Abstract

Windows® HPC Server 2008 R2 was designed to provide ease of use, lower ownership costs, and high performance. This white paper provides an overview of Windows HPC Server 2008 R2 with Service Pack 2 (SP2), including a list of new features, an overview of the Microsoft® HPC solution (including new hardware and software requirements), and an overview of key capabilities in the areas of deployment, system management, job scheduling, the runtime for Service-Oriented Architectures (SOA), message-passing interface (MPI) and networking, big data workloads, Windows Azure™ Integration, Microsoft Excel® 2010 support, and security.

1 Includes Service Packs 1 and 2
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>What’s New in Windows HPC Server 2008 R2</td>
<td>3</td>
</tr>
<tr>
<td>Elements of the Microsoft HPC Solution</td>
<td>8</td>
</tr>
<tr>
<td>General Characteristics</td>
<td>9</td>
</tr>
<tr>
<td>Hardware Requirements</td>
<td>10</td>
</tr>
<tr>
<td>Software Requirements</td>
<td>11</td>
</tr>
<tr>
<td>Supported Network Topologies</td>
<td>12</td>
</tr>
<tr>
<td>Deployment</td>
<td>15</td>
</tr>
<tr>
<td>Graphical Deployment Tools</td>
<td>15</td>
</tr>
<tr>
<td>Windows Deployment Services</td>
<td>15</td>
</tr>
<tr>
<td>Node Templates</td>
<td>15</td>
</tr>
<tr>
<td>Windows Server 2008 and Windows Server 2008 R2 Compute Nodes</td>
<td>16</td>
</tr>
<tr>
<td>iSCSI Deployment</td>
<td>16</td>
</tr>
<tr>
<td>Enhanced Diagnostic Tests</td>
<td>16</td>
</tr>
<tr>
<td>Deployment Environment Validator</td>
<td>16</td>
</tr>
<tr>
<td>Preinstallation Validation</td>
<td>17</td>
</tr>
<tr>
<td>Workstation Nodes</td>
<td>17</td>
</tr>
<tr>
<td>Windows Azure Worker Nodes</td>
<td>17</td>
</tr>
<tr>
<td>Virtual Machine Roles in Windows Azure</td>
<td>18</td>
</tr>
<tr>
<td>Node Image Capture</td>
<td>18</td>
</tr>
<tr>
<td>System Management</td>
<td>19</td>
</tr>
<tr>
<td>Graphical Administration Console</td>
<td>19</td>
</tr>
<tr>
<td>Windows PowerShell and Command-Line Scripts</td>
<td>20</td>
</tr>
<tr>
<td>Remote Database Configuration</td>
<td>21</td>
</tr>
<tr>
<td>Monitoring at Scale</td>
<td>21</td>
</tr>
<tr>
<td>Extensible Reporting</td>
<td>24</td>
</tr>
<tr>
<td>Diagnostics</td>
<td>26</td>
</tr>
<tr>
<td>Software Updates Wizard</td>
<td>28</td>
</tr>
<tr>
<td>Job Scheduling</td>
<td>29</td>
</tr>
<tr>
<td>Interfaces</td>
<td>29</td>
</tr>
<tr>
<td>Policies</td>
<td>30</td>
</tr>
<tr>
<td>Job Workflows</td>
<td>31</td>
</tr>
<tr>
<td>Web Portals for Job Submission</td>
<td>32</td>
</tr>
<tr>
<td>Other Scheduling Enhancements</td>
<td>32</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Windows Azure Integration</td>
<td>34</td>
</tr>
<tr>
<td>Integrated Administration of Windows Azure Worker Nodes</td>
<td>35</td>
</tr>
<tr>
<td>Enabling Workloads on Windows Azure Nodes</td>
<td>36</td>
</tr>
<tr>
<td>SOA Runtime</td>
<td>38</td>
</tr>
<tr>
<td>When SOA Can Be Useful—and How It Works on a Cluster</td>
<td>38</td>
</tr>
<tr>
<td>Building SOA-Based HPC Applications</td>
<td>38</td>
</tr>
<tr>
<td>Running SOA-Based HPC Applications</td>
<td>40</td>
</tr>
<tr>
<td>Managing SOA-Based HPC Applications</td>
<td>41</td>
</tr>
<tr>
<td>Networking and MPI</td>
<td>44</td>
</tr>
<tr>
<td>Microsoft Message Passing Interface (MS-MPI)</td>
<td>44</td>
</tr>
<tr>
<td>NetworkDirect</td>
<td>45</td>
</tr>
<tr>
<td>Easier Troubleshooting of MPI Applications</td>
<td>46</td>
</tr>
<tr>
<td>Tuning Wizard for LINPACK</td>
<td>47</td>
</tr>
<tr>
<td>LINQ to HPC</td>
<td>48</td>
</tr>
<tr>
<td>Microsoft Office Excel 2010 Support</td>
<td>49</td>
</tr>
<tr>
<td>Speeding Up Excel 2010 Workbooks Using an HPC Cluster</td>
<td>49</td>
</tr>
<tr>
<td>Additional Benefits of Running Excel 2010 Calculations on a Cluster</td>
<td>51</td>
</tr>
<tr>
<td>Diagnostic Tests for Microsoft Excel 2010</td>
<td>52</td>
</tr>
<tr>
<td>Security</td>
<td>53</td>
</tr>
<tr>
<td>Conclusion</td>
<td>54</td>
</tr>
<tr>
<td>Additional Resources</td>
<td>54</td>
</tr>
</tbody>
</table>
INTRODUCTION

Windows® HPC Server 2008 R2 is the third version of the Microsoft® solution for high-performance computing (HPC). Built on Windows Server® 2008 R2 64-bit technology, Windows HPC Server 2008 R2 efficiently scales to thousands of nodes and integrates seamlessly with Windows-based IT infrastructures, providing ease of use, lower ownership costs, and high performance. This version adds even more functionality to the HPC solution, including integration with the Windows Azure™ Service Platform. Windows Azure’s public, flexible cloud computing platform offers on-demand, pay-as-you-go access to highly scalable compute and storage resources with 99.99 percent uptime.

For more information about Windows Azure, visit: http://www.microsoft.com/windowsazure/windowsazure.

Compared to previous versions, Windows HPC Server 2008 R2 and Windows HPC Server 2008 R2 with SP2 deliver significant improvements in several areas, including:

- **Improved scalability**—with out-of-the-box support for deploying, running, and managing clusters of 1,000 nodes or more.

- **New configuration and deployment options**—such as iSCSI (diskless) deployment, mixed-operating system version (Windows Server 2008 and Windows Server 2008 R2) clusters, and support for remote head node databases.

- **Windows Azure integration**—such as integrated administration, Virtual Machine (VM) Roles for Windows Azure, running MPI jobs, transparent scheduling, and transparent job submission. In addition, resources running within a Windows Azure data center can be used as temporary “compute nodes” to extend the capabilities of on-premise Windows HPC clusters.

- **Improved system management, diagnostics, and reporting**—including an enhanced heat map with multiple customizable tabs, an extensible diagnostic framework, and the ability to create richer custom reports.

- **Improved support for service-oriented architecture (SOA) workloads**—including a new fire-and-recollct programming model, finalization hooks, improved Java interoperability, automatic restart and failover of broker nodes, and improved management, monitoring, diagnostics, and debugging.

- **Message-passing interface (MPI) and networking enhancements**—including optimizations for new Intel “Nehalem”-based processors, enhanced support for RDMA over Ethernet and Infiniband, improved MPI debugging, and a LINPACK optimization wizard.

- **LINQ to HPC**—LINQ to HPC enables the development and deployment of big data applications.
- **New ways to speed up Microsoft Excel® 2010 workbooks**—such as support for running Excel user-defined functions and Excel workbooks on the nodes of an HPC cluster.

Through these enhancements, Windows HPC Server 2008 R2 makes it easier than ever for companies to benefit from high-performance computing. System administrators can more easily deploy and manage powerful HPC solutions, developers can more easily build applications, and end users can more easily access those solutions from their Windows-based desktops.

The remainder of this document provides a technical overview of Windows HPC Server 2008 R2, with topics addressed in the following order:

- A list of new features in Windows HPC Server 2008 R2 (including Service Packs 1 and 2).
- Solution architecture, hardware requirements, software requirements, and supported network topologies.
- An overview of each of the key functional areas of the product, including:
  - Deployment
  - System management
  - Job scheduler
  - SOA runtime
  - MPI and networking
  - Excel support
  - Security
WHAT’S NEW IN WINDOWS HPC SERVER 2008 R2

Released in 2008, Windows HPC Server 2008 included several new features, including Windows Deployment Services-based node provisioning, support for NetworkDirect, an enhanced Job Scheduler, a management console based on the Microsoft System Center 2007 user interface, and support for Windows PowerShell® scripting. Windows HPC Server 2008 R2 builds on many of those innovations, and Windows HPC Server 2008 R2 with SP2 delivers significant new features and functionality in the following areas:

- **Deployment**
  - Efficient deployment of more than 1,000 nodes
  - Support for deploying Windows Azure worker nodes in a Windows HPC Server Cluster
  - Support for VM Roles in Windows Azure (SP2)
  - iSCSI (diskless) network boot of compute nodes
  - Backwards compatibility with Windows HPC Server 2008 node templates
  - Support for compute nodes based on Windows Server 2008 and Windows Server 2008 R2 (including mixed-version clusters)
  - Support for remote HPC databases
  - Support for using idle workstations as compute nodes
  - A Deployment Environment Validator for finding common problems that can affect node deployment
  - Microsoft HPC Pack 2008 R2 Installation Preparation Wizard, a preinstallation validation tool (SP2)
  - Support for capturing a Windows Imaging Format (WIM) file for a node to use in deploying other nodes

