MPI: A Message-Passing Interface Standard

Version 3.1

Message Passing Interface Forum

September 21, 2012

This document describes the Message-Passing Interface (MPI) standard, version 3.0. The MPI standard includes point-to-point message-passing, collective communications, group and communicator concepts, process topologies, environmental management, process creation and management, one-sided communications, extended collective operations, external interfaces, I/O, some miscellaneous topics, and a profiling interface. Language bindings for C and Fortran are defined.

Historically, the evolution of the standards is from MPI-1.0 (June 1994) to MPI-1.1 (June 12, 1995) to MPI-1.2 (July 18, 1997), with several clarifications and additions and published as part of the MPI-2 document, to MPI-2.0 (July 18, 1997), with new functionality, to MPI-1.3 (May 30, 2008), combining for historical reasons the documents 1.1 and 1.2 and some errata documents to one combined document, and to MPI-2.1 (June 23, 2008), combining the previous documents. Version MPI-2.2 (September 2009) added additional clarifications and seven new routines. This version, MPI-3.0, is an extension of MPI-2.2.

Comments. Please send comments on MPI to the MPI Forum as follows:

1. Subscribe to http://lists.mpi-forum.org/mailman/listinfo.cgi/mpi-comments
2. Send your comment to: mpi-comments@mpi-forum.org, together with the URL of the version of the MPI standard and the page and line numbers on which you are commenting. Only use the official versions.

   Your comment will be forwarded to MPI Forum committee members for consideration. Messages sent from an unsubscribed e-mail address will not be considered.

©1993, 1994, 1995, 1996, 1997, 2008, 2009, 2012 University of Tennessee, Knoxville, Tennessee. Permission to copy without fee all or part of this material is granted, provided the University of Tennessee copyright notice and the title of this document appear, and notice is given that copying is by permission of the University of Tennessee.

Version 3.0: September 21, 2012. Coincident with the development of MPI-2.2, the MPI Forum began discussions of a major extension to MPI. This document contains the MPI-3 Standard. This draft version of the MPI-3 standard contains significant extensions to MPI functionality, including nonblocking collectives, new one-sided communication operations, and Fortran 2008 bindings. Unlike MPI-2.2, this standard is considered a major update to the MPI standard. As with previous versions, new features have been adopted only when there were compelling needs for the users. Some features, however, may have more than a minor impact on existing MPI implementations.

Version 2.2: September 4, 2009. This document contains mostly corrections and clarifications to the MPI-2.1 document. A few extensions have been added; however all correct MPI-2.1 programs are correct MPI-2.2 programs. New features were adopted only when there were compelling needs for users, open source implementations, and minor impact on existing MPI implementations.

Version 2.1: June 23, 2008. This document combines the previous documents MPI-1.3 (May 30, 2008) and MPI-2.0 (July 18, 1997). Certain parts of MPI-2.0, such as some sections of Chapter 4, Miscellany, and Chapter 7, Extended Collective Operations, have been merged into the Chapters of MPI-1.3. Additional errata and clarifications collected by the MPI Forum are also included in this document.

Version 1.3: May 30, 2008. This document combines the previous documents MPI-1.1 (June 12, 1995) and the MPI-1.2 Chapter in MPI-2 ¡(July 18, 1997). Additional errata collected by the MPI Forum referring to MPI-1.1 and MPI-1.2 are also included in this document.

Version 2.0: July 18, 1997. Beginning after the release of MPI-1.1, the MPI Forum began meeting to consider corrections and extensions. MPI-2 has been focused on process creation and management, one-sided communications, extended collective communications, external interfaces and parallel I/O. A miscellany chapter discusses items that do not fit elsewhere, in particular language interoperability.

Version 1.2: July 18, 1997. The MPI-2 Forum introduced MPI-1.2 as Chapter 3 in the standard "MPI-2: Extensions to the Message-Passing Interface", July 18, 1997. This section contains clarifications and minor corrections to Version 1.1 of the MPI Standard. The only new function in MPI-1.2 is one for identifying to which version of the MPI Standard the implementation conforms. There are small differences between MPI-1 and MPI-1.1. There are very few differences between MPI-1.1 and MPI-1.2, but large differences between MPI-1.2 and MPI-2.

Version 1.1: June, 1995. Beginning in March, 1995, the Message-Passing Interface Forum reconvened to correct errors and make clarifications in the MPI document of May 5, 1994, referred to below as Version 1.0. These discussions resulted in Version 1.1. The changes from Version 1.0 are minor. A version of this document with all changes marked is available.

Version 1.0: May, 1994. The Message-Passing Interface Forum (MPIF), with participation from over 40 organizations, has been meeting since January 1993 to discuss and define a set

of library interface standards for message passing. MPIF is not sanctioned or supported by any official standards organization.

The goal of the Message-Passing Interface, simply stated, is to develop a widely used standard for writing message-passing programs. As such the interface should establish a practical, portable, efficient, and flexible standard for message-passing.

This is the final report, Version 1.0, of the Message-Passing Interface Forum. This document contains all the technical features proposed for the interface. This copy of the draft was processed by LaTeX on May 5, 1994.

Contents

Acknowledgments **ix**

1 Introduction to MPI **1**
1.1 Overview and Goals . . . 1
1.2 Background of MPI-1.0 . . . 2
1.3 Background of MPI-1.1, MPI-1.2, and MPI-2.0 . . . 2
1.4 Background of MPI-1.3 and MPI-2.1 . . . 3
1.5 Background of MPI-2.2 . . . 4
1.6 Background of MPI-3.0 . . . 4
1.7 Who Should Use This Standard? . . . 4
1.8 What Platforms Are Targets For Implementation? . . . 5
1.9 What Is Included In The Standard? . . . 5
1.10 What Is Not Included In The Standard? . . . 6
1.11 Organization of this Document . . . 6

2 MPI Terms and Conventions **9**
2.1 Document Notation . . . 9
2.2 Naming Conventions . . . 9
2.3 Procedure Specification . . . 10
2.4 Semantic Terms . . . 11
2.5 Data Types . . . 12
2.5.1 Opaque Objects . . . 12
2.5.2 Array Arguments . . . 14
2.5.3 State . . . 14
2.5.4 Named Constants . . . 15
2.5.5 Choice . . . 16
2.5.6 Absolute Addresses and Relative Address Displacements . . . 16
2.5.7 File Offsets . . . 16
2.5.8 Counts . . . 17
2.6 Language Binding . . . 17
2.6.1 Deprecated and Removed Names and Functions . . . 17
2.6.2 Fortran Binding Issues . . . 19
2.6.3 C Binding Issues . . . 19
2.6.4 Functions and Macros . . . 20
2.7 Processes . . . 20
2.8 Error Handling . . . 20
2.9 Implementation Issues . . . 21

