WHITEPAPER

February 2009

Accelerating Automotive Design with InfiniBand

1.0 Abstract

CAE simulation and analysis are highly sophisticated applications which enable engineers to get insight into complex phenomena and to virtually investigate physical behavior. In order to produce the best results possible these simulation solutions require high-performance compute platforms. In this paper we investigate the optimum usage of high-performance clusters for maximum efficiency and productivity, for CAE applications, and for automotive design in particular.

1.1 Introduction

High-performance computing is a crucial tool for automotive design and manufacturing. It is used for computer-aided engineering (CAE) from component-level to full vehicle analyses: crash simulations, structure integrity, thermal management, climate control, engine modeling, exhaust, acoustics and much more. HPC helps drive accelerated speed to market, significant cost reductions, and tremendous flexibility. The strength in HPC is the ability to achieve best sustained performance by driving the CPU performance towards its limits.

The motivation for high-performance computing in the automotive industry has long been its tremendous cost savings and product improvements. A total cost of a real vehicle crash-tests in order to determine its safety characteristics, is of the range of $250,000 or more. On the other hand, the cost of a high-performance compute cluster can be just a fraction of the price of a single crash test, while providing a system that can be used for every test simulation going forward.

HPC is used for many other aspects than just crash simulations. Compute-intensive systems and applications are used to simulate everything from airbag deployment to brake cooling, exhaust systems, thermal comfort and windshield washer nozzles. HPC-based simulations and analyses empower engineers and designers to create vehicles that are more ready and safer for real-life environments.

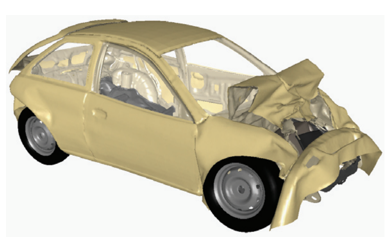

Figure 1: Car crash analysis

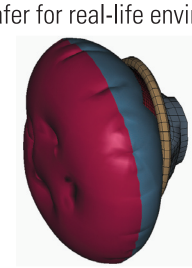

Figure 2: Airbag CFD simulation

1.2 Automotive Crash Simulations

One of the most demanding applications of automotive design is crash simulation (full-frontal, offset-frontal, angle-frontal, side-impact, rear-impact and more). Crash simulations, while performed very early in the development process, are validated very late in the development process once the vehicle is completely built. The more sophisticated and complex the simulation, the more parts and details can be analyzed. Computer-based analyses provide an early insight into phenomena that are difficult to be gathered experimentally, and if so, only at a later stage and at a substantial cost. Time and money is saved without having to build costly prototypes.

Automotive makers increase their dependency for car crash simulations throughout the design process while reducing the need for real prototypes, thus achieving faster time to market with less cost associated with the design phase. HPC clusters enable the vision of pure virtual development and having physical prototypes for verification only. For the Volvo S80 car model, 1993-1998, Volvo performed 1000 simulations and used 15 prototypes during the design. For the S40/V40 series, 1999-2003, Volvo increased their simulations to 6000 and used only 5 prototypes. For their V70N model, 2005-2007, Volvo performed 10,000 crash simulations during the product development without any real prototypes, and the crash simulations included a large variety of landscapes, such as pedestrian, slide impact, rollover and much more. The performance of HPC clusters have helped Volvo to perform the complex crash simulations, and the flexibility and scalability of HPC clusters enabled Volvo to increase the compute power in order to perform a greater numbers of simulation and faster time to market.

HPC Clusters consist of of-the-shelf servers, a high speed interconnect and a storage solution. The interconnect has a great influence on the total cluster performance and scalability. A slow interconnect will cause delays in data transfers between servers and between servers and storage, causing poor utilization of the compute resources and slow execution of simulations. An interconnect that requires CPU cycles as part of the networking process will decrease the compute resources available to the applications and therefore will slow down and limit the numbers of simulations that can be executed on a given cluster. Furthermore, this will limit the cluster scalability as when the amount of CPUs increases, the higher the burden that will be enforced on the CPUs to handle the networking.

InfiniBand, a high-speed interconnect distributed by Mellanox Technologies, provides the lowest latency and the highest throughput between servers, so the communication between the compute nodes will be fast enough to feed the CPUs with data and eliminating idle times. Moreover, InfiniBand was designed to be fully offloaded, meaning all the communications are being handled within the interconnect, with no involvement from the CPU. This further guarantees the ability to scale up with linear performance, when more compute resources are required.

