

Open MPI Collective Operation Performance on Thunderbird

Jelena Pješivac–Grbović

Innovative Computing Laboratory
The University of Tennessee, Knoxville
pjesa@cs.utk.edu

April 13, 2007

1 Introduction

The Open MPI project was allocated time on Thunderbird system at Sandia National Laboratory to perform the large-scale tests of our library. Tests took place from March 22nd through 23rd. The goal of the testing was threefold: analyze scalability of ORTE (Open MPI Run Time Environment), measure performance of collective operations, and analyze Infiniband resource utilization.

The collected data was instrumental in improving the large-scale stability and performance of production releases of Open MPI library. Experience in using and debugging Open MPI at scale is necessary to ensure the stability and scalability of the production quality library.

This document provides the summary and analysis of these performance results for different barrier, broadcast, reduce, and allreduce algorithms available in Open MPI. Jeff Squyres collected the performance of results using SKaMPI-5.0.1[1] benchmark. In addition, results of NetPIPE[2] and LogP[3] benchmarks were obtained. Most of the algorithm performance data was collected on up to 1024 processes, although some of the measurements went up to 4096 processes.

The results reported here are the first large-scale performance tests of tuned collective module in Open MPI. The measurements provided us with invaluable information about the scalability of different algorithms. At the same time, it allowed us to obtain the first sets of necessary data that can enable us to utilize more modeling techniques for future performance tuning of collectives on large-scale systems.

Section 2 covers the point-to-point tests results. Section 3 shows performance results for the mentioned collectives. Section 4 discusses problems we encountered during the measurements. Section 5 provides the list of proposed measurements for the future testing. Section 6 wraps up the report. Appendix A provides details about hardware and software used during the tests.

2 Point-to-point performance

We measured point-to-point communication performance using two standardized benchmarks: NetPIPE and LogP benchmark.

2.1 NetPIPE

The NetPIPE[2] benchmark is used to measure the latency and throughput of point-to-point communication. The results of this benchmark can be directly used as parameters for Hockney parallel communication model [4]. Figure 1 shows the latency and throughput measurements on this system.

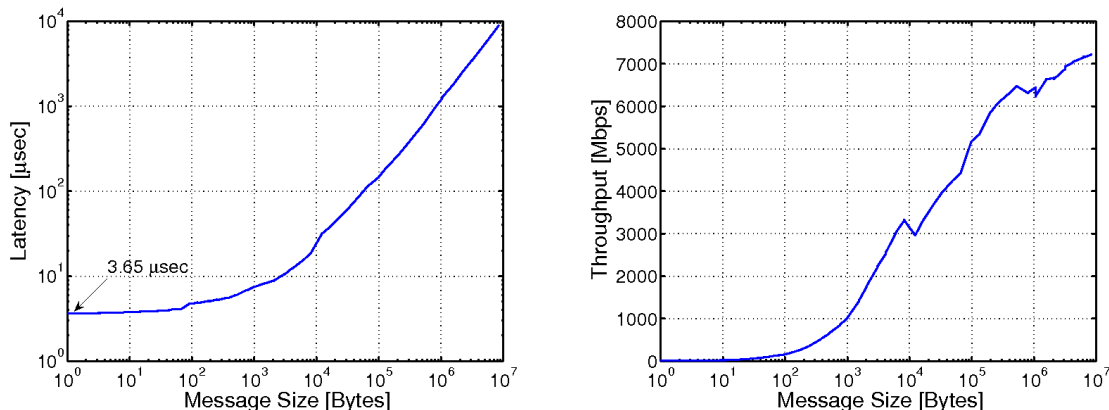


Figure 1: NetPIPE benchmark results.

2.2 LogP benchmark

The LogP[3] benchmark is used to directly measure parameters of PLogP parallel communication model[5]. The PLogP model can be used to predict parallel algorithm performance directly, or the values of its parameters can be used to estimate the values of LogP and LogGP model parameters [6, 7].

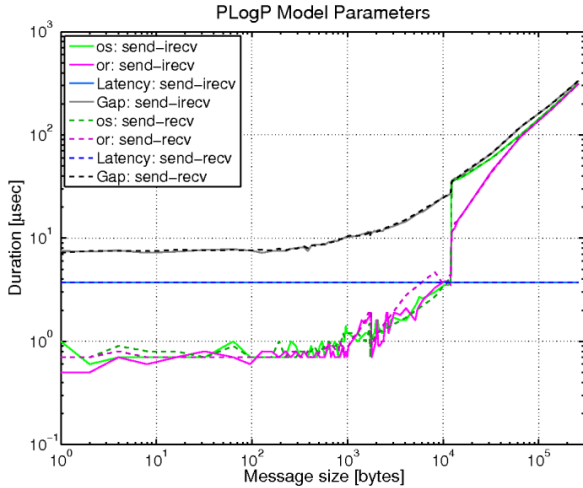
Figure 2 shows the measured values of PLogP model parameters and values of LogP and LogGP model computed from these values. In the original measurement, the latency value was not reported because benchmark output format rounded approximately 3×10^{-6} value to 0.0. Thus we used the NetPIPE results to estimate latency value.

3 Collective operation performance

The collective operation algorithm performance was collected using SKaMPI 5.0.1 benchmark[1]. SKaMPI benchmark measurements are governed by the SKaMPI input scripts.

With no previous information about the algorithm performance on large-scale systems, we decided to limit sets of measurements to the ones we believed would perform the best. In addition, we limited the maximum number of repetitions per data point to 35. We hoped this restriction will not cause large variability in measured data. These two restrictions allowed us to keep measurement time under control.

In addition, we had to modify SKaMPI benchmark to limit the verbosity of its output. By default, SKaMPI reports “measured time”, “standard deviation of the measurement”, “number of measurements”, and individual times measured on each of the processes involved in the



LogP/LogGP model parameters

Latency	3.65	μsec
Overhead	0.65	μsec
Gap	6.7	μsec
Gap per byte	0.0013	$\frac{\mu sec}{Byte}$

Figure 2: PLogP and LogP/LogGP model parameters.

measurement. This amounts to rather large files even on 256 process test. Thus, we modified SKaMPI not to report individual timers. To activate this feature, SKaMPI must be compiled with `-DPRODUCE_SPARSE_OUTPUT` flag.

Most of the tests were split in 9 different groups based on communicator and message size:

- communicator sizes
 - small: 64, 128
 - intermediate: 256, 512, 768, 1024
 - large: 1536, 2048, 2560, 3072, 3684, 4096, 4400
- message sizes
 - small: 8B, 128B, 1KB, 8KB
 - intermediate: 64KB, 256KB
 - large: 1MB, 8MB

We selected a separate set of methods (algorithm and segment size combinations) for each of the communicator \times message size combinations. The following subsections detail the measurement results for each of the collectives we considered.

3.1 Barrier

The performance of the following barrier algorithms was measured:

- Linear algorithm - up to 1024 processes
- Recursive doubling - up to 2048 processes
- Bruck algorithm - up to 2048 processes

The default Open MPI decision function for barrier selects the bruck algorithm for non-power-of-two communicator sizes and the recursive doubling for the power-of-two communicator sizes.

Figure 3 shows the absolute algorithm performance with error bars denoting the standard deviation of the measurement. As expected, the linear algorithm performance is drastically worse

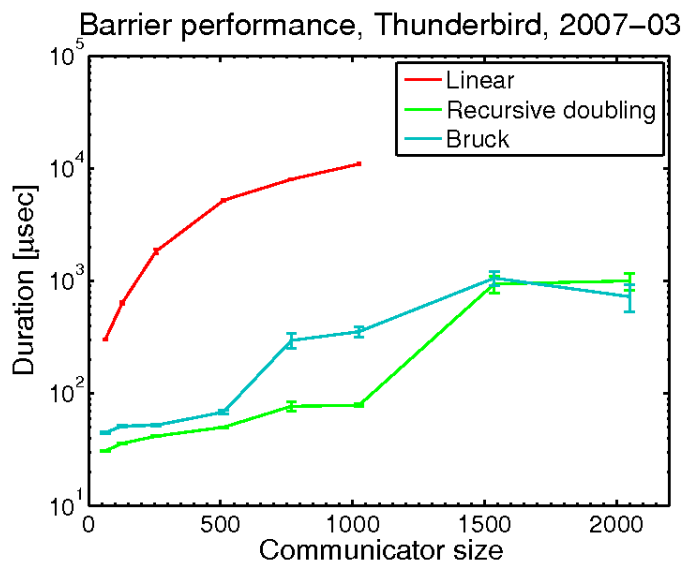


Figure 3: Performance of different barrier algorithms with error bar denoting standard deviation of measurement.

than the performance of bruck and recursive doubling algorithms. However, the performance of the recursive doubling algorithm is surprisingly better than the performance of the bruck algorithm for intermediate communicator sizes (350% for 1024 communicator size). For large communicator sizes (1536 and 2048), the performance of two algorithms is within the standard deviation of the measurement. Even so, the bruck algorithm seems to outperform recursive doubling on 2048 processes. Based on the results, it is hard to determine which algorithm would be better option. Additional (repeated) measurements are necessary to clarify the obtained results.

3.2 Broadcast

We collected the performance of the following broadcast methods:

- Linear without segmentation - up to 1024 processes and for small and intermediate message sizes.
- Binomial without segmentation - up to 1024 processes for small and intermediate message sizes.
- Binomial with 64KB and 256KB segments up to 1024 processes for large message sizes.
- Binary without segmentation - up to 4096 processes for small message sizes, and up to 1024 processes for intermediate message sizes.

- Binary with 64KB and 256KB segments up to 1024 processes for large message sizes.
- Splitbinary without segmentation - up to 4096 processes for small message sizes and up to 1024 processes for intermediate message sizes.
- Splitbinary with 64KB and 256KB segments - up to 4096 processes for small message sizes.
- Pipeline with 16KB, 64KB, and 256KB segments up to 1024 processes with large message sizes.

The performance data for large communicator sizes and some intermediate messages and all large messages is missing for some of the algorithms.

Figure 4 shows the measured “decision map” for broadcast. The color denotes the method (algorithm and segment size) to be used for particular communicator \times message size combination of input parameters. We can see that the Open MPI default and measured decision function differ quite a bit. Part of the problem is the fact that for large communicator sizes (2048 and up) we were only able to collect data for binary and split-binary algorithms.

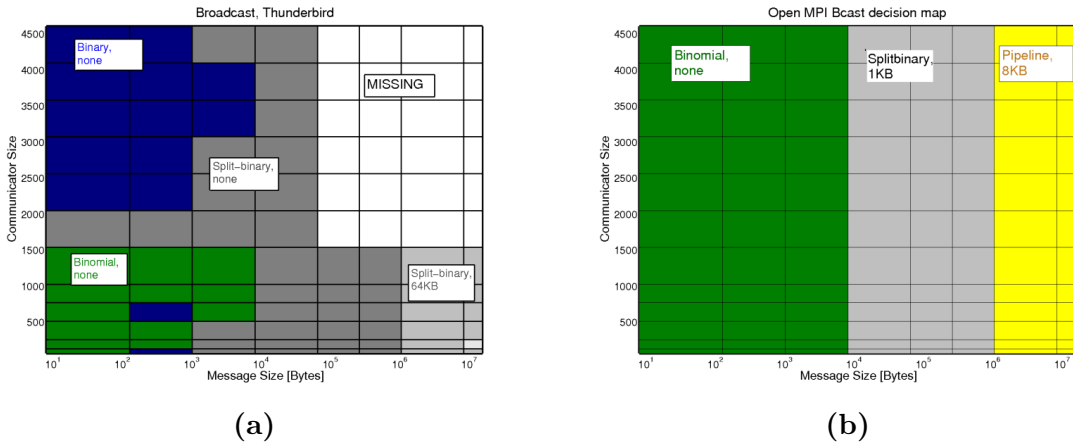


Figure 4: Broadcast decision maps for Thunderbird system: (a) measured, (b) Open MPI default.

