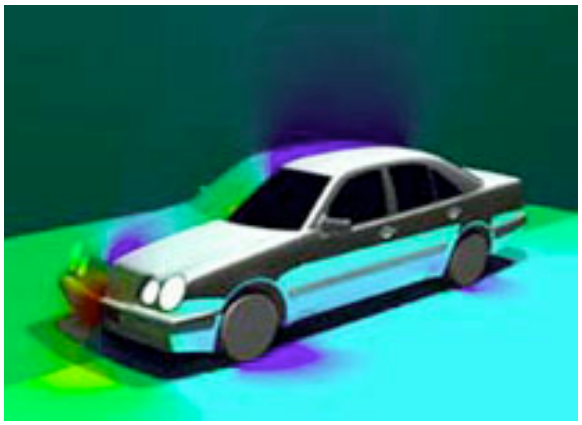

Industrial CFD Application Performance on SGI® Altix® ICE 8200EX

Challenges in CFD Applications

Parallel processing on high-performance computing environments has enabled much larger and complex CFD problems to be addressed. Higher resolution meshes to capture the multi-scale phenomena of a flow field induced by complex physical and geometrical features, transient simulations to capture unsteady effects, multiphase and non-equilibrium chemical reactions phenomena, advanced turbulence modeling such as with the combination of RANS (Reynolds Averaged Navier Stokes) and LES (Large Eddy Simulations) are just a few examples of what more and affordable computing power has enabled over the last years. As users address much larger and more complex CFD applications it naturally leads to a dramatic increase in grid resolution and simulation degrees-of-freedom. Performing simulations with such high resolution demands substantially more potent computing platforms with higher level of parallelism.

The majority of CFD applications, with few exceptions, are implemented in a distributed memory paradigm and rely heavily on the use of explicit message passing technology (MPI). The driving force behind such an approach is the desire to reach global portability over all possible computer architectures. The cluster interconnect and topology have a great influence on the overall application performance. The cluster communication fabric needs to handle I/O requirements from each of the processing cores while not imposing any networking overhead on the computational performance. The modern CFD solvers are using sparse data structures and are demanding memory systems with low latency and high memory bandwidth, When running in parallel on a high number of cores, the communication becomes prohibitive.

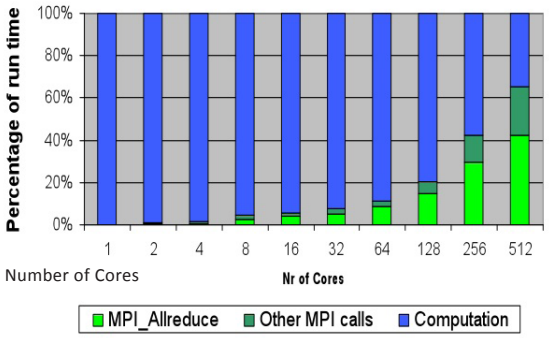

Figure 1. Increased communication when running on a high core count. Profile of Fluent v12.0.5 for the FL5L3 case, a turbulent flow of air through a duct, model size is 10, 000,000 cells, in weak scalability mode.

As shown in Figure 1, CFD communication patterns rely on collective MPI operations that stress the interconnect. In other words, a cluster with a better interconnect will improve the overall application performance and will improve its scalability.

"Using the new Intel® Xeon® Processor 5500 Series (codenamed Nehalem) we put emphasis on improving the byte/flop ratio (also known as system balance) that is a critical factor for performance in CFD applications, as they are dependent on memory bandwidth and latency (unstructured grids)."

How CFD Applications Benefit from SGI Altix ICE 8200EX Architecture

The Altix ICE 8200EX series is the latest HPC hardware offering from SGI for scale out HPC applications. This system combines the advantages of lower entry-level cost with global scalability in processors, memory, InfiniBand connectivity and I/O. Each 42U SGI Altix ICE rack holds from one to four 10U-high individual rack units (IRUs) that support up to 16 blades. Each blade supports two processor sockets that can have two or four processor cores. A maximum system size of 64 compute blades (512 cores) per rack is supported. The IRU enclosure supports a maximum of 16 single-wide compute/memory blades, eight power supplies, one chassis manager interface and two or four InfiniBand architecture I/O fabric switch interface blades. Each IRU comes with two or four InfiniBand fabric switch blades. Figure 2, shows an IRU and the switch blades that build the dual-rail interconnect fabric.

