

Overview

HPE Message Passing Interface (MPI)

v1.1

The HPE Message Passing Interface (MPI) is an MPI development environment designed to enable the development and optimization of high performance computing (HPC) applications. The HPE Message Passing Interface (MPI) leverages a scalable MPI library and boosts performance of existing MPI applications on HPE HPC clusters without requiring recompilation.

The HPE Message Passing Interface (MPI) accelerates application performance by:

- Tuning applications at runtime without requiring you to recompile your code.
- Providing access to profiling tools that identify performance bottlenecks and load imbalances for MPI applications, as well as provide guided placements for threads to improve application performance.
- Enabling applications developed with other MPI implementations to work seamlessly with HPE Message Passing Interface (MPI), and its many features, without needing to be recompiled.

HPE Message Passing Interface: Specifications at a Glance:

What's New

- Support for HPE Apollo 70 systems
 - Run multi-host jobs more securely with RHEL 7 SELinux
 - Increase application performance by creating fewer memory copies and non-blocking activity with Mellanox InfiniBand™ tag matching
-

Overview

Standards	MPI 3.1 OpenSHMEM 1.4
Interoperability	Intel MPI IBM Spectrum MPI Open MPI Mellanox HPC-X MPI Cray MPI MPICH MVAPICH
Programming Languages	Native support for Fortran, C and C++
Operating Systems	SUSE® Linux Enterprise Server 12 Red Hat® Enterprise Linux 7 CentOS 7
Fabric Support	Multi-rail Intel® Omni-Path Multi-rail Mellanox Infiniband™ HPE Superdome Flex Grid TCP/IP
Supported Hardware	HPE SGI 8600 HPE Apollo Systems, including the Apollo 2000, 6000 and 6500 (Both Gen9 and Gen10) HPE Apollo 20 and 40 systems ARM processor-based Apollo 70 system HPE ProLiant DL360, DL380 Integrity MC990X, Superdome Flex
Third Party Software Integration	<p>Software development tools, including:</p> <ul style="list-style-type: none"> • <u>Intel Parallel Studio XE*</u> • PGI Professional Edition • gcc • GFortran • Allinea Forge* • Rogue Wave TotalView* • TAU • Vampir <p>Workload schedulers, including:</p> <ul style="list-style-type: none"> • <u>Altair PBS Professional*</u> • Adaptive Computing Moab/ TORQUE* • SLURM • IBM Spectrum LSF • Univa Grid Engine <p>Fabric software, including:</p> <ul style="list-style-type: none"> • Mellanox Fabric Collective Accelerator (FCA)

Standard Features

Models

Licensing and Media Options

HPE Performance Message Passing Interface for 1 Socket 3yr Support LTU

Q2A48A

NOTE: One license per socket.

NOTE: Includes three (3) years of HPE Pointnext support.

NOTE: Please contact HPE Pointnext for 4 and 5 year support options.

HPE Message Passing Interface (MPI) Media Kit

Q5U69A

NOTE: One media kit per solution.

HPE Performance Message Passing Interface FIO

Q8K21A

NOTE: For factory installation only; this SKU does not include the software license. Please order with PN Q2A48A.

NOTE: Order one PN per node.

NOTE: Only available for the HPE SGI 8600, Apollo 20 and Apollo 40 systems.

Distribution Media and Software Documentation

HPE Message Passing Interface (MPI) software and documentation is available on a single DVD. Please select HPE Performance Message Passing Interface Media Kit (Part Number Q5U69A) to order the DVD.

HPE Message Passing Interface (MPI) software is also available for download. Upon receiving their order, customers will receive a physical document, which includes the website URL and additional license information. Software will be available for download via the Software Updates and Licensing Portal.

Patches may be required and will be delivered via the Software Download Repository. Release notes will be made available for any patches issued to customers.

Factory installation is available for the HPE SGI 8600, Apollo 20 and Apollo 40 systems. Additional platforms are not supported at this time. Please select PN Q8K21A for factory installation.

Customers may also download user guides and other documentation at www.hpe.com/software/mpi

Standards

The Message Passing Interface (MPI) standard supports C and Fortran programs with a library and supporting commands. MPI operates through a technique known as “message passing,” which is the use of library calls to request data delivery from one process to another, or between groups of processes. MPI also supports parallel file I/O and remote memory access (RMA).