In addition, the results on intermediate communicator sizes indicate that we overestimated good segment sizes for pipeline algorithm (See Figure 5). The segmented pipeline algorithm has unique property that for large enough message sizes, it asymptotically achieves constant running time. We can see this phenomenon occurring on 768 and 1024 processes, however, the level of the line is too high. Smaller segment size (such as 8KB) would possibly achieve better performance and become method of choice instead of split-binary algorithm for at least some cases.

The quality of collected data for large communicator sizes is important. Unfortunately, as we can see in Figure 6, some of the measurements had rather large standard deviation of the measurement (as high as 40+%). On the positive side, the measurements with high level of noise happened in the regions where other methods seem to outperform the problematic records.

In order to fully understand the performance of broadcast on this system we would need to repeat some of the collected measurements and collect the missing data. In addition, we would need to consider additional segment sizes: such as 8KB, as the ones we considered seem to be too large.

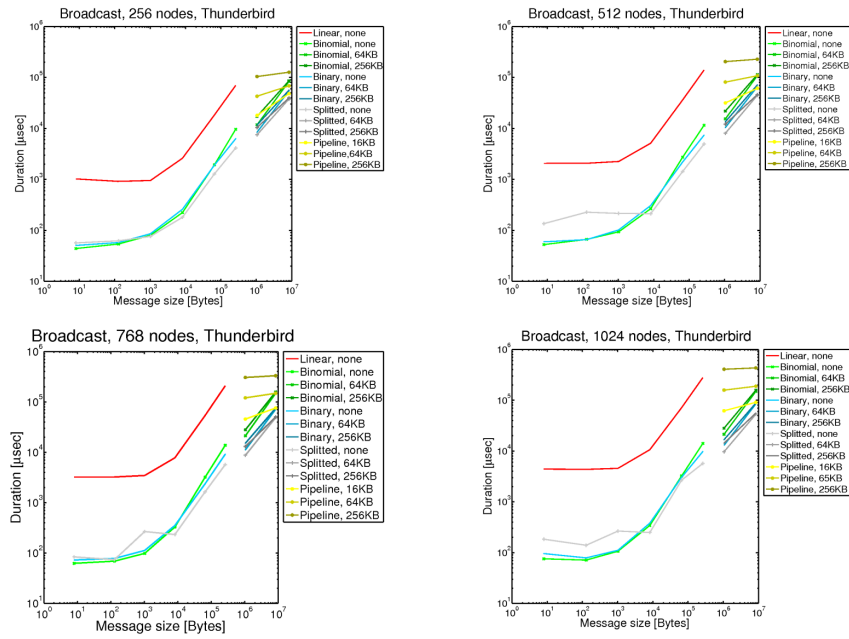


Figure 5: Broadcast performance on intermediate communicator sizes: 256, 512, 768, and 1024.

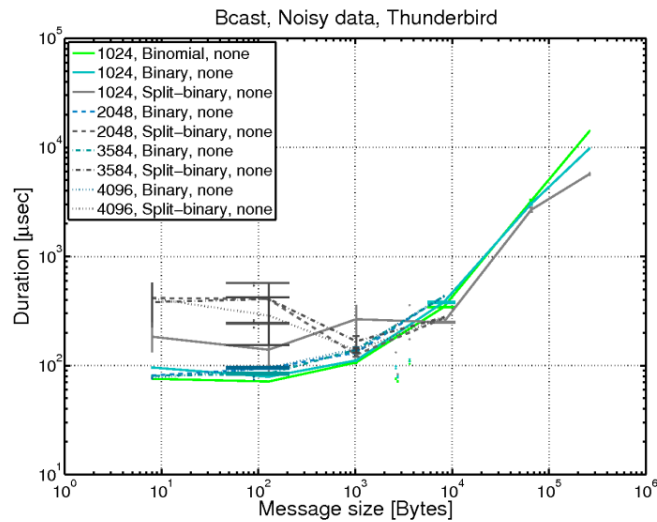


Figure 6: Performance of different broadcast algorithms with error bar denoting standard deviation of measurement.

3.3 Reduce

The reduce performance was measured for small and intermediate communicator sizes (up to and including 1024 processes). We collected performance information for the following reduce methods:

- Linear without segmentation - for small and intermediate message sizes.
- Binary without segmentation - for small and intermediate message sizes.
- Binomial without segmentation - for small and intermediate message sizes.
- Binary with 32KB, 128KB, and 512KB segments - for large message sizes.
- Binomial with 32KB, 128KB, and 512KB segments - for large message sizes.
- Pipeline with 32KB, 128KB, and 512KB segments - for large message sizes.

Figure 7 shows the measured “decision map” for reduce. The color denotes the method (algorithm and segment size) to be used for particular communicator \times message size combination of input parameters. We can see that the Open MPI default and measured decision function differ – but far less than the broadcast decisions (Figure 4). Also, in the regions where the differences occur we did not have measurements for the methods selected by the default function (Binary with 1KB segments, and Pipeline with 1KB segments).

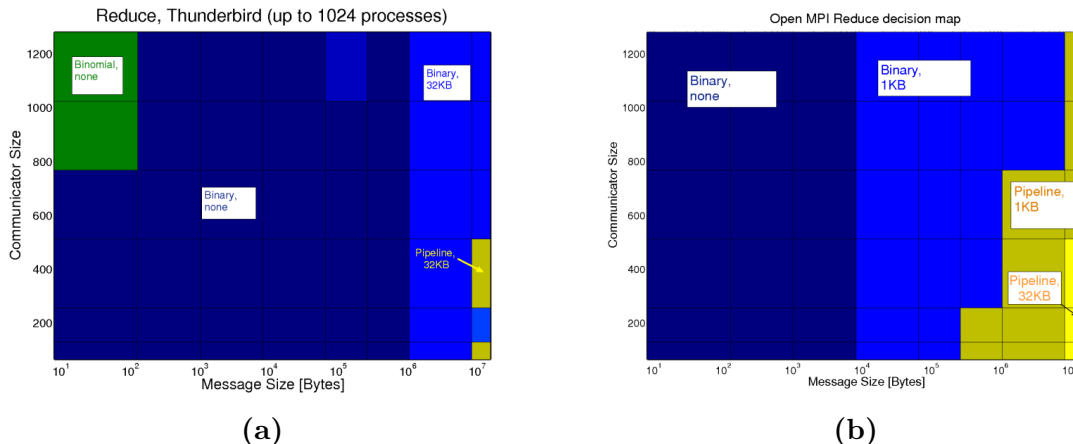


Figure 7: Reduce decision maps for Thunderbird system: (a) measured, (b) Open MPI default.

Similarly to the broadcast case, the absolute performance data indicates that we overestimated the segment sizes for segmented algorithms (See Figure 8). The reduce algorithm measurements were less noisy than the broadcast ones, with only 6 data points out of 218 with standard deviation higher than 10%.

3.4 Allreduce

Due to time constraints and some of the problems discussed in the section 4 we were only able to benchmark performance of recursive doubling algorithm for allreduce.

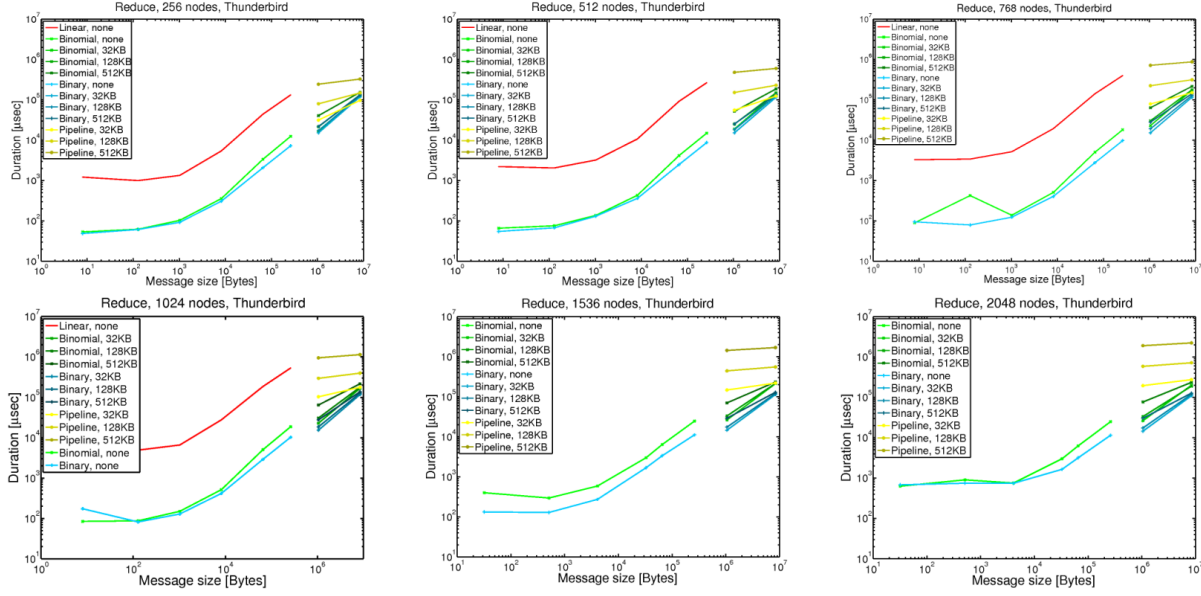


Figure 8: Absolute performance of reduce methods we considered.

The measurements covered all message sizes for small and intermediate communicator sizes, and only large message sizes for large communicator sizes. Figure 9 shows the absolute performance of tests we measured.

4 Problems which occurred during the testing

The following is the list of problems which occurred during the test:

- Segmented algorithms for broadcast and reduce (binary, binomial, split-binary, and pipeline) would have problems for segment size 8KB. We were not able to determine what caused this problem.
- Non-overlapping algorithm for allreduce failed to run. Due to problems which occurred with trunk version of ORTE at scale, the tests were switched to using Open MPI branch v1.2. In order to test the most recent tuned collective module, the whole module was copied to branch v1.2 source tree. In the process, communicator data structure was affected, such that all internal pointers to the default collective functions were shifted by one. Luckily, this was the only test case which was affected by this problem.
- Ring algorithm for allreduce deadlocked. In the time frame we had, we could not identify the problem was.