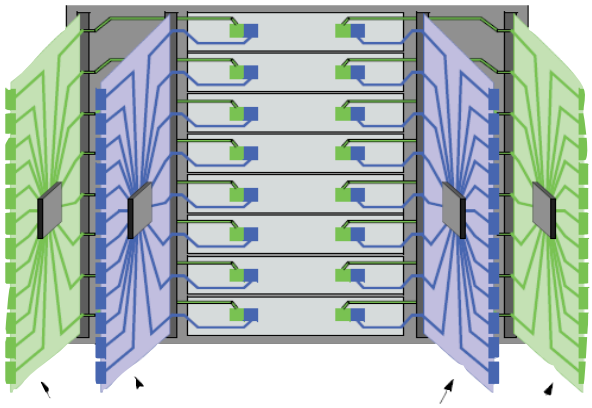

Figure 2. Altix ICE 8200EX four switch blades of an IRU backplane view.

Using the new Intel® Xeon® Processor 5500 Series (codenamed Nehalem) we put emphasis on improving the byte/flop ratio (also known as system balance) that is a critical factor for performance in CFD applications, as they are dependent memory bandwidth and latency (unstructured grids). Intel® QuickPath Interconnect (Intel® QPI) and the integrated memory controller lead to unprecedented amounts of aggregate bandwidth. With the Intel® Xeon® Processor 5500 Series as its multi-core engine, the performance of the Altix ICE 8200EX improved significantly at the node level compared to the Xeon® Processor 5400 Series (Figure 3).

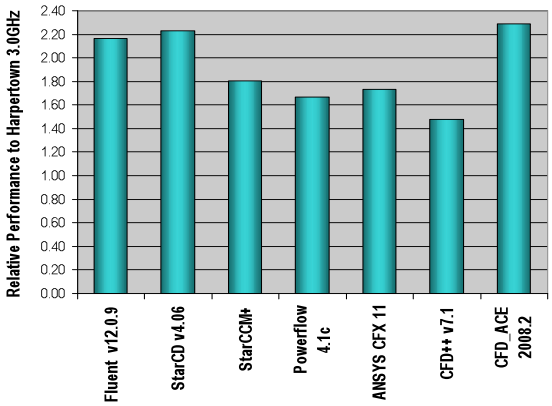

Figure 3. CFD application performance Intel® Xeon® X5570 2.93GHz compared to Intel® Xeon® X5472 3.0GHz at the node level in Altix ICE 8200EX.

One way to optimize communication traffic is through multi-rail networks. These networks can improve MPI (Message Passing Interface) communication performance by dividing large messages into chunks and distributing those chunks across multiple independent InfiniBand (IB) networks, or rails. With the dual-plane network topology of SGI Altix ICE clusters, MPI communications can make use of both IB rails. A multi-rail network may not automatically benefit an application performance. It does depend on several details and importantly, on the relations between HCA and PCIe bus capabilities. For example, each compute blade of an SGI Altix ICE 8200EX cluster has a single dual-port ConnectX HCA on a PCIe x8 bus. So using dual-rail to stripe a single large message across both ports can almost double

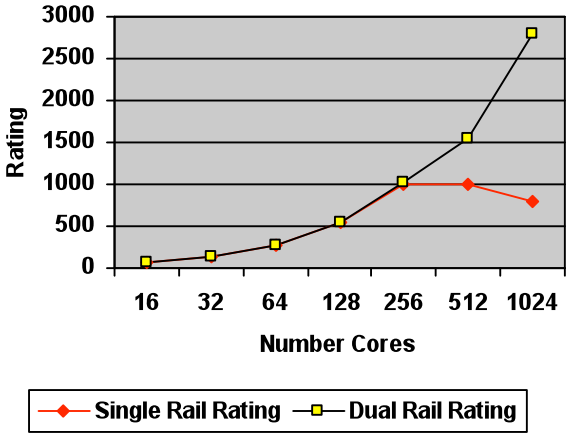

Figure 4. Dual-rail benefit vs. single-rail for Fluent v12.0.9 in Altix ICE 800EX with Xeon 5570 2.93GHz.

the MPI communication bandwidth, depending on the application communication patterns and limited by the PICe x8 bus bandwidth. The other benefit of the dual-plane InfiniBand network is that the system can separate the MPI communication from the I/O traffic, dedicating separate networks for MPI and I/O.

