
White Paper

# Maximizing Success in Exploration and Production with SGI® Altix® ICE

Table of Contents

## 1.0 Executive Summary

High performance computing (HPC) has become vital to oil and gas exploration and production activities. The increasingly difficult task of locating productive energy supplies and maximizing their extraction dictates the use of powerful compute resources to handle critical applications like seismic imaging and reservoir simulation. Intensifying the drive for ever more powerful systems is the growing shortage of geoscientists and the need to leverage their efforts effectively. Cluster-based systems predominate in the exploration and production arena. With excellent price/performance, they represent a good choice for small- to mid-sized HPC needs. As exploration and production issues escalate, however, clusters have difficulty scaling to meet the challenge.

Inefficiencies inherent to cluster-based systems, manageable on a small-scale, emerge as major impediments to deploying larger-scale systems. The key issues impeding large-scale clusters are 1) power and cooling inefficiencies that lead to stress on data center resources; 2) the complexities of managing larger-scale systems that result in high IT personnel costs; 3) bottlenecks in I/O bandwidth that lower system efficiency because of a mismatch between today's powerful processors and relatively slow interconnect; and 4) proliferation of reliability issues that increase system management costs and impact achievable performance.

The SGI® Altix® ICE integrated blade platform surmounts the difficulties that advanced seismic imaging and reservoir simulation codes experience with conventional cluster systems. With SGI Altix ICE, SGI introduces a new standard of achievement in scale-out high performance computing, delivering a platform that far exceeds the capabilities of clusters by incorporating unprecedented power and cooling efficiencies, hierarchical system management software, high-speed interconnect enabling dramatic I/O bandwidth improvements, and key reliability enhancements. At the same time, SGI Altix ICE realizes strong price/performance value, even when scaling to thousands of compute nodes. SGI Altix ICE offers an ideal platform for maximizing exploration and production success through expansion of compute-intensive activities.

## 2.0 Supercomputing Resources Critical to Exploration and Production Activities

As worldwide energy demand intensifies, the need to find new oil and gas reserves and fully exploit existing fields becomes urgent. Finding new fields and maximizing existing ones requires more extensive seismic surveying and more intensive seismic imaging. As demand outpaces the discovery of new reserves , a new emphasis is being placed on reservoir simulation to evaluate secondary and tertiary recovery strategies that improve extraction ratios and increase the size of proven reserves. Compounding the difficulties in exploration and extraction is a growing shortage of experienced geoscientists and engineers, resulting in the need to leverage the efforts of those who remain. By moving to more accurate seismic imaging algorithms enabled by sizable increases in HPC capability, oil and gas asset teams eliminate the need for multiple process/analyze cycles and significantly increase their productivity.

The increasing reliance on seismic imaging and reservoir simulation is driving the quest for ever more powerful compute resources to run ever more complex algorithms. The energy industry now employs some of the world's most powerful super-computers in its exploration and production activities. To maintain their competitive edge in oil and gas, companies must continue to rely on the latest advancements in supercomputing technology.

Figure 1. Accurately drilling oil and gas wells requires the latest HPC solutions

### 3.0 Where Conventional Clusters Fall Short

Clusters represent the majority of high performance computing systems for exploration and production in the oil and gas industry. They offer the advantages of excellent price/performance and seeming ease of scalability. However, as companies attempt to scale their clusters to address increasingly complex problems, they find themselves hitting a wall on both price/performance and scalability. At larger installation sizes, particularly systems of 500 cores or more, a number of hidden factors emerge that subvert the initial advantages of clusters and lead to unanticipated costs and performance shortfalls:

- Power and cooling inefficiencies
- System management complexity
- I/O bottlenecks
- Reliability problems

With new algorithms like reverse time migration (RTM) requiring increases of 30x or more in computing power, the performance and productivity gains essential to supporting exploration and production activities dictate systems that offer more than just high processor counts. Systems must be designed from the ground up to address the key scalability issues.

### 3.1 Power and Cooling Inefficiencies

Today's exploration and production data centers are in a state of crisis. AFCOM's Data Center Institute predicts power failures and power availability will halt IT operations at more than 90% of companies over the next few years (AFCOM, 2006). Built years ago for a very different computing environment, many data centers today suffer from limited power, cooling, and space capacity. The environmental inefficiencies of traditional clusters (low density, combined with inefficient power and cooling utilization), barely noticeable on a small scale, become prohibitive when scaling larger installations consisting of thousands of servers. These inefficiencies lead not only to high operating costs, but to high capital costs as well, as data centers overhaul their HVAC systems or even construct new facilities to accommodate higher capacity systems.