5 Future Testing

The following is the list of tests that would enable us to understand Open MPI performance on Thunderbird better:

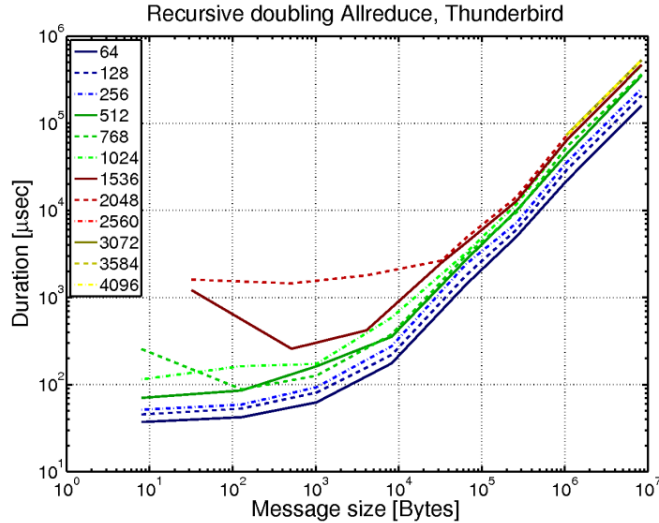


Figure 9: Performance of recursive doubling allreduce algorithm.

- Collect barrier performance results up to full system size. Increase the number of measurements from 35 to at least 50 to ensure higher quality results.¹
- Collect missing broadcast data and run tests with smaller segment size (such as 8KB). Repeat the low quality measurements (split-binary without segmentation for small message sizes). Cover all algorithms selected by the default decision function.
- Collect missing reduce data. There is most likely no need to repeat any of the already collected reduce tests. Cover all algorithms selected by the default decision function.
- Collect missing allreduce performance data.
- Measure performance of different alltoall algorithms.
- Measure performance of different allgather algorithms.

6 Discussion

The performance measurements on Thunderbird provided us with invaluable information about the scalability of different collective algorithms. Due to time constraints, we were not able to complete all the measurements of interest and some of the collected data suffers from high noise. However, the data we have provides us with an excellent insight into what to expect on large-scale system. Moreover, having this information allows us to plan future tests with higher accuracy and detail.

The large-scale tests are necessary for continual improvement of scalability and performance of Open MPI library. They are also the way to ensure that our large-scale users will have access to the best possible production quality MPI library on their systems.

¹SKaMPI benchmark actually executes only the necessary number of measurements to keep the standard deviation of the measurement within predetermined limit (say 5%).

References

- [1] SKaMPI - Special Karlsruher MPI Benchmark, <http://liinwww.ira.uka.de/~skampi/>
- [2] NetPipe - A Network Protocol Independent Performance Evaluator
<http://www.scl.ameslab.gov/netpipe/>
- [3] The MPI LogP Benchmark ftp://ftp.cs.vu.nl/pub/kielmann/logp_mpi.tar.gz
- [4] R.W. Hockney, “The Communication Challenge for MPP: Intel Paragon and Meiko CS-2”, *Parallel Computing*, March, 1994.
- [5] Thilo Kielmann, Henri E. Bal, Kees Verstoep, “Fast Measurement of LogP Parameters for Message Passing Platforms”, *IPDPS '00: Proceedings of the 15 IPDPS 2000 Workshops on Parallel and Distributed Processing*, 2000.
- [6] David Culler, Richard Karp, David Patterson, Abhijit Sahay, Klaus Erik Schauser, Eunice Santos, Ramesh Subramonian, Thorsten von Eicken, “LogP: Towards a realistic model of parallel computation”, *Proceedings of the fourth ACM SIGPLAN symposium on Principles and practice of parallel programming*, 1993.
- [7] Albert Alexandrov, Mihai F. Ionescu, Klaus E. Schauser, Chris Scheiman, “LogGP: Incorporating long messages into the LogP model”, *Proceedings of the seventh annual ACM symposium on Parallel algorithms and architectures*, 1995.

A Thunderbird system specification

This section provides details about the hardware and software used during the Thunderbird tests.

A.1 Hardware

CPU info # cat /proc/cpuinfo

```
processor : 0
vendor_id : GenuineIntel
cpu family : 15
model : 4
model name : Intel(R) Xeon(TM) CPU 3.60GHz
stepping : 3
cpu MHz : 3591.371
cache size : 2048 KB
fpu : yes
fpu_exception : yes
cpuid level : 5
wp : yes
flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht
bogomips : 7187.16
clflush size : 64
cache_alignment : 128
address sizes : 36 bits physical, 48 bits virtual
power management:
```

```
processor : 1
vendor_id : GenuineIntel
cpu family : 15
model : 4
model name : Intel(R) Xeon(TM) CPU 3.60GHz
stepping : 3
cpu MHz : 3591.371
cache size : 2048 KB
fpu : yes
fpu_exception : yes
cpuid level : 5
wp : yes
flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht
bogomips : 7182.53
clflush size : 64
cache_alignment : 128
address sizes : 36 bits physical, 48 bits virtual
power management:
```

Memory # cat /proc/meminfo

```
MemTotal: 6114080 kB
MemFree: 5460424 kB
Buffers: 7212 kB
Cached: 38732 kB
SwapCached: 149148 kB
Active: 440056 kB
Inactive: 58628 kB
HighTotal: 0 kB
HighFree: 0 kB
LowTotal: 6114080 kB
LowFree: 5460424 kB
SwapTotal: 13309844 kB
SwapFree: 13154072 kB
Dirty: 32 kB
```

Writeback: 0 kB
Mapped: 379716 kB
Slab: 93484 kB
CommitLimit: 16366884 kB
Committed_AS: 858520 kB
PageTables: 2760 kB
VmallocTotal: 536870911 kB
VmallocUsed: 286680 kB
VmallocChunk: 536582727 kB
HugePages_Total: 0
HugePages_Free: 0
Hugepagesize: 2048 kB

IBV devinfo # ibv_devinfo

hca_id: mthca0
fw_ver: 5.1.925
node_guid: 0005:ad00:0004:ce4c
sys_image_guid: 0005:ad00:0100:d050
vendor_id: 0x05ad
vendor_part_id: 25218
hw_ver: 0xA0
board_id: HCA.LionCub.A0.MemZero
phys_port_cnt: 2
port: 1
state: PORT_ACTIVE (4)
max_mtu: 2048 (4)
active_mtu: 2048 (4)
sm_lid: 977
port_lid: 1104
port_lmc: 0x00

port: 2
state: PORT_DOWN (1)
max_mtu: 2048 (4)
active_mtu: 512 (2)
sm_lid: 0
port_lid: 0
port_lmc: 0x00

PCI devices # lspci

00:00.0 Host bridge: Intel Corporation E7520 Memory Controller Hub (rev 09)
00:02.0 PCI bridge: Intel Corporation E7525/E7520/E7320 PCI Express Port A (rev 09)
00:04.0 PCI bridge: Intel Corporation E7525/E7520 PCI Express Port B (rev 09)
00:05.0 PCI bridge: Intel Corporation E7520 PCI Express Port B1 (rev 09)
00:06.0 PCI bridge: Intel Corporation E7520 PCI Express Port C (rev 09)
00:1d.0 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #1 (rev 02)
00:1d.1 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #2 (rev 02)
00:1d.2 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #3 (rev 02)
00:1d.7 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB2 EHCI Controller (rev 02)
00:1e.0 PCI bridge: Intel Corporation 82801 PCI Bridge (rev c2)
00:1f.0 ISA bridge: Intel Corporation 82801EB/ER (ICH5/ICH5R) LPC Interface Bridge (rev 02)
01:00.0 PCI bridge: Intel Corporation 80332 [Dobson] I/O processor (A-Segment Bridge) (rev 06)
01:00.2 PCI bridge: Intel Corporation 80332 [Dobson] I/O processor (B-Segment Bridge) (rev 06)
02:05.0 SCSI storage controller: LSI Logic / Symbios Logic 53c1030 PCI-X Fusion-MPT Dual Ultra320 SCSI (rev 08)
05:00.0 PCI bridge: Intel Corporation 6700PXH PCI Express-to-PCI Bridge A (rev 09)
05:00.2 PCI bridge: Intel Corporation 6700PXH PCI Express-to-PCI Bridge B (rev 09)
06:07.0 Ethernet controller: Intel Corporation 82541GI/PI Gigabit Ethernet Controller (rev 05)
07:08.0 Ethernet controller: Intel Corporation 82541GI/PI Gigabit Ethernet Controller (rev 05)
08:00.0 InfiniBand: Mellanox Technologies MT25208 InfiniHost III Ex (rev a0)
09:0d.0 VGA compatible controller: ATI Technologies Inc Radeon RV100 QY [Radeon 7000/VE]

lspci -t

```

-[0000:00]--00.0
  +-02.0-[0000:01-03]---+00.0-[0000:02]----05.0
  |           \-00.2-[0000:03]--
  +-04.0-[0000:04]--
  +-05.0-[0000:05-07]---+00.0-[0000:06]----07.0
  |           \-00.2-[0000:07]----08.0
  +-06.0-[0000:08]----00.0
  +-1d.0
  +-1d.1
  +-1d.2
  +-1d.7
  +-1e.0-[0000:09]----0d.0
  \-1f.0

```

lspci -tvvv

```

-[0000:00]--00.0 Intel Corporation E7520 Memory Controller Hub
  +-02.0-[0000:01-03]---+00.0-[0000:02]----05.0 LSI Logic / Symbios Logic 53c1030 PCI-X Fusion-MPT Dual Ultra32
  |           \-00.2-[0000:03]--
  +-04.0-[0000:04]--
  +-05.0-[0000:05-07]---+00.0-[0000:06]----07.0 Intel Corporation 82541GI/PI Gigabit Ethernet Controller
  |           \-00.2-[0000:07]----08.0 Intel Corporation 82541GI/PI Gigabit Ethernet Controller
  +-06.0-[0000:08]----00.0 Mellanox Technologies MT25208 InfiniHost III Ex
  +-1d.0 Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #1
  +-1d.1 Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #2
  +-1d.2 Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #3
  +-1d.7 Intel Corporation 82801EB/ER (ICH5/ICH5R) USB2 EHCI Controller
  +-1e.0-[0000:09]----0d.0 ATI Technologies Inc Radeon RV100 QY [Radeon 7000/VE]
  \-1f.0 Intel Corporation 82801EB/ER (ICH5/ICH5R) LPC Interface Bridge

```

lspci -vvv

```

00:00.0 Host bridge: Intel Corporation E7520 Memory Controller Hub (rev 09)
Subsystem: Dell: Unknown device 016c
Control: I/O- Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR+ FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0
Capabilities: <available only to root>

```

```

00:02.0 PCI bridge: Intel Corporation E7525/E7520/E7320 PCI Express Port A (rev 09) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=00, secondary=01, subordinate=03, sec-latency=0
I/O behind bridge: 0000e000-0000efff
Memory behind bridge: fe800000-feafffff
Prefetchable memory behind bridge: 00000000f8000000-00000000f8f00000
Secondary status: 66Mhz- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort+ <SERR+ <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