Figure 4 shows an example of the effectiveness of a multi-rail IB network for high-end parallel simulations. The metrics of this example is rating, which means number of jobs/24 hours for a 14M cells case of external flow around a truck configuration (higher numbers are better). SGI Altix ICE systems also allow users to choose between Hypercube (best for larger node count MPI jobs) or non-blocking Fat Tree network topology (suited for smaller node count MPI jobs). Figure 8 shows how each topology affects the performance of key HPC applications.

Unprecedented Scalability in Fluent

For example the commercial Fluent® CFD application implements a specific parallel communication layer. In combination with an advanced multilevel linear solver it is supposed to provide almost ideal parallel scaling for a large number of computational threads. Unfortunately until recent times constraints that were imposed by hardware and software technologies did not allow those Fluent features to be realized. The new approach and the interconnect topology of SGI Altix ICE 8200EX driven by application communication demands, allowed Fluent to achieve new breakthrough level of parallel scalability.

A large truck model of 111M cells of mixed type is used to understand the benefits of multi-rail network and appropriate tuning of the hardware features. The problem involves external flow over the truck body and uses a discrete-event simulation (DES) model with the segregated implicit solver. SGI reported the first ever 2048-core results running Fluent v12.09 on Altix ICE 8200EX, with a measured efficiency of 70% on 2046 cores. The results presented in Figure 5 are also available on the ANSYS benchmark website at http://www.fluent.com/software/fluent/fl6bench/fl6bench_6.4.x/problems/truck_111m.htm

"Altix ICE 8200EX is 2.27x faster than the Intel X5470 (codenamed Harpertown) 3.0GHz Hapertown on 32 cores..."

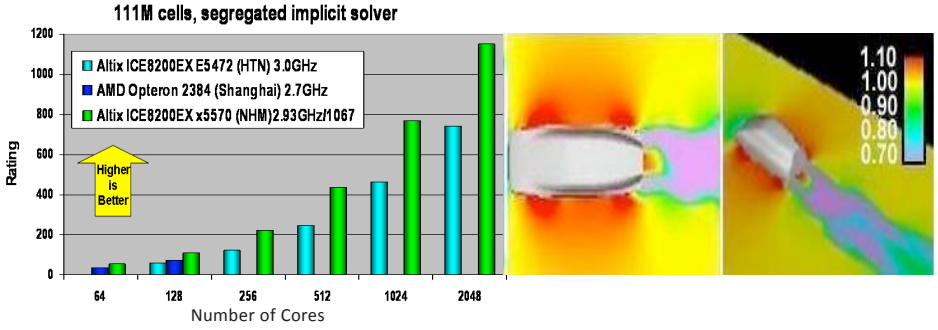

Figure 5. Unprecedented scalability measured with Fluent v12.09 simulating the external flow over a truck body with 111M cells.

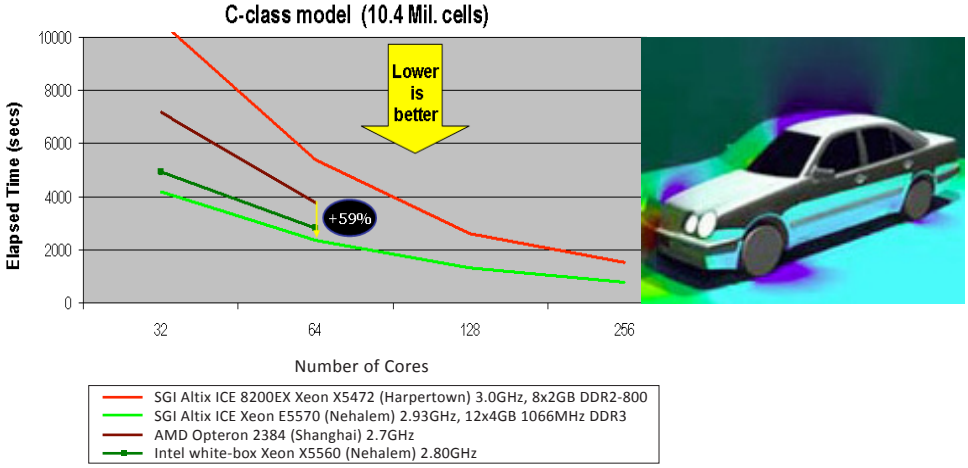

Figure 6. StarCD on ICE 8200EX with Intel Xeon X5570 (code-named Nehalem) 2.93GHz .

