

Comparing Cluster Interconnects: LS-DYNA Performance Benchmarks for Gigabit Ethernet and Cisco Double Data Rate InfiniBand

Overview

The ability of a high-performance cluster (HPC) to positively affect application performance is one of the most important factors to consider in selecting a switching fabric type and interconnect. Though the characteristics of an application affect performance results, you also should consider node configuration, interconnect and switching fabric type and configuration, protocol offloading technology, storage, and file system type

This document presents a comparative analysis of the performance of LS-DYNA, a finite element analysis application, using two interconnect and fabric technologies: Gigabit Ethernet and Cisco® Double Data Rate (DDR) InfiniBand.

HPC Interconnects

Today's computing environments use several different interconnects and technologies for interprocess communication (IPC) traffic. The interconnects and fabrics used in the LS-DYNA application performance and scalability testing discussed here are Gigabit Ethernet and a new technology called Cisco DDR InfiniBand.

Gigabit Ethernet is an established standard technology that is well understood and popular in HPC environments. Cisco DDR InfiniBand is a new advancement that is based on the well-established standard single-data-rate (SDR) technology. DDR aggregates four 4-Gbps links, resulting in a 16-Gbps physical connection.

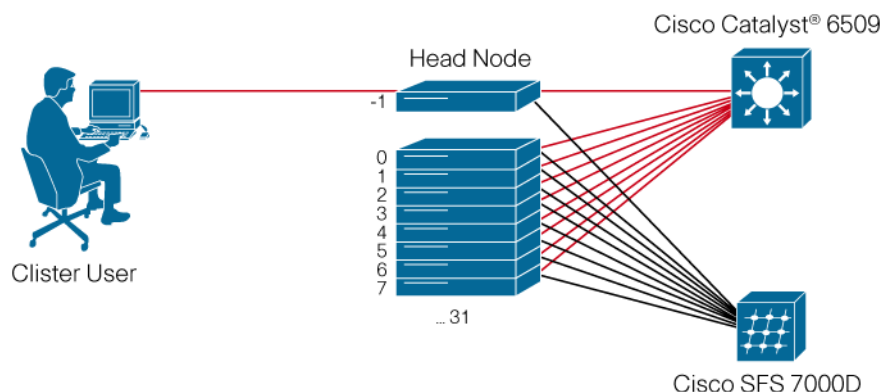
Application Performance

Cluster performance is judged by the total run time, or wall-clock time, of the application. The shorter the wall-clock time, the more application processing can be performed, and the lower the cost per application process. The main areas affecting application performance are message latency, CPU use and contention, and I/O latency for data set access.

Test Methodology

This document focuses on actual application performance metrics by measuring the performance of an application as the number of nodes in the cluster changes. For this test, the LS-DYNA crash simulation application was used. This document contains real application performance data for clusters scaling to 64 CPUs, comparing the widely used Gigabit Ethernet and Cisco DDR InfiniBand interconnects. Figure 1 shows the cluster setup used in this test.

Figure 1. Cluster System Setup with Two Fabric



Each test case was run serially as a single process on a single processor and then in parallel with increasing numbers of processors. Note that each server node in the cluster has two processors in a symmetric multiprocessing (SMP) configuration. Jobs were run with a 1:1 process-to-processor ratio. Each test case was run three times, and the results were averaged to improve accuracy.

LS-DYNA

LS-DYNA (<http://www.lstc.com/>) is a recognized commercial finite element analysis application that is used for several types of analysis. Some of many common uses include automobile crash analysis, building collapse analysis, material formation and failure analysis, and heat transfer analysis.

For the analysis in this test, two recognized vehicle crash benchmarks were used: the Three-Car Crash Test simulation and the Neon-Refined Crash Test simulation. The latter is specific to the Dodge Neon automobile.

Information on the use of LS-DYNA by various organizations can be found at <http://www.dynalook.com/>.

Testing Setup

Hardware for Gigabit Ethernet Tests

- Servers: Dual-core and dual-socket Intel 5160 processor (3.0 GHz) with 8 GB RAM
- Network interface cards (NICs): Onboard Broadcom BCM5721 Gigabit Ethernet NICs
- Ethernet switch: Cisco Catalyst 6509 Switch with Cisco Catalyst 6500 Series Supervisor Engine 720 and Cisco Catalyst 6748 Ethernet Line Cards with Distributed Forwarding Card 3B (DFC3B)
- File system: Panasas Parallel File System

Software for Gigabit Ethernet Tests

- Cisco IOS® Software Release 12.2(14r)S9
- Message Passing Interface (MPI) library used: HP-MPI 2.2.5
- LS-DYNA Version mpp970 compiled for HP-MPI and amd64 build
- Command line used to run LS-DYNA

