THE CLIPPER GROUP

Navigator™

Navigating Information Technology Horizons℠

Published Since 1996 Report #TCG2004054 June 28, 2004

High Performance Computing and AMD Opteron — Marriage or Flirtation?

Analyst: David Reine

Management Summary

Mankind rarely achieves success in any venture because of one single quality or condition. We usually realize success because of a careful blending of two or more important factors. In sports, for example, we usually see the necessity of having a good offense in support of a good defense in order to win a championship. Whether it is baseball, football, basketball or any other sport, it requires teamwork. You need to be able to both score points when you are on offense and stop the other team from scoring when you are on defense. The skills necessary are often quite different. The manager of a professional team needs to be able to put the right mix of athletes on the field. He needs players who have the capability to perform at a high level in both capacities, or, at least, excel in one while being at least adequate in the other.

These same qualities are required to put together a high-performance computing (HPC) environment. **The system developer looking to maximize throughput in a computing system needs to achieve a precise balance between processor speed, I/O throughput, memory bandwidth and operating environment.** Throughout the 1990's, we saw an imbalance created. CPU performance adhered to *Moore's Law* with processor performance doubling every 18 to 24 months. Based upon RISC architectures such as IBM's *PowerPC*, HP's *PA-RISC*, and Sun's *SPARC*, with their respective *UNIX* implementations, CPU capability far outstripped the improvements made to memory and I/O subsystems. Calculations could be processed faster than the data could be moved into or out of the CPU. Thus, **improvements made to the CPU may have had little or no effect in improving the overall performance of the system**. In order to achieve an overall improvement, a new paradigm for microprocessor development had to evolve and then, system manufacturers would have to adapt that technology to the fabrication of newer, faster computer systems.

With AMD's introduction of its *Opteron* microprocessor, we might just have that paradigm shift that we have been waiting for. All of the leading RISC manufacturers are now dipping their toes into the vast ocean of x86 commodity processing. They are doing this to produce a better HPC node in a scale-out architecture, while continuing to improve their legacy scale-up solutions. Please read on to see what HP, IBM and Sun are doing to integrate the best possible HPC engine into the best possible HPC environment.

IN THIS ISSUE

HPC in the Data Center

Over the past decade, some of the most common solutions found in data centers for scientific and technical computing problems have been focused upon three proprietary 64-bit architectures: Hewlett-Packard's *PA-RISC*, IBM's *POWER*, and Sun's *SPARC*. These are all highly mature architectures tuned to perform very well in traditional scale-up environments such as the HP *Superdome*, the IBM *eServer p690*, and the Sun *e25K*, all operating under a variation of the UNIX operating system. Over this same period, the CPUs intrinsic to each of these symmetrical multiprocessing (SMP) systems have experienced significant improvements in processing capability, often playing leapfrog with each other in the commodity benchmarks that dot the landscape. Unfortunately, the other components critical to overall system performance have not kept pace with the advancements in microprocessor design. Specifically, I refer to the blending, or communications, between the CPU and memory, the CPU and I/O.

As in nature, computers require the correct mixture of elements in just the right quantities. In nature, we see an example of that precise blending by combining 2 parts of Hydrogen and 1 part of Oxygen to form water. Without the exact mixture, you have nothing but gas. **An HPC design, typically a scale-out multi-node environment, similarly, must contain sufficient memory and I/O access to resolve high-level technical and scientific workloads.** The overall HPC environment relies upon being able to transfer information to and from memory and I/O devices at a rate that will maintain a steady stream of data for the processing unit. This eliminates any wait-states that simply put the CPU into idle mode. HPC requires a high level of integration within the microprocessor to achieve the desired performance levels. This does not exist currently in the proprietary architectures.

The move by HP from PA-RISC to the 64-bit Intel *Itanium-2* microprocessor in the *Integrity* environment[1] succeeded in transitioning their product line from a proprietary architecture to one with wider commercial availability. Itanium-2, however, did not resolve the long-standing discrepancy between the processor compute and memory access capabilities that existed within PA-RISC. It did initiate a domino effect in the industry, however, with other vendors following HP's lead to an open 64-bit environment with the introduction of other scale-up Itanium-2 servers, notably, Groupe Bull's *NovaScale*[2] and NEC's *Express5800*, among others.