- **System management – including diagnostics and reporting**
  - A customizable monitoring heat map, with at-a-glance viewing of system status for upwards of 1,000 nodes
  - Location-based node grouping, which allows administrators to view, categorize and batch operations on compute nodes based on their locations
  - Color-coded heat map overlays, with prioritized parameter display
Multiple, customizable tabs for different views of heat maps and other system data

An extensible diagnostic framework for implementing custom diagnostics

The ability to change diagnostic parameters at runtime

A richer reporting database and enhanced application programming interface (API) for building custom reports

A wizard for finding and applying software updates to node templates and nodes

Separate representation of node state and node health in the Node Management pane

User activity detection for workstation nodes

**Job scheduling**

Support for larger clusters, more jobs, and larger jobs—including improved scheduling and task throughput at scale

Just-in-time parametric sweep expansion, which improves performance for creating large parametric task sweeps

Web portal job submission and status (SP2)

New service-balanced scheduling policy optimized for SOA and other dynamic workloads

A new user interface and API for viewing and reporting job progress

New job scheduler policies for lending/borrowing resources pools and over/under subscribe nodes (SP2)

Command line and API support for prep and release tasks that run before and after a job to prepare and cleanup nodes

New service tasks, which start as many copies of the task as there are cores assigned to the job

A simplified user interface and experience for troubleshooting jobs

An improved job and node Template Editor

Support for defining 400 different job priority levels

E-mail notifications when users’ jobs start or complete

Enhanced activation filters for jobs
- Support for excluding specific nodes from a job
- Dynamic node groups, with which changes to node groups immediately impact queued jobs
- Mark tasks as critical
- Submit new jobs from a running job
- Submit jobs using Smartcard Authentication (SP2)

- **Windows Azure integration**
  - Add Windows Azure worker nodes (SP1)
  - Start and stop Windows Azure worker nodes on schedule (SP1)
  - Deploy SOA services, XLLs, and batch job assemblies to Windows Azure worker nodes (SP1)
  - Support for running some SOA services on Windows Azure nodes with no changes to the service or the client (SP1)
  - Node management using Cluster Manager (SP1)
  - Support for the VM Role (SP2)

- **Service-Oriented Architecture (SOA) runtime**
  - Fire and recollect programming model support
  - Cleanup interface for SOA services, which enable developers to add logic to perform cleanup before a service exits.
  - Improved interoperability with Java client applications.
  - New capabilities for managing SOA applications—in the areas of setup and configuration, monitoring, diagnostics, and event tracing
  - Broker node auto-restart and persistent storage of calculation results
  - Fail-over across broker nodes in the event of a hardware failure
  - Common data for sessions (SP2)
  - Support for multiple clients in SOA sessions
  - One to one mapping between jobs and SOA sessions, eliminating the need for separate session broker and SOA service jobs
- Support for canceling SOA service requests without canceling the current session
- Cluster-SOA debugging capabilities (provided as add-ons for Visual Studio® 2010)
- Enhanced tracing for SOA applications
- A diagnostic test for verifying that SOA services can be initialized and started

### Networking and Message-Passing Interface (MPI)
- Support for new networking options—including RDMA over Ethernet (iWARP) from Intel and RDMA over Infiniband quad data rate (40 Gbps) hardware
- Optimization of shared memory implementations for new Intel "Nehalem"-based processors
- Run MPI jobs in Windows Azure (SP2)
- Improved MPI debugging capabilities
- A LINPACK optimization wizard tool kit

### LINQ to HPC (SP2)
- Enables the development and deployment of big data workloads on Windows HPC Server 2008 R2 SP2 clusters
- LINQ to HPC includes a new runtime that is responsible for scheduling, executing, and managing big data analysis jobs
- LINQ to HPC allows developers to define distributed queries against distributed data and the Distributed Storage Catalog (DSC) on a cluster infrastructure

### Excel support
- Support for running Excel user-defined functions (UDFs) on a cluster
- Support for running Excel workbooks on a cluster
- Diagnostic tests for determining if Office Excel 2010 is installed and properly licensed, and for verifying that the UDF container service is loaded and ready
- Support for macros that partition iterative calculations into a fork-join pattern
Because it's based on the latest version of the Windows Server operating system, Windows HPC Server 2008 R2 also enables organizations to take advantage of the many enhancements in Windows Server 2008 R2, including:

- Performance improvements in the areas of Server Messaging Block (SMB) support, input/output, and thread scheduling
- Support for more logical processors (256 versus 64)
- Deployment related enhancements, such as support for multiband multicast in the Windows Deployment Services transport
Windows HPC Server 2008 R2 combines the underlying stability and security of Windows Server 2008 R2 with the features of Microsoft HPC Pack 2008 R2 to provide a robust, scalable, cost-effective, and easy-to-use HPC solution. A basic Windows HPC Server 2008 R2 solution is composed of a cluster of servers, with a single head node (or a primary and backup head node in a highly available configuration) and one or more compute nodes (see Figure 1). The head node controls and mediates all access to the cluster resources and is the single point of management, deployment, and job scheduling for the cluster. Windows HPC Server 2008 R2 can integrate with an existing Active Directory® directory service infrastructure for security and account management; and can use Microsoft System Center Operations Manager for data center monitoring.

**Note:** Windows HPC Server 2008 R2 is designed solely for use with HPC applications and should not be used as a general-purpose infrastructure server.
General Characteristics

General characteristics of the Windows HPC Server 2008 R2 solution are provided in Table 1.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Implementation</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system</td>
<td>Windows Server 2008 and/or Windows Server 2008 R2 (Head node is R2 only, compute nodes can be both)</td>
<td>Inherits security and stability features from Windows Server 2008 and Windows Server 2008 R2.</td>
</tr>
<tr>
<td>Processor type</td>
<td>x64 (AMD64 or Intel EM64T)</td>
<td>Large memory model and processor efficiencies of x64 architecture.</td>
</tr>
<tr>
<td>Node deployment</td>
<td>Windows Deployment Services</td>
<td>Image-based deployment, with full support for multicasting and iSCSI deployment.</td>
</tr>
<tr>
<td>Head node redundancy</td>
<td>Windows Failover Clustering and SQL Server® Failover Clustering</td>
<td>Provides a fully redundant head node and scheduler (requires Windows Server 2008 R2 Enterprise and SQL Server Standard Edition).</td>
</tr>
<tr>
<td>Management</td>
<td>Integrated Administration Console</td>
<td>Provides a single user interface for all aspects of node and job management, grouping, monitoring, diagnostics, and reporting.</td>
</tr>
<tr>
<td>Network topology</td>
<td>Network Configuration Wizard</td>
<td>Fully automated Network Configuration Wizard for configuring the desired network topology.</td>
</tr>
<tr>
<td>Application network</td>
<td>MS-MPI</td>
<td>High-speed application network stack using NetworkDirect. Shared memory implementation for multicore processors. Highly compatible with existing MPICH2 implementations.</td>
</tr>
<tr>
<td>Scheduler</td>
<td>Job Manager Console</td>
<td>GUI is integrated into the Administration Console or can be used standalone. Command line interface supports Windows PowerShell scripting and legacy command-line scripts from Windows Compute Cluster Server. Greatly improved speed and scalability. Support for SOA applications.</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Integrated into Administration Console</td>
<td>New heat map provides at-a-glance view of cluster performance and status for up to 1,000 nodes.</td>
</tr>
<tr>
<td>Reporting</td>
<td>Integrated into Administration Console</td>
<td>Standard, prebuilt reports and historical performance charts. Additional reports can be created using SQL Server Reporting Services.</td>
</tr>
<tr>
<td>Diagnostics</td>
<td>Integrated into Administration Console</td>
<td>Out-of-the-box verification and performance tests, with the ability to store, filter, and view test results and history. An extensible diagnostic framework for creating custom diagnostics and reports.</td>
</tr>
<tr>
<td>Parallel runtime</td>
<td>Enterprise-ready SOA infrastructure</td>
<td>Windows HPC Server 2008 R2 provides enhanced support for SOA workloads, helping organizations more easily build interactive HPC applications, make them more resilient to failure, and more easily manage those applications.</td>
</tr>
<tr>
<td>HPC Services for Excel</td>
<td>Parallel execution of workbooks and cluster offload of UDFs</td>
<td>Windows HPC Server 2008 R2 enables organizations to take advantage of HPC clusters to reduce calculation times for Excel workbooks.</td>
</tr>
</tbody>
</table>

Table 1. Windows HPC Server 2008 R2 features.
Hardware Requirements
Hardware requirements for Windows HPC Server 2008 R2 are similar to those for the x64-based editions of the Windows Server 2008 R2 Standard operating system.

Supported Processors
Supported processors include:

- AMD Athlon 64
- AMD Opteron
- AMD Phenom
- Intel Pentium with Intel EM64T
- Intel Core 2 Duo with Intel EM64T
- Intel Xeon with Intel EM64T

Multiprocessor Support
The number of sockets supported depends on which operating system is used:


Memory Requirements
The head node and compute nodes require a minimum of 512 megabytes (MB) of RAM, and support a maximum of 128 gigabytes (GB) of RAM.

Disk Space and Volumes
Minimum disk space required for setup is 50 GB. A single system volume is required for head and compute nodes. Redundant array of independent disks (RAID) is supported but not required. The system volume must be Master Boot Record (MBR). Additional volumes can be MBR or GUID Partition Table (GPT).

Note: Diskless nodes will have different requirements, the details of which are beyond the scope of this document.

