Evaluating Scalability and Power Benefits of Ninth-Generation Dell PowerEdge Servers in an HPC Environment

Energy efficiency and scalability have become increasingly important to many enterprises. This article discusses the benefits of Intel® Xeon® 51xx processors for high-performance computing cluster environments by comparing performance/watt and cluster scalability results using compute-intensive applications.

The high-performance computing (HPC) community is becoming increasingly aware of the impact that power consumption has on the total cost of an HPC cluster. Operational costs incurred in running the cluster and cooling it in a large room, or maintaining specialized buildings to prevent system failures, increase as server power consumption levels rise.

The Intel Core microarchitecture provides an interesting design choice for HPC applications. A team of Dell engineers in July 2006 used HPC benchmarks—Linpack; NAS (NASA Advanced Supercomputing) Parallel Benchmarks (NPB); MIMD (multiple instruction, multiple data) Lattice Computation (MILC); FLUENT; and OOCORE—to measure performance/watt

---

1 For more information about the Intel Core microarchitecture, visit www.intel.com/technology/architecture/coremicro/index.htm.
and cluster scalability of Dell PowerEdge 1950 servers using Intel Xeon 51xx processors.

**Benchmarks for evaluating cluster performance**

The following synthetic and application benchmarks were used in the Dell tests. These benchmarks and applications represent a broad spectrum of HPC workloads.

**Linpack.** This is a popular benchmark for HPC environments. The High-Performance Linpack (HPL)\(^2\) implementation is commonly used to rank supercomputers on the TOP500 Supercomputer Sites list.

**NAS Parallel Benchmarks.** NPB is an application-centric suite of benchmarks that has been widely used to measure and compare the performance of parallel-processing computers.\(^3\) The Dell team used the IS (Integer Sort) and LU (Lower-Upper Diagonal) B Class programs, and then calculated the sum of the IS and LU results to evaluate performance/watt.

**MILC.** The code developed by the MILC Collaboration is used in high-energy physics for simulations of 4-D special unitary (SU) lattice gauge theory on MIMD parallel-processing systems.\(^4\)

**FLUENT.** A popular computational fluid dynamics (CFD) application suite, FLUENT is commonly used in HPC environments. The FLUENT applications allow users to perform CFD analysis around their particular models.\(^5\) Several benchmark data sets (workloads) available from Fluent Inc. were used in the Dell tests.

**OOCORE.** An out-of-core matrix solver, OOCORE handles matrix equations that are too large for the cache. This benchmark writes large amounts of data to the disk and thus also tests the disk I/O performance of the server.\(^6\)

**Test environment for Dell HPC cluster**

The performance/watt test environment was based on a Dell PowerEdge 1850 server and PowerEdge 1950 server using different Intel Xeon processors. The scalability test environment was based on a cluster comprising four Dell PowerEdge 1950 servers using non-blocking Gigabit Ethernet and InfiniBand interconnects. The cluster was installed using Platform Open Cluster Stack (OCSS).\(^7\) Figures 1 and 2 describe the hardware and software the Dell team used for testing.

**Results and analysis of performance/watt and cluster interconnect tests**

The study focused on two sets of tests: one comparing performance/watt for Intel Xeon 51xx processors against previous-generation Intel

---

\(^1\) For more information about HPL, visit www.netlib.org/benchmark/hpl.
\(^2\) For more information about NAS Parallel Benchmarks, visit www.nasa.gov/Software/NPB.
\(^3\) For more information about the MILC code, visit www.physics.utah.edu/~data/milc/milc67.html.
\(^4\) For more information about FLUENT, visit www.fluent.com/software/fluentsite/index.htm.
Xeon processors, and one comparing cluster performance using Woodcrest processors with Gigabit Ethernet interconnects against InfiniBand interconnects.

Performance/watt tests
The tests compared performance/watt for eighth- and ninth-generation Dell PowerEdge servers using different types of Intel Xeon processors. The processor cores in eighth-generation servers are based on the Intel NetBurst microarchitecture. The Intel Xeon 50xx (Dempsey) processors used in ninth-generation servers are also based on the Intel NetBurst microarchitecture, which allows for higher processor frequencies compared to previous architectures while incurring a high penalty on power consumption. The Intel Xeon 51xx (Woodcrest) processors, which are also used in ninth-generation servers, are based on the Intel Core microarchitecture, which is designed to provide higher performance and lower power consumption compared to the Intel NetBurst microarchitecture.

Figure 3 compares the measured performance/watt on eighth- and ninth-generation Dell servers using these different processor architectures. A higher performance/watt metric indicates that the system is providing higher performance while consuming less power. The baseline is an eighth-generation server (PowerEdge 1850) using the Intel Xeon Nocona processor at 3.6 GHz. The next two bars show performance/watt for an eighth-generation server using the Intel Xeon Irwindale processor and the dual-core Intel Xeon 7030 Paxville processor, respectively; the remaining two bars show performance/watt for a ninth-generation server (PowerEdge 1950) using the dual-core Dempsey and Woodcrest processors, respectively.