Itanium-2, however, does not resolve the processor-to-memory bandwidth issue. Designed with a front-side bus, multiple Itanium-2 processors share a common access to the memory controller chip set, and thus, a common access to memory. This works fine as long as there is only one request at a time. When a second request occurs, one CPU has to wait.

While Intel and HP were putting together their Itanium strategy, AMD was developing their own 64-bit microprocessor. Unlike Itanium, however, the **AMD *Opteron* processor uses a direct connect architecture, instead of a front-side bus, to increase application performance by dramatically reducing memory latency**. IBM and Sun have quickly adopted this new processor for high-performance computing with the introduction of IBM's *eServer 325* and SUN'S *Sun Fire V20z*, used as standalone servers or as nodes in multi-node clusters. Further, in addition to their obvious commitment to Itanium, HP has also introduced two AMD Opteron processor based solutions, the *ProLiant DL145 and ProLiant DL585*, designed for both HPC and general enterprise environments.

AMD Opteron Technology

The AMD Opteron processor, based on AMD64 technology, has made a bigger splash in a bigger pool than was initially expected. It allows users to maintain their investment in their existing installed base of 32-bit x86 applications, running with excellent 32-bit performance, and provides a transition path to a 64-bit environment. With a larger Level-1

[1] See *HP takes First Step Toward Product Line Consolidation* in **The Clipper Group Navigator** dated July 29, 2003, and available at http://www.clipper.com/research/TCG2003034.pdf.

[2] See *Groupe Bull Mans the High Performance Point While Leading Migration Charge to Itanium* in **The Clipper Group Navigator** dated August 28, 2003, and available at http://www.clipper.com/research/TCG2003040.pdf.

Exhibit 1 –
Additional AMD Opteron Features

- 64-bit wide key data and address paths that incorporate a 48-bit virtual address space and a 40-bit physical address space;
- ECC protection for L1 cache data, L2 cache data and tags;
- .13 micron Silicon on Insulator process technology for lower thermal output levels and improved frequency scaling;
- Support for all instructions necessary to be fully compatible with SSE2 technology; and
- Two more pipeline stages for increased performance and frequency scalability.

and Level-2 cache[3], AMD Opteron has made inroads in the 64-bit commercial market that was initially expected to belong to Itanium. In fact, AMD Opteron now provides an alternative to Intel's Xeon, Xeon MP, and Itanium in the server market due to its 32- and 63-bit computing capabilities. Published figures, in fact, show that, between them, there were over 37,000 Itanium and AMD Opteron servers sold in 1Q04, with AMD Opteron supporting the lion's share with about 80% of that total. **To some, this may appear to be a niche market, but there is nothing wrong with that if the niche is large enough.**

Moreover, the design of this AMD chip makes it a serious competitor in the high-performance computing space, dealing with floating-point and memory-intensive workloads. IBM was the initial major adopter, filling that void in their product set with e325 nodes with ccNUMA[4] architecture in an *eServer Cluster 1350.* The 1350 uses Linux with robust systems management, PathScale Compilers designed specifically for the AMD Opteron processor, and services to deliver highly scalable, integrated cluster offerings. Sun and HP soon followed suit with the Sun Fire V20z, the ProLiant DL145, and the ProLiant DL585. The significant architectural features that make AMD Opteron suitable for a high-performance environment are:

- **Direct Connect Architecture** – Memory is directly connected to the CPU optimizing memory performance. I/O is directly connected to the CPU for more balanced throughput. CPUs are connected directly to each other allowing for more linear symmetrical processing.
- **Integrated DDR DRAM Memory Controller** – The memory controller is integrated into the microprocessor changing the way that the processor accesses memory, resulting in increased bandwidth, reduced memory latency, and increased CPU performance.
- **HyperTransport Technology** – Supports up to three coherent HyperTransport links providing up to 19.2 GB/s peak bandwidth per processor.