Network Interface Cards
At least one network interface card (NIC) is required on the head node and each of the compute nodes. If a private network is used, the head node requires at least two NICs, and compute nodes require at least one NIC. Each node may also require a high-speed NIC for an application network.
Software Requirements
Windows HPC Server 2008 R2 is a two-DVD package. The first DVD contains the setup for Windows Server 2008 R2 HPC Edition (a 64-bit version of Windows Server 2008 R2 that is restricted to HPC workloads), and the second DVD contains Microsoft HPC Pack 2008 R2, which provides the additional interfaces, tools, and management infrastructure.

Head Node and Broker Nodes
Building the head node for a Windows HPC Server 2008 R2 solution involves installing Windows Server 2008 R2 on the server, joining the server to an Active Directory domain, and then installing HPC Pack 2008 R2. The operating system for the head node can be any of the following:

- Windows Server 2008 R2 HPC Edition
- Windows Server 2008 R2 Standard
- Windows Server 2008 R2 Enterprise
- Windows Server 2008 R2 Datacenter

Broker nodes also require one of the above Windows Server 2008 R2 editions.

Compute Nodes
Compute nodes can run either Windows Server 2008 or Windows Server 2008 R2. Specific versions of Windows Server that are supported include:

- Windows Server 2008 R2 HPC Edition
- Windows Server 2008 R2 Standard
- Windows Server 2008 R2 Enterprise
- Windows Server 2008 HPC Edition
- Windows Server 2008 Standard x64 Edition
- Windows Server 2008 Enterprise x64 Edition
- Windows Server 2008 R2 Datacenter

Database
Windows HPC Server 2008 R2 uses Microsoft SQL Server 2008 SP1 as a data repository for the head node. An existing SQL Server 2008 installation can be used, or the HPC Pack 2008 installer will install SQL Server 2008 Express SP1 (which is included on the HPC Pack 2008 R2 DVD). Windows HPC Server 2008 R2 SP2 adds support for SQL Server™ 2008 R2.

Note: If cluster size is greater than 100 nodes, Microsoft recommends using SQL Server 2008 Standard Edition or higher instead of SQL Server 2008 Express.
Head Node Failover Clustering

Windows HPC Server 2008 R2 can take advantage of the failover clustering capabilities provided in Windows Server 2008 R2 Enterprise and some editions of Microsoft SQL Server to provide high-availability failover clustering for the head node. With clustering, in the event of a head node failure, the Job Scheduler will automatically—or manually, if preferred—fail over to a second server. Job Scheduler clients see no change in the head node during the failover and fail-back processes, helping to ensure uninterrupted cluster operation. SQL Server Standard Edition or Enterprise Edition is required for head node failover clustering.

Remote Head Node Databases

Windows HPC Server 2008 R2 adds support for remote head node databases, enabling organizations to take advantage of an existing enterprise database.

Administrator Tools

During installation of the HPC Server 2008 R2 Pack, HPC Cluster Manager and HPC Job Manager are automatically installed on the head node. These same components can also be installed on other computers, as may be desired to support remote cluster management or job submission. Supported operating systems for installation of the remote components include:

- Windows Server 2003 with Service Pack 2 (SP2) or Windows Server 2003 R2 (32-bit or x64 versions)
- Windows Server 2008 or Windows Server 2008 R2 (32-bit or x64 versions)
- Windows XP Professional with SP3 (32-bit edition)
- Windows XP Professional with SP2 (x64 edition)
- Windows Vista® Business, Enterprise, and Ultimate editions with SP1
- Windows 7 Professional, Enterprise, and Ultimate editions

Supported Network Topologies

Windows HPC Server 2008 R2 supports five different network topologies, each with one to three network interface cards per node, enabling organizations to configure their HPC solutions to suit their unique needs. The Network Configuration Wizard in Figure 2 simplifies the process of configuring the desired network topology.
The five network topologies that are supported include:

1. **Compute nodes isolated on a private network.** The head node has two network interface cards (NICs), one connected to the enterprise network and the other to the private network, and may provide network address translation (NAT) for the compute nodes. The compute nodes have a single NIC, and are connected only to the private network.

2. **All nodes on both enterprise and private networks.** The head nodes and compute nodes each have two NICs, one connected to the enterprise network and one connected to the private network.

3. **Compute nodes isolated on private and application networks.** The head node has three NICs, one connected to the enterprise network, one connected to the private network, and one connected to the application network. Each compute node has one NIC connected to the private network and one connected to the application network. The head node may perform NAT between compute nodes and the enterprise network.

4. **All nodes on enterprise, private, and application networks.** Each node has three NICs. One NIC is connected to the enterprise network, one is connected to a private, dedicated cluster-management network, and one is connected to a high-speed, dedicated application network.
5. **All nodes only on enterprise network.** In this limited networking scenario, where each node has only a single NIC, the use of Windows Deployment Services to deploy compute nodes is not supported, and each compute node must be manually installed and activated.

*Note:* In topologies 1 and 3 above, where the compute nodes are isolated on a private network, network address translation (NAT) between the private and enterprise networks can be performed by the head node or a separate server.
One challenge to the adoption of HPC solutions stems from the time and effort required to deploy large clusters. It's not difficult or time consuming to deploy and configure a four-node cluster—someone just has to insert the DVD and press ENTER at each node and then add the nodes to the cluster. When deploying clusters of a few hundred or few thousand nodes, manual installation can become tedious and error-prone. As cluster scalability continues to improve, the challenges associated with manual deployment will inevitably increase.

With a design goal of supporting the deployment of 1,000 nodes in less than an hour, Windows HPC Server 2008 R2 builds on the capabilities provided by the Windows Deployment Services transport to simplify and streamline the deployment and updating of cluster nodes. Node templates are available in HPC Cluster Manager, for easily defining the configuration of compute nodes. New features in Windows HPC Server 2008 R2—such as support for Windows Server 2008-based mixed version clusters, and iSCSI deployment—provide additional flexibility, enabling organizations to easily deploy solutions that are optimized to meet their needs.

Graphical Deployment Tools
Graphical deployment tools are integrated into the Administration Console, allowing system administrators to quickly select nodes for deployment and easily monitor deployment progress. A to-do list shows the steps necessary to complete the configuration of a compute cluster, including defining the network topology, configuring automatic deployment (including image creation), and adding compute nodes to the cluster.

Windows Deployment Services
Windows HPC Server 2008 R2 uses the Windows Deployment Services transport in Windows Server 2008 R2, using WIM files and multiband multicast to rapidly deploy compute nodes in parallel. Other advantages of this approach include driver integration and updating, as well as support for non-uniform compute node deployment.

Node Templates
Node templates in Windows HPC Server 2008 R2 provide an easy way to define the desired configuration of compute nodes, with each node template including the base operating system image, drivers, configuration parameters, and, if desired, additional software. A Node Template Generation Wizard guides the administrator through the process of creating node templates, including support for injecting drivers into images. An improved template editor
provides advanced configuration capabilities, including configuring node templates for automatic application deployment.

Here’s how the process works: Windows Deployment Services running on the head node discovers compute nodes as they are turned on. Preinstalled compute nodes can be added to the cluster by importing a node list, or the administrator can choose to select nodes interactively as the Administration Console displays the discovered nodes. During either unattended or interactive deployment, the node is imaged and configured and applications are deployed according to the node template.

**Windows Server 2008 and Windows Server 2008 R2 Compute Nodes**

Windows HPC Server 2008 R2 supports the deployment of compute nodes and broker nodes based on Windows Server 2008 or Windows Server 2008 R2, including mixed-version clusters.

**iSCSI Deployment**

A new feature in Windows HPC Server 2008 R2, iSCSI deployment can deploy nodes on clusters that boot over the network by using an iSCSI connection. This new feature helps to centralize storage, and to deploy diskless nodes, computers that do not run the operating system from a local hard disk drive or that do not have a hard disk drive installed.

**Enhanced Diagnostic Tests**

The Administration Console includes diagnostic tests that can be used post-deployment to detect common problems, monitor node loading, and view job status across the cluster. In addition, the HPC Pack 2008 R2 Tool Pack enables administrators to heavily load the cluster—thereby providing an efficient mechanism for detecting issues related to configuration and deployment, networking, power, cooling, and so on.

**Deployment Environment Validator**

A new set of diagnostic tests help find common problems that can affect node deployment. These new tests verify connectivity with the Active Directory domain controller, availability of the DHCP server and the DNS server, and that the provided installation credentials have appropriate privileges to perform the node deployment tasks.
Preinstallation Validation
A free, stand-alone executable that will be made available during the Windows HPC Server 2008 R2 with SP2 timeframe on the tool pack download page. Microsoft HPC Pack 2008 R2 Installation Preparation Wizard is a tool that identifies environment and operating systems configuration settings that can cause HPC Pack installation, configuration, and node deployment issues.

Microsoft HPC Pack 2008 R2 Installation Preparation Wizard performs many similar and identical checks to the Deployment Environment Validator diagnostic tests that were introduced in HPC Server 2008 R2; the primary difference is that with this new wizard they can now be run before the HPC Pack is installed, and from any computer. The ability to run the check prior to HPC Pack installation helps identify issues that could block the installation, helping administrators better prepare for deployment.

Workstation Nodes
System administrators can add workstations running Windows 7 as compute nodes and use them to run jobs. Workstation nodes do not need to be dedicated cluster nodes and can be used for other tasks. They can be brought online manually as compute nodes or can be configured to come online automatically according to a weekly availability policy—for example, every night on weekdays and all day on weekends.

In Windows HPC Server 2008 R2, the client computer availability policy can be scheduled based on configurable user activity detection settings. By defining the days and times during which to use cycles, the number of minutes without keyboard or mouse input, and the CPU usage threshold for the computer to be considered idle, desktop cycles during working and non-working hours can be used without affecting the user.

