Message Passing Toolkit: Release Notes

004–3689–001

Contents

Introduction [1]

This document provides an overview of the Cray Message Passing Toolkit and Message Passing Toolkit for UNICOS and UNICOS/mk 1.4 releases. MPT is a software package that supports interprocess data exchange for applications that use concurrent, cooperating processes on a single host or on multiple hosts. Data exchange is done through *message passing*, which is the use of library calls to request data delivery from one process to another or between groups of processes.

The MPT 1.4 release is supported on the following platforms:

- Cray systems running UNICOS release 10.0.0.3 or later.

- Cray T3E systems running UNICOS/mk release 1.5 or later.

- Cray systems running Programming Environment release 3.3 or later.

  Note: If UNICOS is upgraded to a new major release after MPT 1.4 has been installed (for example, from UNICOS 9.0 to UNICOS 10.0.0.3), the `mpt_config` configuration script must be run to reconfigure some of the MPT 1.4 software. Reconfigure using the following instructions:

  ```
  % cd /opt/ctl/mpt/mpt
  ```
  or:

  ```
  % cd /opt/ctl/mpt/version
  ```

  ```
  % ./mpt_config -u
  % ./mpt_config -i
  ```

These release notes include information about the following:

- New features

- Compatibilities and differences

- Documentation

- Customer services

- Release package contents

The remainder of this chapter discusses the following topics:

- MPT components

- MPT software

- Product description

- Distribution of these release notes

- Reader comments

1.1 MPT components

The MPT 1.4 package contains the following components and accompanying documentation:

- Parallel Virtual Machine (PVM)

- Message Passing Interface (MPI)

- Logically shared, distributed memory (SHMEM) message passing routines

Caution: The MPT 1.4 release is the last release that will support PVM on Cray Inc. systems.

The PVM and MPI components of MPT are based on industry standards for the message passing programming. For a description of MPI, see the MPI man pages, especially MPI(1). For a description of PVM, see the PVM man pages, especially pvm(1) and pvm_intro(1).

The SHMEM library is packaged with MPT on UNICOS. On UNICOS/mk systems, the SHMEM library is delivered with the CrayLibs package. See the SHMEM man pages, especially intro_shmem(3) for more information.

1.2 MPT software

Software included in MPT was designed to be used with the Cray Programming Environment 3.3 release or later.

MPT software is self-configuring, based on the operating system configuration in effect at the time of installation. You are not required to do any configuration at the initial installation of the product. If, however, you upgrade your operating system level to a new major release or change the system host name, you will need to reconfigure the MPT software. This reconfiguration can be done as follows:

```
cray# cd /opt/ctl/mpt/mpt
cray# ./mpt_config -u
cray# ./mpt_config -i
```

> **Note:** You can also change to the /opt/ctl/mpt/*version* directory and do the configuration from there.

The Modules software package is used to support the installation of both the Programming Environment and MPT. To use the MPT software, load the mpt module in addition to loading the Programming Environment module. For information on using modules, see *Installing Programming Environment Products*, or, if the Programming Environment has already been installed on your system, see the online ASCII file /opt/ctl/doc/README. After you have initialized the modules, enter the following command to access the MPT software:

```
module load mpt
```

To unload the mpt module, enter the following command:

```
module unload mpt
```

1.3 Product description

The MPT package contains the following components:

- Optimized implementations of the Message Passing Interface (MPI) for UNICOS systems and UNICOS/mk systems. These implementations are based on the MPI 1.2 specification from the Message Passing Interface Forum. They support high-speed communications among the processors on each system. The implementation on UNICOS/mk also supports a subset of the MPI-2 specification. The implementation on UNICOS systems supports high-speed communications between UNICOS systems of like architecture.

- An optimized version of Parallel Virtual Machine (PVM) for UNICOS systems and UNICOS/mk systems. This version is based on Oak Ridge National Laboratories (ORNL) version 3.3.10. It supports high-speed communications among the processors on each system. It also supports high-speed communications between heterogeneous systems.

