EVALUATING THIRD-GENERATION AMD OPTERON PROCESSORS FOR HPC WORKLOADS

By Jacob Liberman
Mausmi Kotecha
Vishvesh Sahasrabudhe

Third-generation AMD Opteron™ processors are designed to optimize multi-threaded application performance, and include multiple architectural enhancements over second-generation AMD Opteron processors. To help organizations understand the performance increases possible when upgrading to third-generation processors, Dell engineers performed benchmark tests against a variety of high-performance computing (HPC) workloads.

Related Categories:

AMD

AMD Opteron processors

High-performance computing (HPC)

Performance

Visit DELL.COM/PowerSolutions for the complete category index.

AMD Opteron processors have enjoyed considerable success in high-performance computing (HPC) during the last several years. According to the TOP500 list—one indicator of HPC trends—AMD Opteron processors at one point powered 23 percent of the world's fastest supercomputers.[1] An innovative dual-core design enhanced the rapid adoption of these processors: by integrating the memory controller directly on the chip, second-generation AMD Opteron processors bypassed the memory contention inherent to systems using multi-core processors based on legacy technology. Additionally, high-speed bidirectional HyperTransport™ links provided a scalable bandwidth interconnect between the computing cores and the I/O subsystem. These enhancements were collectively known as the Direct Connect Architecture. The balanced processing capability and memory bandwidth of second-generation AMD Opteron processors translated into increased efficiency and scalability for HPC applications.

Third-generation AMD Opteron processors include key architectural enhancements designed to boost performance and energy efficiency over the second-generation processors without sacrificing memory bandwidth. These enhancements include the following:

- An increased number of processor cores per socket (four instead of two) designed to operate without an increase in power envelope
- An innovative cache structure that includes a 512 KB dedicated level 2 (L2) cache per core and a 2 MB shared L3 cache
- 128-bit Streaming SIMD (single instruction, multiple data) Extensions (SSE) execution width that allows four floating-point operations per clock cycle
- 128-bit integrated memory controller divided into two independent 64-bit channels
- Reduced power consumption through multiple-core voltage and frequency scaling[2]

Figure 1 illustrates the architectural differences between second- and third-generation AMD Opteron processors. Ninth- and tenth-generation Dell™ PowerEdge™ servers support both second- and

[1] As listed in the November 2006 TOP500 list. For more information, visit www.top500.org.
[2] For a detailed description of third-generation AMD Opteron processors, visit www.amd.com/us-en/Processors/ProductInformation/0,,30_118_8796_15223,00.html.

third-generation AMD Opteron processors, which are both socket-F compatible, helping simplify upgrades for HPC systems.

To demonstrate the performance increases made possible by upgrading from second-generation to third-generation AMD Opteron processors, in July 2008 Dell engineers compared the performance of each type of processor in an HPC test environment. The results indicate that third-generation AMD Opteron processors can provide significant performance increases across a broad spectrum of HPC workloads.

TEST ENVIRONMENT

The test team carried out two rounds of testing. In the first round, they compared the performance of second- and third-generation AMD Opteron processors using a single Dell PowerEdge M605 blade server, an energy-efficient two-socket server that supports up to 64 GB of double data rate 2 (DDR2) RAM. In the second round, the test team compared processor performance across a 16-node cluster of PowerEdge M605 blade servers. Figure 2 shows the hardware and software used in the test environment for each round of testing.

SINGLE-SERVER TEST RESULTS

The single-server tests compared the performance of second- and third-generation AMD Opteron processors in a single Dell PowerEdge M605 blade server using the following synthetic and application benchmarks:

- **Double Precision General Matrix Multiply (DGEMM):** A Basic Linear Algebra Subprograms (BLAS) subroutine from the AMD Core Math Library (ACML) that performs matrix multiplication in double precision
- **Basic Local Alignment Search Tool (BLAST):** A tool that provides algorithms for rapidly searching nucleotide and protein databases
- **STREAM:** A synthetic benchmark that measures sustainable memory bandwidth across four threaded operations: COPY, SCALE, ADD, and TRIAD
- **ANSYS Benchmarks:** A popular structural dynamics application that evaluates performance across codes with varying data access patterns and problem sizes using benchmark data sets[3]

The test team ran each benchmark three times using the maximum number of threads or processes for each processor—four for the second-generation dual-core processors and eight for the third-generation quad-core processors. The test team then took the mean score for each processor type and normalized the results to show performance relative to the second-generation processors.