Multi-Domain Support
In Windows HPC Server 2008 R2 with SP2, nodes from any trusted domain can be added to a cluster of workstations, providing that each domain is configured for two-way trust with the domain where the Head Node is joined. In previous versions of Windows HPC Server, all nodes in a cluster had to be joined to the same domain as the Head Node.

Windows Azure Worker Nodes
Systems administrators can supplement their on-premise cluster as needed by deploying additional Windows Azure worker nodes. Worker nodes can be brought online manually as compute nodes or can be configured to come online automatically according to a weekly availability policy—for example, every night on weekdays and all day on weekends.
Virtual Machine Roles in Windows Azure

With Windows HPC Server 2008 R2 with SP2, customers can run their HPC applications in Windows Azure using the Azure VM roles. Many HPC applications that can run on on-premise clusters can now be packaged with a Virtual Hard Disk (VHD) and then uploaded to Windows Azure, where they can run in the VM role. These applications do not have to be recertified for Windows Azure worker role VMs. Instead, they can be run in Windows Server 2008 R2 virtual machines.

Node templates for Windows Azure nodes have been updated to support VM roles. From a monitoring and management perspective, the support for VM roles is seamless and the HPC Cluster Manager Heatmap can be used to monitor Azure nodes running HPC Jobs in VM roles, just like worker roles.

Node Image Capture

System administrators can create a WIM file of an existing node, which can then be used to deploy other nodes in a cluster.
Another major challenge that organizations can face is the management and administration of HPC clusters. This has traditionally been a departmental or organizational-level challenge, requiring one or more dedicated IT professionals to manage and deploy nodes. At the same time, users submitting batch jobs are competing for limited HPC resources.

Windows HPC Server 2008 R2 is designed to facilitate ease-of-management. It provides a graphical Administration Console that puts all the tools required for system management at an administrator’s fingertips, including the ability to easily drill down on node details such as metrics, logs, and configuration status. Support for the Windows PowerShell scripting language facilitates the automation of system administration tasks. An enhanced "heat map" view provides system administrators with an “at a glance” view of cluster status, including the ability to define tabs with different views of system health and resource usage. Other new management-related features in Windows HPC Server 2008 R2 include additional criteria for filtering views, support for location-based node grouping, a richer reporting database for building custom reports, and an extensible diagnostic framework.

**Graphical Administration Console**

Based on the Microsoft System Center 2007 user interface, the Administration Console in Windows HPC Server 2008 R2 integrates every aspect of cluster management. The Administration Console has five navigation panes:

- **Configuration**, which includes the To Do List, Network Configuration Wizard, and Node Template Generation Wizard.
- **Node Management**, which is used to monitor node status and initiate node-specific actions such as deployment, bringing nodes offline or online, and adding or removing nodes.
- **Job Management**, which provides control of job scheduling and status.
- **Diagnostics**, which allows administrators to select a node or group of nodes and run diagnostic tests to validate network connectivity, job execution, configurations, performance, and so on—including the ability to view the progress of tests and view past test results.
- **Charts and Reports**, which displays standard reports in support of both scheduled and on-demand reporting.
The Administration Console also provides support for pivoting between these navigation panes for effective contextual monitoring. For example, when using the Job Management pane to view a job that has failed, an administrator can easily pivot to the Node Management Pane to view the status of the nodes upon which that job ran. Similarly, from the Node Management pane, the administrator can select a set of nodes and pivot to the Diagnostics pane to view the diagnostics that have run on the nodes. Furthermore, operations that are performed using the graphical Administration Console also can also be performed from the command line using Windows PowerShell.

Windows PowerShell and Command-Line Scripts

Windows HPC Server 2008 R2 supports Windows PowerShell as a scripting language and shell for managing clusters and jobs, providing more than 80 cmdlets for job submission and monitoring, deployment, diagnostics, and node management functions. An admin-focused scripting language with consistent syntax and utilities, Windows PowerShell accelerates automation of system administration tasks and helps improve an organization's ability to address its unique system management challenges.
Windows PowerShell works with Microsoft .NET objects and supports Windows Management Instrumentation (WMI), combining the power of an object-oriented language with the ease-of-use of an interactive shell. For example, a system administrator could use simple PowerShell script to get a list of all nodes with more than 1 GB of free disk space, or to group together all nodes with at least eight processors and submit a job to that group of nodes.

Windows HPC Server 2008 R2 also maintains compatibility with existing Windows Compute Cluster Server 2003 command-line scripts.

**Remote Database Configuration**

During the installation process of HPC Pack 2008 R2, Microsoft® SQL Server™ 2008 SP1 databases can be configured on computers that are not the head node of a HPC cluster, and then used host cluster management, job scheduling, reporting, and diagnostics information.

**Monitoring at Scale**

The Node Management pane within the Administration Console is used to monitor node status and initiate node-specific actions. New node management-related features in Windows HPC Server 2008 R2 include an enhanced heat map with overlay view, additional filtering criteria, customizable tabs, and location-based node grouping.

**Enhanced Heat Map with Overlay View**

The Heat Map view, illustrated in Figure 4, resides within the Node Management pane. In Windows HPC Server 2008 R2, the heat map has been enhanced to provide an at-a-glance view of system health and performance for clusters upwards of 1,000 nodes, without scrolling. System administrators can define and prioritize up to three metrics (as well as minimum and maximum thresholds for each metric) to build customized views of cluster health and status.
Each node in the cluster is represented individually in the heat map for up to three performance counters. Each counter is assigned a color range (for example, light red to dark red) to represent the performance counter's value at that instance. With the "stacked" heat map view, up to three performance counters can be viewed for each node at once. With the "overlay" heat map view, only one performance counter's color is shown at any one time, as determined by the priority ordering for the counters assigned by the administrator. In both viewing modes, the administrator can identify nodes with outlier performance counters at-a-glance, thereby helping to facilitate effective monitoring.
**Additional Filtering Criteria**

When managing large clusters, it can be helpful for system administrators to filter views of compute nodes by certain criteria. Windows HPC Server 2008 R2 adds to the filtering options provided in the previous version with additional filtering options, including advanced filtering by Node Group, Node State, and a node’s physical location.

**Customizable Tabs**

Administrators now can create multiple, customizable tabs within the Node Management pane—as may be desired to view different types of parameters. For example, a system administrator could configure one tab to provide a heat map view of networking parameters, a second tab to provide a heat map view of job usage, and a third tab to provide a list view of failed operations.

![Figure 5. Customizable tabs.](image)

**Location-based Grouping**

System administrators sometimes need to view the nodes in a cluster according to their physical location. For example, an administrator may want to
see the heat map for a set of nodes contained within a single rack, to view performance metrics that may indicate a power or cooling issue. Windows HPC Server 2008 R2 enables administrators to define up to three levels of hierarchy within the physical layout of a HPC cluster—for example, grouping together nodes in different chassis—and then view compute nodes based on their locations within that hierarchy.

Enhanced Node State and Node Health View

In the Node Management pane, node state and node health are represented separately, using different sets of icons. Nodes that are in an offline state are no longer marked with a warning by default. A health tab shows the specific health issues for each node, including connectivity and service health.

Extensible Reporting

Windows HPC Server 2008 R2 provides both built-in reports and an extensible reporting framework, which allows for the development of custom reports using tools such as SQL Server Reporting Services.
Built-in Reports
Windows HPC Server 2008 R2 provides a set of prebuilt reports and charts to help system administrators understand system status, usage, and performance. Accessed through the Reports and Charts tab on the Administrator Console, these prebuilt reports span four main categories:

- Node availability
- Job Resource usage
- Job throughput
- Job turnaround

Reporting Extensibility
In response to customer demand for more detailed reports, Windows HPC Server 2008 R2 collects far more data than its predecessor, including:

- Node attributes
- Node event history
- Job configurations
- Job task summaries
- Job state and statistics
- Job allocation history

The data is stored in the reporting database and can be exported to a separate database or data warehouse for analysis. Tools such as SQL Server Reporting Services or Excel 2010 can be used to create custom reports and graphs for scenarios such as daily operational analysis, charge-back reporting, cluster utilization and analysis, and capacity planning. The data is made available through predefined database views allowing administrators to easily build their own reports. Node performance metric history can also be retrieved using PowerShell commands.
Diagnostics

In the past, cluster administrators often had to write and maintain custom scripts for troubleshooting. Windows HPC Server 2008 R2 eases this pain through a set of commonly-used diagnostics tests, as well as a new extensible diagnostic framework to give system administrators a comprehensive, easy-to-use set of diagnostic tools.

Built-in Diagnostics

Windows HPC Server 2008 R2 provides a set of prebuilt diagnostic reports to help system administrators verify that their clusters are working properly, along with a systematic way of running the tests and storing and viewing results. This significantly improves an administrator’s experience in verifying deployment, troubleshooting failures, and detecting performance degradation. Cluster administrators can view a list of these diagnostic tests, run them, change diagnostic parameters at runtime, and view the results using the Diagnostics tab in the Administration Console or by using Windows PowerShell™ commands.

Figure 7. Diagnostic Pane.
Extensible Diagnostic Framework
New in Windows HPC Server 2008 R2, an extensible diagnostic framework enables cluster administrators, developers, and HPC industry partners to easily create custom diagnostic tests—as may be required to verify that custom and/or third-party hardware or software is working correctly. Independent hardware and software vendors can use this capability to create their own diagnostic tests, which cluster administrators can add to the list of out-of-the-box diagnostic tests in Windows HPC Server 2008 R2 and run in the same way as the built-in diagnostic tests—thereby helping to reduce support calls and increase customer satisfaction.