```

```

00:04.0 PCI bridge: Intel Corporation E7525/E7520 PCI Express Port B (rev 09) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=00, secondary=04, subordinate=04, sec-latency=0
I/O behind bridge: 0000f000-0000ffff
Memory behind bridge: fff00000-000fffff
Prefetchable memory behind bridge: 00000000fff00000-0000000000000000
Secondary status: 66Mhz- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort+ <SERR- <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

```

```

00:05.0 PCI bridge: Intel Corporation E7520 PCI Express Port B1 (rev 09) (prog-if 00 [Normal decode])

```

Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=00, secondary=05, subordinate=07, sec-latency=0
I/O behind bridge: 0000c000-0000dfff
Memory behind bridge: fe300000-fe7fffff
Prefetchable memory behind bridge: 00000000fff00000-0000000000000000
Secondary status: 66Mhz- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort+ <SERR+ <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

00:06.0 PCI bridge: Intel Corporation E7520 PCI Express Port C (rev 09) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=00, secondary=08, subordinate=08, sec-latency=0
I/O behind bridge: 0000b000-0000bfff
Memory behind bridge: fe100000-fe2fffff
Prefetchable memory behind bridge: 00000000f9000000-00000000f9700000
Secondary status: 66Mhz- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort+ <SERR- <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

00:1d.0 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #1 (rev 02) (prog-if 00 [UHCI])
Subsystem: Dell: Unknown device 016c
Control: I/O+ Mem- BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B-
Status: Cap- 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0
Interrupt: pin A routed to IRQ 169
Region 4: I/O ports at 9ce0 [size=32]

00:1d.1 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #2 (rev 02) (prog-if 00 [UHCI])
Subsystem: Dell: Unknown device 016c
Control: I/O+ Mem- BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B-
Status: Cap- 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0
Interrupt: pin B routed to IRQ 177
Region 4: I/O ports at 9cc0 [size=32]

00:1d.2 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB UHCI Controller #3 (rev 02) (prog-if 00 [UHCI])
Subsystem: Dell: Unknown device 016c
Control: I/O+ Mem- BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B-
Status: Cap- 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0
Interrupt: pin C routed to IRQ 185
Region 4: I/O ports at 9ca0 [size=32]

00:1d.7 USB Controller: Intel Corporation 82801EB/ER (ICH5/ICH5R) USB2 EHCI Controller (rev 02) (prog-if 20 [EHCI])
Subsystem: Dell: Unknown device 016c
Control: I/O- Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR+ FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0
Interrupt: pin D routed to IRQ 193
Region 0: Memory at feb00000 (32-bit, non-prefetchable) [size=1K]
Capabilities: <available only to root>

00:1e.0 PCI bridge: Intel Corporation 82801 PCI Bridge (rev c2) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR+ FastB2B-
Status: Cap- 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0
Bus: primary=00, secondary=09, subordinate=09, sec-latency=32
I/O behind bridge: 0000a000-0000afff
Memory behind bridge: fdf00000-fe0fffff
Prefetchable memory behind bridge: f0000000-f7ffffff
Secondary status: 66Mhz- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort+ <SERR- <PERR-

BridgeCtl: Parity+ SERR+ NoISA- VGA+ MAbort- >Reset- FastB2B-

00:1f.0 ISA bridge: Intel Corporation 82801EB/ER (ICH5/ICH5R) LPC Interface Bridge (rev 02)
Control: I/O+ Mem+ BusMaster+ SpecCycle+ MemWINV- VGASnoop- ParErr+ Stepping- SERR+ FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0

01:00.0 PCI bridge: Intel Corporation 80332 [Dobson] I/O processor (A-Segment Bridge) (rev 06) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=01, secondary=02, subordinate=02, sec-latency=64
I/O behind bridge: 0000e000-0000efff
Memory behind bridge: fe900000-feafffff
Prefetchable memory behind bridge: 00000000f8000000-00000000f8f00000
Secondary status: 66Mhz+ FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort+ <SERR- <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

01:00.2 PCI bridge: Intel Corporation 80332 [Dobson] I/O processor (B-Segment Bridge) (rev 06) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=01, secondary=03, subordinate=03, sec-latency=32
I/O behind bridge: 0000f000-0000ffff
Memory behind bridge: fff00000-000fffff
Prefetchable memory behind bridge: 00000000fff00000-0000000000000000
Secondary status: 66Mhz+ FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort+ <SERR- <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

02:05.0 SCSI storage controller: LSI Logic / Symbios Logic 53c1030 PCI-X Fusion-MPT Dual Ultra320 SCSI (rev 08)
Subsystem: Dell PowerEdge 1850 MPT Fusion SCSI/RAID (Perc 4)
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV+ VGASnoop- ParErr- Stepping- SERR+ FastB2B-
Status: Cap+ 66Mhz+ UDF- FastB2B- ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 72 (4250ns min, 4500ns max), Cache Line Size 10
Interrupt: pin A routed to IRQ 201
Region 0: I/O ports at ec00 [size=256]
Region 1: Memory at fe9f0000 (64-bit, non-prefetchable) [size=64K]
Region 3: Memory at fe9e0000 (64-bit, non-prefetchable) [size=64K]
Expansion ROM at fea00000 [disabled] [size=1M]
Capabilities: <available only to root>

05:00.0 PCI bridge: Intel Corporation 6700PXH PCI Express-to-PCI Bridge A (rev 09) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=05, secondary=06, subordinate=06, sec-latency=32
I/O behind bridge: 0000d000-0000dfff
Memory behind bridge: fe600000-fe7fffff
Prefetchable memory behind bridge: 00000000fff00000-0000000000000000
Secondary status: 66Mhz+ FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort+ <SERR- <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-
Capabilities: <available only to root>

05:00.2 PCI bridge: Intel Corporation 6700PXH PCI Express-to-PCI Bridge B (rev 09) (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Bus: primary=05, secondary=07, subordinate=07, sec-latency=32
I/O behind bridge: 0000c000-0000cfff
Memory behind bridge: fe400000-fe5fffff
Prefetchable memory behind bridge: 00000000fff00000-0000000000000000
Secondary status: 66Mhz+ FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort+ <SERR- <PERR-
BridgeCtl: Parity+ SERR+ NoISA+ VGA- MAbort- >Reset- FastB2B-

Capabilities: <available only to root>

06:07.0 Ethernet controller: Intel Corporation 82541GI/PI Gigabit Ethernet Controller (rev 05)
Subsystem: Dell: Unknown device 016d
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV+ VGASnoop- ParErr- Stepping- SERR+ FastB2B-
Status: Cap+ 66Mhz+ UDF- FastB2B- ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 32 (63750ns min), Cache Line Size 10
Interrupt: pin A routed to IRQ 209
Region 0: Memory at fe6e0000 (32-bit, non-prefetchable) [size=128K]
Region 2: I/O ports at dcc0 [size=64]
Capabilities: <available only to root>

07:08.0 Ethernet controller: Intel Corporation 82541GI/PI Gigabit Ethernet Controller (rev 05)
Subsystem: Dell: Unknown device 016d
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV+ VGASnoop- ParErr- Stepping- SERR+ FastB2B-
Status: Cap+ 66Mhz+ UDF- FastB2B- ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 32 (63750ns min), Cache Line Size 10
Interrupt: pin A routed to IRQ 217
Region 0: Memory at fe4e0000 (32-bit, non-prefetchable) [size=128K]
Region 2: I/O ports at ccc0 [size=64]
Capabilities: <available only to root>

08:00.0 InfiniBand: Mellanox Technologies MT25208 InfiniHost III Ex (rev a0)
Subsystem: Mellanox Technologies MT25208 InfiniHost III Ex
Control: I/O- Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 0, Cache Line Size 10
Interrupt: pin A routed to IRQ 169
Region 0: Memory at fe100000 (64-bit, non-prefetchable) [size=1M]
Region 2: Memory at f9000000 (64-bit, prefetchable) [size=8M]
Capabilities: <available only to root>

09:0d.0 VGA compatible controller: ATI Technologies Inc Radeon RV100 QY [Radeon 7000/VE] (prog-if 00 [VGA])
Subsystem: Dell: Unknown device 016c
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop+ ParErr- Stepping+ SERR+ FastB2B-
Status: Cap+ 66Mhz- UDF- FastB2B+ ParErr- DEVSEL=medium >TAbort- <TAbort- <MAbort- >SERR- <PERR-
Latency: 32 (2000ns min), Cache Line Size 10
Interrupt: pin A routed to IRQ 185
Region 0: Memory at f0000000 (32-bit, prefetchable) [size=128M]
Region 1: I/O ports at ac00 [size=256]
Region 2: Memory at fdff0000 (32-bit, non-prefetchable) [size=64K]
Capabilities: <available only to root>

A.2 Software

System information # `uname -a`

```
Linux en384 2.6.9-42.0.2.EL_lustre.1.4.7.3smp #1 SMP \  
Fri Nov 10 12:50:31 MST 2006 x86_64 x86_64 x86_64 GNU/Linux
```

Kernel Red Hat Enterprise Linux WS release 4 (Nahant Update 4)
Kernel \r on an \m

OpenFabrics Enterprise Distribution info OFED-1.0

```
OFED-1.0  
  
openib-1.0 (REV=8031)  
# User space  
https://openib.org/svn/gen2/branches/1.0/src/userspace
```



```

# Kernel space
https://openib.org/svn/gen2/branches/1.0/ofed/tags/1.0/linux-kernel
Git:
ref: refs/heads/for-2.6.17
commit 959eb39297e8c82f61fbfc283ad4ff11c883bf1e

# MPI
mpi_osu-0.9.7-mlx2.1.0.tgz
openmpi-1.1b1-1.src.rpm
mpitests-1.0-0.src.rpm

