Solutions Brief

SGI® High Performance Computing

Large-node clusters

Shattering the limits of application performance

Why larger nodes?

- Breakthrough performance and scalability on applications requiring unpredictable, wide-ranging, and high-intensity access to data
- Improved ROI via ability to run both distributed-memory and shared-memory applications, with enough throughput to run multiple jobs at once
- Simpler and less costly administration due to fewer copies of Linux
- More flexible interconnect topologies with faster local interconnects within nodes and less costly, commodity interconnects between nodes
- Faster data I/O between memory, disks, and other components

The Large-Node Advantage

Linux® operating system (OS) based clusters represent a dominant and ever-growing sector in today's High Performance Computing (HPC) market. This popularity is largely driven by the cost advantage of cluster configurations, coupled with the ability to increase processing power by scaling out available compute resources. Most cluster solutions today are comprised of several small, often dual -processor, nodes networked together via commodity interconnect technologies.

However, another class of cluster solutions is emerging today, and is specifically designed to offer the advantage of scale-out clusters with the ability to address more complex processing needs by scaling CPUs and memory at each node. These "large-node clusters" offer distinct advantages that can lead to breakthrough performance. Compared to small-node "white box" clusters, they provide a more natural fit for applications with a proven a thirst for memory or for applications that are not structured to run across multiple nodes.

Small-node clusters are well suited to running applications that scale effectively over parallel systems and have highly predictable processing requirements. These applications, often referred to as "embarrassingly parallel," execute instructions in parallel, with minimal data sharing across nodes. But many application workloads are more complex than this, consisting of jobs with volatile resource requirements that require accessibility to larger amounts of data to process efficiently. Many of these applications demand far more accessibility to all data (and thus to all memory) than typical small-node, distributed-memory clusters can provide.

Large-node clusters

The ability to efficiently scale processing on the node, coupled with the ability to scale out across nodes, can be a key to achieving breakthrough application performance. By scaling at the node instead of adding new nodes every time computing needs increase, users spend less on server administration and interconnect fabric, with fewer nodes to connect, manage and provision.

Every day, engineers, scientists and researchers around the world realize the large-node advantage. And increasingly, they look to servers from SGI to deliver that advantage.

Getting to Breakthrough Results

By balancing the properties of both capacity and capability computing via powerful large-node clusters, SGI users are driving even the most complex applications to new heights of real-world performance.

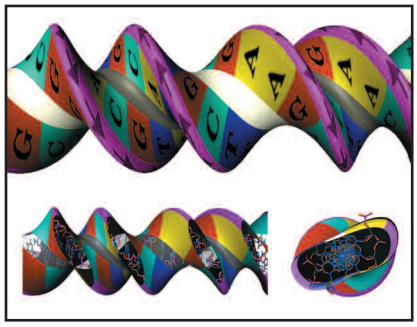

Image courtesy of Mike Carson UAB/CMC

Example #1: Genomics research institute

A European genomics research institute is able to meet growing requests for information from a DNA data bank that it provides free to the worldwide health community.

Challenge: One of the world's leading centers for analyzing and understanding the human genome, the institute maintains a DNA databank available to researchers around the world. The use of genomic data helps fight epidemics such as malaria or typhoid fever, which kill nearly 2 million people each year. But as the battle against these diseases grows more sophisticated, use of the institute's data bank is growing exponentially. In 2000, institute administrators saw 30,000 DNA queries conducted each week by health professionals and researchers. By 2005, the number of weekly searches had soared to more than 8 million.

Administrators determined their existing 1,500-node commodity cluster could not meet the growing need for their search applications – including BLAST and another, even more powerful user-developed application – to provide sufficient memory access. On the institute's existing platforms, their applications would not live up to the test of ever-increasing search loads.

Solution: Because hashing algorithms like those used in the institute's applications require large-memory systems, an SGI® Altix® cluster solution includes a pair of dual-processor SGI® Altix® 350 servers in addition to an SGI® Altix® 3000 system equipped with 192GB of memory. The particularly dense Altix 3000 configuration is aimed at solving the institute's chronic memory-access problems.