Software Updates Wizard
A new Add Software Updates Wizard helps system administrators search for software updates based on a node template, update the template with available updates, and optionally install the software updates on nodes that require them.
The Job Scheduler queues jobs and their associated tasks, allocates resources to the jobs, initiates the tasks on the compute nodes, and monitors the status of jobs and tasks. In Windows HPC Server 2008 R2, the Job Scheduler has been enhanced to support larger clusters, more jobs, and larger jobs—including improved scheduling and task throughput at scale. It includes new policies for greater flexibility and resource utilization, and is built to address both traditional batch jobs as well as newer service-oriented applications.

With the Windows HPC Server 2008 R2 Job Scheduler, jobs are considered to be resource requests and tasks specify what to do with those resources once they are obtained. Using the graphical Administration Console or any of several other interfaces (see Interfaces below), users can easily schedule jobs, allocate resources needed for the job, and change the tasks and properties associated with the job. Jobs can be single tasks or multiple tasks, can span multiple cores or nodes, and can specify whether or not they require exclusive access to nodes.

The Job Scheduler also includes built-in support for parametric sweeps, customized policies for filtering jobs at submission and activation times, and heterogeneous clusters. The Job Scheduler is Non-Uniform Memory Architecture (NUMA)-aware and multicore-aware, allowing for the intelligent scheduling of jobs on large clusters of multicore nodes at the processor core, processor socket, and compute node levels.

Interfaces

The Job Scheduler supports both command-line and graphical interfaces. The graphical interface is provided through the Job Scheduling tab of the Administration Console or through the HPC Job Manager, an interface for use by end-users submitting and managing jobs. Other supported interfaces include:

- Command line (cmd.exe).
- Windows PowerShell 2.0.
- COM and .NET application programming interfaces to support a variety of languages, including VBScript, Perl, Fortran, C/C++, C#, and Java.
- The Open Grid Forum’s HPC Basic Profile Web Services Interface, which supports job submission and monitoring from many platforms and languages.

The Windows HPC Server 2008 R2 interfaces are fully backwards compatible, allowing job submission and management from Microsoft Compute Cluster Server and Windows HPC Server 2008 interfaces.
As demonstrated in Figure 8, Windows HPC Server 2008 R2 also provides a new user interface for showing job progress, and an enhanced API that enables developers to report more detailed job progress status to their HPC applications.

![Job progress display](image)

**Figure 8. Job progress display.**

**Policies**

Scheduling policies determine how resources are allocated to jobs. Windows HPC Server 2008 R2 provides the ability to switch between traditional first-come, first-serve scheduling and a new service-balanced scheduling policy designed for SOA/dynamic (grid) workloads, with support for preemption, heterogeneous matchmaking (targeting of jobs to specific types of nodes), growing and shrinking of jobs, backfill, exclusive scheduling, and task dependencies for creating workflows.
In Windows HPC Server 2008 R2 with SP2, new job scheduler policies have been introduced:

- **Lend/Borrow Resource Pools.** A job scheduler policy that allows administrators to allocate—based on user demand and subject to availability of idle resources—a specified number of cores to groups of users. These cores constitute the “pool of resources” available to that group of users.
  
  - **Lending:** If the current demand for cores by a user group is smaller than their allocation, then the scheduler can lend the remaining cores to one or more other groups.
  
  - **Borrowing:** If the current demand for resource by a user group is larger than their guaranteed allocation, then the usage over the allocation can be borrowed if other groups are not using their full allocation.

- **Over-subscription/under-subscription nodes.** Over-subscription or under-subscription allows administrators to more finely control how tasks are scheduled on nodes and how much resources are consumed by the tasks.

**Job Workflows**

The Job Scheduler supports basic tasks, such as those that are executed from the command line, as well as parametric sweeps. Just-in-time parametric sweep expansion, a new feature in Windows HPC Server 2008 R2, improves performance for creating large parametric task sweeps by creating individual parameter sweep steps as needed rather than at submission time.

Windows HPC Server 2008 R2 also introduces prep and release tasks—tasks that run before and after a job. Prep tasks are guaranteed to run once on each node in a job before any other tasks, as may be required to support setup or validation of a node before the job is run. Release tasks are guaranteed to run once on each node in a job after all other tasks, as may be required to clean up or transfer files after the job.

When a running task is cancelled, the HPC Node Manager Service stops the task by sending a CTRL+BREAK event to the application, which then has a configurable period of time to exit gracefully. The application can use this time to save state information, write a log message, create or delete files, or finish the current service call.
Windows HPC Server 2008 R2 provides a redesigned user interface for displaying job progress, as well as changes to the application programming interface to help applications report more detailed job progress. The user interface for job and troubleshooting also has been improved to make it easier to review errors. Detailed information about a failed job can be viewed by double-clicking on that job in the HPC Cluster Manager or HPC Job Manager.

In addition, Windows HPC Server 2008 R2 includes improvements to the way that user credentials are handled—including the ability to update user credentials once for all of the user’s jobs that are in the queue.

Web Portals for Job Submission
With a web portal, users can easily submit jobs, monitor progress, view results, and troubleshoot their jobs on Windows HPC clusters from any workstation, including those not running a Windows client operating system, through a standard web browser. Users do not need HPC client software, or to know a cluster name to submit their jobs.

Administrators can customize what information and options are available to different users and for different types of jobs. Independent software vendors (ISVs) and system integrators (SIs) can customize the job portal to include application-specific information, input (parameters, environment variables, localized content), and workflow controls.

Other Scheduling Enhancements
Windows HPC Server 2008 R2 also includes the following scheduling-related enhancements:

- **Smartcard Authentication.** This feature enables users who are required to use smart cards to submit jobs using their smartcard certifications without having to enter a user name and password combination.

- **Dynamic Node Groups.** Changes to node groups immediately impact queued jobs. This enables the creation of tools that automatically move nodes between groups to handle different computational loads, or that move nodes based on time of day.

- **Service Balanced Scheduling.** This release includes a new scheduling mode that optimizes the process of starting jobs, and balances in real time the resources that are assigned to jobs, according to their priority. To enable this functionality, run the following command on the head node of your HPC cluster: `cluscfg setparams schedulingmode=balanced`. 


- **Service Tasks.** A new method to conclude service tasks is now available. This method tells the job scheduler to stop creating new instances of the service task. Additionally, the method includes an option to indicate if existing instances of the task should be allowed to complete or should be canceled.

- **More Granular Job Priority Levels.** Cluster administrators that need finer control over job priorities can now define 400 priority levels. The 5 priority levels that were available with previous releases are still available in this release.

- **E-mail Notifications.** Cluster users can choose to receive e-mail notifications when their jobs start or complete.

- **Enhanced activation filters.** Activation filters can implement additional exit codes to block the queue until the job can start, reserve resources for the job without blocking the queue, put the job on hold, or reject the job.

- **Node Exclusion Listing.** In this release, specific nodes can be excluded from running a job. This new feature can help avoid nodes that have a particular configuration or other characteristics that are not appropriate to run the job, or that are known to have intermittent problems running specific types of jobs. The list of excluded nodes is defined in the ExcludedNodes property of each job, and is empty by default.

- **Mark tasks as critical.** In Windows HPC Server 2008 R2, job owners can mark a specific task in a job as critical. If that task fails, then the whole job should fail. To mark a task as critical, set the task property failJobOnFailure to True. This property can be set using the API, the command line, or a Windows PowerShell script.

- **Submit new jobs from a running job.** In Windows HPC Server 2008 R2, job owners can configure and submit new jobs from running tasks without having to supply their password for each new job. A job owner can run cluscfg setcreds on the client computer to specify that the credentials can be reused for later job submission. Cluster administrators can disable this behavior across the cluster by setting the cluster property DisableCredentialReuse to True.
Cloud platforms offer a new approach to deploying, using, and paying for computing resources. Windows Azure provides on-demand access to computational resources and storage. Windows HPC Server 2008 R2, allows administrators to supplement their on-premise cluster as needed by deploying additional Azure nodes. Since Windows Azure provides on-demand access to computational resources and storage, costs for the additional nodes only accrue while they are running. Once the jobs are complete, the Azure worker nodes can be stopped through HPC Cluster Manager.

Windows HPC Server 2008 R2 should be installed and fully configured on an HPC head node computer that has Internet connectivity, allowing it to connect to Windows Azure services over the Internet. Adding worker nodes requires that administrators obtain a Windows Azure subscription account or have an existing Windows Azure subscription in which they can configure hosted services and a storage account.

Windows HPC Server 2008 R2 includes node templates for Windows Azure worker nodes that can be used to specify Windows Azure subscription information in addition to an availability policy for the worker nodes. The node template can be used to deploy a set of Azure nodes from HPC Cluster Manager.

To enable Windows HPC Server 2008 R2 to authenticate properly with the Windows Azure subscription, the Windows Azure subscription must be configured with a public subscription certificate (also called an API certificate).
Integrated Administration of Windows Azure Worker Nodes

For the most part, HPC administrators can treat the Windows Azure worker nodes as regular compute nodes of an on-premise cluster using the same administration tools. For instance, administrators can:

- View the integrated heat map with Windows Azure on-premise and client computer nodes in the same node management pane.
- Run diagnostics on Windows Azure nodes like they would be run on the on-premise nodes.
- Create job templates and node groups with Windows Azure nodes as with on-premise nodes.
- View and create new reports for Windows Azure worker nodes.
In addition to the existing administrator capabilities, for Windows Azure worker nodes, administrators can:

- Specify size, subscription, service, and storage accounts when adding Windows Azure nodes using node templates.
- Choose to have a preset scheduler for when the Windows Azure nodes are available or choose to start and stop the Windows Azure worker node instances manually.