Figure 3 shows the results. As this figure indicates, in most cases the third-generation AMD Opteron processors provided significantly increased performance over the second-generation processors. The DGEMM benchmark resulted in the greatest performance increase, exceeding the baseline performance by

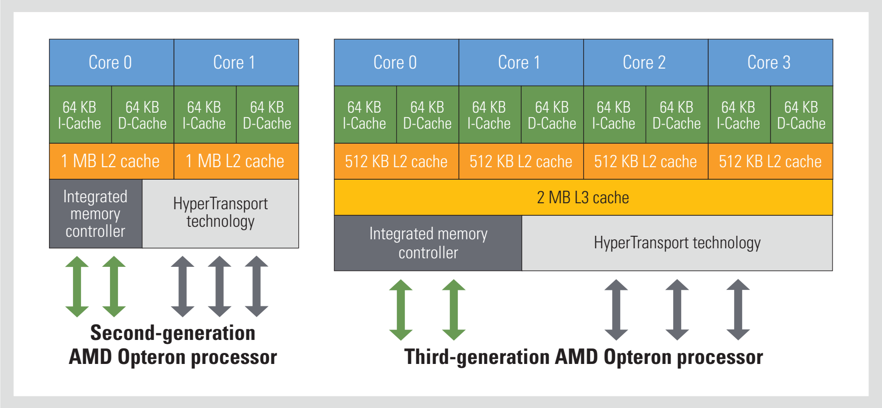

Figure 1. *Second- and third-generation AMD Opteron processor architectures*

	Single-server tests	**Cluster tests**
Servers	1 Dell PowerEdge M605 blade server	16 Dell PowerEdge M605 blade servers
Processors	▪ Two second-generation dual-core AMD Opteron 2218 processors at 2.6 GHz ▪ Two third-generation quad-core AMD Opteron 2356 processors at 2.3 GHz	
Memory	Eight 2 GB DDR2 dual in-line memory modules (DIMMs) at 667 MHz	
Disk	One 36 GB, 15,000 rpm Serial Attached SCSI (SAS) drive	
OS	Red Hat® Enterprise Linux® 5 Update 1 OS	
Compilers	GCC 4.2.1 and PGI 7.2.1	
Message Passing Interface (MPI)	N/A	Open MPI 1.2.5
Interconnect	N/A	Mellanox MT25408 ConnectX DDR InfiniBand with the OpenFabrics Enterprise Distribution (OFED) 1.3 software stack

Figure 2. *Single-server and cluster benchmark configurations*

[3] For more information on these benchmarks, visit developer.amd.com/cpu/libraries/acml/pages/default.aspx, www.ncbi.nlm.nih.gov/Education/BLASTinfo/information3.html, www.cs.virginia.edu/stream, and www.ansys.com/special/news-images/high_per_computing.htm.

nearly 250 percent. This increase over the second-generation processors was a result of the additional cores and wider SSE register, which allowed each core to complete four floating-point operations per clock cycle. Memory bandwidth, as measured by the STREAM benchmark, doubled in the case of the SCALE operation and increased by narrower margins across the COPY, ADD, and TRIAD operations. Taken together, the results indicate that the third-generation processors' expanded floating-point capability did not come at the expense of the memory subsystem.