### 3.2 System Management Complexity

Effective use of large-scale seismic imaging clusters depends on the coordination of hundreds to tens of thousands of processors to handle resource-hungry algorithms. Such a complex environment requires careful management of system resources. Without well-designed management software, the cost of IT personnel to deal with system and data issues can spiral out of control.

### 3.3 I/O Bottlenecks

With today's advanced processors enabling powerful multi-core compute nodes, many seismic imaging applications are becoming I/O-bound on traditional clusters, because those clusters rely on low-bandwidth, ten-year-old Gigabit Ethernet technology. Without sufficient I/O bandwidth, transmission speeds are unable to match the quantity of data generated, and processing power goes to waste. Today's eight-core nodes are already saturating Gigabit Ethernet I/O channels, leaving nodes I/O-starved and limiting overall system performance. As the number of cores per socket continues to increase, this trend will only get worse, requiring wholesale replacement of I/O infrastructure.

### 3.4 Reliability Problems

Reliability is an area for concern when venturing into larger-scale clusters. Compute nodes and other components inevitably malfunction over time, and cluster installations typically lack sufficient redundancy to deal robustly with component failures. The network cables tying the cluster nodes together can also suffer from reliability issues that grow exponentially as clusters scale. While clusters can handle node failures for some seismic imaging algorithms by reassigning processes to other nodes, other seismic algorithms and reservoir simulation applications have multiple levels of parallelism that utilize MPI and are sensitive to node-level interruptions. In MPI-based applications, node failures can require the restart of long-running and costly jobs, significantly reducing efficiency and asset-team productivity. Reliability problems also increase service costs – an issue of particular note for installations in difficult-to-service locations like ships.

## 4.0 SGI Altix ICE: No-Compromise Computing for Critical Applications

To counter the complexity and costs associated with large-scale clusters and provide the capability to handle critical seismic imaging and reservoir simulation applications, SGI offers a revolutionary new type of supercomputer infrastructure – the SGI Altix ICE integrated blade platform. SGI Altix ICE achieves unprecedented scalability, with a design that focuses on overcoming the weaknesses inherent in conventional clusters, particularly for installations exceeding 500 cores.

SGI Altix ICE includes these key features:

- Scalability from 128 to 100,000+ CPU cores per system. Built using the latest Quad-Core Intel® Xeon® processor and chip sets, these solutions can tackle the full range of industry needs – from the small seismic consulting firm to the large national oil company or the global super major.

- Power, cooling, and density optimizations that minimize operational costs and expensive changes to data center infrastructure.
- Easy-to-manage scalability with a hierarchical management node structure that enables node, chassis, rack, and system level management.
- Enhanced hardware reliability and serviceability based on hot-swappable diskless nodes, shared RAID arrays, N+1 redundant fans, and N+1 redundant power supplies.
- Reduced system complexity with integrated, cable-free blade enclosures and InfiniBand 4X DDR switches that offer blazingly fast I/O bandwidth.
- Maximum interconnect performance with dual-plane InfiniBand 4X DDR and an optimized dual-plane MPI implementation designed to maximize throughput and minimize latency.
- Standards-based, all-Linux® integrated software solution stack.
- Power-up-and-go design, with factory integration and testing, for easy deployment and immediate productivity.

Figure 2. SGI Altix ICE Platform

The result is an integrated high performance computing solution that provides almost unlimited expansion of computational power with increased reliability. No matter the size of the problem, asset teams can utilize the latest time saving but computationally intensive seismic imaging and reservoir simulation algorithms that result in an increased understanding of subsurface structure, higher oil extraction ratios, and increased team productivity.