**Enabling Workloads on Windows Azure Nodes**

All SOA jobs, UDF offloading, and parametric sweep–type jobs can be enabled for Windows Azure as long as they can run on the Windows Azure worker node.

*Note:* A Windows Azure worker role instance runs a guest operating system that is substantially compatible with Windows Server 2008 R2.

An administrator has the choice of running jobs:

- Entirely on on-premise clusters.
- Entirely in Windows Azure.
- Partially in Windows Azure and partially on on-premise clusters at the same time.

The third choice enables the “cloud burst” scenario, where jobs are primarily on the on-premise cluster but expand to Windows Azure during peak loads.

*Note:* In all cases, the head node has to run on premise.

**Job Submission and Scheduling on Windows Azure Worker Nodes**

After Windows Azure worker nodes added to the HPC cluster have been started and are online, they are available to run cluster jobs. All Windows Azure nodes when online are available through the default job template. An administrator can restrict access to these nodes by removing them from the default template and adding them to another job template.

- **Transparent job submission.** To a user, submitting a Windows HPC Server job that runs partly on premises and partly on a Windows Azure worker node looks no different from one that runs entirely in an on-premise cluster.

- **Transparent scheduling.** To schedule jobs on Windows Azure worker nodes, administrators simply specify (using a graphical wizard) how many Windows Azure nodes are needed and at what times. Windows HPC Server starts the Windows Azure instances, installs the required software, and schedules the jobs on the Windows Azure nodes. Administrators also have
the option of manually starting and stopping the Windows Azure worker
nodes.

**Deploying SOA Services, XLLs, and Batch Job Assemblies to Windows 
Azure Worker Nodes**

Windows HPC Server 2008 R2 includes hpcPack and hpcSync to help
administrators deploy files to Windows Azure worker nodes. Administrators can
use hpcPack to package and upload files to a Windows Azure storage account.
When provisioning a set of Windows Azure worker nodes from HPC Cluster
Manager, any applications or files on the storage account are automatically
deployed to the worker nodes. If an administrator uploads file packages to
storage after the worker nodes are started, the administrator can use hpcSync
to manually deploy them to the worker nodes.

**Running Some SOA Services on Windows Azure Nodes with No Changes 
to the Service or the Client**

In Windows HPC Server 2008 R2, services that get data through message
requests can run on Windows Azure worker nodes with no changes to the
service or the client. Services that require access to databases or other
external data sources must include code that uses the Windows Azure APIs to
access data. Windows Azure worker nodes cannot access on-premise nodes or
shares directly.

**Running MPI Jobs in Windows Azure**

With Windows HPC Server 2008 R2 with SP2, users who need to run multiple
simulations (or runs) of their MPI applications sequentially on their workstations
can now run them in parallel in Windows Azure. More information about running
MPI jobs in Windows Azure is provided in the Networking and MPI section of
this document.
With the number and size of problems being tackled on ever-larger clusters continuing to grow, organizations face increased challenges in developing HPC applications. Not only must these applications be built quickly, but they must run efficiently and be managed in a way that optimizes application performance, reliability, and resource utilization.

One approach to meeting these challenges is a service-oriented architecture (SOA)—an approach to building distributed, loosely coupled applications in which functions are separated into distinct services that can be distributed over a network, combined, and reused. Windows HPC Server 2008 R2 provides enhanced support for SOA workloads, helping organizations more easily build interactive HPC applications, make them more resilient to failure, and more easily manage those applications—capabilities that open the door to new application scenarios in areas such as financial trading and risk management.

When SOA Can Be Useful—and How It Works on a Cluster

HPC applications submitted to compute clusters are typically classified as either message intensive or embarrassingly parallel. While message-intensive applications comprise sequential tasks, embarrassingly parallel problems can be easily divided into large numbers of parallel tasks, with no dependency or communication between them. To solve these embarrassingly parallel problems without having to write low-level code, developers need to encapsulate core calculations as software modules. A SOA approach to development makes this encapsulation not only possible but easy, effectively hiding the details of data serialization and distributed computing.

With Windows HPC Server 2008 R2, tasks can run interactively as SOA applications. For interactive SOA applications, in addition to a head node and one or more compute nodes, the cluster also includes one or more Windows Communication Foundation broker nodes. The broker nodes act as intermediaries between the client application and the Windows Communication Foundation hosts running on compute nodes, load-balancing the client application’s requests and returning the results to it.

Building SOA-Based HPC Applications

One attractive aspect of SOA applications is the ability to develop them quickly, without having to write a lot of low-level code. To achieve this, developers need to be able to easily encapsulate core calculations as software modules that can be deployed and run on the cluster. These software modules identify and marshal the data required for each calculation and optimize performance by minimizing the data movement and communication overhead.
Microsoft Visual Studio provides easy-to-use Windows Communication Foundation service templates and service referencing utilities to help software developers quickly prototype, debug, and unit-test SOA applications, with Windows Communication Foundation effectively hiding the complexity of data serialization and distributed computing.

**Fire-and-Recollect Programming Model**
A fire-and-recollect programming model—sometimes called *file-and-forget*—is a common approach to building long-running SOA applications. Windows HPC Server 2008 R2 supports offline SOA applications, including client disconnect (SOA batch) and client resilience.

**Durable Sessions**
Another new feature in the Windows HPC Server 2008 R2 is the ability to implement durable sessions, where the SOA runtime persists requests and their corresponding responses on behalf of the client.

**Finalization Hooks**
The SOA runtime in Windows HPC Server 2008 R2 also adds support for finalization hooks, enabling developers to add logic to perform cleanup before a service exits—for example, as may be required to release a remote file, database connection, or COM object.

**Cleanup Interface for SOA Services**
In this release, a new `OnExiting()` interface is available for services. Developer code can now register this event and clean up resources before the calculation is cancelled (for example, release a remote file, database connection, or COM object).

**SOA Service Versioning Support**
SOA service versioning allows multiple versions of the same Windows Communication Foundation (WCF) service to be installed independently on a HPC cluster. SOA clients can now query the runtime for the available versions of a given service, and then specify which version of the service should be used when creating a SOA session. Alternatively, a SOA client can create a session without specifying a service version, and the latest version of the service is used for the session.

**Windows Web Services API**
The Windows Web Services application programming interface (WWSAPI) is a new framework in Windows HPC Server 2008 R2 that enables the development of native SOA services. Windows HPC Server 2008 R2 supports the ability to build fully native WSSAPI services that run across a HPC cluster.
**Improved Java Interoperability**
With Java sample code provided in the Windows HPC Server 2008 R2 Software Development Kit (SDK), developers can more easily write Java-based client applications that communicate with .NET services—and enjoy the same level of functionality provided with clients based on the .NET Framework and Windows Communication Foundation.

**Support for Multiple Clients in SOA Sessions**
The session API supports tagging every batch of computation with a GUID, enabling multiple client applications to share the same SOA session by identifying a GUID when sending computation requests and retrieving results.

**Cancellation of Running SOA requests**
A new `Cancel()` interface provides the ability to cancel SOA service requests without canceling the current session, thereby saving calculation resources.

**C# Cluster–SOA Debugger Add-in for Visual Studio**
The SOA debugger is now available as an add-in to the Visual Studio C# project system. It extends the remote debugger functionality and simplifies the process of debugging cluster SOA services in a Windows HPC cluster.

**Running SOA-Based HPC Applications**
In addition to developing SOA applications quickly, organizations must be able to run those applications efficiently, securely, and reliably. The SOA runtime in Windows HPC Server 2008 R2 helps organizations meet those needs through features such as low-latency round-trips for efficiently distributing short calculation requests, end-to-end Kerberos authentication with Windows Communication Foundation transport-level security, and dynamic allocation of resources to service instances. Windows HPC Server 2008 R2 also provides several new features to help organizations more reliably run their SOA applications, including support for broker restart/failover and message persistence.

**Message Resilience**
In the case of a temporary broker node failure or a catastrophic failure of the cluster, the SOA broker nodes will persist calculation requests and results. The session can continue without lost requests or results after the cluster recovers and the broker nodes are restarted.
**Built-in Flow Control**
The new SOA session API and the new broker Web service interface in this release include a built-in flow control. With this feature, there is no need to implement throttling behavior, as was necessary in previous releases.

**Custom Binding Support**
By default, the SOA broker in Windows HPC Server 2008 R2 exposes client-facing endpoints that use standard WCF NetTcpBinding and BasicHttpBinding bindings. In Windows HPC Server 2008 R2, a SOA broker can also be configured to expose endpoints using a custom binding. This enables clients to interact with your HPC cluster by using a variety of protocols.

**Improved Broker Node Failover**
In this release, when a broker node is configured in a failover cluster, a SOA session can continue to run if the broker node fails because it will be migrated to the failover broker node.

**Single-Job SOA Sessions**
It is no longer necessary to have two jobs running for each SOA session (one for the SOA session broker, and one for the service). This one-to-one mapping between jobs and sessions can make monitoring and reporting simpler.

**Exclusion of Compute Nodes that Fail SOA Tasks**
SOA sessions will exclude compute nodes that keep failing SOA service tasks or requests. This functionality is based on the new node exclusion listing feature for job scheduling (discussed above under Job Scheduling).

**Common Data**
The Common Data feature in SP2 improves performance by offering APIs that allow HPC applications to send—ahead of time—data required by SOA service codes running on compute nodes. The data can be stored in a file system and the administrator may choose the storage location. Windows HPC Server can also manage the lifecycle, or retention period, for the data.