```

Open MPI info # ompinfo -param all all

```

      Open MPI: 1.2
Open MPI SVN revision: r14027
      Open RTE: 1.2
Open RTE SVN revision: r14027
      OPAL: 1.2
OPAL SVN revision: r14027
  MCA backtrace: execinfo (MCA v1.0, API v1.0, Component v1.2)
  MCA memory: ptmalloc2 (MCA v1.0, API v1.0, Component v1.2)
  MCA paffinity: linux (MCA v1.0, API v1.0, Component v1.2)
  MCA maffinity: first_use (MCA v1.0, API v1.0, Component v1.2)
  MCA maffinity: libnuma (MCA v1.0, API v1.0, Component v1.2)
  MCA timer: linux (MCA v1.0, API v1.0, Component v1.2)
  MCA allocator: basic (MCA v1.0, API v1.0, Component v1.0)
  MCA allocator: bucket (MCA v1.0, API v1.0, Component v1.0)
  MCA coll: basic (MCA v1.0, API v1.0, Component v1.2)
  MCA coll: self (MCA v1.0, API v1.0, Component v1.2)
  MCA coll: sm (MCA v1.0, API v1.0, Component v1.2)
  MCA coll: tuned (MCA v1.0, API v1.0, Component v1.3)
  MCA io: romio (MCA v1.0, API v1.0, Component v1.2)
  MCA mpool: openib (MCA v1.0, API v1.0, Component v1.2)
  MCA mpool: sm (MCA v1.0, API v1.0, Component v1.2)
  MCA pml: cm (MCA v1.0, API v1.0, Component v1.2)
  MCA pml: ob1 (MCA v1.0, API v1.0, Component v1.2)
  MCA bml: r2 (MCA v1.0, API v1.0, Component v1.2)
  MCA rcache: rb (MCA v1.0, API v1.0, Component v1.2)
  MCA rcache: vma (MCA v1.0, API v1.0, Component v1.2)
  MCA btl: openib (MCA v1.0, API v1.0.1, Component v1.2)
  MCA btl: self (MCA v1.0, API v1.0.1, Component v1.2)
  MCA btl: sm (MCA v1.0, API v1.0.1, Component v1.2)
  MCA btl: tcp (MCA v1.0, API v1.0.1, Component v1.0)
  MCA topo: unity (MCA v1.0, API v1.0, Component v1.2)
  MCA osc: pt2pt (MCA v1.0, API v1.0, Component v1.2)
  MCA errmgr: hnp (MCA v1.0, API v1.3, Component v1.2)
  MCA errmgr: orted (MCA v1.0, API v1.3, Component v1.2)
  MCA errmgr: proxy (MCA v1.0, API v1.3, Component v1.2)
  MCA gpr: null (MCA v1.0, API v1.0, Component v1.2)
  MCA gpr: proxy (MCA v1.0, API v1.0, Component v1.2)
  MCA gpr: replica (MCA v1.0, API v1.0, Component v1.2)
  MCA iof: proxy (MCA v1.0, API v1.0, Component v1.2)
  MCA iof: svc (MCA v1.0, API v1.0, Component v1.2)
  MCA ns: proxy (MCA v1.0, API v2.0, Component v1.2)
  MCA ns: replica (MCA v1.0, API v2.0, Component v1.2)
  MCA oob: tcp (MCA v1.0, API v1.0, Component v1.0)
  MCA ras: dash_host (MCA v1.0, API v1.3, Component v1.2)
  MCA ras: gridengine (MCA v1.0, API v1.3, Component v1.2)
  MCA ras: localhost (MCA v1.0, API v1.3, Component v1.2)
  MCA ras: slurm (MCA v1.0, API v1.3, Component v1.2)
  MCA ras: tm (MCA v1.0, API v1.3, Component v1.2)
  MCA rds: hostfile (MCA v1.0, API v1.3, Component v1.2)
  MCA rds: proxy (MCA v1.0, API v1.3, Component v1.2)
  MCA rds: resfile (MCA v1.0, API v1.3, Component v1.2)

```

```

MCA rmaps: round_robin (MCA v1.0, API v1.3, Component v1.2)
MCA rmgr: proxy (MCA v1.0, API v2.0, Component v1.2)
MCA rmgr: urm (MCA v1.0, API v2.0, Component v1.2)
MCA rml: oob (MCA v1.0, API v1.0, Component v1.2)
MCA pls: gridengine (MCA v1.0, API v1.3, Component v1.2)
MCA pls: proxy (MCA v1.0, API v1.3, Component v1.2)
MCA pls: rsh (MCA v1.0, API v1.3, Component v1.2)
MCA pls: slurm (MCA v1.0, API v1.3, Component v1.2)
MCA pls: tm (MCA v1.0, API v1.3, Component v1.2)
MCA sds: env (MCA v1.0, API v1.0, Component v1.2)
MCA sds: pipe (MCA v1.0, API v1.0, Component v1.2)
MCA sds: seed (MCA v1.0, API v1.0, Component v1.2)
MCA sds: singleton (MCA v1.0, API v1.0, Component v1.2)
MCA sds: slurm (MCA v1.0, API v1.0, Component v1.2)
Prefix: /scratch3/jmsquyr/local/openmpi-1.2-gnu
Bindir: /scratch3/jmsquyr/local/openmpi-1.2-gnu/bin
Libdir: /scratch3/jmsquyr/local/openmpi-1.2-gnu/lib
Incdir: /scratch3/jmsquyr/local/openmpi-1.2-gnu/include
Pkglibdir: /scratch3/jmsquyr/local/openmpi-1.2-gnu/lib/openmpi
Sysconfdir: /scratch3/jmsquyr/local/openmpi-1.2-gnu/etc
Configured architecture: x86_64-unknown-linux-gnu
Configured by: jmsquyr
Configured on: Thu Mar 22 09:05:06 MDT 2007
Configure host: blogin1.sandia.gov
Built by: jmsquyr
Built on: Thu Mar 22 09:20:37 MDT 2007
Built host: blogin1.sandia.gov
C bindings: yes
C++ bindings: yes
Fortran77 bindings: yes (all)
Fortran90 bindings: yes
Fortran90 bindings size: small
C compiler: gcc
C compiler absolute: /usr/bin/gcc
C char size: 1
C bool size: 1
C short size: 2
C int size: 4
C long size: 8
C float size: 4
C double size: 8
C pointer size: 8
C char align: 1
C bool align: 1
C int align: 4
C float align: 4
C double align: 8
C++ compiler: g++
C++ compiler absolute: /usr/bin/g++
Fortran77 compiler: gfortran
Fortran77 compiler abs: /usr/bin/gfortran
Fortran90 compiler: gfortran
Fortran90 compiler abs: /usr/bin/gfortran
Fort integer size: 4
Fort logical size: 4
Fort logical value true: 1
Fort have integer1: yes
Fort have integer2: yes
Fort have integer4: yes
Fort have integer8: yes
Fort have integer16: no
Fort have real4: yes
Fort have real8: yes
Fort have real16: no
Fort have complex8: yes

```

```

Fort have complex16: yes
Fort have complex32: no
Fort integer1 size: 1
Fort integer2 size: 2
Fort integer4 size: 4
Fort integer8 size: 8
Fort integer16 size: -1
Fort real size: 4
Fort real4 size: 4
Fort real8 size: 8
Fort real16 size: -1
Fort dbl prec size: 4
Fort cplx size: 4
Fort dbl cplx size: 4
Fort cplx8 size: 8
Fort cplx16 size: 16
Fort cplx32 size: -1
Fort integer align: 4
Fort integer1 align: 1
Fort integer2 align: 2
Fort integer4 align: 4
Fort integer8 align: 8
Fort integer16 align: -1
Fort real align: 4
Fort real4 align: 4
Fort real8 align: 8
Fort real16 align: -1
Fort dbl prec align: 4
Fort cplx align: 4
Fort dbl cplx align: 4
Fort cplx8 align: 4
Fort cplx16 align: 8
Fort cplx32 align: -1
C profiling: yes
C++ profiling: yes
Fortran77 profiling: yes
Fortran90 profiling: yes
C++ exceptions: no
Thread support: posix (mpi: no, progress: no)
Build CFLAGS: -O3 -DNDEBUG -finline-functions -fno-strict-aliasing -pthread
Build CXXFLAGS: -O3 -DNDEBUG -finline-functions -pthread
Build FFLAGS:
Build FCFLAGS:
Build LDFLAGS: -export-dynamic
Build LIBS: -lnsl -lutil -lm
Wrapper extra CFLAGS: -pthread
Wrapper extra CXXFLAGS: -pthread
Wrapper extra FFLAGS: -pthread
Wrapper extra FCFLAGS: -pthread
Wrapper extra LDFLAGS:
Wrapper extra LIBS: -ldl -Wl,--export-dynamic -lnsl -lutil -lm -ldl
Internal debug support: no
MPI parameter check: runtime
Memory profiling support: no
Memory debugging support: no
libltdl support: yes
Heterogeneous support: yes
mpirun default --prefix: no
MCA mca: parameter "mca_param_files" (current value: "/home/jmsquyr/.openmpi/mca-params.conf:/scratch3/jmsquyr/.openmpi/mca-params.conf")
Path for MCA configuration files containing default parameter values
MCA mca: parameter "mca_component_path" (current value: "/scratch3/jmsquyr/local/openmpi-1.2-gnu/lib/openmpi-1.2-gnu/lib")
Path where to look for Open MPI and ORTE components
MCA mca: parameter "mca_verbose" (current value: <none>)
Top-level verbosity parameter
MCA mca: parameter "mca_component_show_load_errors" (current value: "1")

```

Whether to show errors for components that failed to load or not

MCA mca: parameter "mca_component_disable_dlopen" (current value: "0")
Whether to attempt to disable opening dynamic components or not

MCA mpi: parameter "mpi_param_check" (current value: "1")
Whether you want MPI API parameters checked at run-time or not. Possible values are 0 (no check)

MCA mpi: parameter "mpi_yield_when_idle" (current value: "0")
Yield the processor when waiting for MPI communication (for MPI processes, will default to 1 when not specified)

MCA mpi: parameter "mpi_event_tick_rate" (current value: "-1")
How often to progress TCP communications (0 = never, otherwise specified in microseconds)

MCA mpi: parameter "mpi_show_handle_leaks" (current value: "0")
Whether MPI_FINALIZE shows all MPI handles that were not freed or not

MCA mpi: parameter "mpi_no_free_handles" (current value: "0")
Whether to actually free MPI objects when their handles are freed

MCA mpi: parameter "mpi_show_mca_params" (current value: "0")
Whether to show all MCA parameter value during MPI_INIT or not (good for reproducibility of MPI)

MCA mpi: parameter "mpi_show_mca_params_file" (current value: <none>)
If mpi_show_mca_params is true, setting this string to a valid filename tells Open MPI to dump the MCA parameters to the file

MCA mpi: parameter "mpi_paffinity_alone" (current value: "0")
If nonzero, assume that this job is the only (set of) process(es) running on each node and bind to the local node

MCA mpi: parameter "mpi_keep_peer_hostnames" (current value: "1")
If nonzero, save the string hostnames of all MPI peer processes (mostly for error / debugging purposes)

MCA mpi: parameter "mpi_abort_delay" (current value: "0")
If nonzero, print out an identifying message when MPI_ABORT is invoked (hostname, PID of the process)

MCA mpi: parameter "mpi_abort_print_stack" (current value: "0")
If nonzero, print out a stack trace when MPI_ABORT is invoked

MCA mpi: parameter "mpi_preconnect_all" (current value: "0")
Whether to force MPI processes to create connections / warmup with *all* peers during MPI_INIT

MCA mpi: parameter "mpi_preconnect_oob" (current value: "0")
Whether to force MPI processes to fully wire-up the OOB system between MPI processes.

MCA mpi: parameter "mpi_leave_pinned" (current value: "0")
Whether to use the "leave pinned" protocol or not. Enabling this setting can help bandwidth performance

MCA mpi: parameter "mpi_leave_pinned_pipeline" (current value: "0")
Whether to use the "leave pinned pipeline" protocol or not.