The third-generation processors delivered substantial performance increases while running the synthetic DGEMM and STREAM benchmarks; however, because these benchmarks represent idealized workloads, they often fail to accurately predict application performance. The Dell team therefore performed additional tests using authentic applications. In these tests, changes in application performance across processor generations varied according to the application characteristics. Four of the seven ANSYS benchmarks showed performance increases in the range of 30–40 percent. These results were favorable, because the ANSYS benchmark suites do not typically scale well. The ANSYS benchmarks with larger problem sizes—including bmd-4 and bmd-5—did not fare as well. The bmd-5 benchmark showed some performance increase; bmd-4 showed slight performance degradation. And although bmd-5 can make effective use of the processor cache, which generally results in good parallel performance, a larger problem size does not decompose into cacheable elements, which in bmd-4 increases the amount of I/O relative to computation. This increased I/O had the overall effect of minimizing performance gains related to the processing subsystem. BLAST, which also includes a substantial I/O component, showed similar results.

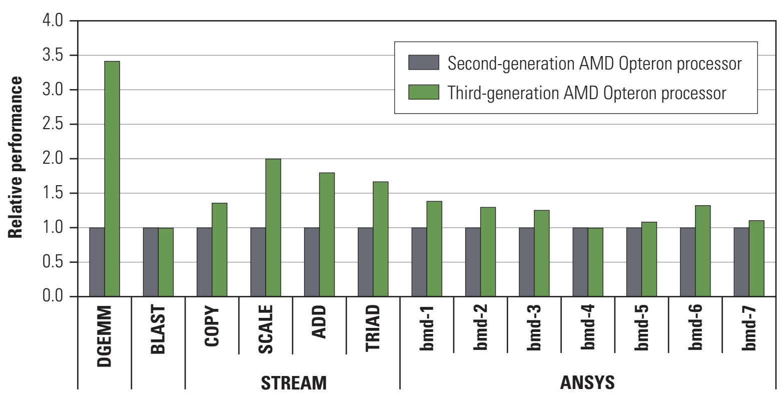

Figure 3. *Relative single-server performance results for each AMD Opteron processor*

CLUSTER TEST RESULTS

Although the performance of the third-generation AMD Opteron processors exceeded that of the second-generation processors across most of the single-node benchmarks, single-server performance enhancements do not always translate to the cluster level. Certain applications may not scale well beyond a single node because of their lack of inherent parallelism or overhead introduced by communication. To evaluate performance at the cluster level, the Dell team used the following synthetic and application benchmarks:

- **FLUENT:** A popular commercial computational fluid dynamics application for simulating fluid flow and heat transfer; the Dell team used three FLUENT benchmark data sets of increasing size
- **High-Performance Linpack (HPL):** A highly configurable and floating-point-intensive benchmark used to rank the world's fastest computers; HPL solves a dense linear system in double precision and represents an extreme HPC workload because it involves very little communication and I/O
- **NASA Advanced Supercomputing (NAS) Parallel Benchmarks (NPB):** An application-centric set of benchmarks that tests the performance of multiprocessor computers; the Dell team ran the Class C Embarrassingly Parallel (EP), Fast Fourier Transform (FT), Integer Sort (IS), Lower-Upper Symmetric Gauss-Seidel (LU), and Multigrid (MG) applications from the NPB suite
- **Weather Research and Forecasting (WRF):** A numerical weather prediction system[4]

Figure 4 compares the measured performance of the second- and third-generation AMD Opteron processors running benchmarks across the cluster, with the average of the three benchmark results again normalized relative to the performance of the second-generation processors. The FLUENT results in Figure 4 show a positive correlation between relative performance gains for third-generation AMD Opteron processors and increasing problem size. The larger benchmark data sets—truck and truck_poly, at 13 million cells

[4] For more information on these benchmarks, visit www.fluent.com/software/fluent/fl6bench, www.netlib.org/benchmark/hpl, www.nas.nasa.gov/Resources/Software/npb.html, and www.mmm.ucar.edu/wrf/users/download/get_source.html.