> **Caution:** MPT 1.4 will be the last release in which PVM will be supported.

- Logically shared, distributed memory (SHMEM) message passing capability using the shmem_get(3) and shmem_put(3) functions on UNICOS systems. (SHMEM on UNICOS/mk systems is delivered with CrayLibs.) This

high-performance library supports communications among the processors on each system.

MPI is a standard specification for message passing libraries, allowing portable message passing programs in the Fortran and C languages. MPI was created by the Message Passing Interface Forum (MPIF). MPIF is not sanctioned or supported by any official standards organization. Its goal was to develop a widely used standard for writing message passing programs. The implementation of this standard includes a library (`libmpi.a`), a run-time command (`mpirun(1)`), and a library that allows profiling of message passing applications (`libpmpi.a`). On UNICOS systems, the profiling functionality is included in `libmpi.a`.

PVM is a software project that was developed jointly by Oak Ridge National Laboratories (ORNL), the University of Tennessee, and Emory University. PVM consists of a main library, a user-level daemon (`pvmd3(1)`), a console (`pvm(1)`), and some additional commands and libraries. The main PVM library name is `libpvm3.a`.

The PVM and MPI libraries provide communications and synchronization functions that are necessary for writing distributed applications. For example, you can add calls that cause one task to send a message to another, or to receive a message, or to wait until another task is finished. The PVM software supports heterogeneous systems by automatically converting data. The MPI software is supported between UNICOS systems of the same architecture and within a partition on UNICOS/mk systems.

SHMEM message passing is another form of distributed programming. It differs from PVM and MPI message passing in that it uses one-sided communication (that is, one processing element (PE) on a Cray T3E system can send or receive data from another PE without the knowledge of that PE).

To use PVM, MPI, or SHMEM message passing directly, you must change your source code to add the appropriate calls.

MPI support for Cray T3D systems is provided through a third-party product available from Edinburgh Parallel Computing Centre (EPCC).

1.4 Distribution of these release notes

You can access ASCII and PostScript versions of these release notes electronically, as follows:

- Through the CRInform system, which is an online information and problem-reporting system for customers. For more information on CRInform, see Section 5.5, page 15.

- Through the `craypark` system in the `/home/craypark/release_docs` directory. The `craypark` system is available only to Cray Inc. service personnel.

1.5 Reader comments

If you have comments about the technical accuracy, content, or organization of this document, please tell us. Be sure to include the title and part number of the document with your comments.

You can contact us in any of the following ways:

- Send email to the following address:

  `pubs@cray.com`

- Send a fax to the attention of "Software Publications" at: +1 651 683 5599.

- Call the Software Publications Group, through the Technical Assistance Center, using one of the following numbers:

  1 800 950 2729 (toll free from the United States and Canada) or
  +1 651 683 5600

- Send mail to the following address:

  Software Publications
  Cray Inc.
  655F Lone Oak Drive
  Eagan, MN, 55121–1560, USA

We value your comments and will respond to them promptly.

New Features [2]

This chapter describes MPT 1.4 features and features available in earlier update releases.

2.1 Changes to Cray T3E systems

The new features described in the following sections apply to Cray T3E systems only.

2.1.1 Latency optimizations

Enhancements provided by the Edinburgh Parallel Computing Centre (EPCC) reduce the system's minimum latency by a factor of 2. The optimizations to the MPI library make use of the E-register hardware message queues in the critical MPI communication path. Latency is the time between when a processor requests data and when it is able to use that data. This feature was first available in the MPT 1.3.0.5 release.

2.1.2 Improved MPI barrier performance

This feature, provided by EPCC, improves the performance of the `MPI_Barrier` synchronization routine using `MPI_COMM_WORLD`. This feature was first available in the MPT 1.3.0.5 release.

