Windows Compute Cluster Server 2003 and Windows Compute Cluster Edition Operating Systems for IBM System x and BladeCenter blade servers

At a glance

Windows Compute Cluster Server (WCCS) provides a single high-performance computing (HPC) package in which all of the components were architected and developed to run together.

Overview

IBM System x™ and BladeCenter® server blade models can now be shipped with Windows® Compute Cluster Server 2003 (WCCS) and Windows Compute Cluster Edition (WCCE) operating systems. These offerings bring together the value and computing power of x64 (64-bit x86) servers, the ease of use and security of Active Directory service, and the Windows operating system to provide a security-enhanced and affordable high-performance computing (HPC) solution.

HPC clusters are implemented primarily to increase performance by splitting a computational task across many different nodes in the cluster. They are commonly used in scientific computing and are often run on custom programs designed to exploit parallel processing. HPC clusters are optimized for workloads that require jobs to run on separate nodes, which communicate actively during the computation. In these computations, intermediate results from one node’s calculations will affect future calculations on the other nodes.

Many software programs running on HPC clusters use libraries such as Message Passing Interface (MPI). These libraries are specifically designed for writing scientific applications for HPC computers. The WCCS operating system is specifically designed for HPC workloads.

The WCCS pack contains two CDs.

1. The first CD contains Windows Server 2003 Compute Cluster Edition (CCE), which is a full Windows Server operating system. It is identical to Windows Server 2003 64-bit Standard Edition, however it is strictly licensed to be used only for HPC workloads.

2. The second CD contains the Microsoft® Compute Cluster Pack, which contains a Microsoft MPI pack, a scheduler, and a deployment wizard.

The Windows Server 2003 CCE DIB offering from IBM is for those customers running HPC clusters and using alternative MPI stacks and scheduler.

The WCCS licensing model consists of a Windows Server operating system license for each server, or node, in the cluster. For example, a four-node cluster requires four Compute Cluster Server licenses. An eight-node cluster requires eight licenses for Compute Cluster Server. You do not need to purchase client access licenses (CALs). Compute Cluster Server is licensed as a server license with no incremental CALs.

Licensing for WCCS is restricted to computers that are used as dedicated computational servers running only HPC applications. HPC
applications are those that use several servers as a group to solve a single computational problem or a single set of closely related computational problems. Applications that run on a single server are not considered HPC applications. Applications that are distributed across multiple servers are not considered HPC applications unless they are working on a set of closely-related computational problems.

If you want to use a server for both HPC and non-HPC applications, you can use another 64-bit edition of Windows Server 2003.

The Compute Cluster Pack (CCP) is the application stack that provides the services, interfaces, and supporting software needed to create and configure the cluster nodes. It also provides the utilities and the management infrastructure. The CCP creates directories and assigns permissions under a user-specified installation directory (%CCP_HOME%).

The CCP allows you to build, run, and manage your HPC cluster.

• **Build**: Microsoft provides an integrated development environment for MPI applications. The MPI stack uses a Windows port of the MPICH2 reference implementation called MS MPI, which is compatible with MPICH2. Your MPICH2 applications should run on MS MPI with little, if any, modification. MS MPI can utilize any Ethernet interconnect that is supported on Windows Server 2003. It can also utilize low-latency and high-bandwidth interconnects, such as InfiniBand or Myrinet, through the Winsock Direct drivers provided by hardware manufacturers.

• **Run**: The Job Scheduler lets you schedule jobs, allocate the resources needed for the job, and change the tasks and properties associated with the job. The main features of the Job Scheduler include:
  – Error Recovery: Automatic retry of failed tasks and jobs and automatic routing around unresponsive nodes
  – Automated Cleanup: Helps prevents runaway processes on nodes
  – Security: Each job or task runs in the context of the submitting user

• **Manage**: The CCP integrates with Active Directory to provide easy, role-based cluster management. Compute Cluster Manager is the console used for job submission and management. You can also use a command-line interface for administering the cluster and for managing jobs.

Orders for WCCS and WCCE operating systems are accepted only in conjunction with a valid hardware order.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM makes no representation or warranty regarding third-party products.