1.3 Multi-core Cluster Environments

Compute cluster solutions consist of multi-core servers. A multi-core environment introduces higher demands on the cluster components, especially on the cluster connectivity. Each CPU core imposes a separate demand on the network during simulation job execution, and therefore the cluster interconnect needs to be able to handle those multiple data streams simultaneously and at the same time guarantee fast and reliable data transfer for each of the streams.

In a multi-core environment, it is essential to avoid overhead processing in the CPU cores. By providing low-latency, high-bandwidth and extremely low CPU overhead, InfiniBand provides a balanced compute system and maximizes application performance: a large factor for why InfiniBand is emerging as the most deployed high-speed interconnect, replacing proprietary or low-performance solutions.

1.4 SMP versus MPI

A common multi-core environment consists of 8 to 16 CPU cores in a single server. In a typical one server environment, application jobs can be executed in a Shared Memory Processing (SMP) fashion, or with a Message Passing Interface (MPI) protocol. In order to compare between the two options, we have used Livermore Software Technology Corporation (LSTC) LS-DYNA benchmarks.

LS-DYNA is a general purpose structural and fluid analysis simulation software package capable of simulating complex real world problems. It is widely used in the automotive industry for crashworthiness, occupant safety and metal forming and also for aerospace, military and defense and consumer products. There are three main LS-DYNA benchmarks used for evaluating a platform's performance, efficiency and scalability: 3 Vehicle Collision (a van crashes into the rear of a compact car, which, in turn, crashes into a midsize car), neon_refined (frontal crash with initial speed at 31.5 miles/hour) and car2car (NCAC minivan model). Recently, a revised version of neon_refined was introduced, named neon_refined_revised.

The platform used for this performance evaluation is the Mellanox Technologies "Helios" cluster. It is part of the Mellanox Cluster Center, a compute resource available for performance testing and application development. The Helios cluster consists of 32 server nodes, connected with gigabit Ethernet and 20Gb/s InfiniBand. Each server node has dual-socket, Quad-core 2.66GHz Intel Xeon CPUs (code name Clovertown). The MPI used in the test was Scali's MPI Connect.

In order to compare between an SMP and MPI approach, a single server was used. The comparison metric was the amount of jobs that can be achieved per 24 hours. According to figure 3, the usage of MPI improves the system's efficiency and parallel scalability, and as more cores were used, the MPI approach performed better versus the traditional SMP.

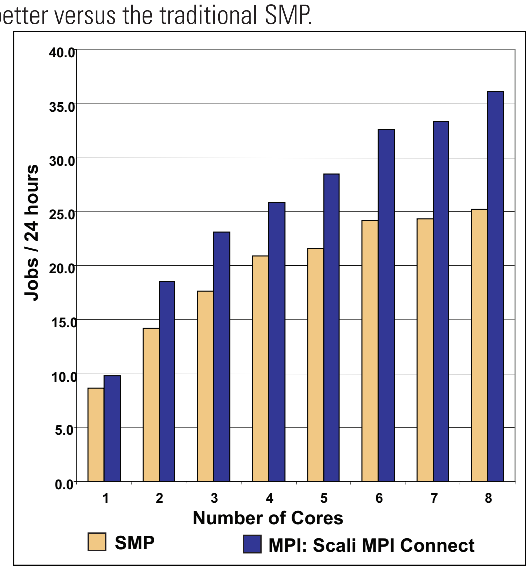

Figure 3: SMP versus MPI (LS-DYNA neon_refined_revised benchmark)

The usage of MPI in a single server not only provides better performance and efficiency, it also enables a smooth integration of a single server into a cluster environment. In addition, when more compute power is needed, no software changes are required.

1.5 Scaling in a cluster environment: the importance of interconnects

Clusters are scalable performance compute solutions based on commodity hardware, on a private system network. The main benefits of clusters are scalability, availability, and high performance. A cluster uses the combined compute power of the compute sever nodes to form a high-performance solution for cluster-based applications such as a MCAD or CAE applications. When more compute power is needed, it can be simply achieved by adding server nodes to the cluster.