### 4.1 System-Integrated Storage Options

SGI Altix ICE uses a “diskless node” architecture that removes storage from the compute blades, a design with numerous power and reliability benefits. In addition, by moving the storage off the compute blades, SGI Altix ICE allows customers to choose the storage option that best fits their computing environment. Many seismic imaging algorithms do not require intermediate storage. For these applications, the SGI Altix ICE design increases reliability and reduces power consumption. Applications that require intermediate storage can utilize high-performance RAID-based solutions over the InfiniBand infrastructure and achieve both higher reliability and transfer rates that are 5x to 10x that of local disks. Another advantage of moving the storage “off-blade” is the centralization of storage resources, promoting better capacity balancing, easier management, and on-line scaling of capacity.

SGI offers a range of storage solutions for SGI Altix ICE. Key storage options integrate with a customer’s existing storage infrastructure and support high-performance SGI Altix ICE installations. These include the SGI® InfiniteStorage high-performance, ultra-dense scalable RAID arrays, fully integrated Lustre storage environments, Panasas Parallel Storage Clusters, and SGI InfiniteStorage NEXIS NAS appliances. SGI provides full integration services for all storage solutions, ensuring a power-up-and-go experience for the customer.

## 5.0 Meeting the Scalability Challenge

SGI Altix ICE provides a robust solution to the challenges of scaling oil and gas compute resources. Its design tackles the critical issues inherent to large-scale computing by offering unprecedented capabilities in the areas of:

- Power and cooling efficiencies
- Simplified management and scalability
- High-speed unified interconnect and I/O infrastructure
- Reliability and serviceability

### 5.1 Power and Cooling Efficiencies

The design of SGI Altix ICE takes into account the environmental constraints of today’s exploration and production data centers and offers integrated solutions to meet both operational and infrastructure challenges. SGI has long been a technological leader in power and cooling optimization for HPC systems and is now applying that experience to the needs of scalable cluster systems.

SGI Altix ICE incorporates the latest power conversion and distribution architecture, with redundant, single-stage power supplies that convert AC directly into 12VDC at 90% efficiency. This represents an average savings of up to 33% when compared to conventional cluster implementations. With the cost of electricity spiraling upwards, this already represents a savings of $21k annually for every 10 TFLOPS of compute power (based on an electricity cost of $0.092/kWh), or $42k if traditional facility overhead is also considered.

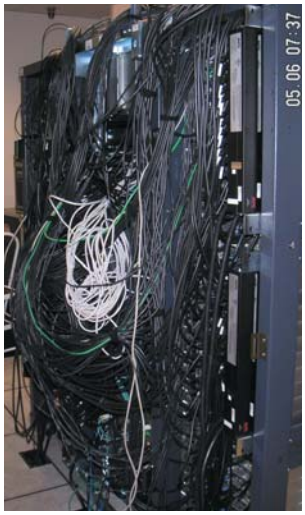

Figure 3: Traditional Cluster Cabling

Figure 4: Multiple SGI Altix ICE Racks, Reduced Cabling

SGI Altix ICE also integrates the latest cooling technologies, combining high-efficiency redundant blowers with optional water-cooled rear doors that remove up to 95% of the heat. While this has been shown to save up to 17% in cooling costs when compared to conventional air conditioning (or up to an addition $11k annually for every 10 TFLOPS), the real savings occur by dramatically increasing computer room capacity while avoiding costly modifications to HVAC systems or even the building of completely new computing centers. With new exploration and production data centers costing $10M to $30M and taking two to three years to build, extending the life of existing centers represents enormous savings and logistics simplification.

### 5.2 Simplified Management and Scalability

With its emphasis on component and software integration, SGI Altix ICE sets a new standard for simplicity in the world of scale-out environments. On the hardware side, the clean design of the blade enclosure, with its integrated blades, switches, and interconnect, stands in welcome contrast to the ad hoc maze of many cluster-based systems. SGI's integrated approach continues on the software side with an integrated system management solution stack that builds on the power and familiarity of standards-based environments to ensure rapid deployment and simplified administration.

SGI Altix ICE employs a hierarchical management architecture that provides a high degree of modularity, allowing SGI Altix ICE installations to scale to very large sizes while enabling direct management at every level of the system: node, blade, blade enclosure, rack, and system. Administrators can provision, monitor, and service resources at any of these levels, while ensuring consistent operating system images and user accounts across all nodes.

The entire SGI Altix ICE system undergoes factory integration and testing. SGI Altix ICE arrives at the customer site ready for out-of-the-box deployment. SGI's more than 25 years of experience in delivering "power-up-and-go" systems ensures immediate productivity for your asset teams.