**Planned availability date**

January 25, 2008

**Description**

WCCS provides a single HPC package in which all of the components have been architected and developed to run together. While you can use a variety of tools to create an HPC solution on Windows, a major advantage of CCS is that it is a single, standardized solution developed and supported by a single organization.

CCS features include:

• Task-based configuration
• CLI-based/GUI-based node and job management
• Monitoring with Microsoft Operations Manager (MOM), third-party tools, and Performance Monitor (PerfMon)
• Active Directory service and other Windows tools and technologies
• Integrated application stack
• Extensible job scheduler

WCCS leverages several different components, some required and some optional. The cluster design that you select depends on your intended computational goal and on the configuration of the servers that make up your cluster.

The main components of a typical CCS network include the head node, the compute nodes, and the client.

**Head node**

The head node provides deployment and administration user interfaces (UIs) and management services for the compute cluster. The UIs include the Compute Cluster Administrator, the Compute Cluster Job Manager, and a Command-Line Interface (CLI). Management services include job scheduling, as well as job and resource management. You can use Remote Installation Services (RIS), Automated Deployment Services (ADS), or Windows Deployment Services (WDS) to support automated compute node deployment.

If compute nodes do not have interfaces to the public network (private interfaces only), Internet Connection Sharing (ICS) is used to configure Network Address Translation (NAT) between the nodes and the public network. In this case, the head node acts as a gateway between the public and the private networks that make up the cluster and also provides limited Dynamic Host Configuration Protocol (DHCP) and Domain Name System (DNS) services to the private network. These limited DHCP and DNS services are integral to ICS, and you can configure them when ICS is implemented. You can also assign static addresses to the compute nodes.

The head node includes:

• **Active Directory service:** Active Directory is required by CCS. Each node of the cluster must be a member of an Active Directory domain because Active Directory provides the authorization and authentication services for CCS. The Active Directory domain can be independent of the cluster. For example, if you have an Active Directory infrastructure in your organization, you can simply add the CCS cluster to the existing Active Directory domain. Active Directory can also run exclusively within the cluster if the cluster is isolated, if it is self-contained, or if there is no other Active Directory domain for the cluster to join. In this case, Active Directory must be installed on the head node.

• **Job Scheduler:** The Job Scheduler is a service that runs on the head node and manages the job queue, resource allocation, and job execution by communicating with the Node Manager Service that runs on each compute node.

• **Systems management tools:** The CCP offers a set of management tools that enables the cluster administrator to deploy and manage the compute nodes. This infrastructure of the CCP provides the administrative, user, and command-line interfaces used to administer the cluster, submit jobs, and manage the job queue.

**Compute nodes**

A compute node is any computer configured to provide computational resources for the compute cluster. Compute nodes must run supported operating systems, but they do not all have to have the same operating system or even the same hardware configuration. A similar configuration does, however, simplify deployment, administration, and resource management. Compute clusters with different hardware configurations limit the cluster’s capabilities. Jobs running in parallel mode and using nodes with different capabilities will only be able to run at the speed of the slowest processor in the selected nodes. Compute nodes include:

• **Node manager:** The node manager interacts with the Compute Cluster Scheduler Service on the head node and is responsible for job execution. If the head node is also configured as a compute node, it will act as the node manager.

• **Microsoft MPI:** MS MPI is a key networking component of the compute cluster. The MS MPI implementation is compatible with the reference MPICH2. One of the most important decisions in compute cluster design is selecting the right network topology for MPI traffic. WCCS supports five different network topologies.

**Client**

• **Tools for job submission:** Several interfaces are available for job submission, including the Compute Cluster Job Manager, the CLI, a Web service designed for interoperability with non-Microsoft schedulers, and a COM interface for integration with C or C++ custom interfaces and for scripting support. The CLI also supports a variety of scripting languages,
including the Perl, FORTRAN, C/C++, C#, and Java™ languages.

• **Systems management**

The CCS works as follows:

1. The user stages the data into a shared file system or into the database.
2. The user submits jobs to the Job Scheduler through one of the following:
   - CLI
   - Job Manager UI
   - Integrated desktop applications through the COM API
3. The Job Scheduler allocates idle resources for the job. Multiple tasks can be dispatched to allocated nodes to accelerate execution.
4. The tasks can use the built-in MPI stack or can run independently to complete a parametric sweep.
5. The tasks are run inside the users’ security context, and are validated against Active Directory.

