Solutions Brief

MCAE Application Productivity with
SGI® Altix™ 3000 Technology

SGI® Altix™ 3700 Supercluster
Scalable to hundreds of processors, with up to 64-processors and 512GB memory per node

SGI® Altix™ 3300 Server
4, 8, or 12 processors, up to 96GB memory

The combined forces of mechanical computer-aided engineering (MCAE) application software and HPC system technology provide engineers with an increasingly competitive advantage in today's global manufacturing market of product development. Manufacturers and suppliers in automotive, aerospace, and a variety of general manufacturing sectors benefit from MCAE applications that enable reduced design-cycle time and costs and overall improvements in design quality.

This article examines the MCAE productivity benefits of a new HPC system technology, the SGI Altix 3000 family of servers and superclusters, which was developed to advance the current capabilities of technical HPC. The design of SGI Altix 3000 combines the open-source 64-bit Linux® operating system, the Intel® Itanium® 2 microprocessor, and the SGI® NUMAflex™ shared-memory system architecture design. With this new server technology introduction, SGI offers the MCAE community a dual-platform roadmap based on UNIX® and Linux.

SGI Altix 3000 Introduction

Recent developments in HPC technology continue to rapidly advance the MCAE simulation capabilities of engineers across all disciplines. SGI is an established and leading supplier of HPC technology for MCAE simulation with the company's SGI® Origin® server family, based on the company's proprietary MIPS® microprocessors and IRIX® operating system. On January 7, 2003, SGI expanded its HPC offering with the introduction of SGI Altix 3000, the industry's first global shared-memory cluster that combines SGI NUMAflex supercomputing architecture with Intel Itanium 2 processors and the 64-bit Linux operating system.

The SGI® NUMA (nonuniform memory access) architecture was introduced in the SGI® Origin® 2000 server in 1995 and later advanced with the SGI NUMAflex modular design concept of the SGI® Origin® 3000 servers. This is the same high-bandwidth and low-latency NUMA architecture that is available in SGI Altix 3000, yet with a significant cost-performance advantage for MCAE applications. This advantage is achieved by, among other factors, contribution to and leverage of investments in Linux by the open-source community, and the Itanium 2 microprocessor roadmap from Intel.

SGI Altix 3000 is recognized by the Linux community as the first Linux cluster that scales to 64 processors and 512GB of shared memory within each single Linux OS-image node and the first cluster of any variety to allow global shared-memory access across nodes. These 64-processor single nodes can be clustered with a choice of scalable interconnect networks, including the proprietary SGI® NUMAlink™ interconnect technology, to much larger system configurations—up to 2,048 processors and a total of 16TB of memory—that make up an SGI Altix 3000 supercluster. The high-bandwidth SGI NUMAlink interconnect fabric of SGI Altix 3000 superclusters delivers information between cluster nodes up to 200 times faster than conventional clustering switches.

SGI Altix 3000 is binary compatible with the industry-standard 64-bit Linux distribution, currently based on the 2.4.19 kernel. In addition, SGI offers differentiated middleware and other functionality to enhance demanding HPC workloads in a bundle called SGI ProPack™. SGI ProPack is a set of user tools that ride on top of Linux and is similar to other commercial software packages. SGI ProPack is used to boost the performance of Linux and user applications on the Altix 3000, not to alter Linux itself. Such enhancements are commonplace in the industry and are also offered by other Linux system providers.

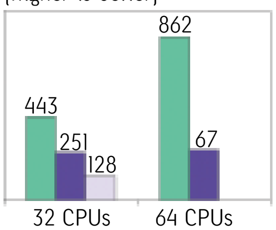
The performance of the SGI Altix 3000 family has set new records with many industry-standard benchmarks. In particular, the SPECfp®_rate_base2000 benchmark, a measure of a systems compute performance, and STREAM Triad benchmark, a measure of memory bandwidth performance, show SGI Altix 3000 easily outperforming all of the top-end UNIX OS-based systems that are popular with MCAE applications today.

MCAE Software and Industry Practice

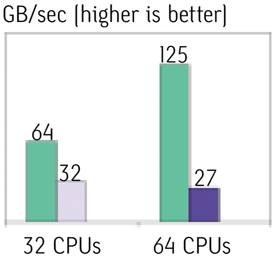
Rapid progress in MCAE simulation performance has been influenced by advanced developments in both application software algorithms and HPC hardware systems. From a software algorithm and hardware perspective, there are three MCAE disciplines to consider for their requirement on HPC resources: implicit and explicit finite element analyses (FEA) for structural analysis and computational fluid dynamics (CFD) for fluid flow simulation.

The three MCAE disciplines exhibit a range of HPC resource demands and highlight the importance of a balanced HPC system architecture. The features desired most of a balanced system include (1) high-speed processors with large cache; (2) large addressable memory; (3) high memory-to-processor bandwidth

SPECfp_rate_base2000 benchmark scores (higher is better)



STREAM Triad Benchmark results in GB/sec (higher is better)



- SGI Altix 3000
- HP Superdome™
- IBM® pSeries™ p690

rates; (4) high disk-to-memory I/O rates; and (5) a low-latency interconnect that provides efficient parallel scalability to hundreds of processors.