### 5.3 High-Speed Unified Interconnect and I/O Infrastructure

The SGI Altix ICE network architecture contributes to significant application-level performance boosts. The dual-plane network runs on InfiniBand 4X DDR, providing maximum bandwidth for both MPI and I/O traffic. With its high-speed, low-latency hardware characteristics and low-overhead drivers, InfiniBand enhances the performance of the network, and thus the system as a whole.

| | Gigabit Ethernet | InfiniBand 4X DDR | Dual-Plane InfiniBand 4X DDR |
|---|---|---|---|
| **Link Rate (Bidirectional)** | 1 Gbit/sec | 20 Gbit/sec | 20 Gbit/sec |
| **Bandwidth (Measured MPI)** | 85 MByte/sec | 1390 MByte/sec | 2410 MByte/sec |
| **Latency (Measured MPI)** | 30 uSec | 3.4 uSec | 3.4 uSec |
| **Latency Variability** | High | Medium | Low |

Table 1. Interconnect Performance Characteristics

Compared to the Gigabit Ethernet found on most cluster systems, InfiniBand provides 20 times the amount of I/O bandwidth, enabling applications to exploit fully the processing power of eight-core compute nodes. In addition, high-performance file servers built on NAS/NFS or Lustre directly attach to the InfiniBand network, allowing data to stream at much faster rates than from low-end disks or disk arrays.

### 5.4 Reliability and Serviceability

SGI Altix ICE, with its focus on component integration along with other innovative features, achieves a new standard of reliability for scale-out configurations. Key reliability features include:

- Diskless blades with hot-swap capabilities, leading to low node failure rates
- Redundant storage, using RAID-based solutions to store intermediate results
- Cable-free blade enclosures, for reduced potential points of failure
- Redundant, hot-swappable power supplies and blowers
- High-efficiency power architecture, for reduced heat dissipation
- Optimal thermal design, for more efficient and consistent cooling across all components
- Fully buffered DIMMs, to reduce transient errors
- Dual-plane InfiniBand backplane, for high signal reliability

Along with reliability, serviceability is a key SGI differentiator in the high performance computing marketplace. The SGI Altix ICE modular blade design enables administrators to access and replace components easily. The diskless blades are hot-swappable, maximizing serviceability and system availability in case of component failure. To quicken resolution of service issues, all Altix systems also offer the SGI Embedded Support Partner program – an integrated suite of services that work together to monitor and manage systems and report potential problems to system administrators and SGI for preventive maintenance. In addition, the SGI customer service organization offers a broad range of support offerings, up to and including mission-critical 24x7 system support. Dedicated to excellence, the SGI support team consistently ranks among the top in the industry, according to the SatMetrix™ third-party evaluation.

## 6.0 SGI Modular Seismic Imaging Solution

The SGI Modular Seismic Imaging Solution, incorporating SGI Altix ICE, realizes these design objectives:

- Process more and higher-resolution seismic surveys with more accurate imaging algorithms
- Maximize throughput for mixtures of algorithms and ranges of survey sizes
- Optimize costs, asset-team productivity, and time-to-results
- Maintain control over system and data management costs

### 6.1 A Fully Customizable Solution

The modular solution couples multiple SGI Altix ICE racks with storage solutions built on fully integrated Lustre storage environments, and SGI InfiniteStorage high-performance RAID arrays. Panasas and other storage options are also available. Each modular unit provides:

- 10 to 12 TFLOPS via 2 SGI Altix ICE racks using the latest Quad-Core Intel Xeon processors
- 1TB to 4TB of memory, 100TB to 400TB of storage, and 1.5GB/sec I/O
- Full factory integration of all components

Modules can be combined, allowing the solution to scale to any size necessary – with a 120 TFLOPS system with 15 GB/sec of I/O and 4PB of storage requiring 10 modules, or 30 racks of computing and storage capability. The solution is completely customizable, so that optimal memory, I/O, storage capacity, processing power, and software can be tailored to the need.