WCCS is a cluster of servers that includes a single head node and one or more compute nodes. 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 compute cluster. The head node is connected to the public network, and you can choose to have the compute nodes directly connected to the public network as well.

WCCS requires Active Directory. WCCS uses the Active Directory infrastructure for security, for account management, and for helping to facilitate overall operations management with tools such as MOM 2005, System Center Operations Manager 2007, and Microsoft Systems Management Server (SMS) 2003. Though you may have an existing LDAP server, you are still required to have Active Directory for a Compute Cluster Server installation. Active Directory can be installed on the head node.

Job scheduling is critical in running a computer cluster. Efficient job scheduling allows work to be carried out in accordance with the needs of the cluster administrators and of management.

Jobs can be single task or multiple tasks and can specify the number of processors required for the job and whether those processors are needed exclusively or can be shared with other jobs or tasks. Jobs are the activities scheduled to perform on the computer cluster, while tasks are the discrete commands that the jobs execute.

The job scheduler (or batch scheduler) starts jobs, stops jobs, and determines they will run on a system. The scheduler:

• Allocates all the resources that have been requested.
• Dynamically partitions the large scale-out system into a smaller system tailored to users’ requests.
• Tracks which processors are associated with the job.
• De-allocates these resources when the job has terminated. The scheduler provides a mechanism for users, as well as administrators, to monitor the progress of their jobs.

The WCCS 2003 Job Scheduler uses four scheduling policies:

• Priority-based first come, first served (FCFS): Based on the priority setting of the job itself.
• Backfilling: Resources are reserved based on the time and resources required to execute a job; can select smaller jobs to fill available open windows.
• Nonexclusive scheduling: Resources can be shared.
• License-aware scheduling

You can access the Job Scheduler with the Compute Cluster Job Manager, the WIN32 GUI that enables you to create, submit, and monitor jobs. Features of the Job Scheduler are listed in earlier text.

Other job schedulers can also run on Windows Server CCE, however, you cannot use the MS MPI without the CCP. High-performance schedulers include Cluster CoNTrroller by MPI Software
Technology, Load Sharing Facility (LSF) by Platform Computing, and Portable Batch System Pro (PBS Pro) by Altair Grid Technologies. Cluster CoNTroller and the latest version of PBS Pro (5.3) also run on the Windows platform.

The Job Scheduler provides integrated queue monitoring, integrated node add and remove operations, and integrated administrative command and job scheduling. It is programmatically accessible through the PerfMon.

The Compute Cluster Administrator is a Microsoft Management Console (MMC) 3.0 snap-in named Computeclusteradmin.msc. It is installed on the head node of a cluster during setup and also when the Software Development Kit (SDK) and client utilities are installed. It can also be installed on a remote workstation and used for remote access and management of the head node and the cluster. It includes five major pages:

- A start page: Overview
- The To Do List: Helps quickly deploy the compute cluster
- Node management: Displays information about head and compute nodes
- Remote desktop sessions: Can create and close remote desktop sessions to one or more nodes
- Systems monitor: Can select Start Perfmon Task to select a session for each node

In order to run the CCP, you must install additional software on your compatible 64-bit Windows Server operating system. The CCP can be run on a 64-bit supported operating system. Likewise, Windows Server Compute Cluster Edition can be used for other Windows-based HPC solutions.

If you have Windows Server 2003 Service Pack 1 installed, there are some additional hotfixes required in order to install the CCP. Windows Server 2003 Service Pack 2 includes these hotfixes:

- ICS hotfix
  http://go.microsoft.com/fwlink/?linkid=55166
  Install this update to correct an issue in which Internet Connection Sharing does not appear in the properties of the active network connection after you install Active Directory to configure a computer that is running Windows Server 2003 with Service Pack 1 as a domain controller.

- RIS hotfix
  http://go.microsoft.com/fwlink/?linkid=55167
  Install this HPC update to Windows Remote Installation Services (RIS) to obtain changes to RIS specifically designed to operate in the Windows 2003 CCS environment.

- MMC 3.0
  http://go.microsoft.com/fwlink/?linkid=62400
  MMC 3.0 hosts a variety of operating system tools and third-party administrative tools.