**Managing SOA-Based HPC Applications**
Organizations running SOA applications on a cluster need to be able to manage and monitor those applications in one place—in a way that maximizes application performance, reliability, and resource usage. Windows HPC Server 2008 R2 provides comprehensive tools for managing SOA applications through service resource usage reports and runtime monitoring of performance counters, including the number and status of outstanding resource calls and the resources used by those services.
Windows HPC Server 2008 R2 also includes new capabilities designed to facilitate system management for SOA applications, including: easier setup and configuration; simpler monitoring; new diagnostic tests for common problems; and correlated tracing.

**Easier Setup and Configuration**
Windows HPC Server 2008 R2 provides several new features that aid in setup and configuration, including:

- Out-of-the-box setup of the cluster head node as a broker node.
- Support for broker nodes in Node Templates.
- Service configuration in a single place using the Administrator Console.

**Enhanced Monitoring**
Monitoring-related enhancements provided by Windows HPC Server 2008 R2 include:

- The ability to monitor a session as a single job rather than managing separate broker and service jobs—including through the command line interface.
- A job manager progress bar, which eliminates the need to track counters to view session progress.
- A Session page in the job dialog, showing session progress in detail.
- More counters for monitoring broker node performance.
- Easily enabling full capture of service host output.

**Improved Diagnostics**
Windows HPC Server 2008 R2 provides new diagnostic capabilities for SOA applications, helping administrators more easily detect issues and determine whether they are environmental or application related—and helping developers more easily diagnose application issues. Those capabilities include:

- Enhanced environmental tests, including out-of-the-box support for diagnosing common environmental issues.
- The ability to implement custom diagnostic tests using the new diagnostic framework (previously described under System Management).
**Enhanced Tracing**

Windows HPC Server 2008 R2 provides a new interface for service code to write a user-level trace. HPC Cluster Manager includes a new user interface to configure service tracing. New user interface and new PowerShell cmdlets are provided to collect and remove traces, and trace output has been modified so that it is easier to review in the Service Trace Viewer.

**SOA Service Loading Test**

System administrators can run a new diagnostic test that attempts to load a SOA service and verify that it can be initialized and started. This can help detect common configuration and environment issues that may cause errors (such as incorrect firewall configuration, network issues, msvcrt errors, or problems with the service registration file).
A message-passing interface (MPI) is the application interface between the nodes of an HPC cluster, providing a portable and powerful mechanism for interprocess communication among hundreds or thousands of processors working in parallel.

**Microsoft Message Passing Interface (MS-MPI)**

Windows HPC Server 2008 R2 uses the Microsoft Message Passing Interface (MS-MPI), a portable, flexible, interconnect-independent API for messaging within and between HPC nodes. MS-MPI is based on the Argonne National Laboratory open-source MPICH2 implementation, and is compatible with the MPI2 standard.

MS-MPI can run over Gigabit Ethernet, 10 Gigabit Ethernet, and high-performance networking hardware such as Infiniband, iWARP Ethernet, and Myrinet—or any other type of interconnect that provides a Winsock Direct, NetworkDirect, or TCP/IP interface. MS-MPI includes application support (bindings) for the C, Fortran77, and Fortran90 programming languages. With Windows HPC Server 2008 R2, organizations also can take advantage of new interconnect options, such as support for RDMA over Ethernet (iWARP) from Intel and new RDMA over Infiniband quad data rate (40 Gbps) hardware.

MS-MPI is optimized for shared memory communication to benefit the multicore systems prevalent in today’s HPC clusters. MS-MPI in Windows HPC Server 2008 R2 introduces optimization of shared memory implementations for new Intel “Nehalem”-based processors, with internal testing by Microsoft showing up to a 20 to 30 percent performance improvement on typical commercial HPC applications.

**MPI Runtime in Windows Azure**

In SP2, users who need to run multiple simulations (or runs) of their MPI applications sequentially on their workstations can now run them in parallel in Windows Azure.

This feature is best suited for jobs that complete relatively quickly, can fit in a single physical instance, do not require or produce large amounts of data, and for users who are currently running their MPI applications on a personal workstation or very small clusters.

*Note: At the time of writing this document, MPI support is available in VM role only. Writing to STDIN, STDOUT and STDERR is not supported.*

More information about using the Microsoft Message Passing Interface is available at [http://go.microsoft.com/fwlink/?LinkId=55930](http://go.microsoft.com/fwlink/?LinkId=55930).
**NetworkDirect**

MS-MPI can take advantage of NetworkDirect—a remote direct memory access (RDMA)-based interface—for superior networking performance and CPU efficiency. As shown in Figure 9 (next page), NetworkDirect uses a more direct path from MPI applications to networking hardware, resulting in very fast and efficient networking. Speeds and latencies are similar to those of custom, hardware-native interfaces from hardware providers.

---

**Figure 10. NetworkDirect architecture.**
Easier Troubleshooting of MPI Applications

MS-MPI integrates with Event Tracing for Windows to facilitate performance-tuning, providing a time-synchronized log for debugging MPI system and application events across multiple computers running in parallel. In addition, Microsoft Visual Studio® 2008 includes a MPI Cluster Debugger that works with MS-MPI. Developers can launch their MPI applications on multiple compute nodes from within the Visual Studio environment, and Visual Studio will automatically connect the processes on each node, enabling developer individually pause and examine program variables on each node.

Improvements in the upcoming Visual Studio 2010 release will further help developers build parallel programs that efficiently take advantage of HPC clusters—especially when using Microsoft’s latest Visual Studio 2010 programming models for node-level concurrency. Key enhancements include:

- **F5 Experience for MPI Cluster Debugger.** Applications built using MS-MPI can be deployed and debugged on the cluster via the usual F5 keystroke in Visual Studio, without having to do any prep work on the cluster nodes. Using a simple interface in the Project Properties page and a rich Node Selector dialog, developers can pick the deployment configuration they prefer and debug multiple processes concurrently on multiple nodes, just like when debugging a single process on a client machine. When done debugging, cleanup on the cluster’s compute nodes is also done automatically for the developer without them having to manually connect to each node to perform such tasks. A video showing the MPI Cluster Debugger in action can be found at [http://channel9.msdn.com/posts/DanielMoth/VS2010-MPI-Cluster-Debugger-launch-integration/](http://channel9.msdn.com/posts/DanielMoth/VS2010-MPI-Cluster-Debugger-launch-integration/).

- **Hybrid parallel applications that take advantage of node-level concurrency.** Applications that use the Parallel Patterns Library (PPL) built on top of the Concurrency Runtime can benefit from increased debugging support for the task-based programming model through two new windows, which support both managed and native task models as well as traditional threading programming models:

  - **Parallel Tasks,** which shows a list of all tasks and their properties (such as thread assignment and status, including deadlock detection). A video showing the Parallel Tasks window in action can be found at [http://channel9.msdn.com/posts/DanielMoth/Parallel-Tasks--new-Visual-Studio-2010-debugger-window/](http://channel9.msdn.com/posts/DanielMoth/Parallel-Tasks--new-Visual-Studio-2010-debugger-window/).
- **Parallel Stacks**, which shows multiple call stacks in a single view, coalescing the call stack segments that are common across threads—including the ability to switch to any stack frame with a single click. A video showing the Parallel Stacks window in action can be found at [http://channel9.msdn.com/posts/DanielMoth/Parallel-Stacks--new-Visual-Studio-2010-debugger-window/](http://channel9.msdn.com/posts/DanielMoth/Parallel-Stacks--new-Visual-Studio-2010-debugger-window/).

**Tuning Wizard for LINPACK**

The HPC Pack 2008 R2 Tool Pack is a standalone executable that enables administrators to easily measure computational performance and efficiency for an HPC cluster. Furthermore, because it heavily loads the cluster, the tool pack can be a valuable tool for break-in and detecting issues related to configuration and deployment, networking, power, cooling, and so on.

The tool pack calculates the performance and efficiency of an HPC cluster by automatically running the LINPACK Benchmark several times, analyzing the results of each run and automatically adjusting the parameters used for the subsequent LINPACK run. Eventually, the tool pack determines the parameters that provide optimal LINPACK performance, which is measured in terms of billions of floating-point operations per second (GFLOPS) and percentage efficiency that was achieved at peak performance. After running the tool pack, administrators can review the LINPACK results and save both the results and the parameters that were used to achieve them to a file.

Administrators can run the tool pack either in express tuning mode or in advanced tuning mode. In express tuning mode, the tool pack starts the tuning process immediately, using default values for LINPACK parameters. In advanced tuning mode, administrators can provide specific values to use when the tuning process starts, and can also configure how the tuning process is run.

For more information or to download the tool pack, visit: [http://go.microsoft.com/fwlink/?LinkID=183004](http://go.microsoft.com/fwlink/?LinkID=183004)

*Note: The tool pack does not tune the performance of the cluster itself—it only optimizes the parameters that determine LINPACK results on the cluster.*
LINQ to HPC is a set of technologies that supports big data workloads that run on a Windows HPC Server 2008 R2 with SP2 cluster. LINQ to HPC enables the development and deployment of large-scale data-intensive applications on Windows–based clusters.

- Includes a new runtime that is responsible for scheduling, executing, and managing HPC Services for distributed data jobs.
- Includes a programming model and API based on LINQ (Language Integrated Query) technology. LINQ to HPC allows developers to define distributed queries against distributed data and DSC on a cluster infrastructure.
- Provides data management functionality, including replication and data locality. Distributed Storage Catalog (DSC) and NTFS together provide the storage capability underlying LINQ to HPC.

LINQ to HPC allows the creation and running of applications that process large amounts of unstructured data on a cluster of commodity servers. It also provides data replication, failure detection, and recovery for the commodity server clusters. LINQ to HPC integrates with Windows HPC Server 2008 R2 management tools, allowing big data analysis jobs to be viewed, inspected, modified, and monitored using the HPC Cluster Manager.