HPE Message Passing Interface (MPI) supports the MPI 3.1 standard.

In addition, HPE Message Passing Interface (MPI) supports the OpenSHMEM 1.4 standard. The OpenSHMEM standard describes a low-latency library that supports RMA on symmetric memory in parallel environments. The OpenSHMEM programming model is a partitioned global address space (PGAS) programming model that presents distributed processes with symmetric arrays that are accessible via PUT and GET operations.

Support for the MPI and OpenSHMEM standards is built on top of a message passing toolkit, which is available within the HPE Message Passing Interface (MPI).

Standard Features

High Performance MPI Environment

The message passing toolkit is a high-performance communications middleware software that is a component of the HPE Message Passing Interface (MPI). This toolkit is the core of the HPE Message Passing Interface (MPI) performance engine as it provides the flexibility users need to develop and run MPI programs.

HPE Message Passing Interface (MPI) includes a tool that uses a wrapper library to run applications compiled against other MPI implementations through the message passing toolkit, without requiring the user to relink or recompile their code. This allows users with applications compiled against Intel MPI, IBM Spectrum MPI, OpenMPI, Cray MPI and Mellanox HPC-X MPI to still run their applications through the HPE Message Passing Interface (MPI) engine and increase overall application performance.

The message passing toolkit increases the speed of an application from launch. The environment includes HPE-proprietary software that enables parallel applications to run on multiple hosts in a cluster, resulting in the ability to launch 100,000 cores in a matter of seconds.

HPE Message Passing Interface (MPI) also supports Checkpoint/Restart. The Checkpoint/Restart implementation allows applications to periodically save a copy of their state. Applications can resume from the set checkpoint if the application crashes or if the job is aborted to free resources for higher-priority jobs. This functionality saves the user time by eliminating the need to restart applications from the beginning.

Application Performance Optimization

HPE Message Passing Interface (MPI) accelerates application performance through tools designed to tune applications at runtime without recompiling code, as well as libraries which optimize performance through specialized algorithms.

These tools and libraries allow customers to:

- Restrict the number of processor and memory resources used for a specific process, or set of processes to avoid oversubscribing the system and possible interference between applications.
- Improve performance for I/O intensive applications without recompiling or retooling the software logic.
- Enable data placement to specific memory locations to limit communications overhead.

Users also gain access to profiling and analysis tools with HPE Message Passing Interface (MPI). One such tool captures communication analytics, pinpoints bottlenecks and identifies load imbalances during the application run. Once the tool is implemented and the application completes its run, users will receive a file outlining key statistics from the run.

This information includes:

- Time spent in computation, communication and file I/O
- The size and number of each data request
- The size of the communicator used for collectives
- The number of times each rank was a root in a collective
- Size distributions
- Collective wait time and send late time

Users also have the option to analyze subsets of the application run by inserting start and stop commands into their program. The tool will begin analyzing the application run when the start command is executed and will generate the statistics file after the stop command is executed.

The profiling and analysis tools available within HPE Message Passing Interface (MPI) can be used for applications compiled against the Intel MPI, OpenMPI, Cray MPI and Mellanox HPC-X MPI.

Additional features within the HPE Message Passing Interface (MPI) environment include a runtime tool to optimize processes and data placement on a set of nodes.

HPE Message Passing Interface (MPI) includes tools to optimize the placement of worker processes within the cluster. Users first run their application with HPE MPI's profiling tool to record the communication pattern between the workers. In future runs, HPE MPI's placement optimization tool takes this pattern information and combines it with information about where the application will be running to create a placement file assigning workers to specific locations in the cluster. This improves performance by putting workers that intercommunicate a lot close to each other, minimizing transfer costs. This feature is not supported on Apollo 70 at this time.

Operating System Compatibility

The HPE Message Passing Interface (MPI) is operational on Red Hat Enterprise Linux (RHEL) 7, SUSE Linux Enterprise Server (SLES) 12, and CentOS 7.