The client utilities are supported on the following operating systems:

- Windows Server 2003 with Service Pack 1 (SP1) or R2
- Windows XP Professional with Service Pack 2 (SP2)
- Windows XP Professional x64 Edition
- Windows Vista® with Compute Cluster Server Service Pack 1

**Business Partner information**

If you are a Direct Reseller - System Reseller acquiring products from IBM, you may link directly to Business Partner information for this announcement. A PartnerWorld ID and password are required (use IBM ID).

BP Attachment for Announcement Letter 208-002

Trademarks

System x is a trademark of International Business Machines Corporation in the United States or other countries or both.

BladeCenter is a registered trademark of International Business Machines Corporation in the United States or other countries or both.

Windows, Microsoft, and Windows Vista are registered trademarks of Microsoft Corporation.

Java is a trademark of Sun Microsystems, Inc.

Other company, product, and service names may be trademarks or service marks of others.

Offering Information

Product information is available via the Offering Information Web site

http://www.ibm.com/common/ssi

Publications

No publications are shipped with these programs.

Technical information

Hardware requirements: For the latest list of systems supported, visit

http://www-03.ibm.com/servers/eserver/serverproven/compat/us/

Planning information

Direct customer support: When it comes to remote software support, you expect fast, accurate answers that help keep your IT staff productive. But in your complex, multivendor operating environment, it’s not always easy to find the level of expertise you need, much less a solution that also gives you flexible options designed to meet your unique requirements. The IBM portfolio of remote support offerings provides world class usage and defect support for a broad range of products running on System x™ servers, including IBM Director, Microsoft®, Linux™, VMware, clustering, SANs, disk, and tape.

Remote assistance is available through toll-free telephone access and electronic access (where available). For all eligible software and systems, we will help you with usage and installation questions, product compatibility and interoperability questions, interpretation of product documentation, problem diagnosis, and software defect support.

IBM Remote Technical Support Services — ServicePac®

End-to-end hardware and software support for your System x or BladeCenter® solution. Coverage includes IBM Director, Microsoft Windows®, and Linux operating systems. Support can also be added for VMware. Access to the answers you need is available 24 hours a day, 365 days a year for severity 1 problems, and Monday through Friday, 8:00 a.m. to 5:00 p.m. in your local time zone, for all other questions and problems. With unlimited calls and unlimited callers, anyone at your company can call as often as needed and receive quick and efficient responses.

Highlights of the offering include:

• Access to live technical experts
• Unlimited calls and unlimited callers
• Fast, precise answers and problem resolution
Hardware, software, and services support from a single source
Competitively priced, predictable support costs
One-time fee for one year of support
Packaged on a “per system” basis

For more information on Remote Technical Support, visit

IBM Operational Support Services — Support Line
Support Line is an Annuity service that provides unlimited calls and unlimited callers at a fixed price. You are provided with the flexibility to choose which “support groups” (select operating systems, software and hardware products) meet your business coverage needs. Basic support includes prime-shift coverage, defined as normal business hours, Monday through Friday, excluding national holidays. If you require support beyond prime-shift hours, IBM offers extended/full-shift coverage for support around the clock.

Highlights of the offering include:

- Telephone and electronic access
- Unlimited calls and unlimited callers
- Packaged in ranges of servers or processors
- Supported product groups include:
  - Microsoft (operating system and applications)
  - Linux (operating system and applications)
  - Linux Clusters (operating system and applications)
  - SANs and Network Attached Storage (NAS)
  - Disk and tape
- Support for operating systems and applications running on OEM hardware

For more information, visit
http://www.ibm.com/services/us/index.wss/so/its/a1000030

Security, auditability, and control
The customer is responsible for evaluation, selection, and implementation of security features, administrative procedures, and appropriate controls in application systems and communication facilities.

Software Services
IBM Software Services has the breadth, depth, and reach to manage your services needs. You can leverage the deep technical skills of our lab-based, software services team and the business consulting, project management, and infrastructure expertise of our IBM Global Services team. Also, we extend our IBM Software Services reach through IBM Business Partners to provide an unmatched portfolio of capabilities. Together, we provide the global reach, intellectual capital, industry insight, and technology leadership to support any critical business need.