| |  |  |  |  |  |
|---|---|---|---|---|---|
| **TFLOPS** | 11.4 | 22.8 | 34.3 | 45.6 | 57 |
| **Memory** | 2TB | 4TB | 6TB | 8TB | 10TB |
| **Storage Capacity** | 120–480TB | 240–960TB | 360–1,440 TB | 480–1,920TB | 600–2,400TB |
| **I/O Rates** | 1.5GB/sec | 3GB/sec | 4.5GB/sec | 6.0GB/sec | 7.5GB/sec |
| **Floor Space** | 29.9 sq ft | 53.1 sq ft | 76.4 sq ft | 106 sq ft | 130 sq ft |
| **Typical Power** | 63.4 kW | 127 kW | 190 kW | 254 kW | 317 kW |
| **HVAC Load*** | 44.6 kBTU | 89.2 kBTU | 134 kBTU | 179 kBTU | 223 kBTU |

| | |
|---|---|
| **Operating System** | Red Hat® Enterprise Linux® 5 and above or SUSE® Linux® Enterprise Server 10 and above |
| **Performance Optimization** | SGI ProPack 5 |
| **Platform Management** | SGI® Tempo or Platform™ Scali Manage |
| **Workload Manager** | Altair® PBS Professional™ 8.0 and others |
| **MPI** | Multiple options available, including SGI MPT with dual-rail optimizations, Intel MPI Library, Scali MPIConnect, MVAPICH-2, and OpenMPI |
| **IB Fabric and Subnet Management** | SGI InfiniBand Fabric Subnet Management (based on OFED and OpenSM) |
| **Development Tools** | Intel C++ and Intel Fortran compilers, Intel VTune, Intel Math Kernel Library, Intel Trace Analyzer and Collector, Intel Thread Checker, and more |

Table 2. SGI Altix ICE Software Solution Stack

## 7.0 SGI Altix ICE Software Solution Stack

The SGI Altix ICE software stack is standards-based, with SGI-engineered extensions to maximize performance and manageability and ease development efforts. The result is a powerful and cost-effective solution, designed to ensure that SGI Altix ICE users become productive immediately. Table 2 lists the main components of the solution stack.

SGI Altix ICE runs on Red Hat Enterprise Linux Server or SUSE Linux Enterprise Server. The platform builds on SGI's leadership position in the Linux community. SGI has been, and continues to be, a major contributor to the Linux standard, and brings a wealth of experience and expertise to resolve customers' kernel-level issues quickly and efficiently. With the combination of SGI and Linux, SGI Altix ICE offers a scalable, robust, and standards-based software platform.

SGI ProPack 5, SGI's optimization software package, extends the Linux standard with tools to enhance system administration, development, and performance. These tools include linkless FFIO to accelerate I/O calls, resulting in dramatic performance enhancement for many I/O-intensive applications, and SGI MPT for compute-intensive applications.

SGI Altix ICE also ships with its own system management tool, SGI Tempo. Tempo has been designed to optimize management of SGI Altix ICE, enabling centralized control over SGI Altix ICE's unique advantages, including its hierarchical management framework and its diskless booting capability. For workload management, Altair PBS Professional allows scheduling and management of jobs and resources, resulting in a higher total utilization of compute resources.

## 8.0 Summary: An Ideal Platform for Oil and Gas Computing

SGI Altix ICE, with its highly scalable design, provides a manageable, efficient, and reliable platform that offers unbeatable value. Backed by SGI's extensive experience in the energy computing sector, SGI Altix ICE significantly augments the ability to perform the analyses necessary for success in oil and gas. SGI Altix ICE provides your staff with the power they need to accelerate exploration and production activities to the maximum.

## 9.0 About SGI

SGI is a leader in Linux high performance computing, with over 25 years of experience in solving the most demanding compute- and data-intensive problems. SGI delivers a complete range of high performance server and storage solutions along with industry-leading Professional Services and Support, enabling its customers to overcome the challenges of complex data-intensive workflows and accelerate breakthrough discoveries, innovation, and information transformation. With offices worldwide, the company is headquartered in Sunnyvale, California, and can be found on the Web at www.sgi.com.

## 10.0 References

AFCOM, 2006. "Five Bold Predictions for the Data Center Industry That Will Change Your Future", AFCOM's Data Center Institute, Data Center World Conference, Atlanta, GA, March 2006.



Corporate Office
1140 E. Arques Avenue
Sunnyvale, CA 94085
(650) 960-1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5185.2860
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +61 2.9448.1463



**4103 [05.28.2008]** **J15371**