Standard Features

Third Party Software Integration

The HPE Message Passing Interface (MPI) (MPI) allows for tight integration with third-party debugging and profiling tools, as well as workload schedulers and fabric software. Such applications include:

- **Intel Parallel Studio XE***
- PGI Professional Edition
- **Allinea Forge***
- **Rogue Wave TotalView***
- Tuning and Analysis Utilities (TAU)
- Vampir
- **Altair PBS Professional***
- SLURM
- Univa Grid Engine
- IBM Spectrum LSF
- **Adaptive Computing Moab/ TORQUE***
- Mellanox Fabric Collective Accelerator

Hardware Compatibility

The HPE Message Passing Interface (MPI) (MPI) is supported on the HPE SGI 8600, Apollo 2000, Apollo 6000, Apollo 6500, Apollo 40 and Apollo 70 systems and HPE ProLiant DL360 and DL380 systems. For more information, please contact your Hewlett Packard Enterprise sales representative.

The HPE Message Passing Interface includes fabric support and corresponding features that make it the ideal MPI library for HPE high performance computing systems. Fabric support includes:

- Data transfer optimization for HPE-proprietary HPE Superdome Flex Grid where available, including single-copy data transfer.
- Multi-rail Intel Omni-Path Architecture support.
- Multi-rail Mellanox InfiniBand support.
- TCP/IP

Software Licensing Information

For the Software to be valid on an HPE Cluster, each socket in the HPE Cluster must have a valid HPE Message Passing Interface (MPI) license. Software licensing information can be found at <https://www.hpe.com/us/en/software/licensing.html>. Subject to the terms and conditions of this Agreement and the payment of any applicable license fee, HPE grants a non-exclusive, non-transferable license to use (as defined below), in object code form, one copy of the Software on one device (socket) at a time for internal business purposes, unless otherwise indicated above or in applicable Transaction Document(s). "Use" means to install, store, load, execute and display the Software in accordance with the Specifications. Use of the Software is subject to these license terms and to the other restrictions specified by Hewlett Packard Enterprise in any other tangible or electronic documentation delivered or otherwise made available with or at the time of purchase of the Software, including license terms, warranty statements, Specifications, and "readme" or other informational files included in the Software itself. Such restrictions are hereby incorporated in this Agreement by reference. Some Software may require license keys or contain other technical protection measures. HPE reserves the right to monitor compliance with Use restrictions remotely or otherwise. Hewlett Packard Enterprise may make a license management program available which records and reports license usage information. If so supplied, customer agrees to install and run such license management program beginning no later than one hundred and eighty (180) days from the date it is made available and continuing for the period that the Software is used.

Other terms of the HPE Software License are provided on the license agreement that is delivered with the HPE Message Passing Interface (MPI) software.

Service and Support

Warranty

Hewlett Packard Enterprise will replace defective delivery media for a period of 90 days from the date of purchase. This warranty applies to all HPE Message Passing Interface (MPI) products found on the delivery media.

HPE Software Support

- **HPE Pointnext** leverages our breadth and depth of technical expertise and innovation to help accelerate digital transformation with Advisory, Professional, and Operational Services. There is a full range of services to complement HPE Message Passing Interface (MPI) (MPI) software from advisory and design, benchmarking and tuning services, factory pre-installation, configuration, and acceptance as well as training and operational services.
 - Advisory Services includes design, strategy, road map, and other services to help enable the digital transformation journey, tuned to IT and business needs. Advisory Services helps customers on their journey to Hybrid IT, Big Data, and the Intelligent Edge.
 - Professional Services helps integrate the new solution with project management, installation and startup, relocation services, and more. In addition, Factory Express installs the software in the factory when building the system. HPE Education Services helps train staff using and managing the software and other technology. We help mitigate risk to the business, so there is no interruption when new technology is being integrated into the existing IT environment.
-

Operational Services

- HPE Flexible Capacity is a new consumption model to manage on-demand capacity, combining the agility and economics of public cloud with the security and performance of on-premises IT.
- HPE Datacenter Care offers a tailored operational support solution built on core deliverables. It includes hardware and software support, a team of experts to help personalize deliverables and share best practices, as well as optional building blocks to address specific IT and business needs. HPE Datacenter Care for Hyperscale gives customers access to the Hyperscale Center of Excellence with technical experts who understand how to manage IT at scale including the software.
- HPE Proactive Care is an integrated set of hardware and software support including an enhanced call experience with start to finish case management helping resolve incidents quickly and keeping IT reliable and stable.
- HPE Foundation Care helps when there is a hardware or software problem offering several response levels dependent on IT and business requirements.