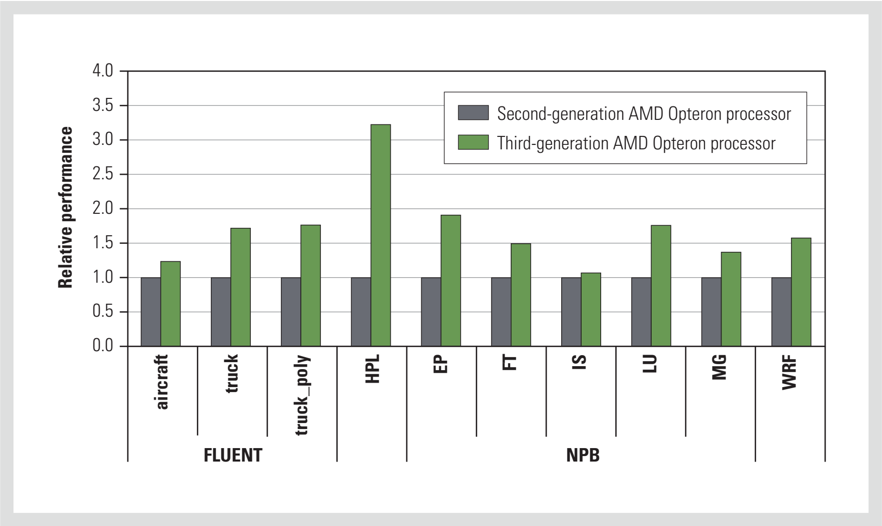

Figure 4. *Relative cluster performance results for each AMD Opteron processor*

each—showed performance increases of 71 percent and 74 percent, respectively. The smaller data set—aircraft, at 1.8 million cells—showed a performance increase of only 22 percent. These results indicate that the cluster was able to effectively utilize the additional cores of the third-generation processors, providing increased performance as the size of the data set grew.

HPL showed the largest performance increase of the benchmarks used. With a large problem size and a data distribution scheme that can make effective use of the processor cache, HPL is optimized to exploit both the additional cores and the increased floating-point operations per clock cycle, which is reflected in a more than threefold increase in HPL performance. The results also suggest that the DDR InfiniBand network was able to handle the increased traffic to provide near-optimal scaling.

The five NPB results show performance increases across a series of tests with wide variation in inter-node communication, latency, and floating-point requirements. The EP benchmark showed the largest increase because of a lack of communication between nodes combined with the increased number of cores. The IS benchmark, which involves intensive global communication with mixed message sizes, showed the smallest increase. This benchmark is communication bound, providing a counterexample to the EP results, and does not take advantage of the additional SSE width because integer sorting does not involve floating-point calculations. The remaining NPB benchmarks, which feature balanced communication and floating-point calculation, showed increases in the range of 40–70 percent.

WRF showed a 63 percent performance increase with a constant problem size. Although both HPL and WRF are floating-point intensive, WRF fell short of the increase when using HPL. This disparity results from the larger memory footprint and I/O overhead of WRF compared with HPL.

SCALABLE PERFORMANCE FOR HPC WORKLOADS

Dell clusters equipped with third-generation AMD Opteron processors can offer substantial performance increases over those equipped with previous-generation processors. In the Dell tests, the third-generation processors demonstrated increased performance across a broad spectrum of HPC workloads with varying I/O and communication patterns, with memory-bound and floating-point-intensive application performance improving by more than 80 percent in most cases. Because 9th- and 10th-generation Dell PowerEdge servers support both processor generations within the same power envelope, organizations have a clear upgrade path that can provide tangible performance gains in their environments.

Jacob Liberman is a development engineer in the Dell Scalable Systems Group. He has a master's degree in Instructional Technology from the University of Texas at Austin and holds multiple industry certifications.

Mausmi Kotecha is an engineering analyst in the Enterprise Solutions Group at the Dell Bangalore Development Center. Her areas of interest include HPC clustering packages, performance analysis of parallel applications, and cluster file systems. She has a bachelor's degree in Computer Science from Atmiya Institute of Technology and Science.

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 and an M.S. in Computer Science and Engineering from the Ohio State University.

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell HPC cluster solutions:
DELL.COM/HPCC

Dell PowerEdge blade servers:
DELL.COM/Blades

AMD Opteron processors:
www.amd.com/opteron