Leading Performance for StarCD 4.06

StarCD is a CFD process integrated software environment from CD-ADAPCO Group designed to simplify the analysis of a wide range of flows. Figure 6 illustrates the leading performance of Altix ICE 8200EX running StarCD. Using the Intel® Xeon® X5570 processor (codenamed Nehalem) 2.93GHz, Altix ICE 8200EX is 2.27x faster than the Intel X5470 (codenamed Harpertown) 3.0GHz on 32 cores, and respectively 1.93x on 256 cores (as data size per core decreases the benefit of higher memory bandwidth in Xeon X5570 diminishes). Compared to the AMD Shanghai 2.70GHz cluster, Altix ICE 8200EX with Xeon X5570 1.59x is faster on 64 cores.

For the latest performance data for StarCD on Altix ICE 8200EX please visit the CD-adapco website at http://www.cd-adapco.com/products/STAR-CD/performance/320/largeaclass.html

Summary: An Ideal Platform for CFD

When combined with the memory bandwidth intensive design of the Intel® Xeon® Processor 5500 Series (codenamed Nehalem), SGI Altix ICE systems enable today's users to achieve unprecedented application performance and scalability. This standards-based, industry-leading platform is a core component of a sustainable IT infrastructure. In addition to innovations in I/O architecture and scalability, the SGI Altix ICE integrated blade platform offers other unique benefits important to any HPC organization or enterprise:

- Integration. SGI Altix ICE systems are designed for HPC. Altix ICE systems feature performance-dense architecture that combines an innovative board design, cable-free board enclosures, integrated switches, and a high-performance, InfiniBand-focused interconnect architecture that can efficiently scale to thousands of nodes.
- Reliability. Altix ICE systems feature an advanced reliability architecture with redundant power and fans, diskless blades and hot-swappable blades. The platform's proven power/cooling architecture delivers rack power efficiency of greater than 75%, and the optional water-cooled door design lowers cooling costs and extends the life of dense configurations.
- Immediate Productivity. The SGI Altix ICE system's "Power Up and Go" design, with systems that are fully integrated and tested in the factory, enables customers to deploy scalable cluster solutions in hours, rather than weeks. The standards-based Linux system also features a comprehensive cluster solution stack that includes SGI® Tempo management software for quick deployment and easy management.

SGI Altix ICE 8200EX, with its highly scalable and efficient design, offers an ideal platform for CFD simulations. Its innovative design provides a highly manageable, efficient, and reliable platform with unbeatable price/performance value. Backed by SGI's extensive experience supporting CFD application needs, Altix ICE vastly expands the capability for organizations to perform the analyses necessary to solve today's complex problems in fluid dynamics. SGI Altix ICE speeds time-to-solution and overall productivity by freeing the engineering staff to focus on solving engineering problems, rather than resolving IT issues.

SGI collaborates directly with most commercial CFD software providers, as well as with designers of many independent open-source and corporate-specific solutions, to provide an optimized platform for their CFD modeling tools. SGI is a long-time partner with vendors of all the major CFD platforms, such as Fluent® and CFX® from ANSYS®, CFD++™ from Metacomp, Exa PowerFLOW®, and CD-adapco's STARCD and StarCCM+ and OpenFOAM.

For more information on SGI Altix ICE systems and Intel Xeon Processor 5500 Series (codenamed Nehalem) processors, visit http://www.sgi.com/products/servers/altix/ice/

"SGI Altix ICE 8200EX, with its highly scalable and efficient design, offers an ideal platform for CFD simulations. Its innovative design provides a highly manageable, efficient, and reliable platform with unbeatable price/performance value."

Corporate Office
46600 Landing Parkway
Fremont, CA 94538
tel 510.933.8300
fax 408.321.0293
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