The SGI Altix cluster is suited to meeting the institute's needs because Altix features modular building blocks that make it easy to scale memory, I/O, and CPUs. This means the institute can scale up in terms of CPUs, or scale out memory and I/O in an easy, cost-effective way.

Results: Software developers at the institute now have systems capable of helping them develop a new generation of sequence assembly tools. Productivity has increased because the larger overall memory made available by the Altix cluster can easily accomodate the applications and data. The Altix servers' memory availability was so beneficial that the institute purchased another large-node Altix server – this one driven by 16 processors and equipped with another 192GB of memory. The new clustered system will not only serve internal users, but will provide a powerful resource for external developers and researchers as well.

Image courtesy of CEI

Example #2: Aerospace manufacturer

An aerospace company increases its competitiveness by accelerating computational fluid dynamics workflows and reducing development costs with a data centric solution from SGI, thus accelerating new business wins.

Challenge: A major aerospace company is under contract to deliver a new generation of cabin ventilation for an airliner that will increase air filtration, reduce weight, and cut operational costs. The new ventilation design must be delivered at a fixed price and on a given date two years in the future.

The successful completion of the design requires new global optimization of the fluid flow around the sub-assembly and through the cabin, as well as the exchange of air with the external environment. To meet stringent goals laid down by the aircraft manufacturer, 100 percent of the design must be simulated prior to fabrication, and failure will eliminate the company from future multi-billion-dollar-per- year opportunities.

To succeed, the company must:

- Increase the complexity of flow simulation from 5M cells to over 20M cells for internal operations, and over 200M cells for several different cabin configurations

Large-node clusters

- Optimize designs to reduce weight and maximize filtration
- Analyze time-varying results that range from 50GB up to 4TB in size
- Enable multiple remote teams (including the aircraft manufacturer and launch customers) to collaborate on the design and analysis
- Minimize capital investments while accelerating schedules

Solution: The solution for this large data analysis challenge is adoption of a data-centric approach to computing and visualization that gives all team members access to the data and visual results but does not move the raw data itself.

This is accomplished with an SGI® Altix® 1350 cluster powered by a total of 128 Intel® Itanium® 2 Processors and 512GB of memory for running STAR-CD from CD-adapco. The large-node Altix cluster is integrated with a Silicon Graphics Prism™ visualization system, a 20TB SGI® Storage Area Network for sharing results between HPC and visualization systems, and SGI® Visual Area Networking for delivering the high-end results generated on the Silicon Graphics Prism visualization system to local and remote end-users.

Results: Overall, the developer projects significantly accelerated project delivery and operational cost reduction. This is the result of a variety of improvements in the workflow:

- Critical path development stays on track, even as model sizes expand
- HPC results are immediately available for visualization, thus eliminating data copying and accelerating design and analysis cycles by up to 50 percent

Example #3: Defense system integrator

A large defense system integrator (SI) drastically reduces cycle time for Radar Signature analysis on airplane prototypes that can elude radar detection with an integrated SGI solution for compute, visualization, and storage that includes a large-node cluster configuration.

Challenge: The SI is in a competitive fly-off where the performance of the aircraft and its radar signature are critical evaluation criteria. In this competitive environment, the schedule for the design of this next-generation aircraft is highly compressed. The engineering team must optimize the aircraft geometry, develop new radar-absorptive coatings, and then simulate the radar cross section for all weapons configurations and operating conditions, as well as for current and future radar technologies.

The existing UNIX® OS based systems employed for these tasks are functional but do not allow the team enough computer simulation time within project deadlines for adequate design optimization.

Solution: Integrating cluster capabilities with large shared-memory and visualization into a single solution tied together by a single, shared-file system, the company maximizes computational throughput and enables its engineering team to understand the effects of various design and process changes.

This is accomplished with:

- An SGI Altix 1350 cluster system with a total of 96 CPUs and 192GB of memory to simulate a large number of different Xpatch scenarios
- An SGI® Altix® 3700 server with 64 Itanium 2 Processors and 128GB of memory for producing high-resolution simulations of promising design candidates using MM3D
- Two Silicon Graphics Prism systems for complex visualization and a shared 50TB storage infrastructure that gives all compute and visualization systems equal access to high-performance storage, further reducing wait times and increasing productivity
- Visual Area Networking (VAN) that allows engineers throughout the organization to use normal desktop systems to access the total power of the Silicon Graphics Prism

Results: Computation time for equivalent problem sizes is accelerated 300 percent. A full aircraft scattering simulation at 2GHz is conducted in hours rather than days. Using accelerated computational electromagnetics simulation with SGI technology, the customer improves aircraft geometries and coating materials without having to perform experimentation on costly physical prototypes.