2.9.1 Independence of Basic Runtime Routines . . . 22
2.9.2 Interaction with Signals . . . 22
2.10 Examples . . . 22

3 Point-to-Point Communication 23
3.1 Introduction . . . 23
3.2 Blocking Send and Receive Operations . . . 24
3.2.1 Blocking Send . . . 24
3.2.2 Message Data . . . 25
3.2.3 Message Envelope . . . 27
3.2.4 Blocking Receive . . . 28
3.2.5 Return Status . . . 30
3.2.6 Passing MPI_STATUS_IGNORE for Status . . . 32
3.3 Data Type Matching and Data Conversion . . . 33
3.3.1 Type Matching Rules . . . 33
Type MPI_CHARACTER . . . 34
3.3.2 Data Conversion . . . 35
3.4 Communication Modes . . . 37
3.5 Semantics of Point-to-Point Communication . . . 40
3.6 Buffer Allocation and Usage . . . 44
3.6.1 Model Implementation of Buffered Mode . . . 46
3.7 Nonblocking Communication . . . 47
3.7.1 Communication Request Objects . . . 48
3.7.2 Communication Initiation . . . 48
3.7.3 Communication Completion . . . 52
3.7.4 Semantics of Nonblocking Communications . . . 56
3.7.5 Multiple Completions . . . 57
3.7.6 Non-destructive Test of status . . . 63
3.8 Probe and Cancel . . . 64
3.8.1 Probe . . . 64
3.8.2 Matching Probe . . . 67
3.8.3 Matched Receives . . . 69
3.8.4 Cancel . . . 71
3.9 Persistent Communication Requests . . . 73
3.10 Send-Receive . . . 78
3.11 Null Processes . . . 80

4 Datatypes 83
4.1 Derived Datatypes . . . 83
4.1.1 Type Constructors with Explicit Addresses . . . 85
4.1.2 Datatype Constructors . . . 85
4.1.3 Subarray Datatype Constructor . . . 94
4.1.4 Distributed Array Datatype Constructor . . . 96
4.1.5 Address and Size Functions . . . 101
4.1.6 Lower-Bound and Upper-Bound Markers . . . 104
4.1.7 Extent and Bounds of Datatypes . . . 106
4.1.8 True Extent of Datatypes . . . 108
4.1.9 Commit and Free . . . 109

4.1.10 Duplicating a Datatype . . . 111
4.1.11 Use of General Datatypes in Communication . . . 111
4.1.12 Correct Use of Addresses . . . 115
4.1.13 Decoding a Datatype . . . 116
4.1.14 Examples . . . 122
4.2 Pack and Unpack . . . 131
4.3 Canonical MPI_PACK and MPI_UNPACK . . . 138

5 Collective Communication 141
5.1 Introduction and Overview . . . 141
5.2 Communicator Argument . . . 144
5.2.1 Specifics for Intracommunicator Collective Operations . . . 144
5.2.2 Applying Collective Operations to Intercommunicators . . . 145
5.2.3 Specifics for Intercommunicator Collective Operations . . . 146
5.3 Barrier Synchronization . . . 147
5.4 Broadcast . . . 148
5.4.1 Example using MPI_BCAST . . . 149
5.5 Gather . . . 149
5.5.1 Examples using MPI_GATHER, MPI_GATHERV . . . 152
5.6 Scatter . . . 159
5.6.1 Examples using MPI_SCATTER, MPI_SCATTERV . . . 162
5.7 Gather-to-all . . . 165
5.7.1 Example using MPI_ALLGATHER . . . 167
5.8 All-to-All Scatter/Gather . . . 168
5.9 Global Reduction Operations . . . 173
5.9.1 Reduce . . . 174
5.9.2 Predefined Reduction Operations . . . 176
5.9.3 Signed Characters and Reductions . . . 178
5.9.4 MINLOC and MAXLOC . . . 179
5.9.5 User-Defined Reduction Operations . . . 183
Example of User-defined Reduce . . . 186
5.9.6 All-Reduce . . . 187
5.9.7 Process-Local Reduction . . . 189
5.10 Reduce-Scatter . . . 190
5.10.1 MPI_REDUCE_SCATTER_BLOCK . . . 190
5.10.2 MPI_REDUCE_SCATTER . . . 191
5.11 Scan . . . 193
5.11.1 Inclusive Scan . . . 193
5.11.2 Exclusive Scan . . . 194
5.11.3 Example using MPI_SCAN . . . 195
5.12 Nonblocking Collective Operations . . . 196
5.12.1 Nonblocking Barrier Synchronization . . . 198
5.12.2 Nonblocking Broadcast . . . 199
Example using MPI_IBCAST . . . 199
5.12.3 Nonblocking Gather . . . 200
5.12.4 Nonblocking Scatter . . . 202
5.12.5 Nonblocking Gather-to-all . . . 204
5.12.6 Nonblocking All-to-All Scatter/Gather . . . 206

5.12.7 Nonblocking Reduce . . . 209
5.12.8 Nonblocking All-Reduce . . . 210
5.12.9 Nonblocking Reduce-Scatter with Equal Blocks . . . 211
5.12.10 Nonblocking Reduce-Scatter . . . 212
5.12.11 Nonblocking Inclusive Scan . . . 213
5.12.12 Nonblocking Exclusive Scan . . . 214
5.13 Correctness . . . 214

6 Groups, Contexts, Communicators, and Caching 223
6.1 Introduction . . . 223
6.1.1 Features Needed to Support Libraries . . . 223
6.1.2 MPI's Support for Libraries . . . 224
6.2 Basic Concepts . . . 226
6.2.1 Groups . . . 226
6.2.2 Contexts . . . 226
6.2.3 Intra-Communicators . . . 227
6.2.4 Predefined Intra-Communicators . . . 227
6.3 Group Management . . . 228
6.3.1 Group Accessors . . . 228
6.3.2 Group Constructors . . . 230
6.3.3 Group Destructors . . . 235
6.4 Communicator Management . . . 235
6.4.1 Communicator Accessors . . . 235
6.4.2 Communicator Constructors . . . 237
6.4.3 Communicator Destructors . . . 248
6.4.4 Communicator Info . . . 248
6.5 Motivating Examples . . . 250
6.5.1 Current Practice #1 . . . 250
6.5.2 Current Practice #2 . . . 251
6.5.3 (Approximate) Current Practice #3 . . . 251
6.5.4 Example #4 . . . 252
6.5.5 Library Example #1 . . . 253
6.5.6 Library Example #2 . . . 255
6.6 Inter-Communication . . . 257
6.6.1 Inter-communicator Accessors . . . 259
6.6.2 Inter-communicator Operations . . . 260
6.6.3 Inter-Communication Examples . . . 263
Example 1: Three-Group "Pipeline" . . . 263
Example 2: Three-Group "Ring" . . . 264
6.7 Caching . . . 265
6.7.1 Functionality . . . 266
6.7.2 Communicators . . . 267
6.7.3 Windows . . . 272
6.7.4 Datatypes . . . 275
6.7.5 Error Class for Invalid Keyval . . . 279
6.7.6 Attributes Example . . . 279
6.8 Naming Objects . . . 281
6.9 Formalizing the Loosely Synchronous Model . . . 285