By far the most important HPC advancement in recent years for MCAE applications is the parallel scalability of MCAE software. Most commercial MCAE software employs a distributed-memory parallel (DMP) technique for compatibility across the range of available HPC architectures. Other techniques include shared-memory parallel (SMP) and hybrid parallel schemes that take advantage of both DMP and SMP within a single computation.

While each discipline has inherent complexities with regard to efficient parallel scaling depending upon the particular parallel scheme and HPC system architecture, CFD can scale efficiently to hundreds of CPUs, explicit FEA can scale to more than 50 CPUs, and implicit FEA can scale to up to 10 CPUs.

Single-job turnaround of MCAE simulations on SGI Altix 3000 has been impressive, with many ISV applications showing a performance level that exceeds high-end UNIX systems. For some MCAE applications, Altix 3000 demonstrates efficient parallel scalability as high as 64 CPUs. This stand-alone job performance is critical for success of any new server today, but true commercial success requires the additional capability of job throughput, which more closely captures the industry's true MCAE practice. Industry practice most often combines the use of moderate single-job scalability with multi-job throughput. That is, a single job typically uses 12 CPUs on average, in a mix that combines several disparate jobs that require a throughput level of productivity.

Throughput is the domain and responsibility of hardware vendors, and SGI leads the industry in this critical HPC industry requirement. The NUMAflex system architecture was designed for a combination of single-job turnaround and multi-job throughput. Based on this architecture, SGI Altix 3000 offers MCAE simulations a high-availability, nondegrading, and efficient application environment to ensure that turnaround and throughput are delivered in support of hundreds of simultaneous users with a demanding mix of disciplines.

Multi-job throughput has been demonstrated on SGI Altix 3000 with a range of MCAE commercial software applications and industrial-sized customer jobs that exhibit an average throughput degradation of 10% compared with their turnaround times. This means that a single job, on average, requires just 10% more time to complete when that job is included in the mix of jobs.

SGI Altix 3000 Technology Benefits for MCAE

SGI continues to invest in the MCAE community's shared vision with the introduction of SGI Altix 3000, an HPC server that will enable further advancements for MCAE applications. SGI Altix 3000 delivers the existing application advantages of NUMA with complete 64-bit Linux compatibility, but in a shared memory capability that is not available in conventional clusters.

Unlike conventional clusters, the SGI Altix 3000 scalable memory architecture can conduct memory-resident MCAE analyses with the data as a whole entity, without breaking the data into smaller partitions to be handled by individual processors. Consequently, a programmer or user does not need to spend time developing rules to divide the data into smaller sets, as is required with domain decomposition techniques for distributed parallel MCAE.

The primary reason for efficient multi-job throughput capability with SGI Altix 3000 is the latest Linux kernel scheduler that performs well with large memory, an advantage SGI Altix 3000 has over commodity clusters. The trend of cheap and abundant commodity DRAM means this advantage will grow over time.

Another important advantage is I/O performance, a requirement for throughput of implicit FEA applications that simulate dynamic response of a structure. SGI Altix 3000 offers I/O rates of more than 2GB per second, well beyond the typical Linux barrier of 500MB per second.

MCAE Technology Directions

The design breakthroughs of SGI Altix 3000 offer the industry a shared-memory choice for open-source computing with the cost-effectiveness of conventional clusters. With SGI Altix 3000, the entire suite of MCAE applications can now achieve capability levels that are equivalent to the most mature UNIX platforms, but at less than half the cost. This capability, along with its economic benefits, will further expand the use of MCAE to include a variety of new applications and practices that will become routine for product development.

The capabilities of SGI Altix 3000 create opportunities for increased innovation, such as expanded use of time-dependent and time-accurate MCAE simulations. Additionally, there are opportunities for more production-level coupling of applications for multidisciplinary design optimization. The shared-memory, multipurpose architecture of SGI Altix 3000 is a platform that is well suited to the simultaneous demands of all three MCAE disciplines. SGI Altix 3000 provides the ability to capture more realism in simulations.

The record-setting performance levels of the SGI Altix 3000 systems will be at the forefront of performance because SGI designed the Altix 3000 servers with the ability to upgrade to the future Itanium 2 architectures code-named Madison and Montecito. Perhaps of even greater significance is the fact that SGI has accomplished these technological breakthroughs by working closely with the Linux open-source community and continues to offer much of its work back to the community, to be incorporated into later versions of the Linux kernel. These investments will ensure success of the new SGI Altix 3000 technology and the company's continued commitment to the MCAE community in delivering valuable leadership in HPC advancements for the manufacturing industry.

sgi®

Corporate Office
1600 Amphitheatre Pkwy.
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America 1(800) 800-7441
Latin America (52) 5267-1387
Europe (44) 118.925.75.00
Japan (81) 3.5488.1811
Asia Pacific (65) 6771.0290

3497 (04/14/2003) J14279