SUPERIOR PERFORMANCE FOR LOW-LATENCY, HIGH-THROUGHPUT CLUSTERS RUNNING WINDOWS® HPC SERVER 2008

Windows HPC Server 2008 and Mellanox InfiniBand interconnect adapters deliver maximum performance and scalability for HPC.

OVERVIEW

High-performance computing (HPC) applications are increasing the demand for CPU processing performance and the amount of data being transferred between subsystems. Off-loading data movement to input/output (I/O) devices increases the amount of CPU resources available for these applications, boosting the system’s performance.

To shorten solution times, HPC clusters require high-performance interconnect technology to provide high-throughput with low-latency, and direct data transfer (CPU off-load) between compute nodes.

Windows HPC Server 2008 is a highly productive platform for deploying, developing, and accessing dedicated HPC clusters. Windows HPC Server 2008 leverages high-speed interconnects by using NetworkDirect—Microsoft’s new Remote Direct Memory Access (RDMA) interface for high-speed, low-latency networks such as InfiniBand.

Microsoft is pleased to be working with Mellanox Technologies to ensure that Windows HPC Server 2008 solutions with high-speed InfiniBand interconnect adapters deliver an outstanding level of performance, scalability and efficiency. Our new networking architecture takes advantage of the advanced InfiniBand capabilities to provide lowest latency and higher throughput for performance-sensitive applications. The extensive collaboration between our organizations is benefiting HPC users worldwide through a highly productive supercomputing platform.

Kyril Faenov, general manager of HPC, Microsoft Corporation

INFINIBAND ADAPTERS

Clusters linked with InfiniBand interconnect deliver maximum performance, scalability, flexibility, and ease of management. HPC applications achieve maximum performance over InfiniBand networks because CPU cycles are available to focus on critical application processing instead of networking functions. Node-to-node latency of less than 3μs and almost 3GByte per second has been demonstrated on Mellanox InfiniBand clusters using the MPI protocol.

With InfiniBand’s proven scalability and efficiency, small and large clusters easily scale up to tens-of-thousands of nodes. The Microsoft HPC team works in collaboration with the OpenFabrics Alliance Windows driver set (WinOF), to ensure that certified, high-performance networking drivers like InfiniBand are supported by Windows HPC Server 2008.

WINDOWS HPC SERVER 2008

Windows HPC Server 2008 builds on the strengths and ecosystem of Windows Compute Cluster Server (WCCS) 2003 to allow Windows-based HPC to support larger deployments. Windows HPC Server 2008 advances every facet of WCCS including: a new job scheduler, better provisioning based on the Windows Server® 2008 Windows Deployment Services technology, a new management interface, and a faster Microsoft Message Passing Interface (MS-MPI) that includes new NetworkDirect support.
ANSYS customers will be very well served by the new performance levels achieved with InfiniBand on Windows HPC Server 2008. Our engineering simulation software solutions, including ANSYS® Mechanical™, ANSYS® CFX®, and FLUENT®, will achieve new levels of performance scalability as a result of this collaboration between Microsoft and Mellanox. The availability of simple, manageable, high-performance Windows-based clusters helps our customers get more value from their product development process, resulting in improved time to market and reduced engineering costs. Support for software solutions from ANSYS on Windows HPC Server 2008 will be an important focus for us in our upcoming product releases.

Chris Reid, Vice President of Marketing, ANSYS, Inc.

FURTHER INFORMATION

For more information about Windows HPC Server 2008 and HPC please visit: http://www.microsoft.com/hpc

For more information about Mellanox technologies please visit: http://www.mellanox.com

For more information about ANSYS technologies please visit: http://www.ansys.com

NetworkDirect AND RDMA-STYLE NETWORKING

NetworkDirect provides an RDMA network architecture that can deliver dramatic network performance improvements for MPI applications. By using an architecture that bypasses TCP/IP overhead, NetworkDirect takes advantage of advanced InfiniBand capabilities and enables applications to realize the benefits of high-speed interconnects.

The Microsoft Message Passing Interface standard (MS-MPI) is a portable, flexible, vendor- and platform-independent standard for messaging within and between HPC nodes. The MS-MPI in Windows HPC Server 2008 uses a NetworkDirect interface for best performance and CPU efficiency. As shown in Figure 1, NetworkDirect uses a more direct path to support networking hardware, providing extremely fast and efficient networking.

NETWORK TOPOLOGY

Windows HPC Server 2008 supports five different network topologies to allow users to configure networking to suit their environment, infrastructure, and clustering needs. The Network Wizard, shown in Figure 2, simplifies configuring a network topology. Network topology options include:

- Compute nodes isolated on a private network
- All nodes on both public and private network
- Compute nodes isolated on private and MPI network
- All nodes on public, private, and MPI networks
- All nodes only on public network

A TYPICAL EXAMPLE

The ANSYS CFX Solver Manager monitors mathematical convergence as the ANSYS CFX solver runs. As shown in Figure 3, this latency-sensitive fluid dynamics application executes faster on the InfiniBand, NetworkDirect-enabled compute cluster nodes than on Gigabit Ethernet (GigE) compute nodes.