6.9.1 Basic Statements . . . 285
6.9.2 Models of Execution . . . 286
Static Communicator Allocation . . . 286
Dynamic Communicator Allocation . . . 286
The General Case . . . 287

7 Process Topologies 289
7.1 Introduction . . . 289
7.2 Virtual Topologies . . . 290
7.3 Embedding in MPI . . . 290
7.4 Overview of the Functions . . . 290
7.5 Topology Constructors . . . 292
7.5.1 Cartesian Constructor . . . 292
7.5.2 Cartesian Convenience Function: MPI_DIMS_CREATE . . . 293
7.5.3 Graph Constructor . . . 294
7.5.4 Distributed Graph Constructor . . . 296
7.5.5 Topology Inquiry Functions . . . 302
7.5.6 Cartesian Shift Coordinates . . . 310
7.5.7 Partitioning of Cartesian Structures . . . 311
7.5.8 Low-Level Topology Functions . . . 312
7.6 Neighborhood Collective Communication . . . 314
7.6.1 Neighborhood Gather . . . 315
7.6.2 Neighbor Alltoall . . . 318
7.7 Nonblocking Neighborhood Communication . . . 323
7.7.1 Nonblocking Neighborhood Gather . . . 324
7.7.2 Nonblocking Neighborhood Alltoall . . . 326
7.8 An Application Example . . . 329

8 MPI Environmental Management 333
8.1 Implementation Information . . . 333
8.1.1 Version Inquiries . . . 333
8.1.2 Environmental Inquiries . . . 334
Tag Values . . . 335
Host Rank . . . 335
IO Rank . . . 335
Clock Synchronization . . . 336
Inquire Processor Name . . . 336
8.2 Memory Allocation . . . 337
8.3 Error Handling . . . 340
8.3.1 Error Handlers for Communicators . . . 341
8.3.2 Error Handlers for Windows . . . 343
8.3.3 Error Handlers for Files . . . 345
8.3.4 Freeing Errorhandlers and Retrieving Error Strings . . . 346
8.4 Error Codes and Classes . . . 347
8.5 Error Classes, Error Codes, and Error Handlers . . . 350
8.6 Timers and Synchronization . . . 354
8.7 Startup . . . 355
8.7.1 Allowing User Functions at Process Termination . . . 361

8.7.2 Determining Whether MPI Has Finished . . . 361
8.8 Portable MPI Process Startup . . . 362

9 The Info Object 365

10 Process Creation and Management 371
10.1 Introduction . . . 371
10.2 The Dynamic Process Model . . . 372
10.2.1 Starting Processes . . . 372
10.2.2 The Runtime Environment . . . 372
10.3 Process Manager Interface . . . 374
10.3.1 Processes in MPI . . . 374
10.3.2 Starting Processes and Establishing Communication . . . 374
10.3.3 Starting Multiple Executables and Establishing Communication . . . 379
10.3.4 Reserved Keys . . . 382
10.3.5 Spawn Example . . . 383
Manager-worker Example Using MPI_COMM_SPAWN . . . 383
10.4 Establishing Communication . . . 385
10.4.1 Names, Addresses, Ports, and All That . . . 385
10.4.2 Server Routines . . . 386
10.4.3 Client Routines . . . 388
10.4.4 Name Publishing . . . 390
10.4.5 Reserved Key Values . . . 392
10.4.6 Client/Server Examples . . . 392
Simplest Example — Completely Portable. . . . 392
Ocean/Atmosphere — Relies on Name Publishing . . . 393
Simple Client-Server Example . . . 393
10.5 Other Functionality . . . 395
10.5.1 Universe Size . . . 395
10.5.2 Singleton MPI_INIT . . . 396
10.5.3 MPI_APPNUM . . . 396
10.5.4 Releasing Connections . . . 397
10.5.5 Another Way to Establish MPI Communication . . . 399

11 One-Sided Communications 401
11.1 Introduction . . . 401
11.2 Initialization . . . 402
11.2.1 Window Creation . . . 403
11.2.2 Window That Allocates Memory . . . 405
11.2.3 Window That Allocates Shared Memory . . . 407
11.2.4 Window of Dynamically Attached Memory . . . 410
11.2.5 Window Destruction . . . 413
11.2.6 Window Attributes . . . 414
11.2.7 Window Info . . . 415
11.3 Communication Calls . . . 417
11.3.1 Put . . . 418
11.3.2 Get . . . 420
11.3.3 Examples for Communication Calls . . . 421

11.3.4 Accumulate Functions . . . 423
Accumulate Function . . . 424
Get Accumulate Function . . . 426
Fetch and Op Function . . . 427
Compare and Swap Function . . . 429
11.3.5 Request-based RMA Communication Operations . . . 430
11.4 Memory Model . . . 435
11.5 Synchronization Calls . . . 436
11.5.1 Fence . . . 440
11.5.2 General Active Target Synchronization . . . 441
11.5.3 Lock . . . 445
11.5.4 Flush and Sync . . . 448
11.5.5 Assertions . . . 450
11.5.6 Miscellaneous Clarifications . . . 451
11.6 Error Handling . . . 452
11.6.1 Error Handlers . . . 452
11.6.2 Error Classes . . . 452
11.7 Semantics and Correctness . . . 452
11.7.1 Atomicity . . . 460
11.7.2 Ordering . . . 461
11.7.3 Progress . . . 461
11.7.4 Registers and Compiler Optimizations . . . 463
11.8 Examples . . . 464

12 External Interfaces **475**
12.1 Introduction . . . 475
12.2 Generalized Requests . . . 475
12.2.1 Examples . . . 479
12.3 Associating Information with Status . . . 482
12.4 MPI and Threads . . . 484
12.4.1 General . . . 484
12.4.2 Clarifications . . . 485
12.4.3 Initialization . . . 487