These features complement the improved AMD Opteron processor performance to ensure rapid movement between CPU and memory as well as between CPUs. Additional AMD Opteron features are enumerated in Exhibit 1. It is also important to note that AMD manufactures multiple versions of the processor specific to the number of processors planned for the server. The AMD Opteron 200 series is for up to 2-way servers and the AMD Opteron 800 series is for 4-8 way servers.

Regardless of the technological advancements to improve performance, no commodity processor can become successful without mass acceptance and distribution. To achieve that status, a server must have superior price/performance. Dual-processor AMD Opteron servers, with expandability to 4-way, are priced between $10K - $12K, while Xeon servers with the same configuration are priced around $16K. Itanium-2 2-way servers are significantly higher, priced > $20K.

Open Systems Performance Results

The only true benchmark to evaluate system performance is the actual application. Unfortunately, the cost to establish an application benchmark for all systems under consideration would be prohibitive for all but the largest companies. In order to have a yardstick to rate system performance, a number of

[3] Opteron has a 128KB Level-1 cache and a 1MB Level-2 cache compared to 32KB L1 and 256KB L2 for Itanium-2.
[4] Cache Coherent Non-Uniform Memory Access

Exhibit 2 – Relative SPEC Positioning*

	Cint	Cfp	Cint_rate – Mono	Cint_rate - Quad	Cfp_rate - Mono	Cfp_rate - Quad
AMD Opteron x50	1655	1694	19.2	68.5	19.1	52.5
Itanium-2	1404	2161	16.0	64.2	25.1	82.7
Xeon	1563	1350	18.6	61.6	15.7	32.4
UltraSPARC III	722	1344	8.38	31.7	13.0	45.5
POWER4	1069	1642		48.4		66.6

* These numbers represent the highest value for a server based upon that technology, regardless of configuration.

"standard" benchmarks have been established by organizations such as the Standard Performance Evaluation Corporation (SPEC) and the Transaction Processing Performance Council. In addition, companies such as SAP have established benchmark suites through which disparate systems can be measured against one another using a standard script.

Because technology is evolving so quickly, SPEC continually upgrades their computer benchmarks. SPEC *CPU2000* is the next-generation industry-standardized CPU-intensive benchmark suite. SPEC designed CPU2000 to provide a comparative measure of compute intensive performance for integer and floating point calculations across the widest practical range of hardware. The implementation resulted in source code benchmarks developed from real user applications. These benchmarks measure the performance of the processor, memory and compiler on the tested system. Unfortunately, it is often difficult to get a true indication of system capability because the configuration of the tested system is not always the same as the proposed configuration. In addition, vendors are forever playing games with the configurations, changing factors such as the amount of memory, the size of cache, and the number of disks.

All of that considered, this bulletin tries to evaluate the latest results of several standard benchmarks that are available on the Internet for our edification. Basing our examination on processor technology, we look at the results for systems from companies such as HP, IBM, and Sun for their commodity and proprietary microprocessors: AMD Opteron, Intel Xeon, Intel Itanium-2, IBM POWER4[5], and Sun UltraSPARC[6].

Using the raw numbers from SPEC CPU2000, (see Exhibit 2 above), we can see that the AMD Opteron x50 processor has a clear lead in integer math while the Itanium-2 is in front with floating point calculations. The *Cint* and *Cfp* measurements define CPU speed, while the *Cint_rate* and *Cfp_rate* figures identify throughput. This is where data center management has to factor in the instruction make-up of the application and the price of the system. Does the application center on integer math? Then the Cint and Cint_rate calculations need to be weighted. If the application involves a significant number of floating-point calculations, then the floating-point measurements need to be considered. Even more, management will review the cost of the systems. As described above, the AMD Opteron nodes tend to be priced lower than Xeon nodes and significantly lower than Itanium nodes. In some instances, the data center could acquire two AMD Opteron nodes for the price of a single Itanium node. In this case, the effective floating-point performance of the two AMD Opteron nodes would exceed that of one Itanium node. The data center staff needs to review the price/performance of each node, not merely the tabular result.