MCA orte: parameter "orte_debug" (current value: "0")
Top-level ORTE debug switch

MCA orte: parameter "orte_no_daemonize" (current value: "0")
Whether to properly daemonize the ORTE daemons or not

MCA orte: parameter "orte_base_user_debugger" (current value: "totalview @mpirun@ -a @mpirun_args@ : fxp")
Sequence of user-level debuggers to search for in orterun

MCA orte: parameter "orte_abort_timeout" (current value: "10")
Time to wait [in seconds] before giving up on aborting an ORTE operation

MCA orte: parameter "orte_timing" (current value: "0")
Request that critical timing loops be measured

MCA opal: parameter "opal_signal" (current value: "6,7,8,11")
If a signal is received, display the stack trace frame

MCA backtrace: parameter "backtrace" (current value: <none>)
Default selection set of components for the backtrace framework (<none> means "use all components")

MCA backtrace: parameter "backtrace_base_verbose" (current value: "0")
Verbosity level for the backtrace framework (0 = no verbosity)

MCA backtrace: parameter "backtrace_execinfo_priority" (current value: "0")

MCA memory: parameter "memory" (current value: <none>)
Default selection set of components for the memory framework (<none> means "use all components")

MCA memory: parameter "memory_base_verbose" (current value: "0")
Verbosity level for the memory framework (0 = no verbosity)

MCA memory: parameter "memory_ptmalloc2_priority" (current value: "0")

MCA paffinity: parameter "paffinity" (current value: <none>)
Default selection set of components for the paffinity framework (<none> means "use all components")

MCA paffinity: parameter "paffinity_linux_priority" (current value: "10")
Priority of the linux paffinity component

MCA paffinity: information "paffinity_linux_have_cpu_set_t" (value: "1")
Whether this component was compiled on a system with the type cpu_set_t or not (1 = yes, 0 = no)

MCA paffinity: information "paffinity_linux_CPU_ZERO_ok" (value: "1")
Whether this component was compiled on a system where CPU_ZERO() is functional or broken (1 = functional, 0 = broken)

MCA paffinity: information "paffinity_linux_sched_setaffinity_num_params" (value: "3")
The number of parameters that sched_set_affinity() takes on the machine where this component was compiled

MCA maffinity: parameter "maffinity" (current value: <none>)
Default selection set of components for the maffinity framework (<none> means "use all components")

MCA maffinity: parameter "maffinity_first_use_priority" (current value: "10")
Priority of the first_use maffinity component

MCA maffinity: parameter "maffinity_libnuma_priority" (current value: "25")
Priority of the libnuma maffinity component

MCA timer: parameter "timer" (current value: <none>)
Default selection set of components for the timer framework (<none> means "use all components")

MCA timer: parameter "timer_base_verbose" (current value: "0")
Verbosity level for the timer framework (0 = no verbosity)

MCA timer: parameter "timer_linux_priority" (current value: "0")

MCA allocator: parameter "allocator" (current value: <none>)
Default selection set of components for the allocator framework (<none> means "use all components")

MCA allocator: parameter "allocator_base_verbose" (current value: "0")
Verbosity level for the allocator framework (0 = no verbosity)

MCA allocator: parameter "allocator_basic_priority" (current value: "0")

MCA allocator: parameter "allocator_bucket_num_buckets" (current value: "30")

MCA allocator: parameter "allocator_bucket_priority" (current value: "0")

MCA coll: parameter "coll" (current value: <none>)
Default selection set of components for the coll framework (<none> means "use all components")

MCA coll: parameter "coll_base_verbose" (current value: "0")
Verbosity level for the coll framework (0 = no verbosity)

MCA coll: parameter "coll_basic_priority" (current value: "10")
Priority of the basic coll component

MCA coll: parameter "coll_basic_crossover" (current value: "4")
Minimum number of processes in a communicator before using the logarithmic algorithms

MCA coll: parameter "coll_self_priority" (current value: "75")

MCA coll: parameter "coll_sm_priority" (current value: "0")
Priority of the sm coll component

MCA coll: parameter "coll_sm_control_size" (current value: "4096")
Length of the control data -- should usually be either the length of a cache line on most SMPs,

MCA coll: parameter "coll_sm_bootstrap_filename" (current value: "shared_mem_sm_bootstrap")
Filename (in the Open MPI session directory) of the coll sm component bootstrap rendezvous mmap

MCA coll: parameter "coll_sm_bootstrap_num_segments" (current value: "8")
Number of segments in the bootstrap file

MCA coll: parameter "coll_sm_fragment_size" (current value: "8192")
Fragment size (in bytes) used for passing data through shared memory (will be rounded up to the

MCA coll: parameter "coll_sm_mpool" (current value: "sm")
Name of the mpool component to use

MCA coll: parameter "coll_sm_comm_in_use_flags" (current value: "2")
Number of "in use" flags, used to mark a message passing area segment as currently being used

MCA coll: parameter "coll_sm_comm_num_segments" (current value: "8")
Number of segments in each communicator's shared memory message passing area (must be >= 2, and

MCA coll: parameter "coll_sm_tree_degree" (current value: "4")
Degree of the tree for tree-based operations (must be >= 1 and <= min(control_size, 255))

MCA coll: information "coll_sm_shared_mem_used_bootstrap" (value: "216")
Amount of shared memory used in the shared memory bootstrap area (in bytes)

MCA coll: parameter "coll_sm_info_num_procs" (current value: "4")
Number of processes to use for the calculation of the shared_mem_size MCA information parameter

MCA coll: information "coll_sm_shared_mem_used_data" (value: "548864")
Amount of shared memory used in the shared memory data area for info_num_procs processes (in bytes)

MCA coll: parameter "coll_tuned_priority" (current value: "30")
Priority of the tuned coll component

MCA coll: parameter "coll_tuned_pre_allocate_memory_comm_size_limit" (current value: "32768")
Size of communicator where we stop pre-allocating memory for the fixed internal buffer used for

MCA coll: parameter "coll_tuned_init_tree_fanout" (current value: "4")
Initial fanout used in the tree topologies for each communicator. This is only an initial guess,

MCA coll: parameter "coll_tuned_init_chain_fanout" (current value: "4")
Initial fanout used in the chain (fanout followed by pipeline) topologies for each communicator.

MCA coll: parameter "coll_tuned_use_dynamic_rules" (current value: "0")
Switch used to decide if we use static (compiled/if statements) or dynamic (built at runtime) d

MCA io: parameter "io_base_freelist_initial_size" (current value: "16")
Initial MPI-2 IO request freelist size

MCA io: parameter "io_base_freelist_max_size" (current value: "64")
Max size of the MPI-2 IO request freelist

MCA io: parameter "io_base_freelist_increment" (current value: "16")
Increment size of the MPI-2 IO request freelist

MCA io: parameter "io" (current value: <none>)
Default selection set of components for the io framework (<none> means "use all components that")

MCA io: parameter "io_base_verbosity" (current value: "0")
Verbosity level for the io framework (0 = no verbosity)

MCA io: parameter "io_romio_priority" (current value: "10")
Priority of the io romio component

MCA io: parameter "io_romio_delete_priority" (current value: "10")
Delete priority of the io romio component

MCA io: parameter "io_romio_enable_parallel_optimizations" (current value: "0")
Enable set of Open MPI-added options to improve collective file i/o performance

MCA mpool: parameter "mpool" (current value: <none>)
Default selection set of components for the mpool framework (<none> means "use all components that")

MCA mpool: parameter "mpool_base_verbosity" (current value: "0")
Verbosity level for the mpool framework (0 = no verbosity)

MCA mpool: parameter "mpool_openib_rcache_name" (current value: "rb")
The name of the registration cache the mpool should use

MCA mpool: parameter "mpool_openib_priority" (current value: "0")

MCA mpool: parameter "mpool_sm_allocator" (current value: "bucket")
Name of allocator component to use with sm mpool

MCA mpool: parameter "mpool_sm_max_size" (current value: "536870912")
Maximum size of the sm mpool shared memory file

MCA mpool: parameter "mpool_sm_min_size" (current value: "134217728")
Minimum size of the sm mpool shared memory file

MCA mpool: parameter "mpool_sm_per_peer_size" (current value: "33554432")
Size (in bytes) to allocate per local peer in the sm mpool shared memory file, bounded by min_size

MCA mpool: parameter "mpool_sm_priority" (current value: "0")

MCA mpool: parameter "mpool_base_use_mem_hooks" (current value: "0")
use memory hooks for deregistering freed memory

MCA mpool: parameter "mpool_use_mem_hooks" (current value: "0")
(deprecated, use mpool_base_use_mem_hooks)

MCA pml: parameter "pml" (current value: <none>)
Default selection set of components for the pml framework (<none> means "use all components that")

MCA pml: parameter "pml_base_verbosity" (current value: "0")
Verbosity level for the pml framework (0 = no verbosity)

MCA pml: parameter "pml_cm_free_list_num" (current value: "4")
Initial size of request free lists

MCA pml: parameter "pml_cm_free_list_max" (current value: "-1")
Maximum size of request free lists

MCA pml: parameter "pml_cm_free_list_inc" (current value: "64")
Number of elements to add when growing request free lists

MCA pml: parameter "pml_cm_priority" (current value: "30")
CM PML selection priority

MCA pml: parameter "pml_ob1_free_list_num" (current value: "4")

MCA pml: parameter "pml_ob1_free_list_max" (current value: "-1")

MCA pml: parameter "pml_ob1_free_list_inc" (current value: "64")

MCA pml: parameter "pml_ob1_priority" (current value: "20")

MCA pml: parameter "pml_ob1_eager_limit" (current value: "131072")

MCA pml: parameter "pml_ob1_send_pipeline_depth" (current value: "3")

MCA pml: parameter "pml_ob1_recv_pipeline_depth" (current value: "4")

MCA bml: parameter "bml" (current value: <none>)
Default selection set of components for the bml framework (<none> means "use all components that")

MCA bml: parameter "bml_base_verbosity" (current value: "0")
Verbosity level for the bml framework (0 = no verbosity)

MCA bml: parameter "bml_r2_show_unreach_errors" (current value: "1")
Show error message when procs are unreachable

MCA bml: parameter "bml_r2_priority" (current value: "0")

MCA rcache: parameter "rcache" (current value: <none>)
Default selection set of components for the rcache framework (<none> means "use all components that")

MCA rcache: parameter "rcache_base_verbosity" (current value: "0")
Verbosity level for the rcache framework (0 = no verbosity)

MCA rcache: parameter "rcache_rb_priority" (current value: "0")

MCA rcache: parameter "rcache_vma_mru_len" (current value: "256")
The maximum size IN ENTRIES of the MRU (most recently used) rcache list

MCA rcache: parameter "rcache_vma_mru_size" (current value: "1073741824")
The maximum size IN BYTES of the MRU (most recently used) rcache list

MCA rcache: parameter "rcache_vma_priority" (current value: "0")

MCA btl: parameter "btl_base_debug" (current value: "0")
If btl_base_debug is 1 standard debug is output, if > 1 verbose debug is output

MCA btl: parameter "btl" (current value: <none>)
Default selection set of components for the btl framework (<none> means "use all components tha

MCA btl: parameter "btl_base_verbose" (current value: "0")
Verbosity level for the btl framework (0 = no verbosity)

MCA btl: parameter "btl_openib_verbose" (current value: "0")
Output some verbose OpenIB BTL information (0 = no output, nonzero = output)

MCA btl: parameter "btl_openib_warn_no_hca_params_found" (current value: "1")
Warn when no HCA-specific parameters are found in the INI file specified by the btl_openib_hca_

MCA btl: parameter "btl_openib_warn_default_gid_prefix" (current value: "1")
Warn when there is more than one active ports and at least one of them connected to the network

MCA btl: parameter "btl_openib_hca_param_files" (current value: "/scratch3/jmsquyr/local/openmpi-1.2-gnu
Colon-delimited list of INI-style files that contain HCA vendor/part-specific parameters