Join the Conversation

The HPE Support Center is a community-based, user-supported tool for Hewlett Packard Enterprise customers to participate in discussions amongst the customer community about Hewlett Packard Enterprise products.

Contact Support

<http://www.hpe.com/services>

Service and Support

HPE Pointnext Services

Flex Support Services for 4 Year Support Plan

HPE 4Y Proactive Care NBD SVC	H1K90A4
HPE 4Y Proactive Care NBD wDMR SVC	H1K91A4
HPE 4Y Proactive Care ADV NBD SVC	H8B33A4
HPE 4Y Proactive Care ADV NBD wDMR SVC	H8B34A4
HPE 4Y Foundation Care 24x7 SVC	H7J34A4
HPE 4Y Foundation Care 24x7 wDMR SVC	H7J35A4
HPE 4Y Proactive Care 24x7 SVC	H1K92A4
HPE 4Y Proactive Care 24x7 wDMR SVC	H1K93A4
HPE 4Y Proactive Care ADV 24x7 SVC	H8B35A4
HPE 4Y Proactive Care ADV 24x7 wDMR SVC	H8B36A4
HPE 4Y Foundation Care CTR SVC	H7J36A4
HPE 4Y Foundation Care CTR wDMR SVC	H7J37A4
HPE 4Y Proactive Care CTR SVC	H1K94A4
HPE 4Y Proactive Care CTR wDMR SVC	H1K95A4
HPE 4Y Proactive Care ADV CTR SVC	H8B37A4
HPE 4Y Proactive Care ADV CTR wDMR SVC	H8B38A4

Flex Support Services for 5 Year Support Plan

HPE 5Y Foundation Care NBD SVC	H7J32A5
HPE 5Y Foundation Care NBD wDMR SVC	H7J33A5
HPE 5Y Proactive Care NBD SVC	H1K90A5
HPE 5Y Proactive Care NBD ADV SVC	H8B33A5
HPE 5Y Proactive Care ADV NBD wDMR SVC	H8B34A5
HPE 5Y Foundation Care 24x7 SVC	H7J34A5
HPE 5Y Foundation Care 24x7 wDMR SVC	H7J35A5
HPE 5Y Proactive Care 24x7 SVC	H1K92A5
HPE 5Y Proactive Care 24x7 wDMR SVC	H1K93A5
HPE 5Y Proactive Care ADV 24x7 SVC	H8B35A5
HPE 5Y Proactive Care ADV 24x7 wDMR SVC	H8B36A5
HPE 5Y Foundation Care CTR SVC	H7J36A5
HPE 5Y Foundation Care CTR wDMR SVC	H7J37A5
HPE 5Y Proactive Care CTR SVC	H1K94A5
HPE 5Y Proactive Care CTR wDMR SVC	H1K95A5
HPE 5Y Proactive Care ADV CTR SVC	H8B37A5
HPE 5Y Proactive Care ADV CTR wDMR SVC	H8B38A5

Summary of Changes

Date	Version History	Action	Description of Change
24-Oct-2018	Version 5	Changed	Overview section was updated and PB ID was corrected in Summary of Changes section. Link was corrected in Standard Features section.
4-Jun-2018	Version 4	Changed	Updated the following sections: Overview, What's new, hardware compatibility
4-Dec-2017	From version 2 to 3	Updated	Updated the following sections: Overview, What's New, Models and Licenses, Hardware Compatibility
25-Sept-2017	From version 1 to 2	Updated	Updated the Models and Licenses section
7-Aug-2017	Version 1	Created	Create QuickSpecs for HPE Performance Software – Message Passing Interface



Sign up for updates



**Hewlett Packard
Enterprise**

© Copyright 2018 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft and Windows NT are US registered trademarks of Microsoft Corporation. Intel, the Intel logo, Xeon and Xeon Inside are trademarks of Intel Corporation in the U.S. and other countries. For hard drives, 1GB = 1 billion bytes. Actual formatted capacity is less.

a00018446enw - 15997 - Worldwide - V5 - 24-October-2018