Compute acceleration allows the team to try more scenarios before arriving at a final design that hides the true size, shape, and speed of the aircraft.

By allowing radar cross section analysts to construct, simulate, and analyze more than three times as many scenarios in a given period of time, the system integrator was able to improve the design by 10 percent, reduce staffing on the project by 20 percent, and eliminate over $4 million in physical testing, resulting in a 220 percent ROI — independent of whether the SI wins the contract.

Break Through with SGI Altix Servers

With the acclaimed SGI® Altix® family of servers and cluster solutions, users can choose from a broad range of large-node, 64-bit Linux servers for cluster implementations – from the SGI® Altix® 330 entry-level server, which scales up to 16 Intel Itanium 2 Processors and 128GB of memory per node, to the SGI Altix 350 line, which scales up to 32 Intel Itanium 2 Processors and 384GB of memory per node, up to the world-class supercomputer capabilities of the SGI Altix 3000 family, which scales up to 512 processors and 24TB of memory per node.

Large-node clusters

The SGI Altix 1350 cluster takes this one step further. This fully integrated cluster solution incorporates compute nodes that scale to 32 processors, along with a complete cluster solution stack that includes industry-leading management, interconnect and storage technology, such as Scali Manage™, Platform LSF®, InfiniBand, 10Gb Ethernet, and SGI® InfiniteStorage storage solution.

Altix 1350 clusters are designed to be easy to deploy and administer, with simple, easy-to-manage configurations and the flexibility to scale up with more processors and out with more nodes to address changing business requirements.

Altix 1350 clusters overcome the limitations of small-node clusters by combining five key features that together create a platform capable of driving exceptional application performance:

1. High-speed processors with large cache
2. Large addressable memory
3. High memory-to-processor bandwidth rates
4. High disk-to-memory I/O rates
5. A low-latency interconnect that can provide efficient parallel scalability to hundreds of processors

Efficiently handles mixed workloads. Altix 1350 clusters enable HPC users to efficiently process a diverse job mix, an increasingly critical requirement for overall HPC and end-user productivity. Most systems today can deliver good single-job parallel speed-up, but industry practice typically demands 12 to 16 CPUs for a single job executed simultaneously with other jobs in a multi-job throughput environment.

Leverages SGI NUMAflex™ shared memory architecture. SGI NUMAflex™ is the industry's fastest interconnect, with I/O rates of more than 6.4GB per second. Data crosses over an SGI NUMAlink™ switch, round-trip, in as little as 50 nanoseconds—less time than it takes a beam of light to travel 50 feet—compared to 10,000 nanoseconds or more with many commodity clustering interconnects. SGI Altix servers are all built on a NUMAflex core architecture, which provides lightening fast communication and data sharing across processors. Developers can leverage this high performance interconnect to run large models and process large data sets on a scalable Altix node. These nodes can then be extended in a scale-out configuration using industry-standard interconnects, to provide the ultimate in configuration flexibility and power. This leads to real-world application performance breakthroughs.

Balanced architecture. SGI Altix 1350 clusters provide maximum flexibility by giving users the option of running both shared-memory and distributed-memory applications on a single node or across an entire cluster. This improved balance between capability and capacity computing, an attribute once exclusive to more expensive UNIX systems, allows HPC applications and disciplines to co-exist in a system environment without competing for CPU cycles, memory, and I/O resources. Combining this power and capability with the ability to scale out in distributed memory configurations using industry standard interconnects offers customers the "best of both worlds".

Get the Altix Advantage

To find out how highly scalable SGI Altix servers and Altix 1350 clusters can help you handle more real work with fewer systems, run both shared- and distributed-memory applications, dramatically reduce overall solution costs, and push application performance to new heights, visit: *www.sgi.com/altix.*

Corporate Office
1500 Crittenden Lane
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5509.1455
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

3844 [07.2005]

J14995