In a cluster setting, each node can be a backup option to each other node in the event of node failure, or when a server node needs to be taken out for a planed service. To the application or the user, this operation is transparent, and no application run time will occur.

The way the cluster nodes are connected together has a great influence on the overall application performance, especially when multi-core servers are used. The cluster interconnect is very critical to efficiency and scalability of the entire cluster, as it needs to handle the I/O requirements from each of the CPU core, while not imposing any networking overhead on the same CPUs.

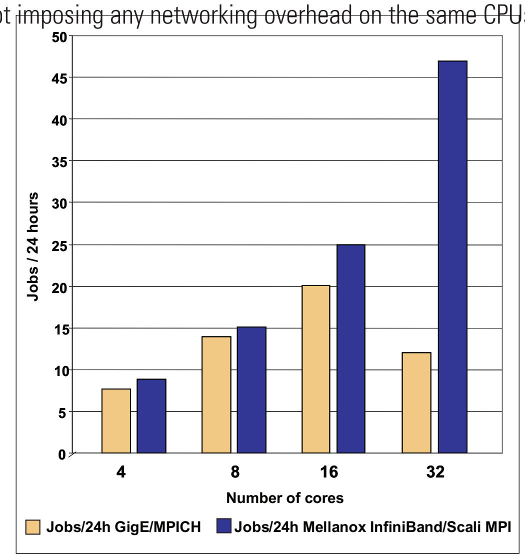

Figure 4: InfiniBand versus GigE (LS-DYNA neon_refined_revised benchmark)

Figure 4 compares gigabit Ethernet using the open source MPICH implementation and 10Gb/s InfiniBand using Scali MPI Connect as an interconnect solution. The cluster consisted of dual-socket, dual core Intel Xeon CPUs (code name Woodcrest) server nodes. For up to 16 cores (4 compute nodes), InfiniBand shows better efficiency than GigE, enabling up to 25% more LS-DYNA jobs per day. When scaling up to 32 cores, or 8 server nodes, GigE failed to provide an increase number of jobs, while also diminishing the overall compute power. InfiniBand continued to provide almost linear scalability and high-efficiency by almost doubling the number of LS-DYNA crash simulations achieved per day.

1.6 Maximizing crash simulations on multi-core InfiniBand cluster

For multi-core cluster platforms, GigE becomes ineffective with cluster size and InfiniBand is required in order to maximize the application performance and the amount of application jobs that can be obtained per day. As more jobs can be performed, the quality of the product increases and the time-to-market reduces.

A typical method for evaluating a compute solution is with the time it takes to run one application benchmark, or one application job - the faster the run-time, the more effective the compute solution.

However, it is not always the best approach for real simulation on multi-core platforms. Multi-core platforms create more demand on the cluster interconnect, but also on the CPU connectivity within a server node and the between the CPUs and memory. While running a single job on the cluster will provide the fastest time run for that specific job, the goal of maximum simulations per day might not be achieved with this manner.

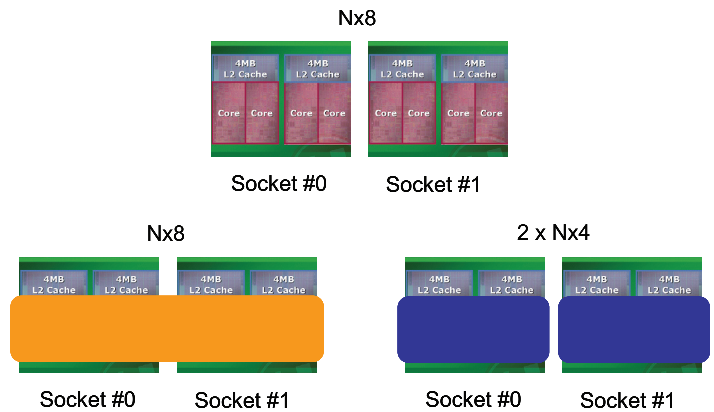

Figure 5: Socket affinity approached

Figure 5 shows the two different methods that will be examined. The platform used for this testing is the “Helios” cluster from the Mellanox Cluster Center, using Scali MPI Connect. Each node consists of dual socket, quad-core CPUs. One method (left side) is to run a job on the entire compute cluster and launch the second job once the first job is completed. The second method (right side) is to run two jobs in parallel, with each using only one socket per node and placing higher demands on the cluster interconnect. Figure 6 illustrate the two methods.