13 I/O **491**
13.1 Introduction . . . 491
13.1.1 Definitions . . . 491
13.2 File Manipulation . . . 493
13.2.1 Opening a File . . . 493
13.2.2 Closing a File . . . 495
13.2.3 Deleting a File . . . 496
13.2.4 Resizing a File . . . 497
13.2.5 Preallocating Space for a File . . . 498
13.2.6 Querying the Size of a File . . . 498
13.2.7 Querying File Parameters . . . 499
13.2.8 File Info . . . 500
Reserved File Hints . . . 502
13.3 File Views . . . 503

13.4 Data Access . . . 506
13.4.1 Data Access Routines . . . 506
Positioning . . . 507
Synchronism . . . 508
Coordination . . . 508
Data Access Conventions . . . 508
13.4.2 Data Access with Explicit Offsets . . . 509
13.4.3 Data Access with Individual File Pointers . . . 514
13.4.4 Data Access with Shared File Pointers . . . 522
Noncollective Operations . . . 522
Collective Operations . . . 525
Seek . . . 526
13.4.5 Split Collective Data Access Routines . . . 527
13.5 File Interoperability . . . 534
13.6 . . . 536
13.6.1 Datatypes for File Interoperability . . . 536
13.6.2 External Data Representation: "external32" . . . 538
13.6.3 User-Defined Data Representations . . . 539
Extent Callback . . . 541
Datarep Conversion Functions . . . 542
13.6.4 Matching Data Representations . . . 544
13.7 Consistency and Semantics . . . 544
13.7.1 File Consistency . . . 544
13.7.2 Random Access vs. Sequential Files . . . 547
13.7.3 Progress . . . 548
13.7.4 Collective File Operations . . . 548
13.7.5 Nonblocking Collective File Operations . . . 548
13.7.6 Type Matching . . . 549
13.7.7 Miscellaneous Clarifications . . . 549
13.7.8 MPI_Offset Type . . . 549
13.7.9 Logical vs. Physical File Layout . . . 550
13.7.10 File Size . . . 550
13.7.11 Examples . . . 550
Asynchronous I/O . . . 553
13.8 I/O Error Handling . . . 555
13.9 I/O Error Classes . . . 555
13.10 Examples . . . 555
13.10.1 Double Buffering with Split Collective I/O . . . 555
13.10.2 Subarray Filetype Constructor . . . 558
13.11 . . . 559

14 Tool Support **561**
14.1 Introduction . . . 561
14.2 Profiling Interface . . . 561
14.2.1 Requirements . . . 561
14.2.2 Discussion . . . 562
14.2.3 Logic of the Design . . . 562
14.2.4 Miscellaneous Control of Profiling . . . 563

14.2.5 Profiler Implementation Example . . . 564
14.2.6 MPI Library Implementation Example . . . 564
Systems with Weak Symbols . . . 564
Systems Without Weak Symbols . . . 565
14.2.7 Complications . . . 565
Multiple Counting . . . 565
Linker Oddities . . . 566
Fortran Support Methods . . . 566
14.2.8 Multiple Levels of Interception . . . 566
14.3 The MPI Tool Information Interface . . . 567
14.3.1 Verbosity Levels . . . 568
14.3.2 Binding MPI Tool Information Interface Variables to MPI Objects . 568
14.3.3 Convention for Returning Strings . . . 569
14.3.4 Initialization and Finalization . . . 570
14.3.5 Datatype System . . . 571
14.3.6 Control Variables . . . 573
Control Variable Query Functions . . . 573
Example: Printing All Control Variables . . . 576
Handle Allocation and Deallocation . . . 577
Control Variable Access Functions . . . 578
Example: Reading the Value of a Control Variable . . . 579
14.3.7 Performance Variables . . . 580
Performance Variable Classes . . . 580
Performance Variable Query Functions . . . 582
Performance Experiment Sessions . . . 585
Handle Allocation and Deallocation . . . 585
Starting and Stopping of Performance Variables . . . 587
Performance Variable Access Functions . . . 588
Example: Tool to Detect Receives with Long Unexpected Message Queues . . . 590
14.3.8 Variable Categorization . . . 592
14.3.9 Return Codes for the MPI Tool Information Interface . . . 596
14.3.10 Profiling Interface . . . 596

15 Deprecated Functions **599**
15.1 Deprecated since MPI-2.0 . . . 599
15.2 Deprecated since MPI-2.2 . . . 602

16 Removed Interfaces **603**
16.1 Removed MPI-1 Bindings . . . 603
16.1.1 Overview . . . 603
16.1.2 Removed MPI-1 Functions . . . 603
16.1.3 Removed MPI-1 Datatypes . . . 603
16.1.4 Removed MPI-1 Constants . . . 603
16.1.5 Removed MPI-1 Callback Prototypes . . . 604
16.2 C++ Bindings . . . 604

17 Language Bindings **605**
17.1 Fortran Support . . . 605
17.1.1 Overview . . . 605
17.1.2 Fortran Support Through the `mpi_f08` Module . . . 606
17.1.3 Fortran Support Through the `mpi` Module . . . 609
17.1.4 Fortran Support Through the `mpif.h` Include File . . . 611
17.1.5 Interface Specifications, Procedure Names, and the Profiling Interface 612
17.1.6 MPI for Different Fortran Standard Versions . . . 617
17.1.7 Requirements on Fortran Compilers . . . 621
17.1.8 Additional Support for Fortran Register-Memory-Synchronization . 622
17.1.9 Additional Support for Fortran Numeric Intrinsic Types . . . 623
Parameterized Datatypes with Specified Precision and Exponent Range 624
Support for Size-specific MPI Datatypes . . . 627
Communication With Size-specific Types . . . 630
17.1.10 Problems With Fortran Bindings for MPI . . . 631
17.1.11 Problems Due to Strong Typing . . . 633
17.1.12 Problems Due to Data Copying and Sequence Association with Subscript Triplets . . . 633
17.1.13 Problems Due to Data Copying and Sequence Association with Vector Subscripts . . . 636
17.1.14 Special Constants . . . 637
17.1.15 Fortran Derived Types . . . 637
17.1.16 Optimization Problems, an Overview . . . 639
17.1.17 Problems with Code Movement and Register Optimization . . . 640
Nonblocking Operations . . . 640
Persistent Operations . . . 641
One-sided Communication . . . 641
MPI_BOTTOM and Combining Independent Variables in Datatypes 641
Solutions . . . 641
The Fortran ASYNCHRONOUS Attribute . . . 643
Calling MPI_F_SYNC_REG . . . 644
A User Defined Routine Instead of MPI_F_SYNC_REG . . . 645
Module Variables and COMMON Blocks . . . 646
The (Poorly Performing) Fortran VOLATILE Attribute . . . 646
The Fortran TARGET Attribute . . . 646
17.1.18 Temporary Data Movement and Temporary Memory Modification . 646
17.1.19 Permanent Data Movement . . . 648
17.1.20 Comparison with C . . . 648
17.2 Language Interoperability . . . 653
17.2.1 Introduction . . . 653
17.2.2 Assumptions . . . 653
17.2.3 Initialization . . . 653
17.2.4 Transfer of Handles . . . 654
17.2.5 Status . . . 656
17.2.6 MPI Opaque Objects . . . 658
Datatypes . . . 659
Callback Functions . . . 660
Error Handlers . . . 660