- **Three-Car Crash Test: Gigabit Ethernet and MPI**

```
opt/hpmpi/bin/mpirun -TCP -np 64 -hostfile $PBS_NODEFILE
./mpp970_s_6763_amd64_linux_hpmpi i=3cars/3cars_shell2_150ms.k ncpu=64
memory=200m
```

- **Neon Refined Test: Gigabit Ethernet and MPI**

```
/opt/hpmpi/bin/mpirun -TCP -np 64 -hostfile $PBS_NODEFILE
./mpp970_s_6763_amd64_linux_hpmpi i=neon_refined/neon_refined01_30ms.k
ncpu=64 memory=200m
```

Hardware for Cisco InfiniBand Tests

- Servers: Dual-core and dual-socket Intel 5160 processor (3.0 GHz) with 8 GB RAM
- Host channel adapter (HCA): Cisco Single-Port 4X DDR
- InfiniBand server switch: Cisco SFS 7000D 24-Port DDR
- File system: Panasas Parallel File System

Software for Cisco InfiniBand Tests

- Cisco InfiniBand Server Switch Release 2.9.0-147
- LS-DYNA Version mpp970 compiled for HP-MPI and amd64 build
- MPI library used: HP-MPI 2.2.5
- Host drivers: OpenFabrics Enterprise Distribution (OFED)
- Command line used to run LS-DYNA

- **Three-Car Crash Test: InfiniBand OFED and MPI**

```
opt/hpmpi/bin/mpirun -ibv -np 64 -hostfile $PBS_NODEFILE
./mpp970_s_6763_amd64_linux_hpmpi i=3cars/3cars_shell2_150ms.k ncpu=64
memory=200m
```

- **Neon Refined Test: InfiniBand OFED and MPI**

```
/opt/hpmpi/bin/mpirun -ibv -np 64 -hostfile $PBS_NODEFILE
./mpp970_s_6763_amd64_linux_hpmpi i=neon_refined/neon_refined01_30ms.k
ncpu=64 memory=200m
```

Results of the Performance Tests

Table 1 shows the results for the standard Three-Car Crash Test simulations.

Table 1. Car Crash Test A: Gigabit Ethernet and Cisco InfiniBand DDR

Performance Test	Number of Nodes	Number of CPUs	Gigabit Ethernet Total Time (in Seconds)	Cisco InfiniBand DDR Total Time (in Seconds)
Three-Car Crash Test	8	32	10,732	6344
Three-Car Crash Test	16	64	14,273	3761

Table 2 shows the results for the standard Neon Refined Test simulations.

Table 2. Neon Refined Test: Gigabit Ethernet and Cisco InfiniBand DDR

Performance Test	Number of Nodes	Number of CPUs	Gigabit Ethernet Total Time (in Seconds)	Cisco InfiniBand DDR Total Time (in Seconds)
Neon Refined	8	32	1276	558
Neon Refined	16	64	5106	398

Understanding the Output Data

When analyzing performance data for applications such as LS-DYNA, the less the total time to job completion, the better.

Tables 1 and 2 show that Cisco InfiniBand DDR scales much better in both the Three-Car Crash Test simulation and the Neon Refined Test simulation.

The total time required for Ethernet jobs increases as the number of processors increases. Job slowdown occurs as the number of CPUs increases likely because IPC overwhelms the computation in the processing. These performance tests used MPI over TCP over Gigabit Ethernet.

In the InfiniBand DDR tests, the total time required for jobs is less than for Gigabit Ethernet, and it scales even better as the number of processors increases. These performance tests used MPI over the InfiniBand verbs interface (an optimized communication interface for InfiniBand).



Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 653-1315 (toll free)
Fax: 408 527-0689

Asia Pacific Headquarters
Cisco Systems, Inc.
165 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7798

Europe Headquarters
Cisco Systems International BV
Hertoforbergweg 13-19
1101 CH Amsterdam
The Netherlands
www.europe.cisco.com
Tel: +31 20 600 020 0/91
Fax: +31 20 657 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

©2007 Cisco Systems, Inc. All rights reserved. CDPV, the Cisco logo, and the Cisco Square Bridge logo are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access, Register, Abolish, BIX, Catalyst, CCDA, CCIP, CCE, CCNP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, GigaStack, HomeLink, Internet Quotient, IQS, iPhone, IPTV, IQ Expertise, the IQ logo, IQ Net, Roadshow, Scoreboard, iQuick Study, iSupport, iStocks, Meeting Place, MGX, Networking Academy, Network Registrar, Packet, PIX, ProConnect, Raptor, SaaS, ScriptShare, SlideCast, SMARTnet, StackWise, The Router, Way to Increase Your Internet Quotient, and Thousand are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (07013)