White Paper

SGI® Altix®: Enhanced Return on Investment for the Datacenter

The Coming Processing Crunch

Nearly all technology analysts predict a significant jump in the need for processing power, storage capacity, and bandwidth in IT during the next few years. This surge in demand will result from the confluence of several emerging trends and technologies:

- Increased computational demands: The consolidation of multiple databases and applications on a single server is a trend that is already well underway and will continue to accelerate as the cost advantages of this approach make its adoption compelling.

  Separately, the increasing need for security is pushing enterprises to employ strong encryption of all data-whether stored in databases, files, or backups or transmitted via wireless or wireline connections.
- Increased data analysis: As Radio Frequency Identification (RFID) is deployed by vendors and suppliers of goods, torrents of detailed information will become available. Agile companies will want not only to track and monitor this data, but also to perform extensive analysis on it to optimize transportation, delivery, and conveyance systems.
- Larger data files: Documents and files will increase substantially in size due to greater use of multimedia formats in documents, presentations, and core corporate files (meeting minutes, engineering whiteboard sessions, customer interactions, VoIP calls, and so forth). These new documents will impose greater processing burdens for the presentation of these audio and visual components. In addition, due to the Sarbanes-Oxley Act of 2002, much of this data must be retained for long periods and must include related e-mails and IM streams.
- Greater demands for bandwidth: As Service-Oriented Architecture (SOA) becomes the default connection among systems, departments, and companies, the processing of multiple, high-volume streams of network packets will place increasing demands on server-side computational resources.
- Greater performance demands: Businesses increasingly are being pushed to implement real-time decision-making to respond to customer demands, internal events, and supply-chain anomalies. This requires high-speed, high-volume business intelligence and real-time predictive analytics. When implemented properly, real-time response is a distinguishing competitive advantage.

Undoubtedly, the processing power of hardware computing resources will need to expand dramatically to adapt to these evolving IT requirements. The big challenge facing IT managers is the lack of guidance on how to scale hardware resources effectively-partly because each situation has its own unique set of needs and partly because the model for scaling processing resources today is limited by a single view of enterprise hardware.

The Limits of Server Scalability Today

Currently, when enterprises examine options for scaling their server platforms, they must choose between two models: The big server approach, which has hard limits on the maximum capacities of memory, processors, and network and disk I/O; or the cluster approach, which weaves together less-expensive systems into something approaching a single cohesive server. Corporate technology buyers must decide which model-big server or cluster-will have the highest return on investment (ROI).

Although the big server approach typically offers sufficient capacity to meet most processing needs, the growth path for these systems is often too rigid to efficiently address changing computational requirements. For example, to expand server memory past the hardwired maximum, the only choice is to migrate the applications to a new server with more memory capacity. If that new server needs a larger complement of processors, then those additional processors often require new software licenses-adding unnecessary software expense. A more logical and cost-effective approach is to purchase only the resources required-whether memory, processors, or I/O-rather than to perform a complete system upgrade.

The alternative to the big server model is the cluster model, which evolved as a solution to this conundrum. With clusters, several inexpensive processing boxes are networked together to perform common work. For certain classes of problems, this solution works well-despite the increased IT complexity. But not all problems are amenable to a cluster solution, especially those that cannot be compartmentalized on small-scale systems. These include data mining, complex database transactions, and any tasks that require substantial shared memory.

What IT buyers require is a third model, one that offers the flexibility to scale up or out at the server level plus finer-grained scalability at the computing resource level-all in an open, standards-based environment.

A New Approach to Operational Efficiency

The SGI® Altix® server line delivers this third model. Altix offers very extensive, independent scalability along any key dimension (processors, memory, network bandwidth, and I/O) in a platform based on Linux® and other industry standards. Altix servers can be customized to meet a particular processing requirement at the lowest possible price and also can grow over time, eliminating the need for whole system hardware replacement. Altix servers provide this remarkable flexibility using industry standard components and off-the-shelf Linux operating distributions.

How big is 24TB of RAM?

According to Winter Corp.'s survey of Top Databases, this much RAM could hold all but the six largest data warehouses in the world.