2.1.3 MPI-2 one-sided communication

The MPI-2 one-sided communication feature is implemented by EPCC. One-sided communication, or remote memory access (RMA), moves data directly between local memory and remote memory. The `MPI_Put`(3), `MPI_Get`(3), and `MPI_Accumulate`(3) routines perform the actual data movement. Three synchronization models are available; see the `MPI_Win`(3) man page for more information.

2.1.4 Improved default internal buffering

This features is a change in the default internal buffering for MPI. The MPI library no longer buffers contiguous data by default on Cray T3E-900 systems or later. Users can enable internal buffering with the `MPI_BUFFER_MAX`

environment variable, which is documented in the `MPI`(1) man page. This feature was first available in the MPT 1.3.0.4 release. (This feature is also activated for Cray T3E-600 systems if hardware streams are disabled.)

Users with erroneous programs that assume an MPI library buffers data internally, can use the `MPI_BUFFER_MAX` environment variable to specify how many bytes of data the MPI library should buffer. This change is also documented in the `MPI`(1) man page under the description of the `MPI_BUFFER_MAX` environment variable. This change in default buffering increases the default peak bandwidth for MPI applications. Peak bandwidth is defined as the maximum number of bytes that can be transferred in a second.

2.2 Changes to Cray SV1 systems

2.2.1 Latency optimizations

Latency optimizations have been implemented. The minimum MPI latency for a 0-byte message within a single system has been reduced from approximately 78 user seconds to 60 user seconds.

Latency is the time between when a processor requests data and when it is able to use that data.

2.2.2 More MPI processes

This release includes support for up to 63 MPI processes within a single Cray SV1 system.

Compatibilities and Differences [3]

This chapter describes the compatibility issues that users should consider when purchasing MPT version 1.4.

3.1 Internal buffering change

Users may have to change the way they run MPI programs because internal buffering is no longer done as of MPT 1.4 on Cray T3E-900 systems or later:

- If a user is accustomed to turn off internal buffering by setting the environment variable `MPI_BUFFER_MAX` to 0, doing so is no longer necessary.

- If an illegal program relies on internal buffering, the user will have to set the environment variable `MPI_BUFFER_MAX` to some value for the program to run.

Documentation [4]

This chapter describes the documentation that supports the MPT 1.4 release. It contains information about the following:

- Online information access

- Documentation included with the release

4.1 Online information access

The following types of online information products are available to MPT 1.4 customers:

- Publications information at the following URL:

  ```
  http://www.cray.com/swpubs/
  ```

  This web site contains information that allows you to browse documents online, order documents, download documents, and send feedback to the Cray publications department.

- The Cray online documentation server, which allows you to view manuals online at your site using a browser such as Netscape. Please see your Cray system administrator for the local URL of the Cray online documentation server.

- The *User Publications Catalog*, describes the availability and content of all Cray hardware and software documents for customers. Customers who subscribe to the Cray Inform (CRInform) program can access this information on the CRInform system.

- Man pages, which describe Cray software commands and routines. To see a detailed description of a particular command or routine, use the man(1) command.

- A message system, which provides explanations of error messages. To see an explanation of a message, use the explain(1) command.

- Cray online glossary, which explains many of the terms used in Cray documentation. To get a definition, use the define(1) command.

4.2 Documentation included with the release

Most documentation included with the MPT 1.4 release is available through the Cray online documentation server and man pages; it may also be available in printed form. For information on accessing these documents, see Section 4.1, page 11.

The man pages for MPI, PVM, and SHMEM are available in online form. SHMEM man pages are printed in the *Application Programmer's Library Reference Manual.* The following publications are available for purchase through the Cray Distribution Center. For ordering information, see Section 6.4.1, page 23.

- *Message Passing Toolkit: Release Notes*

- *Application Programmer's Library Reference Manual*

- *Using MPI: Portable Parallel Programming with the Message-Passing Interface*, by Gropp, Lusk, and Skjellum, publication TPD–0011

Customer Services [5]

This chapter describes the MPT support policy and the following customer services that are available to support the MPT 1.4 release:

- MPT web site

- Training support

- Software problem reporting and resolution process

- Cray CRInform

- *Cray Service Bulletin*

- Additional resources

5.1 Cray support policy

Cray Inc. offers standard support for problems in the following categories:

- Build problems

- Run-time problems specific to Cray computer systems

Fixes to problems will be made available through one of the following mechanisms:

- Upgrade utility

- Major release

- Minor release

5.1.1 Problem resolution

Cray Inc. will address MPI and SHMEM problems in a timely fashion.