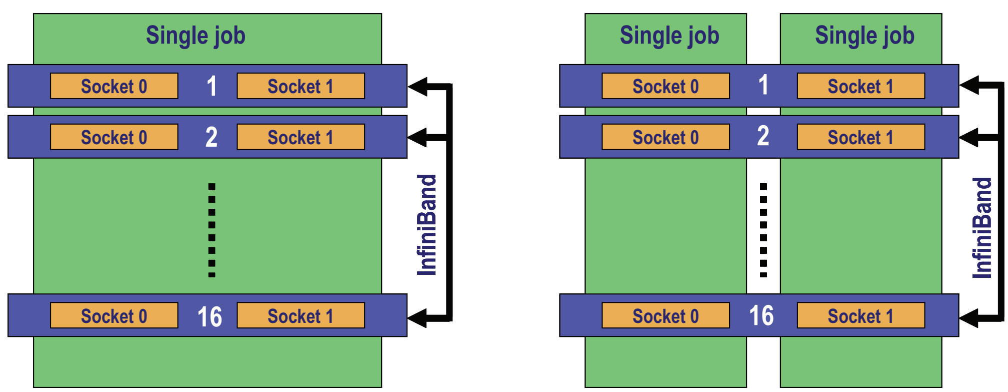

Figure 6: Two ways to distribute the Sockets for an LS-DYNA Simulation Run

Figure 7 shows the performance results of the two options. Although the run time of a single job on the entire cluster is faster, running multiple jobs at the same time using InfiniBand to connect between the servers provides more than 2x LS-DYNA jobs per day.

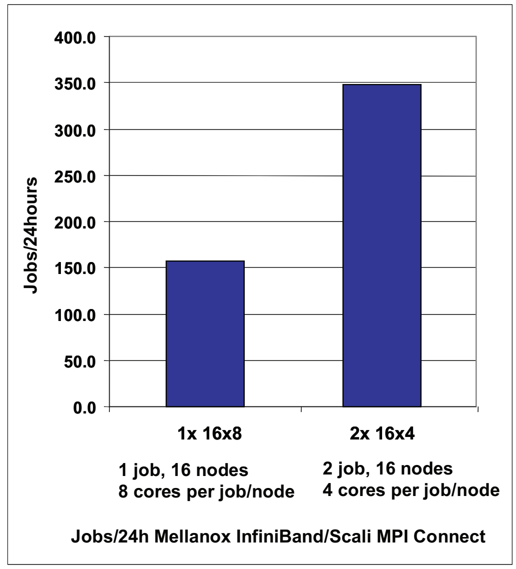

Figure 7: LS-DYNA jobs per day with and without socket affinity

1.7 Accelerating Automotive Design

From concept to engineering, and from design to test and manufacturing, the automotive industry relies on powerful virtual development solutions. CFD and crash simulations are performed in an effort to secure quality and speed up the development process. Cluster solutions maximize the total value of ownership for CAE environments and extend innovation in virtual product development.

Multi-core cluster environments impose high demands for cluster connectivity throughput, low-latency, low CPU overhead, network flexibility and high-efficiency in order to maintain a balanced system and to achieve high application performance and scaling. Low-performance interconnect solutions, or lack of interconnect hardware capabilities will result in degraded system and application performance.

Three cases were investigated. In the first case, it was clear that using applications over MPI will provide more performance versus using an SMP mode - even on a single server. In the second case, we have investigated the importance of using high-speed, low-latency and low CPU overhead interconnects for crash simulations. According to the results, a low-speed interconnect, such as GigE becomes ineffective with cluster size and even reduces the cluster compute power when adding more compute nodes. InfiniBand shows greater efficiency and scalability with cluster size.

The third case shows that CPU affinity and interconnect usage needs to be configured correctly to maximize the cluster efficiency. By reducing the stress on the socket connectivity and memory, while better utilizing the interconnect, more application jobs can be accomplished. Crash simulations are a major part of automotive design, and the ability to run more crash simulations per day will enable more complex simulations, reduce the design phase and the amount of prototypes needed.

Automobile and Airbag models were developed by FHWA/NHTSA National Crash Analysis Center of the George Washington University.

Mellanox TECHNOLOGIES
350 Oakmead Parkway, Suite 100
Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com