MCA btl: parameter "btl_openib_max_btls" (current value: "-1")
Maximum number of HCA ports to use (-1 = use all available, otherwise must be >= 1)

MCA btl: parameter "btl_openib_free_list_num" (current value: "8")
Initial size of free lists (must be >= 1)

MCA btl: parameter "btl_openib_free_list_max" (current value: "-1")
Maximum size of free lists (-1 = infinite, otherwise must be >= 0)

MCA btl: parameter "btl_openib_free_list_inc" (current value: "32")
Increment size of free lists (must be >= 1)

MCA btl: parameter "btl_openib_mpool" (current value: "openib")
Name of the memory pool to be used (it is unlikely that you will ever want to change this

MCA btl: parameter "btl_openib_reg_mru_len" (current value: "16")
Length of the registration cache most recently used list (must be >= 1)

MCA btl: parameter "btl_openib_ib_cq_size" (current value: "1000")
Size of the IB completion queue (will automatically be set to a minimum of (2 * number_of_peers

MCA btl: parameter "btl_openib_ib_sg_list_size" (current value: "4")
Size of IB segment list (must be >= 1)

MCA btl: parameter "btl_openib_ib_pkey_ix" (current value: "0")
InfiniBand pkey index (must be >= 0)

MCA btl: parameter "btl_openib_ib_psn" (current value: "0")
InfiniBand packet sequence starting number (must be >= 0)

MCA btl: parameter "btl_openib_ib_qp_ous_rd_atom" (current value: "4")
InfiniBand outstanding atomic reads (must be >= 0)

MCA btl: parameter "btl_openib_ib_mtu" (current value: "3")
IB MTU, in bytes (if not specified in INI files). Valid values are: 1=256 bytes, 2=512 bytes, ,

MCA btl: parameter "btl_openib_ib_min_rnr_timer" (current value: "5")
InfiniBand minimum "receiver not ready" timer, in seconds (must be >= 1)

MCA btl: parameter "btl_openib_ib_timeout" (current value: "10")
InfiniBand transmit timeout, in seconds(must be >= 1)

MCA btl: parameter "btl_openib_ib_retry_count" (current value: "7")
InfiniBand transmit retry count (must be >= 1)

MCA btl: parameter "btl_openib_ib_rnr_retry" (current value: "7")
InfiniBand "receiver not ready" retry count (must be >= 1)

MCA btl: parameter "btl_openib_ib_max_rdma_dst_ops" (current value: "4")
InfiniBand maximum pending RDMA destination operations (must be >= 1)

MCA btl: parameter "btl_openib_ib_service_level" (current value: "0")
InfiniBand service level (must be >= 0)

MCA btl: parameter "btl_openib_ib_static_rate" (current value: "0")
InfiniBand static rate (must be >= 0; defulat: %d)

MCA btl: parameter "btl_openib_exclusivity" (current value: "1024")
OpenIB BTL exclusivity (must be >= 0)

MCA btl: parameter "btl_openib_rd_num" (current value: "8")
Number of receive descriptors to post to a queue pair (must be >= 1)

MCA btl: parameter "btl_openib_rd_low" (current value: "6")
Low water mark before reposting occurs (must be >= 1)

MCA btl: parameter "btl_openib_rd_win" (current value: "4")
Window size at which generate explicit credit message (must be >= 1)

MCA btl: parameter "btl_openib_use_srq" (current value: "0")
If nonzero, use the InfiniBand shared receive queue ("SRQ")

MCA btl: parameter "btl_openib_srq_rd_max" (current value: "1000")
Maximum number of receive descriptors posted per SRQ (only relevant if btl_openib_use_srq is true)

MCA btl: parameter "btl_openib_srq_rd_per_peer" (current value: "16")
Number of receive descriptors posted per peer in the SRQ (only relevant if btl_openib_use_srq is true)

MCA btl: parameter "btl_openib_srq_sd_max" (current value: "8")
Maximum number of send descriptors posted (only relevant if btl_openib_use_srq is true; must be >= 1)

MCA btl: parameter "btl_openib_use_eager_rdma" (current value: "1")
Use RDMA for eager messages

MCA btl: parameter "btl_openib_eager_rdma_threshold" (current value: "16")
Use RDMA for short messages after this number of messages are received from a given peer (must be >= 1)

MCA btl: parameter "btl_openib_max_eager_rdma" (current value: "16")
Maximum number of peers allowed to use RDMA for short messages (RDMA is used for all long messages)

MCA btl: parameter "btl_openib_eager_rdma_num" (current value: "16")
Number of RDMA buffers to allocate for small messages (must be >= 1)

MCA btl: parameter "btl_openib_btls_per_lid" (current value: "1")
Number of BTLs to create for each InfiniBand LID (must be >= 1)

MCA btl: parameter "btl_openib_max_lmc" (current value: "0")
Maximum number of LIDs to use for each HCA port (must be >= 0, where 0 = use all available)

MCA btl: parameter "btl_openib_buffer_alignment" (current value: "64")
Preferred communication buffer alignment, in bytes (must be >= 0)

MCA btl: parameter "btl_openib_eager_limit" (current value: "12288")
Eager send limit, in bytes (must be >= 1)

MCA btl: parameter "btl_openib_min_send_size" (current value: "32768")
Minimum send size, in bytes (must be >= 1)

MCA btl: parameter "btl_openib_max_send_size" (current value: "65536")
Maximum send size, in bytes (must be >= 1)

MCA btl: parameter "btl_openib_min_rdma_size" (current value: "1048576")
Minimum RDMA size, in bytes (must be >= 1)

MCA btl: parameter "btl_openib_max_rdma_size" (current value: "1048576")
Maximum RDMA size, in bytes (must be >= 1)

MCA btl: parameter "btl_openib_flags" (current value: "54")
BTL flags, added together: SEND=1, PUT=2, GET=4 (cannot be 0)

MCA btl: parameter "btl_openib_bandwidth" (current value: "800")
Approximate maximum bandwidth of network (must be >= 1)

MCA btl: parameter "btl_openib_priority" (current value: "0")

MCA btl: parameter "btl_self_free_list_num" (current value: "0")
Number of fragments by default

MCA btl: parameter "btl_self_free_list_max" (current value: "-1")
Maximum number of fragments

MCA btl: parameter "btl_self_free_list_inc" (current value: "32")
Increment by this number of fragments

MCA btl: parameter "btl_self_eager_limit" (current value: "131072")
Eager size fragment (before the rendez-vous protocol)

MCA btl: parameter "btl_self_min_send_size" (current value: "262144")
Minimum fragment size after the rendez-vous

MCA btl: parameter "btl_self_max_send_size" (current value: "262144")
Maximum fragment size after the rendez-vous

MCA btl: parameter "btl_self_min_rdma_size" (current value: "2147483647")
Maximum fragment size for the RDMA transfer

MCA btl: parameter "btl_self_max_rdma_size" (current value: "2147483647")
Maximum fragment size for the RDMA transfer

MCA btl: parameter "btl_self_exclusivity" (current value: "65536")
Device exclusivity

MCA btl: parameter "btl_self_flags" (current value: "10")
Active behavior flags

MCA btl: parameter "btl_self_priority" (current value: "0")

MCA btl: parameter "btl_sm_free_list_num" (current value: "8")

MCA btl: parameter "btl_sm_free_list_max" (current value: "-1")

MCA btl: parameter "btl_sm_free_list_inc" (current value: "64")

MCA btl: parameter "btl_sm_exclusivity" (current value: "65535")

MCA btl: parameter "btl_sm_latency" (current value: "100")

MCA btl: parameter "btl_sm_max_procs" (current value: "-1")

MCA btl: parameter "btl_sm_sm_extra_procs" (current value: "2")

MCA btl: parameter "btl_sm_mpool" (current value: "sm")

MCA btl: parameter "btl_sm_eager_limit" (current value: "4096")


```

MCA btl: parameter "btl_sm_max_frag_size" (current value: "32768")
MCA btl: parameter "btl_sm_size_of_cb_queue" (current value: "128")
MCA btl: parameter "btl_sm_cb_lazy_free_freq" (current value: "120")
MCA btl: parameter "btl_sm_priority" (current value: "0")
MCA btl: parameter "btl_tcp_if_include" (current value: <none>)
MCA btl: parameter "btl_tcp_if_exclude" (current value: "lo")
MCA btl: parameter "btl_tcp_free_list_num" (current value: "8")
MCA btl: parameter "btl_tcp_free_list_max" (current value: "-1")
MCA btl: parameter "btl_tcp_free_list_inc" (current value: "32")
MCA btl: parameter "btl_tcp_sndbuf" (current value: "131072")
MCA btl: parameter "btl_tcp_rcvbuf" (current value: "131072")
MCA btl: parameter "btl_tcp_endpoint_cache" (current value: "30720")
MCA btl: parameter "btl_tcp_exclusivity" (current value: "0")
MCA btl: parameter "btl_tcp_eager_limit" (current value: "65536")
MCA btl: parameter "btl_tcp_min_send_size" (current value: "65536")
MCA btl: parameter "btl_tcp_max_send_size" (current value: "131072")
MCA btl: parameter "btl_tcp_min_rdma_size" (current value: "131072")
MCA btl: parameter "btl_tcp_max_rdma_size" (current value: "2147483647")
MCA btl: parameter "btl_tcp_flags" (current value: "122")
MCA btl: parameter "btl_tcp_priority" (current value: "0")
MCA btl: parameter "btl_base_include" (current value: <none>)
MCA btl: parameter "btl_base_exclude" (current value: <none>)
MCA btl: parameter "btl_base_warn_component_unused" (current value: "1")
      This parameter is used to turn on warning messages when certain NICs are not used
MCA mtl: parameter "mtl" (current value: <none>)
      Default selection set of components for the mtl framework (<none> means "use all components tha
MCA mtl: parameter "mtl_base_verbose" (current value: "0")
      Verbosity level for the mtl framework (0 = no verbosity)
MCA topo: parameter "topo" (current value: <none>)
      Default selection set of components for the topo framework (<none> means "use all components th
MCA topo: parameter "topo_base_verbose" (current value: "0")
      Verbosity level for the topo framework (0 = no verbosity)
MCA osc: parameter "osc" (current value: <none>)
      Default selection set of components for the osc framework (<none> means "use all components tha
MCA osc: parameter "osc_base_verbose" (current value: "0")
      Verbosity level for the osc framework (0 = no verbosity)
MCA osc: parameter "osc_pt2pt_no_locks" (current value: "0")
      Enable optimizations available only if MPI_LOCK is not used.
MCA osc: parameter "osc_pt2pt_eager_limit" (current value: "16384")
      Max size of eagerly sent data
MCA osc: parameter "osc_pt2pt_priority" (current value: "0")
MCA errmgr: parameter "errmgr" (current value: <none>)
      Default selection set of components for the errmgr framework (<none> means "use all components
MCA errmgr: parameter "errmgr_hnp_debug" (current value: "0")
MCA errmgr: parameter "errmgr_hnp_priority" (current value: "0")
MCA errmgr: parameter "errmgr_orted_debug" (current value: "0")
MCA errmgr: parameter "errmgr_orted_priority" (current value: "0")
MCA errmgr: parameter "errmgr_proxy_debug" (current value: "0")
MCA errmgr: parameter "errmgr_proxy_priority" (current value: "0")
MCA gpr: parameter "gpr_base_maxsize" (current value: "2147483647")
MCA gpr: parameter "gpr_base_blocksize" (current value: "512")
MCA gpr: parameter "gpr" (current value: <none>)
      Default selection set of components for the gpr framework (<none> means "use all components tha
MCA gpr: parameter "gpr_null_priority" (current value: "0")
MCA gpr: parameter "gpr_proxy_debug" (current value: "0")
MCA gpr: parameter "gpr_proxy_priority" (current value: "0")
MCA gpr: parameter "gpr_replica_debug" (current value: "0")
MCA gpr: parameter "gpr_replica_isolate" (current value: "0")
MCA gpr: parameter "gpr_replica_priority" (current value: "0")
MCA ioof: parameter "ioof_base_window_size" (current value: "4096")
MCA ioof: parameter "ioof_base_service" (current value: "0.0.0")
MCA ioof: parameter "ioof" (current value: <none>)
      Default selection set of components for the ioof framework (<none> means "use all components tha
MCA ioof: parameter "ioof_proxy_debug" (current value: "1")
MCA ioof: parameter "ioof_proxy_priority" (current value: "0")