Reduce Operations . . . 661
17.2.7 Attributes . . . 661
17.2.8 Extra-State . . . 665
17.2.9 Constants . . . 665
17.2.10 Interlanguage Communication . . . 666

A Language Bindings Summary **669**
A.1 Defined Values and Handles . . . 669
A.1.1 Defined Constants . . . 669
A.1.2 Types . . . 682
A.1.3 Prototype Definitions . . . 684
C Bindings . . . 684
Fortran 2008 Bindings with the mpi_f08 Module . . . 684
Fortran Bindings with mpif.h or the mpi Module . . . 687
A.1.4 Deprecated Prototype Definitions . . . 689
A.1.5 Info Keys . . . 690
A.1.6 Info Values . . . 690
A.2 C Bindings . . . 692
A.3 Fortran 2008 Bindings with the mpi_f08 Module . . . 693
A.4 Fortran Bindings with mpif.h or the mpi Module . . . 694

B Change-Log **695**
B.1 Changes from Version 3.0 to Version 3.1 . . . 695
B.1.1 Fixes to Errata in Previous Versions of MPI . . . 695
B.1.2 Changes in MPI-3.1 . . . 697
B.2 Changes from Version 2.2 to Version 3.0 . . . 698
B.2.1 Fixes to Errata in Previous Versions of MPI . . . 698
B.2.2 Changes in MPI-3.0 . . . 699
B.3 Changes from Version 2.1 to Version 2.2 . . . 704
B.4 Changes from Version 2.0 to Version 2.1 . . . 706

Bibliography **711**

Examples Index **716**

MPI Constant and Predefined Handle Index **719**

MPI Declarations Index **724**

MPI Callback Function Prototype Index **725**

MPI Function Index **726**

List of Figures

5.1 Collective communications, an overview . . . 143
5.2 Intercommunicator allgather . . . 146
5.3 Intercommunicator reduce-scatter . . . 147
5.4 Gather example . . . 153
5.5 Gatherv example with strides . . . 154
5.6 Gatherv example, 2-dimensional . . . 155
5.7 Gatherv example, 2-dimensional, subarrays with different sizes . . . 156
5.8 Gatherv example, 2-dimensional, subarrays with different sizes and strides . 158
5.9 Scatter example . . . 163
5.10 Scatterv example with strides . . . 163
5.11 Scatterv example with different strides and counts . . . 164
5.12 Race conditions with point-to-point and collective communications . . . 217
5.13 Overlapping Communicators Example . . . 221

6.1 Intercommunicator creation using MPI_COMM_CREATE . . . 242
6.2 Intercommunicator construction with MPI_COMM_SPLIT . . . 246
6.3 Three-group pipeline . . . 263
6.4 Three-group ring . . . 264

7.1 Neighborhood gather communication example. . . . 316
7.2 Set-up of process structure for two-dimensional parallel Poisson solver. . . . 330
7.3 Communication routine with local data copying and sparse neighborhood all-to-all. . . . 331
7.4 Communication routine with sparse neighborhood all-to-all-w and without local data copying. . . . 332

11.1 Schematic description of the public/private window operations in the MPI_WIN_SEPARATE memory model for two overlapping windows. . . . 436
11.2 Active target communication . . . 438
11.3 Active target communication, with weak synchronization . . . 439
11.4 Passive target communication . . . 440
11.5 Active target communication with several processes . . . 444
11.6 Symmetric communication . . . 462
11.7 Deadlock situation . . . 463
11.8 No deadlock . . . 463

13.1 Etypes and filetypes . . . 492
13.2 Partitioning a file among parallel processes . . . 492
13.3 Displacements . . . 505

13.4 Example array file layout 558
13.5 Example local array filetype for process 1 559

17.1 Status conversion routines 657

List of Tables

2.1 Deprecated and Removed constructs 18

3.1 Predefined MPI datatypes corresponding to Fortran datatypes 25
3.2 Predefined MPI datatypes corresponding to C datatypes 26
3.3 Predefined MPI datatypes corresponding to both C and Fortran datatypes . 27
3.4 Predefined MPI datatypes corresponding to C++ datatypes 27

4.1 combiner values returned from MPI_TYPE_GET_ENVELOPE 117

6.1 MPI_COMM_* Function Behavior (in Inter-Communication Mode) 259

8.1 Error classes (Part 1) 348
8.2 Error classes (Part 2) 349

11.1 C types of attribute value argument to MPI_WIN_GET_ATTR and MPI_WIN_SET_ATTR. 414
11.2 Error classes in one-sided communication routines 452

13.1 Data access routines 507
13.2 "external32" sizes of predefined datatypes 540
13.3 I/O Error Classes 556

14.1 MPI tool information interface verbosity levels 568
14.2 Constants to identify associations of variables 569
14.3 MPI datatypes that can be used by the MPI tool information interface . . . 571
14.4 Scopes for control variables 575
14.5 Return codes used in functions of the MPI tool information interface 597

16.1 Removed MPI-1 functions and their replacements 603
16.2 Removed MPI-1 datatypes and their replacements 604
16.3 Removed MPI-1 constants 604
16.4 Removed MPI-1 callback prototypes and their replacements 604

17.1 Specific Fortran procedure names and related calling conventions. MPI_ISEND is used as an example. For routines without choice buffers, only 1A and 2A apply. 613
17.2 Occurrence of Fortran optimization problems 639

Acknowledgments

This document is the product of a number of distinct efforts in three distinct phases: one for each of MPI-1, MPI-2, and MPI-3. This section describes these in historical order, starting with MPI-1. Some efforts, particularly parts of MPI-2, had distinct groups of individuals associated with them, and these efforts are detailed separately.

This document represents the work of many people who have served on the MPI Forum. The meetings have been attended by dozens of people from many parts of the world. It is the hard and dedicated work of this group that has led to the MPI standard.

The technical development was carried out by subgroups, whose work was reviewed by the full committee. During the period of development of the Message-Passing Interface (MPI), many people helped with this effort.