Cray Inc. will help customers resolve PVM problems by forwarding general PVM problems to developers at Oak Ridge National Laboratories (ORNL) and the University of Tennessee, and assisting in their resolution. Problems with Cray enhancements to PVM will be addressed in a timely fashion.

5.2 MPT web site

Current product information regarding MPT is available online through the web site at the following URL:

```
http://www.cray.com/products/software/mpt.html
```

5.3 Training support

For information about Cray Inc. courses, training office locations, current class schedules, and training services listed by country, contact us in one of the following ways:

- On the web at the following URL:

  ```
  http://www.cray.com/products/training/
  ```

- Call one of the following numbers:

  1 800 800 4744 (toll free in the United States or Canada)

  +1 651 683 3825 (from all other locations)

5.4 Software problem reporting and resolution process

If you experience problems with MPT 1.4, contact your service representative. Your service representative will work with you to resolve the problem. If you choose to have full-time or part-time on-site support, your on-site support personnel are your primary contacts for service. If you have elected not to have on-site support, please either call the Customer Support Center and report your problem to them or submit a request for technical assistance (RTA) through the CRInform program (see Section 5.5).

When you report problems, use the product name most appropriate for the problem and refer to MPT for all products. Use any of the following product names:

- PVM

- MPI

- LIBSMA

For current information on the status of reported problems, see the Software Problem Report (SPR) section of CRInform.

5.5 CRInform

Cray Inform (CRInform) is a web-based information and problem-reporting service for customers. Using CRInform, you can do the following:

- Report software problems

- Request technical assistance

- Communicate directly with other customers

- Read about similar software problems reported at other sites

- Learn about solutions to various problems

- Find information about classes

- Read about new products, and more

CRInform automatically logs as news items those events that are pertinent to your site, so you do not have to search through the system for new information. The logged events include changes in SPR or RTA activity, new orderable software, new issues of the *Cray Service Bulletin*, new field notices (FNs), new software release documents, new software problem fix information, new marketing information, and new CRInform information. You can also get automatic email notification of any or all of the news items.

Version 4.0 of CRInform is available through the web. You need access to the CRInform web server and a browser (such as Netscape), which allows you to view information and make service requests. You can use your own site's browser, or use either the Mosaic or the Lynx browser available on the CRInform system. You can access CRInform at the following URL:

```
http://crinform.cray.com/
```

5.6 *Cray Service Bulletin*

Customers who have a support contract receive the *Cray Service Bulletin* (CRSB), the Cray customer newsletter. The *Cray Service Bulletin* provides product and support information about Cray supercomputers. It is published six times a year. The *Cray Service Bulletin* is available in CRInform (`http://crinform.cray.com/`) (see Section 5.5, page 15).

5.7 Additional resources

Along with the standard support described in the preceding sections, users of PVM and MPI have a variety of other resources available to help them use these products. These resources should supplement, not replace, standard Cray support. Users are encouraged to use Cray support when they encounter what may be a PVM or MPI bug. However, these other resources can be helpful in gaining a better understanding of PVM and MPI and how they can be used. The following sections describe some of these resources.

5.7.1 MPI standard

You can access the MPI standard in any of the following ways:

* As online PostScript or hypertext on the web at the following URL:

  `http://www-unix.mcs.anl.gov/mpi`

* As a journal article in the fall 1994 issue of the *Journal of Supercomputing Applications*

5.7.2 News groups

The `comp.parallel.pvm` and `comp.parallel.mpi` Internet news groups focus on PVM and MPI and related products. The PVM and MPI developers use these news groups to announce new releases of the products and to discuss related topics, such as the PVM User Group meeting (see Section 5.7.5). Users are also free to use the news groups to interact with one another by responding to postings made by others.

