [561](#)
PMIX_TOPOLOGY_SIGNATURE
Deprecated, [542](#)
Removed, [561](#)
PMIX_TOPOLOGY_XML
Deprecated, [542](#)
Removed, [561](#)