The Paxville processor exhibited significant improvement over the Nocona and Irwindale processors because of the additional processing cores designed to enhance performance. The Dempsey processor, which is based on the same microarchitecture as the Nocona, Irwindale, and Paxville processors, did not provide any performance gains compared to the Paxville processor except with the NPB benchmark. However, performance/watt for the Woodcrest processor was significantly higher compared with all other processors—the increase ranged from 57 percent (FLUENT) to 138 percent (OOCORE) compared with the Dempsey processor—making it the most energy-efficient choice for HPC cluster environments among the tested processors.

Scalability tests
The first set of tests focused on the performance and energy efficiency derived from the Intel Core microarchitecture; the second set focused on the scalability of the microarchitecture when the same HPC benchmarks were run in a cluster using Gigabit Ethernet and InfiniBand interconnects. The testing environment is described in Figure 2.

To obtain optimal performance from each cluster node, all the physical processors in each server were utilized when running the benchmarks. Each benchmark was run using 4, 8, and 16 processors in one, two, and four servers, respectively. Figure 4 shows the results when running the benchmark with these configurations using either Gigabit Ethernet or InfiniBand interconnects. The 1 × 4 (one node with four processors) configuration results serve as the baseline, and the results for the 2 × 4 and 4 × 4 configurations are shown as relative speedups.

In Figure 4, the set of lines on the left shows the scalability of the benchmarks when Gigabit Ethernet was used as the internode communication fabric. As can be seen from these scalability lines, benchmarks bound by computation and local I/O but less dependent on internode communication scaled better than communication-intensive benchmarks. For example, the HPL and FLUENT (large) benchmarks, which are highly computation intensive and have less internode communication, scaled well. On the other hand, the NPB-IS benchmark involves intensive all-to-all communications with a mixture of small and very large messages. NPB-IS is sensitive to communication latency and bandwidth;
therefore, it experienced a performance degradation when run in a Gigabit Ethernet cluster.

The set of lines on the right in Figure 4 shows the scalability of the same set of benchmarks when InfiniBand was used as the internode communication fabric. NPB-IS, which experienced a degradation in performance when using Gigabit Ethernet, scaled up to approximately 2.5 times the baseline performance when using InfiniBand in a $4 \times 4$ configuration. In addition, scalability for every other benchmark was better when using InfiniBand as the interconnect as compared to Gigabit Ethernet. An interesting observation is that OOCORE exhibited superscalar performance on the $4 \times 4$ configuration. The problem size for all node configurations was the same, and hence the amount of data written to disk decreased significantly when run on four nodes. Thus, OOCORE benefited not only from improved interconnect performance but also from the decreased local I/O traffic on each node, leading to superscalar performance improvement.

As Figure 4 shows, the benchmarks scaled better on a cluster with InfiniBand fabric than on one using Gigabit Ethernet. Moreover, as Figure 5 shows, the actual performance of those benchmarks for the same number of nodes was also better when run using InfiniBand. This figure compares the performance of the FLUENT (large and medium), HPL, OOCORE, and MILC benchmarks for the $4 \times 4$ configuration running over InfiniBand relative to the results for the same problem size running over Gigabit Ethernet. The results for the FLUENT (large) benchmark showed a slight degradation. As stated earlier in this section, this is a compute-intensive benchmark, and hence the performance benefit from using a high-bandwidth, low-latency interconnect is not readily apparent. All the other benchmarks experienced a performance improvement with InfiniBand as the cluster fabric, with OOCORE and MILC showing the largest improvement. When communication-sensitive benchmarks run in a cluster larger than this $4 \times 4$ configuration, the benefits of InfiniBand over Gigabit Ethernet should be more apparent.

**Changing landscape for HPC**

For decades, the focus when choosing an HPC cluster has been on performance and occasionally price/performance. However, the primary focus for enterprise IT organizations has become a substantial reduction in HPC system power consumption to provide the best performance/watt possible. As is evident from the tests described in this article, the Intel Xeon 51xx processors employing the Intel Core microarchitecture are designed to enable significant power savings while providing a higher level of performance compared to previous-generation processors. The enhanced scalability afforded by the Intel Core microarchitecture and the increase in the number of available processors per server necessitate using high-bandwidth, low-latency interconnects such as InfiniBand for communication-sensitive applications.

Rizwan Ali is a member of the Scalable Systems Group at Dell. His current research interests include performance benchmarking, cluster architecture, parallel applications, and high-speed interconnects. He has a B.S. in Electrical Engineering from the University of Minnesota.

Baris Guler is an HPC application specialist in the Scalable Systems Group at Dell. His current research interests are parallel processing, diskless HPC clusters, performance benchmarking, reservoir engineering and simulation, and numerical methods. Baris has a B.S. in Petroleum and Natural Gas Engineering (PNGE) from the Middle East Technical University in Turkey and an M.S. in PNGE from Pennsylvania State University.

Vishvesh Sahasrabudhe is a member of the Scalable Systems Group at Dell. His current research interests include high-speed interconnects and performance benchmarks. He has a B.Tech. in Electrical Engineering from the Indian Institute of Technology in Bombay and an M.S. in Computer Science and Engineering from the Ohio State University.