To access `comp.parallel.pvm` or `comp.parallel.mpi`, you must have access to the Internet and a news reader program. If necessary, contact your site administrator for additional information.

5.7.3 Web servers

Several PVM-related web servers are available. There is one for `netlib` (see Section 5.7.7, page 17) that includes the PVM source code, other material on PVM, and additional software packages and subjects. The URL for this server is as follows:

`http://www.netlib.org/pvm3/index.html`

Another web page specific to PVM can be found at the following URL:

`http://www.epm.ornl.gov/pvm/pvm_home.html`

5.7.4 MPI user group meetings

You can obtain information about upcoming MPI developers conferences and user group meetings through `comp.parallel.mpi` or through the MPI web page listed in Section 5.7.1, page 16.

5.7.5 PVM user group meetings

There have been several PVM user group meetings held in the past, and there are plans to continue them in the future. These meetings have been quite successful, attracting a wide range of presentations from the PVM development team, vendors working on PVM, and users working on PVM-related tools or using PVM in applications.

Announcements of PVM user group meetings are posted in the news group `comp.parallel.pvm`.

5.7.6 PVM email support

The PVM developers have established the following email alias, which they use to answer questions and provide support, as time permits:

`pvm@msr.epm.ornl.gov`

Users first try Cray support or contacts when asking questions. This email address, however, can be used when needed.

The MPI Forum email discussions, and both current and earlier versions of the standard, are available from the `netlib` server. For more information about `netlib`, see the following section.

5.7.7 `netlib` source

Public domain versions of PVM and MPI are available from a source called `netlib`. If your site is using the versions of these products contained in MPT, you will not need to obtain the source code from `netlib`. However, the following additional resources available from `netlib` might be useful:

- Test and sample programs contained in the public domain releases that are not included in the Cray release

- Papers and presentations by the PVM and MPI developers that describe PVM, MPI, and related products

- Presentations from past PVM user group meetings

You can obtain information from `netlib` in the following ways:

- Use the web, which allows you to obtain the files directly. This is perhaps the easiest and most convenient method. The PVM page at `netlib` is located at the following URL:

  ```
  http://www.netlib.org/pvm3/index.html
  ```

  The MPI page at `netlib` is located at the following URL:

  ```
  http://www.netlib.org/mpi/index.html
  ```

- Send email to `netlib@ornl.gov` (or `netlib@research.att.com`), and include the following in the `Subject:` line or message body:

  ```
  send index from pvm3
  ```

  or

  ```
  send index from mpi
  ```

  Once you receive the index, similar `send` messages will return parts of PVM or MPI. This assumes the ORNL or AT&T host can determine how to return your email. If not, use anonymous `ftp`.

- Use anonymous `ftp` from `netlib2.cs.utk.edu`.

- Use anonymous `rcp` from `netlib2.cs.utk.edu`. For example:

  ```
  rcp anon@netlib2.cs.utk.edu:pvm3/your-local-file
  ```

  or

  ```
  rcp anon@netlib2.cs.utk.edu:mpi/your-local-file
  ```

  You can use the following commands to obtain file lists:

  ```
  rsh netlib2.cs.utk.edu -l anon ls pvm3
  ```

  or

```
rsh netlib2.cs.utk.edu -l anon ls mpi
```

- Use the `xnetlib` browsing tool (which you can obtain from `netlib` by using one of the preceding methods).

- For access in Europe, use the duplicate collection in Oslo, as follows:

    Internet: `netlib@nac.no`

    EARN/BITNET: `netlib%nac.no@norunix.bitnet`

    X.400: `s=netlib; o=nac; c=no`

    EUNET/uucp: `nac!netlib`

- For access in the Pacific, use the following address, located at the University of Wollongong, NSW, Australia:

    `netlib@draci.cs.uow.edu.au`

Release Package [6]

This chapter contains the following information about the MPT 1.4 release package:

- Summary of hardware and software requirements

- Licensing information

- List and description of the contents of the release package

- Ordering instructions

6.1 Hardware and software requirements

To build and use MPT software, you must have the following software:

- UNICOS release 10.0.0.3 or later or UNICOS/mk release 1.5 or later.