In transaction performance, HP's implementation of Itanium-2, the *rx5670*, is about 12% faster than the *DL585* (AMD Opteron), when comparing quad-processors. However, the DL585, as tested, is 28% less expensive than the rx5670. Based on price/ performance, the edge would have to go to AMD Opteron. It is also worth noting that using Linux, instead

[5] The only quad-processor POWER4 system listed in SPEC is the pSeries 655

[6] Please note that we have listed the UltraSPARC III processor from Sun, and not the UltraSPARC IV. Sun has not yet submitted any UltraSPARC IV results to SPEC.

Exhibit 3 - *LS-DYNA TopCrunch* Results*

	Quad		Dual	
	Itanium-2	**AMD Opteron**	**Itanium-2**	**AMD Opteron**
N-R	4040	4696	8172	8791
3-V	100,938	65,275	197,422	124,842

* All values are in seconds

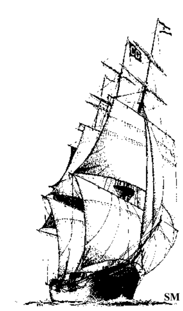
of Windows Server 2003, generated an improvement in performance of more than 10% with a reduction in price of almost 14%.

When we look at application specific benchmarks, we can see the continuation of this performance trend, with Itanium-2 and AMD Opteron leading the way in the SAP SD standard R/3 Enterprise 4.70 application benchmark for quad-processors. In fact, the HP Integrity rx4640-8 (Itanium-2) measures at 266,000-dialog steps/hour, compared to the HP Proliant DL585 (AMD Opteron) which measures at 232,000 dialogs/hour. On that score, you would note that the Itanium-2 system was 12% faster. However, when you peel back the covers and look at the configuration, you realize that the rx4640 was configured with 32GB of main memory and 1620GB of disk. The DL585 used 12GB of memory and only 311GB of disk. The price for the configurations as tested was not listed in the benchmark report, however, it is clear that the configured rx4640 costs significantly more than the DL585. In addition, the rx4640 used Oracle 9i as an RDBMS while the DL585 used Microsoft SQL Server 2003. Combining all of these factors, the price/performance decision would again seem to favor AMD Opteron.

Livermore Software Technology Corp provides another independent view of performance. Their suite, *TopCrunch*, tracks the aggregate performance trends of HPC systems and engineering software. Instead of using a synthetic benchmark, they use actual engineering software applications with real data and run them on HPC systems using SuSE Linux Professional for the AMD Opteron systems and HP-UX 11.23 for the rx2600 Itanium-2 systems. There are two applications. The first, Neon Refined (N-R), simulates a frontal vehicle crash with initial speed at 31.5 miles/hour for a 1996 Plymouth Neon. As shown in Exhibit 3, Itanium-2 has a slight advantage of 7–14% in this test. The second test, 3-Vehicle Collision (3-V), simulates a van crashing into the rear of a compact car, which, in turn, crashes into a midsize car. As is shown, AMD Opteron has a distinct advantage of about 60% for both the dual and quad configurations. This benchmark deals in pure performance. It does not factor in price/performance.

Conclusion

The best way to evaluate HPC systems is to run your application on each, using the same configuration, the same O/S, the same compilers, etc. When that option is too expensive, the data center staff can fall back on the specific benchmarks deemed appropriate for HPC and compare the results as we have done here. While it is impossible to forecast results in all environments, it is evident that the AMD Opteron CPU provides a consistent "bang" for the "bucks" you need to spend to acquire the best HPC system available. How the chip is packaged is equally important. Check out the operating system and compilers as well. Everything has to be in tune, optimized together to achieve the kind of performance only attained through careful integration of the various component pieces. If this "dating" period is successful, the data center staff and their AMD Opteron system will have a long, performant life together.



About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

David Reine **is Director, Enterprise Systems for The Clipper Group.** Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. He joined The Clipper Group after three decades in server and storage product marketing and program management for Groupe Bull, Zenith Data Systems, and Honeywell Information Systems. Mr. Reine earned a Bachelor of Arts degree from Tufts University, and an MBA from Northeastern University.

- ***Reach David Reine via e-mail at dave.reine@clipper.com or at 781-235-0085 Ext. 23. (Please dial "1-23" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** ***Captain's Log***, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.