```

```

MCA iof: parameter "iof_svc_debug" (current value: "1")
MCA iof: parameter "iof_svc_priority" (current value: "0")
MCA ns: parameter "ns" (current value: <none>)
      Default selection set of components for the ns framework (<none> means "use all components that
MCA ns: parameter "ns_proxy_debug" (current value: "0")
MCA ns: parameter "ns_proxy_maxsize" (current value: "2147483647")
MCA ns: parameter "ns_proxy_blocksize" (current value: "512")
MCA ns: parameter "ns_proxy_priority" (current value: "0")
MCA ns: parameter "ns_replica_debug" (current value: "0")
MCA ns: parameter "ns_replica_isolate" (current value: "0")
MCA ns: parameter "ns_replica_maxsize" (current value: "2147483647")
MCA ns: parameter "ns_replica_blocksize" (current value: "512")
MCA ns: parameter "ns_replica_priority" (current value: "0")
MCA oob: parameter "oob" (current value: <none>)
      Default selection set of components for the oob framework (<none> means "use all components tha
MCA oob: parameter "oob_base_verbosity" (current value: "0")
      Verbosity level for the oob framework (0 = no verbosity)
MCA oob: parameter "oob_tcp_peer_limit" (current value: "-1")
MCA oob: parameter "oob_tcp_peer_retries" (current value: "60")
MCA oob: parameter "oob_tcp_debug" (current value: "0")
MCA oob: parameter "oob_tcp_include" (current value: "eth0")
MCA oob: parameter "oob_tcp_exclude" (current value: <none>)
MCA oob: parameter "oob_tcp_sndbuf" (current value: "131072")
MCA oob: parameter "oob_tcp_rcvbuf" (current value: "131072")
MCA oob: parameter "oob_tcp_connect_timeout" (current value: "600")
      connect() timeout in seconds, before trying next interface
MCA oob: parameter "oob_tcp_connect_sleep" (current value: "1")
      Enable (1) /Disable (0) random sleep for connection wireup
MCA oob: parameter "oob_tcp_listen_mode" (current value: "event")
      Mode for HNP to accept incoming connections: event, listen_thread
MCA oob: parameter "oob_tcp_listen_thread_max_queue" (current value: "10")
      High water mark for queued accepted socket list size
MCA oob: parameter "oob_tcp_listen_thread_max_time" (current value: "10")
      Maximum amount of time (in milliseconds) to wait between processing accepted socket list
MCA oob: parameter "oob_tcp_accept_spin_count" (current value: "10")
      Number of times to let accept return EWOULDBLOCK before updating accepted socket list
MCA oob: parameter "oob_tcp_priority" (current value: "0")
MCA ras: parameter "ras" (current value: <none>)
MCA ras: parameter "ras_dash_host_priority" (current value: "5")
      Selection priority for the dash_host RAS component
MCA ras: parameter "ras_gridengine_debug" (current value: "0")
      Enable debugging output for the gridengine ras component
MCA ras: parameter "ras_gridengine_priority" (current value: "100")
      Priority of the gridengine ras component
MCA ras: parameter "ras_gridengine_verbosity" (current value: "0")
      Enable verbose output for the gridengine ras component
MCA ras: parameter "ras_gridengine_show_jobid" (current value: "0")
      Show the JOB_ID of the Grid Engine job
MCA ras: parameter "ras_localhost_priority" (current value: "0")
      Selection priority for the localhost RAS component
MCA ras: parameter "ras_slurm_priority" (current value: "75")
      Priority of the slurm ras component
MCA ras: parameter "ras_tm_priority" (current value: "100")
      Priority of the tm ras component
MCA rds: parameter "rds" (current value: <none>)
MCA rds: parameter "rds_hostfile_debug" (current value: "0")
      Toggle debug output for hostfile RDS component
MCA rds: parameter "rds_hostfile_path" (current value: "/scratch3/jmsquyr/local/openmpi-1.2-gnu/etc/open
      ORTE Host filename
MCA rds: parameter "rds_hostfile_priority" (current value: "0")
MCA rds: parameter "rds_proxy_priority" (current value: "0")
MCA rds: parameter "rds_resfile_debug" (current value: "0")
      Toggle debug output for resfile RDS component
MCA rds: parameter "rds_resfile_name" (current value: <none>)
      ORTE Resource filename

```

MCA rds: parameter "rds_resfile_priority" (current value: "0")

MCA rmaps: parameter "rmaps_base_verbose" (current value: "0")
Verbosity level for the rmaps framework

MCA rmaps: parameter "rmaps_base_schedule_policy" (current value: "unspec")
Scheduling Policy for RMAPS. [slot | node]

MCA rmaps: parameter "rmaps_base_pernode" (current value: "0")
Launch one ppn as directed

MCA rmaps: parameter "rmaps_base_n_pernode" (current value: "-1")
Launch n procs/node

MCA rmaps: parameter "rmaps_base_schedule_local" (current value: "1")
If nonzero, allow scheduling MPI applications on the same node as mpirun (default). If zero, d

MCA rmaps: parameter "rmaps_base_no_oversubscribe" (current value: "0")
If nonzero, then do not allow oversubscription of nodes - mpirun will return an error if there

MCA rmaps: parameter "rmaps" (current value: <none>)
Default selection set of components for the rmaps framework (<none> means "use all components t

MCA rmaps: parameter "rmaps_round_robin_debug" (current value: "1")
Toggle debug output for Round Robin RMAPS component

MCA rmaps: parameter "rmaps_round_robin_priority" (current value: "1")
Selection priority for Round Robin RMAPS component

MCA rmgr: parameter "rmgr" (current value: <none>)
Default selection set of components for the rmgr framework (<none> means "use all components th

MCA rmgr: parameter "rmgr_proxy_priority" (current value: "0")

MCA rmgr: parameter "rmgr_urm_priority" (current value: "0")

MCA rml: parameter "rml" (current value: <none>)
Default selection set of components for the rml framework (<none> means "use all components tha

MCA rml: parameter "rml_base_verbose" (current value: "0")
Verbosity level for the rml framework (0 = no verbosity)

MCA rml: parameter "rml_oob_priority" (current value: "0")

MCA pls: parameter "pls_base_reuse_daemons" (current value: "0")
If nonzero, reuse daemons to launch dynamically spawned processes. If zero, do not reuse daemo

MCA pls: parameter "pls" (current value: <none>)
Default selection set of components for the pls framework (<none> means "use all components tha

MCA pls: parameter "pls_base_verbose" (current value: "0")
Verbosity level for the pls framework (0 = no verbosity)

MCA pls: parameter "pls_gridengine_debug" (current value: "0")
Enable debugging of gridengine pls component

MCA pls: parameter "pls_gridengine_verbose" (current value: "0")
Enable verbose output of the gridengine qrsh -inherit command

MCA pls: parameter "pls_gridengine_priority" (current value: "100")
Priority of the gridengine pls component

MCA pls: parameter "pls_gridengine_orted" (current value: "orted")
The command name that the gridengine pls component will invoke for the ORTE daemon

MCA pls: parameter "pls_proxy_priority" (current value: "0")

MCA pls: parameter "pls_rsh_debug" (current value: "0")
Whether or not to enable debugging output for the rsh pls component (0 or 1)

MCA pls: parameter "pls_rsh_num_concurrent" (current value: "128")
How many pls_rsh_agent instances to invoke concurrently (must be > 0)

MCA pls: parameter "pls_rsh_force_rsh" (current value: "0")
Force the launcher to always use rsh, even for local daemons

MCA pls: parameter "pls_rsh_orted" (current value: "orted")
The command name that the rsh pls component will invoke for the ORTE daemon

MCA pls: parameter "pls_rsh_priority" (current value: "10")
Priority of the rsh pls component

MCA pls: parameter "pls_rsh_delay" (current value: "1")
Delay (in seconds) between invocations of the remote agent, but only used when the "debug" MCA

MCA pls: parameter "pls_rsh_reap" (current value: "1")
If set to 1, wait for all the processes to complete before exiting. Otherwise, quit immediatel

MCA pls: parameter "pls_rsh_assume_same_shell" (current value: "1")
If set to 1, assume that the shell on the remote node is the same as the shell on the local nod

MCA pls: parameter "pls_rsh_agent" (current value: "ssh : rsh")
The command used to launch executables on remote nodes (typically either "ssh" or "rsh")

MCA pls: parameter "pls_slurm_debug" (current value: "0")
Enable debugging of slurm pls

MCA pls: parameter "pls_slurm_priority" (current value: "75")
Default selection priority

MCA pls: parameter "pls_slurm_orted" (current value: "orted")
Command to use to start proxy orted

MCA pls: parameter "pls_slurm_args" (current value: <none>)
Custom arguments to srun

MCA pls: parameter "pls_tm_debug" (current value: "0")
Enable debugging of the TM pls

MCA pls: parameter "pls_tm_verbose" (current value: "0")
Enable verbose output of the TM pls

MCA pls: parameter "pls_tm_priority" (current value: "75")
Default selection priority

MCA pls: parameter "pls_tm_orted" (current value: "orted")
Command to use to start proxy orted

MCA pls: parameter "pls_tm_want_path_check" (current value: "1")
Whether the launching process should check for the pls_tm_orted executable in the PATH before 1

MCA sds: parameter "sds" (current value: <none>)
Default selection set of components for the sds framework (<none> means "use all components tha

MCA sds: parameter "sds_base_verbose" (current value: "0")
Verbosity level for the sds framework (0 = no verbosity)

MCA sds: parameter "sds_env_priority" (current value: "0")

MCA sds: parameter "sds_pipe_priority" (current value: "0")

MCA sds: parameter "sds_seed_priority" (current value: "0")

MCA sds: parameter "sds_singleton_priority" (current value: "0")

MCA sds: parameter "sds_slurm_priority" (current value: "0")