Those who served as primary coordinators in MPI-1.0 and MPI-1.1 are:

- Jack Dongarra, David Walker, Conveners and Meeting Chairs
- Ewing Lusk, Bob Knighten, Minutes
- Marc Snir, William Gropp, Ewing Lusk, Point-to-Point Communication
- Al Geist, Marc Snir, Steve Otto, Collective Communication
- Steve Otto, Editor
- Rolf Hempel, Process Topologies
- Ewing Lusk, Language Binding
- William Gropp, Environmental Management
- James Cownie, Profiling
- Tony Skjellum, Lyndon Clarke, Marc Snir, Richard Littlefield, Mark Sears, Groups, Contexts, and Communicators
- Steven Huss-Lederman, Initial Implementation Subset

The following list includes some of the active participants in the MPI-1.0 and MPI-1.1 process not mentioned above.

Ed Anderson	Robert Babb	Joe Baron	Eric Barszcz
Scott Berryman	Rob Bjornson	Nathan Doss	Anne Elster
Jim Feeney	Vince Fernando	Sam Fineberg	Jon Flower
Daniel Frye	Ian Glendinning	Adam Greenberg	Robert Harrison
Leslie Hart	Tom Haupt	Don Heller	Tom Henderson
Alex Ho	C.T. Howard Ho	Gary Howell	John Kapenga
James Kohl	Susan Krauss	Bob Leary	Arthur Maccabe
Peter Madams	Alan Mainwaring	Oliver McBryan	Phil McKinley
Charles Mosher	Dan Nessett	Peter Pacheco	Howard Palmer
Paul Pierce	Sanjay Ranka	Peter Rigsbee	Arch Robison
Erich Schikuta	Ambuj Singh	Alan Sussman	Robert Tomlinson
Robert G. Voigt	Dennis Weeks	Stephen Wheat	Steve Zenith

The University of Tennessee and Oak Ridge National Laboratory made the draft available by anonymous FTP mail servers and were instrumental in distributing the document.

The work on the MPI-1 standard was supported in part by ARPA and NSF under grant ASC-9310330, the National Science Foundation Science and Technology Center Cooperative Agreement No. CCR-8809615, and by the Commission of the European Community through Esprit project P6643 (PPPE).

MPI-1.2 and MPI-2.0:

Those who served as primary coordinators in MPI-1.2 and MPI-2.0 are:

- Ewing Lusk, Convener and Meeting Chair
- Steve Huss-Lederman, Editor
- Ewing Lusk, Miscellany
- Bill Saphir, Process Creation and Management
- Marc Snir, One-Sided Communications
- Bill Gropp and Anthony Skjellum, Extended Collective Operations
- Steve Huss-Lederman, External Interfaces
- Bill Nitzberg, I/O
- Andrew Lumsdaine, Bill Saphir, and Jeff Squyres, Language Bindings
- Anthony Skjellum and Arkady Kanevsky, Real-Time

The following list includes some of the active participants who attended MPI-2 Forum meetings and are not mentioned above.

Greg Astfalk	Robert Babb	Ed Benson	Rajesh Bordawekar
Pete Bradley	Peter Brennan	Ron Brightwell	Maciej Brodowicz
Eric Brunner	Greg Burns	Margaret Cahir	Pang Chen
Ying Chen	Albert Cheng	Yong Cho	Joel Clark
Lyndon Clarke	Laurie Costello	Dennis Cottel	Jim Cownie
Zhenqian Cui	Suresh Damodaran-Kamal		Raja Daoud
Judith Devaney	David DiNucci	Doug Doefler	Jack Dongarra
Terry Dontje	Nathan Doss	Anne Elster	Mark Fallon
Karl Feind	Sam Fineberg	Craig Fischberg	Stephen Fleischman
Ian Foster	Hubertus Franke	Richard Frost	Al Geist
Robert George	David Greenberg	John Hagedorn	Kei Harada
Leslie Hart	Shane Hebert	Rolf Hempel	Tom Henderson
Alex Ho	Hans-Christian Hoppe	Joefon Jann	Terry Jones
Karl Kesselman	Koichi Konishi	Susan Kraus	Steve Kubica
Steve Landherr	Mario Lauria	Mark Law	Juan Leon
Lloyd Lewins	Ziyang Lu	Bob Madahar	Peter Madams
John May	Oliver McBryan	Brian McCandless	Tyce McLarty
Thom McMahon	Harish Nag	Nick Nevin	Jarek Nieplocha
Ron Oldfield	Peter Ossadnik	Steve Otto	Peter Pacheco
Yoonho Park	Perry Partow	Pratap Pattnaik	Elsie Pierce
Paul Pierce	Heidi Poxon	Jean-Pierre Prost	Boris Protopopov
James Pruyve	Rolf Rabenseifner	Joe Rieken	Peter Rigsbee
Tom Robey	Anna Rounbehler	Nobutoshi Sagawa	Arindam Saha
Eric Salo	Darren Sanders	Eric Sharakan	Andrew Sherman
Fred Shirley	Lance Shuler	A. Gordon Smith	Ian Stockdale
David Taylor	Stephen Taylor	Greg Tensa	Rajeev Thakur
Marydell Tholburn	Dick Treumann	Simon Tsang	Manuel Ujaldon
David Walker	Jerrell Watts	Klaus Wolf	Parkson Wong
Dave Wright			

The MPI Forum also acknowledges and appreciates the valuable input from people via e-mail and in person.

The following institutions supported the MPI-2 effort through time and travel support for the people listed above.

Argonne National Laboratory
Bolt, Beranek, and Newman
California Institute of Technology
Center for Computing Sciences
Convex Computer Corporation
Cray Research
Digital Equipment Corporation
Dolphin Interconnect Solutions, Inc.
Edinburgh Parallel Computing Centre
General Electric Company
German National Research Center for Information Technology
Hewlett-Packard
Hitachi
Hughes Aircraft Company
Intel Corporation

International Business Machines
Khoral Research
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
MPI Software Techology, Inc.
Mississippi State University
NEC Corporation
National Aeronautics and Space Administration
National Energy Research Scientific Computing Center
National Institute of Standards and Technology
National Oceanic and Atmospheric Adminstration
Oak Ridge National Laboratory
Ohio State University
PALLAS GmbH
Pacific Northwest National Laboratory
Pratt & Whitney
San Diego Supercomputer Center
Sanders, A Lockheed-Martin Company
Sandia National Laboratories
Schlumberger
Scientific Computing Associates, Inc.
Silicon Graphics Incorporated
Sky Computers
Sun Microsystems Computer Corporation
Syracuse University
The MITRE Corporation
Thinking Machines Corporation
United States Navy
University of Colorado
University of Denver
University of Houston
University of Illinois
University of Maryland
University of Notre Dame
University of San Fransisco
University of Stuttgart Computing Center
University of Wisconsin

MPI-2 operated on a very tight budget (in reality, it had no budget when the first meeting was announced). Many institutions helped the MPI-2 effort by supporting the efforts and travel of the members of the MPI Forum. Direct support was given by NSF and DARPA under NSF contract CDA-9115428 for travel by U.S. academic participants and Esprit under project HPC Standards (21111) for European participants.