LINQ to HPC understands data locality for data registered with DSC, and it preferentially schedules computations to the servers holding data that the computation needs. This minimizes the need to transfer data across the network. The benefits include reduced network bandwidth consumption and better performance.

LINQ to HPC supports an expressive programming model that is designed to facilitate cluster-based computation. Developers can implement big data applications in managed code using Visual Studio. Applications are most commonly written in C#, but developers can use any managed language that supports LINQ. LINQ to HPC data jobs and traditional HPC Server jobs (MPI, SOA or Parametric Sweep) jobs can run on the same cluster at the same time.

**Note:** Currently LINQ to HPC is not available on Windows Azure.
Microsoft Office Excel is a critical business application across a broad range of industries. With its wealth of statistical analysis functions, support for constructing complex analyses, and virtually unlimited extensibility, Excel has become a popular tool for analyzing business data. However, as calculations and modeling performed in Excel become more complex, Excel workbooks can take longer to calculate, potentially reducing the business value provided.

For example, insurance companies must calculate reserve requirements on an ongoing basis—a task that is often performed using Excel, which provides an excellent framework for constructing complex policy analyses, managing data, and creating reports. However, as the number of policies increases and new jurisdictional and regulatory factors are incorporated into calculations, the time to perform those calculations grows. It's not uncommon for complex reserve requirements calculations involving tens or hundreds of thousands of individual policies to take days or weeks to complete. Companies that perform financial portfolio modeling using Excel face similar challenges.

HPC Services for Excel 2010, available in Windows HPC Server 2008 R2, enables organizations to take advantage of HPC clusters to reduce calculation times for Excel 2010 workbooks by one or more orders of magnitude, scaling close to linearly as nodes or cores are added. Faster calculation times give business users and decision makers more information in less time, enabling more thorough analysis, faster access to important information, and better informed decisions.

Running Excel 2010 workbooks on an HPC cluster provides unique benefits in terms of reliability, resource utilization, and accounting and auditing support.

**Speeding Up Excel 2010 Workbooks Using an HPC Cluster**

HPC Services for Excel 2010 supports three different approaches to calculating Excel 2010 workbooks on an HPC cluster:

- Using Excel as a cluster SOA client
- Running Excel 2010 UDFs on a cluster
- Running Excel 2010 workbooks on a cluster

Using Excel as a cluster SOA client was possible with earlier versions of Windows HPC Server. Running Excel 2010 UDFs and Excel 2010 workbooks on a cluster are new capabilities, both of which require a combination of Windows HPC Server 2008 R2 and Excel 2010.
**Expanded Macros**

In this release, macros can be created to partition iterative calculations into a fork-join pattern. Built-in macros have been added to make implementation simpler and more efficient. Built-in macros include: HPC_GetVersion (for future compatibility with previous versions), HPC_Initialize, and HPC_Finalize.

**Performance and Scale Enhancements**

In this release, 300 million calculations running on 5,000 cores can be completed in approximately 12 days, or 288 hours. That is the equivalent of approximately 1 million calculations per hour, or about 1,700 calculations per minute.

**Using Excel 2010 as a Cluster SOA Client**


**Running Excel 2010 User-Defined Functions on an HPC Cluster**

UDFs are a well-established mechanism for extending Excel, enabling functions that are contained in Excel extension libraries (XLLs) to be called from spreadsheet cells like any standard Excel function. Excel 2010 extends this model to the HPC cluster by enabling UDFs to be calculated on an HPC cluster by one or more compute nodes. If a long-running workbook includes multiple independent calls to defined functions and these functions contribute to the overall processing time, then moving those calculations to the cluster can result in significant overall performance improvement. As far as users are concerned, there is no difference between a desktop function and a function running on the cluster—except for better performance.

**Status Window**

A status window can be displayed while UDFs are calculating on a cluster. The window shows the number of requests and responses in progress, and any errors that are returned.

**Improved Error Handling and Tracing**

Events are categorized into administrative and operational, which is achieved through new client-side logging and improved server-side logging.
Running Excel 2010 Workbooks on an HPC Cluster

Many complex, long-running workbooks run iteratively—that is, they perform a single calculation over and over, using different sets of input data. Such workbooks might include complex mathematical calculations contained in multiple worksheets, or they might contain complex VBA applications. When a workbook runs iteratively, the best option for parallelizing the calculation can be to run the entire workbook on the cluster.

Windows HPC Server 2008 R2 supports running Excel 2010 instances on the compute nodes of an HPC cluster, so that multiple long-running and iterative workbooks can be calculated in parallel to achieve better performance. Many workbooks that run on the desktop can run on the cluster—including workbooks that use Visual Basic for applications, macros, and third-party add-ins. Support for running Excel 2010 workbooks on a cluster also includes features designed to run workbooks without user interaction, providing a robust platform for calculating Excel 2010 models without requiring constant oversight.

Although this approach can be used to calculate many workbooks on a cluster, some development is required. When workbooks run on the desktop, calculation results are inserted into spreadsheet cells. Because running Excel 2010 workbooks on a cluster uses Excel processes running on cluster nodes, the user or developer must define what data is to be calculated and how to retrieve the results. A macro framework is provided that can handle much of this work, and developers can customize the framework or write their own code to manage calculations and results, providing for virtually unlimited flexibility.

Note: Running Excel workbooks on a cluster is limited in functionality with Windows HPC Server 2008 R2, supporting only workbooks with simple VBA, macros, and cell dependencies.


Additional Benefits of Running Excel 2010 Calculations on a Cluster

In addition to faster calculation speeds, running Excel 2010 calculations on an HPC cluster can provide several additional benefits, including:

- **Reliability.** A complex Excel 2010 workbook that calculates in hours or days on a desktop PC could be interrupted by accidents, system update shutdowns, power outages, or hardware failures. On an HPC cluster, if an individual compute node fails, depending on calculation scope, calculations can be run on the remaining nodes—meaning that they will complete even
in the event of a hardware or other kind of failure. (Note: Excel 2010 support is dependent on the SOA runtime and associated features in Windows HPC Server 2008 R2, with the same limitations with respect to failover and nonstop processing.)

- **Shared Resources.** With a HPC cluster installed on an organization’s network, all authorized users in the organization can take advantage of the cluster. A single cluster can support multiple users and Excel 2010 applications while providing a central point for resource management.

- **Scalability.** Compute nodes can be added to an HPC cluster at any time to increase available computing power. As users, calculations, and utilization increases, the cluster can be scaled as needed to meet the needs of an organization.

- **Accounting and Auditing.** Windows HPC Server 2008 R2 provides comprehensive reporting, accounting, and auditing tools, enabling organizations to obtain visibility into which calculation models are being run and insight into the performance of specific applications.

**Diagnostic Tests for Microsoft Excel 2010**

Two new diagnostic tests are provided for Excel 2010. System administrators can run these tests to determine if Excel 2010 is installed and properly licensed on an HPC cluster and to verify that the UDF container service is loaded and ready on the nodes.
Because HPC clusters are being adopted by a broad range of mainstream companies for mission-critical applications, security and integration with existing infrastructure is important. Windows HPC Server 2008 R2 helps meet these needs by using Active Directory to support system administration and running jobs.

Through support for Active Directory, Windows HPC Server 2008 R2 enables systems administrators to apply and audit security policies using existing, familiar Active Directory mechanisms—and enables compute jobs to access network resources such as file or database servers in the security context of the user. The scheduler runs each job under the context and credentials of the submitting user, not as a super user or administrator, with user credentials encrypted and stored with the job only until its completion. Job management-related communications take place over encrypted and authenticated channels, with the user’s credentials known only to the scheduler and the node manager process on the compute nodes—not to processes or applications started on the user’s behalf—to further isolate credentials and protect their integrity. This additional level of security integration with Active Directory is a unique benefit provided by MS-MPI when compared with the reference MPICH2 implementation.

Integrated patch management built into Windows HPC Server 2008 R2 is another key security-related (and stability-related) feature, in that it allows cluster administrators to schedule and deploy updates to groups of nodes with assurance that patch deployment won’t interfere with running jobs. Windows HPC Server 2008 R2 also supports the use of existing enterprise patch deployment tools.

Windows HPC Server 2008 R2 also inherits security features that are a part of Windows Server 2008 R2, such as Network Access Protection, Role Management, Network Policy Management, and an integrated, bidirectional Windows Firewall for enterprise-facing networks.
For organizations seeking the advantages of high-performance computing, Windows HPC Server 2008 R2 efficiently scales to thousands of nodes and integrates seamlessly with Windows-based IT infrastructures. New capabilities provided in Windows HPC Server 2008 R2 can help companies across many industries to more easily take advantage of parallel computing, making it easier than ever to build, run, and manage both traditional HPC workloads and new interactive SOA applications.

Windows Azure integration brings additional scale, power, and economic benefits to high-performance computing. Windows Azure’s pay-for-only-what-you-use model allows administrators to maintain an HPC infrastructure that meets their typical demands while relying on Windows Azure to address periods of peak computational needs.

Additional Resources
For more information about Windows HPC Server 2008 R2 and Windows HPC Server 2008 R2 with SP2, visit the following websites:

- [http://www.microsoft.com/hpc](http://www.microsoft.com/hpc)

For more information about Microsoft System Center, Windows Server 2008 R2, and other Microsoft server products and tools, visit the following Web sites:

- [http://www.microsoft.com/systemcenter](http://www.microsoft.com/systemcenter)
- [http://www.microsoft.com/servers](http://www.microsoft.com/servers)