- Cray Programming Environment release 3.3 or later.

- The `Modules` software package. In addition to having the Cray Programming Environment installed on your system, you must have the `PrgEnv` module loaded before you can use MPT. For information on using modules, see *Installing Programming Environment Products.*

6.2 Licensing information

To order MPT 1.4, customers must pay for the product and agree to the stated terms and conditions that relate to the product. The following sections provide names and addresses of licensing contacts.

6.2.1 Licensing contacts for customers in the U.S. and Canada

For information on the licensing and pricing of MPT 1.4, customers in the United States and Canada should see their account representative or field contract negotiator. If those people are unavailable, or if you are not sure who to call, contact:

Jenny Gross
Cray Inc.
655F Lone Oak Drive

Eagan, MN 55121 USA
Telephone: +1 651 683 5661
Fax: +1 651 683 7482
E-mail: `jennyg@cray.com`

6.2.2 Licensing contacts for customers outside of the U.S. and Canada

Customers outside of the United States and Canada may obtain further licensing and export information by contacting their account representative or Jenny Gross at the following address:

Jenny Gross
Cray Inc.
655F Lone Oak Drive
Eagan, MN 55121 USA
Telephone: +1 651 683 5661
Fax: +1 651 683 7482
E-mail: `jennyg@cray.com`

6.3 Release package contents

The MPT 1.4 release package includes the following:

- Cray Message Passing Toolkit release CD

- *Message Passing Toolkit: Release Notes* (this manual)

- *Installing Programming Environment Products*

MPT 1.4 consists of the following software:

- PVM

- MPI

- SHMEM `libsma` library (on UNICOS systems)

- Header files

- PVM, MPI, and SHMEM man pages

- `Modules` software

6.4 Ordering the MPT 1.4 release package

MPT 1.4 is distributed by order to sites that have agreed to the stated terms and conditions that relate to the product (for licensing details, see Section 6.2, page 21). The most current revision of the release package is supplied.

You can use one of the following methods to order MPT 1.4:

- Customers outside of the United States and Canada should contact their account representative or contract negotiator for ordering information.

- Customers who subscribe to the CRInform program can order software release packages electronically by using the `Order Cray Software` option.

- Customers can contact the order desk at the Minnesota Distribution Center by telephone (+1 651 683 5907 or 1 800 284 2729 extension 35907) or through email (`orderdsk@cray.com`).

Software will be shipped by ground service or 5-day international service unless otherwise requested.

Sites outside the United States may be required to provide a customer-signed Letter of Assurance before this software can be shipped. Address questions about which customers must sign Letters of Assurance or which software requires such a letter to Alan Benfell at the following address:

Alan Benfell
International Administration
655F Lone Oak Drive
Eagan, MN 55121 USA
Telephone: +1 651 683 7476 or 1 800 284 2729
Fax: +1 651 683 7297
Email: `benfa@cray.com`

6.4.1 Obtaining publications

The MPT 1.4 release package includes the publications listed in Section 6.3, page 22.

Cray Inc. maintains publications information at the following URL:

`http://www.cray.com/swpubs/`

This web site contains information that allows you to browse documents online, order documents, and send feedback to Cray.

The *User Publications Catalog* describes the availability and content of all Cray hardware and software documents for customers. Customers who subscribe to CRInform can access this information.

Customers who subscribe to the CRInform program can also order software release packages electronically by using the `Order Cray Software` option.

To order a document, call the Minnesota Distribution Center at +1 651 683 5907. Cray employees can send email to `orderdsk@cray.com`.

Customers outside of the United States and Canada should contact their local service organization for ordering and documentation information.

Index