MPI-1.3 and MPI-2.1:

The editors and organizers of the combined documents have been:

- Richard Graham, Convener and Meeting Chair

- Jack Dongarra, Steering Committee
- Al Geist, Steering Committee
- Bill Gropp, Steering Committee
- Rainer Keller, Merge of MPI-1.3
- Andrew Lumsdaine, Steering Committee
- Ewing Lusk, Steering Committee, MPI-1.1-Errata (Oct. 12, 1998) MPI-2.1-Errata Ballots 1, 2 (May 15, 2002)
- Rolf Rabenseifner, Steering Committee, Merge of MPI-2.1 and MPI-2.1-Errata Ballots 3, 4 (2008)

All chapters have been revisited to achieve a consistent MPI-2.1 text. Those who served as authors for the necessary modifications are:

- Bill Gropp, Front matter, Introduction, and Bibliography
- Richard Graham, Point-to-Point Communication
- Adam Moody, Collective Communication
- Richard Treumann, Groups, Contexts, and Communicators
- Jesper Larsson Träff, Process Topologies, Info-Object, and One-Sided Communications
- George Bosilca, Environmental Management
- David Solt, Process Creation and Management
- Bronis R. de Supinski, External Interfaces, and Profiling
- Rajeev Thakur, I/O
- Jeffrey M. Squyres, Language Bindings and MPI-2.1 Secretary
- Rolf Rabenseifner, Deprecated Functions and Annex Change-Log
- Alexander Supalov and Denis Nagorny, Annex Language Bindings

The following list includes some of the active participants who attended MPI-2 Forum meetings and in the e-mail discussions of the errata items and are not mentioned above.

Pavan Balaji, Purushotham V. Bangalore, Brian Barrett,
Richard Barrett, Christian Bell, Robert Blackmore,
Gil Bloch, Ron Brightwell, Jeffrey Brown,
Darius Buntinas, Jonathan Carter, Nathan DeBardeleben,
Terry Dontje, Gabor Dozsa, Edric Ellis,
Karl Feind, Edgar Gabriel, Patrick Geoffray,
David Gingold, Dave Goodell, Erez Haba,
Robert Harrison, Thomas Herault, Steve Hodson,
Torsten Hoefler, Joshua Hursey, Yann Kalemkarian,
Matthew Koop, Quincey Koziol, Sameer Kumar,
Miron Livny, Kannan Narasimhan, Mark Pagel,
Avneesh Pant, Steve Poole, Howard Pritchard,
Craig Rasmussen, Hubert Ritzdorf, Rob Ross,
Tony Skjellum, Brian Smith, Vinod Tipparaju,
Jesper Larsson Träff, Keith Underwood

The MPI Forum also acknowledges and appreciates the valuable input from people via e-mail and in person.

The following institutions supported the MPI-2 effort through time and travel support for the people listed above.

Argonne National Laboratory
Bull
Cisco Systems, Inc.
Cray Inc.
The HDF Group
Hewlett-Packard
IBM T.J. Watson Research
Indiana University
Institut National de Recherche en Informatique et Automatique (INRIA)
Intel Corporation
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
Mathworks
Mellanox Technologies
Microsoft
Myricom
NEC Laboratories Europe, NEC Europe Ltd.
Oak Ridge National Laboratory
Ohio State University
Pacific Northwest National Laboratory
QLogic Corporation
Sandia National Laboratories
SiCortex
Silicon Graphics Incorporated
Sun Microsystems, Inc.
University of Alabama at Birmingham
University of Houston
University of Illinois at Urbana-Champaign

University of Stuttgart, High Performance Computing Center Stuttgart (HLRS)
University of Tennessee, Knoxville
University of Wisconsin

Funding for the MPI Forum meetings was partially supported by award #CCF-0816909 from the National Science Foundation. In addition, the HDF Group provided travel support for one U.S. academic.

MPI-2.2:

All chapters have been revisited to achieve a consistent MPI-2.2 text. Those who served as authors for the necessary modifications are:

- William Gropp, Front matter, Introduction, and Bibliography; MPI-2.2 chair.
- Richard Graham, Point-to-Point Communication and Datatypes
- Adam Moody, Collective Communication
- Torsten Hoefler, Collective Communication and Process Topologies
- Richard Treumann, Groups, Contexts, and Communicators
- Jesper Larsson Träff, Process Topologies, Info-Object and One-Sided Communications
- George Bosilca, Datatypes and Environmental Management
- David Solt, Process Creation and Management
- Bronis R. de Supinski, External Interfaces, and Profiling
- Rajeev Thakur, I/O
- Jeffrey M. Squyres, Language Bindings and MPI-2.2 Secretary
- Rolf Rabenseifner, Deprecated Functions, Annex Change-Log, and Annex Language Bindings
- Alexander Supalov, Annex Language Bindings

The following list includes some of the active participants who attended MPI-2 Forum meetings and in the e-mail discussions of the errata items and are not mentioned above.

Pavan Balaji	Purushotham V. Bangalore	Brian Barrett
Richard Barrett	Christian Bell	Robert Blackmore
Gil Bloch	Ron Brightwell	Greg Bronevetsky
Jeff Brown	Darius Buntinas	Jonathan Carter
Nathan DeBardeleben	Terry Dontje	Gabor Dozsa
Edric Ellis	Karl Feind	Edgar Gabriel
Patrick Geoffray	Johann George	David Gingold
David Goodell	Erez Haba	Robert Harrison
Thomas Herault	Marc-André Hermanns	Steve Hodson
Joshua Hursey	Yutaka Ishikawa	Bin Jia
Hideyuki Jitsumoto	Terry Jones	Yann Kalemkarian
Ranier Keller	Matthew Koop	Quincey Koziol
Manojkumar Krishnan	Sameer Kumar	Miron Livny
Andrew Lumsdaine	Miao Luo	Ewing Lusk
Timothy I. Mattox	Kannan Narasimhan	Mark Pagel
Avneesh Pant	Steve Poole	Howard Pritchard
Craig Rasmussen	Hubert Ritzdorf	Rob Ross
Martin Schulz	Pavel Shamis	Galen Shipman
Christian Siebert	Anthony Skjellum	Brian Smith
Naoki Sueyasu	Vinod Tipparaju	Keith Underwood
Rolf Vandevaart	Abhinav Vishnu	Weikuan Yu

The MPI Forum also acknowledges and appreciates the valuable input from people via e-mail and in person.

