LS-DYNA Performance and Scalability in the Multi-Core Environment

Gilad Shainer
Mellanox Technologies

Correspondence:
Gilad Shainer
Mellanox Technologies
Phone +1 408-916-0048
Fax +1 408-970-3403
Email shainer@mellanox.con

ABSTRACT:

Efficient data transfer between clustered compute nodes is critical for balanced system performance. In a balanced system, the overall performance is equal to or greater than the sum of its components, while in a non-balanced system, the performance is less than the sum. The challenge of achieving balanced performance becomes more evident in multi-core environments. A multi-core environment introduces high demands on the cluster interconnect and the interconnect must handle multiple I/O streams simultaneously. In this paper we explore the benefits of InfiniBand high-speed connectivity solution for multi-core clusters, and show the scalability and efficiency of LS-DYNA on InfiniBand connected multi-core cluster platforms.

Keywords:
InfiniBand, Multi-Core, Scalability, Efficiency

INTRODUCTION

"In 1978, a commercial flight between New York and Paris cost around $900 and took seven hours. If the principles of Moore's Law had been applied to the airline industry the way they have to the semiconductor industry since 1978, that flight would now cost about a penny and take less than one second." (Source: Intel). In 1965, Gordon Moore predicted that the number of transistors that could be integrated into a single silicon chip would approximately double about every two years. For more than four decades Intel has been transforming that law into reality. The increase in transistor density enables more transistors on a single chip and therefore increases in the CPU performance. However, it is not the only factor driving the CPU performance. The increase of the CPU clock frequency, a bi-product of the transistor density was an important factor in the overall performance improvement. We can expect that Moore's Law will continue to deliver increasing transistor densities for the near future but power consumption and heat generation, which rise exponentially with clock frequency, will limit the increase in the CPU clock frequency.

High-performance computations are rapidly becoming a critical tool for conducting research, creative activity, and economic development. In order to provide intense computing platforms and still maintain the historic rates of performance and price/performance improvements, more execution cores are being integrated into each CPU. With multiple cores executing simultaneously, the CPU clock frequency can be reduced in order to contain heat generation and power consumption, while still increasing total system performance. This transition to multi-core CPUs is a key trend shaping the automotive computing market in addition to clusters of low cost servers and high-performance, industry-standard interconnects.

CONNECTING MULTI-CORE PLATFORMS

Multi-core environment introduces high demands on the cluster components and especially on the cluster connectivity. Each CPU core imposes a separate demand on the network and therefore the cluster interconnect needs to be able to handle those multiple data streams simultaneously and in the same time guarantee fast and reliable data transfer for each of the streams. Efficient data transfer with low CPU core overhead is a key ingredient for balanced systems performance and for the ability to maximize the CPU cycles dedicated to the application.

By providing low-latency, high-bandwidth and extremely low CPU overhead, InfiniBand is emerging as the most deployed high-speed interconnect, replacing proprietary or low-performance solutions. In a multi-core environment, it is essential to avoid interconnect protocol processing in the CPU cores. In order to maximize the overall compute cluster efficiency and to allow LS-DYNA to efficiently utilize the CPU’s core resources, a fully hardware transport-offload solution is needed. Furthermore, unnecessary overhead on the CPU cores reduces the ability of balanced computing between the various cores, leading to higher degradation in real application performance.

Interconnect flexibility is another requirement for multi-core systems. As various cores can perform different tasks, it is necessary to provide the flexibility to use remote direct memory access (RDMA) along with the traditional semantics of Send/Receive. RDMA and Send/Receive in the same network provides the user with a variety of tools that are crucial for achieving the best performance and scalability for LS-DYNA.

APPLICATIONS DEMAND A HIGH-SPEED INTERCONNECT

Finite element analysis (FEA) consists of a computer model of a material or design that is stressed and analyzed for specific results. A company that is able to verify a proposed design will be able to perform to the client's specifications prior to manufacturing or construction and reduce time to market. Modifying an existing product or structure is utilized to qualify the product or structure for a new service condition. In case of structural failure, FEA may be used to help determine the design modifications to meet the new condition.

LS-DYNA by Livermore Software Technology Corporation (LSTC) 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*.

In order to determine the importance of the interconnect architecture for multi-core environments, the same multi-core platform was used to compare between low-speed gigabit Ethernet and high-speed InfiniBand. We have used the dual-core Xeon 5160 3GHz (code name Woodcrest) and the 3 Vehicle Collision benchmark. We have run the benchmark on small cluster sizes, 2 nodes and 4 nodes, in order to determine the point where InfiniBand is needed. Each server includes dual sockets dual core CPUs. The results are shown in figure 1.

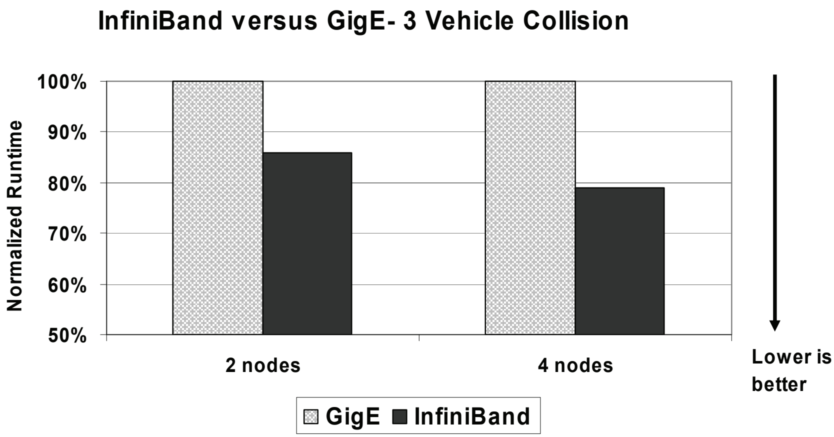

Figure 1: InfiniBand versus GigE performance results on 2 and 4 node clusters

For the 2 node cluster case, InfiniBand achieved 15% reduction in runtime, compare to GigE, and in the 4 node case, the difference between InfiniBand and GigE was more than 27%. In other words, a task that will require 1 hour of runtime on a 4 node cluster connected with GigE, will only need around 40 minutes with InfiniBand. As the number of nodes increase, the performance advantage of InfiniBand increases significantly compared to GigE.