Today, the Altix product family includes systems that can scale to 24TB of global shared memory, 512 processors, and more than 12GB/sec. of I/O bandwidth, running against a single copy of the

Linux operating system. These servers leverage industry standard components, including the powerful Intel® Itanium 2 processor.

In addition to extensive scale-up capabilities, SGI Altix can scale out using industry standard interconnects such as InfiniBand. In this scenario, Altix is unique in being able to support fat node clusters to meet virtually any workflow requirement. This design gives IT managers considerable flexibility in designing hardware configurations to fit their current and future needs. In this way, SGI uniquely maps the solution to the problem, rather than constraining the problem to fit the solution.

Driving Higher ROI

With ever-increasing demands to improve operational and budgetary efficiency, organizations today are under constant pressure to minimize total cost of ownership and maximize return on investment. SGI Altix servers meet the challenge in several ways, enabling you to:

- Keep acquisition cost low: SGI Altix servers start at (USLP) $6,995-an exceptional entry price for a true 64-bit system.
- Buy only what you need: The independent scalability of Altix means you buy only the memory, processors, and I/O bandwidth you need for your application and nothing more. This eliminates the cost of unnecessary hardware and software licenses-all without compromising performance.
- Expand as you grow: Upgrade when and how you wish to achieve needed performance levels without busting your budget.
- Leverage the power of Open Source: Altix is based on industry-standard components and your choice of Novell SUSE or Red Hat Linux distributions, eliminating the need to pay proprietary premiums. Standardization on Linux also means you can leverage existing skill sets to install and run Altix, rather than pay for specialized consulting assistance.
- Reduce cost through consolidation: With unsurpassed scalability and capacity, Altix is the ideal platform for consolidating databases and applications on a single, easily administered, industry-standard platform.

Conclusion

SGI offers customers a new choice for maximizing ROI with a highly scalable platform based on industry standards. SGI Altix servers and clusters eliminate bottlenecks, enable real-time access to complex data sets, and scale with customer needs. Altix provides a unique combination of an open environment, a powerful processor platform, and the flexibility and scalability of the SGI system architecture. The result is a new way for IT to increase productivity and competitiveness by enabling the real-time, flexible enterprise.

Why use Intel® Itanium® 2 processors?

Intel Itanium 2 processors have consistently been adopted in IT wherever high-speed computation needs to be delivered in a scalable design. The processor is ideally suited to highly computational tasks such as in-memory databases, data mining, analytics, and other forms of data analysis. This is in part due to its industry-leading floating-point performance and its enormous on-chip caches-some scaling to 9MB-which are among the largest on-chip caches available anywhere. As a result, database activities run at the highest possible speeds on Itanium platforms. Oracle Database 10g was specifically tuned for the new Dual-Core Intel Itanium 2 processor, resulting in very high levels of performance for Oracle-based applications in the enterprise . SAP's own benchmarks show that 64-bit Itanium 2 processor-based servers had the best average dialog response time for the last three years. And this performance is not limited to strictly database activities: Altix systems based on Itanium 2 processors achieved the best results per processor core on industry-standard Java application-server benchmarks.
Itanium 2 supports mainstream Linux distributions such as those from Novell® SUSE™ and Red Hat®, as well as key enterprise applications such as those from Oracle®, IBM® Information Management, SAP®, and SAS®.

1. In November 2005, Microsoft used Itanium 2 servers to set world records for database performance. See http://www.microsoft.com/sql/prodinfo/compare/tpcc.mspx
2. SD Standard Application Benchmark Results, Three-Tier Internet Configuration, http://www50.sap.com/benchmarkdata/sd3tier.asp
3. See results of the world record through-put SPEC jbb2005 benchmark at http://www.spec.org/jbb2005/.

Corporate Office
SGI
1140 East Arques Avenue
Sunnyvale, CA 94085-4602
650.960.1980

North America +1 800.800.7441
Latin America +55 11.5185.2860
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

Dual-Core Itanium Quotes; Andrew Sutherland, Oracle EMEA

3938 [1.2007]

J15213