The following institutions supported the MPI-2.2 effort through time and travel support for the people listed above.

Argonne National Laboratory
Auburn University
Bull
Cisco Systems, Inc.
Cray Inc.
Forschungszentrum Jülich
Fujitsu
The HDF Group
Hewlett-Packard
International Business Machines
Indiana University
Institut National de Recherche en Informatique et Automatique (INRIA)
Institute for Advanced Science & Engineering Corporation
Intel Corporation
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
Mathworks
Mellanox Technologies
Microsoft
Myricom
NEC Corporation
Oak Ridge National Laboratory

Ohio State University
Pacific Northwest National Laboratory
QLogic Corporation
RunTime Computing Solutions, LLC
Sandia National Laboratories
SiCortex, Inc.
Silicon Graphics Inc.
Sun Microsystems, Inc.
Tokyo Institute of Technology
University of Alabama at Birmingham
University of Houston
University of Illinois at Urbana-Champaign
University of Stuttgart, High Performance Computing Center Stuttgart (HLRS)
University of Tennessee, Knoxville
University of Tokyo
University of Wisconsin

Funding for the MPI Forum meetings was partially supported by awards #CCF-0816909 and #CCF-1144042 from the National Science Foundation. In addition, the HDF Group provided travel support for one U.S. academic.

MPI-3:

MPI-3 is a signficant effort to extend and modernize the MPI Standard.
The editors and organizers of the MPI-3 have been:

- William Gropp, Steering committee, Front matter, Introduction, Groups, Contexts, and Communicators, One-Sided Communications, and Bibliography
- Richard Graham, Steering committee, Point-to-Point Communication, Meeting Convener, and MPI-3 chair
- Torsten Hoefler, Collective Communication, One-Sided Communications, and Process Topologies
- George Bosilca, Datatypes and Environmental Management
- David Solt, Process Creation and Management
- Bronis R. de Supinski, External Interfaces and Tool Support
- Rajeev Thakur, I/O and One-Sided Communications
- Darius Buntinas, Info Object
- Jeffrey M. Squyres, Language Bindings and MPI-3 Secretary
- Rolf Rabenseifner, Steering committee, Terms and Definitions, and Fortran Bindings, Deprecated Functions, Annex Change-Log, and Annex Language Bindings
- Craig Rasmussen, Fortran Bindings

The following list includes some of the active participants who attended MPI-3 Forum meetings or participated in the e-mail discussions and who are not mentioned above.

Tatsuya Abe	Tomoya Adachi	Sadaf Alam
Reinhold Bader	Pavan Balaji	Purushotham V. Bangalore
Brian Barrett	Richard Barrett	Robert Blackmore
Aurelien Bouteiller	Ron Brightwell	Greg Bronevetsky
Jed Brown	Darius Buntinas	Devendar Bureddy
Arno Candel	George Carr	Mohamad Chaarawi
Raghunath Raja Chandrasekar	James Dinan	Terry Dontje
Edgar Gabriel	Balazs Gerofi	Brice Goglin
David Goodell	Manjunath Gorentla	Erez Haba
Jeff Hammond	Thomas Herault	Marc-André Hermanns
Jennifer Herrett-Skjellum	Nathan Hjelm	Atsushi Hori
Joshua Hursey	Marty Itzkowitz	Yutaka Ishikawa
Nysal Jan	Bin Jia	Hideyuki Jitsumoto
Yann Kalemkarian	Krishna Kandalla	Takahiro Kawashima
Chulho Kim	Dries Kimpe	Christof Klausecker
Alice Koniges	Quincey Koziol	Dieter Kranzlmueller
Manojkumar Krishnan	Sameer Kumar	Eric Lantz
Jay Lofstead	Bill Long	Andrew Lumsdaine
Miao Luo	Ewing Lusk	Adam Moody
Nick M. Maclaren	Amith Mamidala	Guillaume Mercier
Scott McMillan	Douglas Miller	Kathryn Mohror
Tim Murray	Tomotake Nakamura	Takeshi Nanri
Steve Oyanagi	Mark Pagel	Swann Perarnau
Sreeram Potluri	Howard Pritchard	Rolf Riesen
Hubert Ritzdorf	Kuninobu Sasaki	Timo Schneider
Martin Schulz	Gilad Shainer	Christian Siebert
Anthony Skjellum	Brian Smith	Marc Snir
Raffaele Giuseppe Solca	Shinji Sumimoto	Alexander Supalov
Sayantan Sur	Masamichi Takagi	Fabian Tillier
Vinod Tipparaju	Jesper Larsson Träff	Richard Treumann
Keith Underwood	Rolf Vandevaart	Anh Vo
Abhinav Vishnu	Min Xie	Enqiang Zhou

The MPI Forum also acknowledges and appreciates the valuable input from people via e-mail and in person.

The MPI Forum also thanks those that provided feedback during the public comment period. In particular, the Forum would like to thank Jeremiah Wilcock for providing detailed comments on the entire draft standard.

The following institutions supported the MPI-3 effort through time and travel support for the people listed above.

Argonne National Laboratory
Bull
Cisco Systems, Inc.
Cray Inc.
CSCS
ETH Zurich
Fujitsu Ltd.

German Research School for Simulation Sciences
The HDF Group
Hewlett-Packard
International Business Machines
IBM India Private Ltd
Indiana University
Institut National de Recherche en Informatique et Automatique (INRIA)
Institute for Advanced Science & Engineering Corporation
Intel Corporation
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
Mellanox Technologies, Inc.
Microsoft Corporation
NEC Corporation
National Oceanic and Atmospheric Administration, Global Systems Division
NVIDIA Corporation
Oak Ridge National Laboratory
The Ohio State University
Oracle America
Platform Computing
RIKEN AICS
RunTime Computing Solutions, LLC
Sandia National Laboratories
Technical University of Chemnitz
Tokyo Institute of Technology
University of Alabama at Birmingham
University of Chicago
University of Houston
University of Illinois at Urbana-Champaign
University of Stuttgart, High Performance Computing Center Stuttgart (HLRS)
University of Tennessee, Knoxville
University of Tokyo

Funding for the MPI Forum meetings was partially supported by awards #CCF-0816909 and #CCF-1144042 from the National Science Foundation. In addition, the HDF Group and Sandia National Laboratories provided travel support for one U.S. academic each.

MPI-3.1:

This is the initial stub for the MPI-3.1 credits. We use this to collect information on the participants and their institutions.

Marc-Andre Hermanns Forschungszentrum Jülich
German Research School for Simulation Sciences
and
Jülich Aachen Research Alliance, High-Performance Computing (JARA-HPC)