In order to explore the scalability of InfiniBand, we have tested the car2car benchmark on a 16 node cluster, again based on Intel dual-core Xeon 5160 CPUs. The first test was executed on a single node only and then we have increased the numbers of nodes. Figure 2 summarize the testing results.

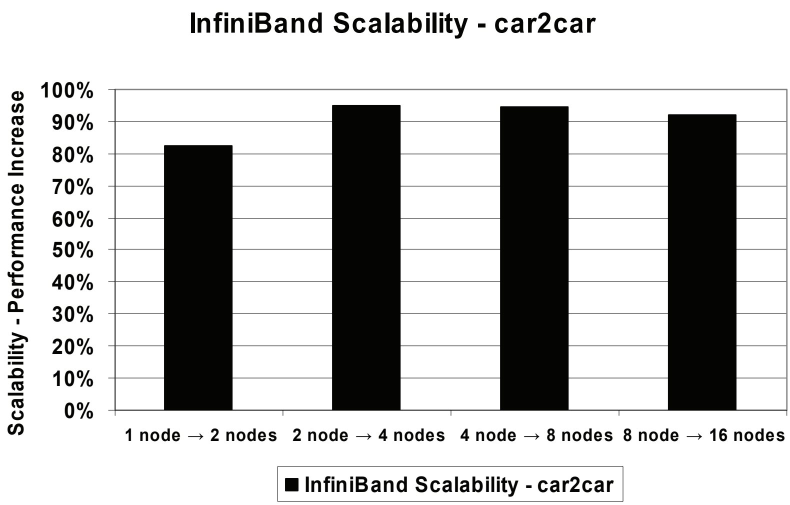

Figure 2: LS-DYNA scalability with InfiniBand

InfiniBand scalability was greater than 90%, providing almost linear scalability when scaling up from 2 nodes to 4, 8 and 16 node clusters. As more nodes exist in the cluster, and more CPU cores exist per node, the need for bandwidth increases. InfiniBand provides 10 and 20Gb/s bandwidth today and 40Gb/s is anticipated in the first half of 2008. InfiniBand is proved to provide the aggregate CPU cores bandwidth demands, while Gigabit Ethernet fails to do so.

Mellanox's InfiniBand network adapter architecture is based on a full offload approach, hardware-based reliable data transfer and RDMA capabilities, reducing the traditional protocol overhead from the CPU and increasing the overall CPU core efficiency. A different architectural approach exist from QLogic, where QLogic adapters are based on an on-load approach, where the CPU needs to drive the data transfer (as needed in the traditional TCP approach). The CPU is required to deal with the transport layer, error handling etc., and therefore increases the overhead on the CPU and reduces processor efficiency, leaving less cycles for useful application processing.

Figure 3 shows comparisons between Mellanox InfiniBand (InfiniHost III Ex) and QLogic InfiniPath adapters, with the revised neon_refined benchmark, a version that is preferred over the previous version for future benchmarks.

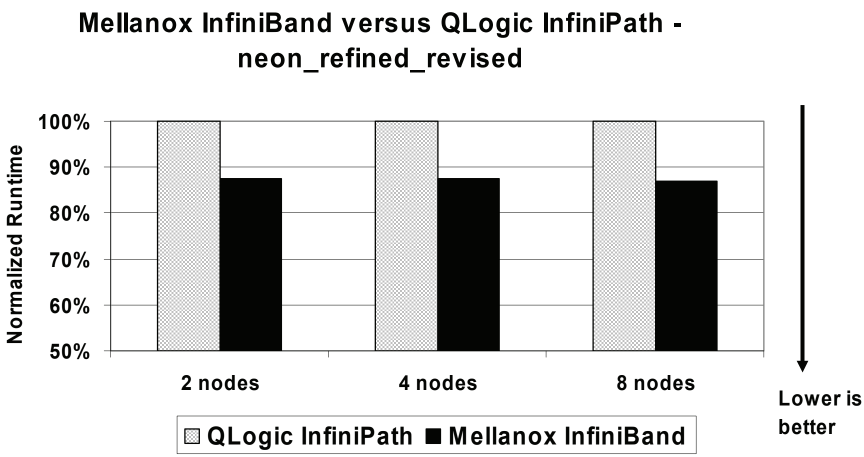

Figure 3: LS-DYNA efficiency with Mellanox InfiniBand versus Qlogic InfiniPath

Mellanox InfiniBand shows higher performance and better scaling and efficiency compared to QLogic InfiniPath. The LS-DYNA is a latency-sensitive application, and while QLogic shows lower latency compared to Mellanox InfiniHost III Ex adapter, Mellanox's architecture delivers higher system performance, delivering more CPU cycles for application use. Mellanox have recently announced their new adapters, providing lower latency compared to QLogic adapters, and the performance gap is anticipated to increase.

SUMMARY AND CONCLUSIONS

Multi-core 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 native hardware support, will result in degraded system and applications performance. As shown in the performance results on all three LS-DYNA benchmarks, Mellanox high-speed InfiniBand meets the current and future multi-core system requirements and provides a balanced compute solution.