To learn more about IBM Software Services or to contact a Software Services sales specialist, visit
http://www.ibm.com/software/sw-services/

Ordering information

Charge metric
IBM United States Announcement 208-002
Server

A server is a computer system that executes requested procedures, commands, or applications to one or more user and/or client devices. A PoE must be obtained for each server regardless of the number of processors and partitions in the server on which the program runs or for each server managed by the program. Where blade technology is employed, each blade is considered a separate server.

New Licensees

License orders for the program numbers associated with the operating systems in this announcement can only be ordered in conjunction with a new system order on which the selected operating system has been announced as an offering.

Shipment of materials associated with operating systems in this announcement when ordered with a new system will be included with the system unit based on available medium delivery options offered; preinstalled, preinstalled with backup media, or physical media only where preinstall is not offered.

Billing for the operating systems will occur under the licensed program order number contained in this announcement.

Program delivery

Except where noted, configuration aids, either IBM Configurator for e-business or IBM Hardware Configurator, must be used for creation of valid software and hardware orders for the offerings in this announcement as well as other referenced and affiliated announcement offerings.

To order the programs described in this announcement, specify the type model number, order type description, one-time charge (OTC) feature, preinstall feature, or drop-in-the-box feature as indicated and listed in the tables below.

Note: A Hardware Installation Productivity Order (HIPO) order must be drop-in-the-box feature specified in the tables below. It must be included when placing an order for the following type models.

<table>
<thead>
<tr>
<th>Type model</th>
<th>Operating system</th>
<th>OTC Billing</th>
<th>Supp Billing</th>
<th>Pre-Install Billing</th>
<th>Drop-in-the-box Billing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5639-S64</td>
<td>Windows 2003 64-bit</td>
<td>0002</td>
<td>N/A</td>
<td>N/A</td>
<td>3523</td>
</tr>
<tr>
<td>Compute Cluster Server Per 1-4 Processor Server</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compute Cluster Edition Per 1-4 Processor Server</td>
<td>0003</td>
<td>N/A</td>
<td>N/A</td>
<td>3523</td>
<td></td>
</tr>
<tr>
<td>Serialization feature: 3444</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For order routing purposes, serialization feature number 3444 must be included on every operating system license program order.

Hardware installation productivity order (HIPO)

5372-SWX HIPO is mandatory for preinstall or shipments of program medium and publications associated with select operating system offerings in this announcement. The translation table below is for reference purposes. The operating system type model number, with offered OTC features numbers, are cross referenced to features that must be included on the 5372-SWX HIPO type model order for the purpose of completing valid (error free) "system solution" orders.

<table>
<thead>
<tr>
<th>Operating system</th>
<th>OTC</th>
<th>Supp</th>
<th>Spec</th>
<th>Spec</th>
<th>Spec</th>
<th>Spec</th>
<th>PI</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>type model</td>
<td>feat</td>
<td>feat</td>
<td>ID</td>
<td>ID</td>
<td>Lang</td>
<td>Proc</td>
<td>H/W</td>
<td>no</td>
</tr>
<tr>
<td>number</td>
<td>num</td>
<td>num</td>
<td>feat</td>
<td>feat</td>
<td>feat</td>
<td>feat</td>
<td>feat</td>
<td>feat</td>
</tr>
</tbody>
</table>
Customization options

Select the appropriate feature numbers to customize your order with delivery options desired. These features can be specified on the initial or MES orders.

Expedite shipment

Feature Description

3445 Expedite

Expedite shipments will be processed to receive 72-hour delivery from the time IBM Software Delivery and Fulfillment (SDF) receives the order.

Terms and conditions

The information provided in this announcement letter is for reference and convenience purposes only. The terms and conditions that govern any transaction with IBM are contained in the applicable contract documents.

Agreement: Supplier's license terms apply.

Limited warranty: Not warranted by IBM. Warranty, if any, provided by supplier.

Volume orders: Not applicable

Prices

<table>
<thead>
<tr>
<th>Type</th>
<th>Model Number</th>
<th>Feature Description</th>
<th>Feature Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>5639-S64</td>
<td>Per 1-4 Processor Server</td>
<td>0002</td>
</tr>
<tr>
<td>Compute</td>
<td>Server</td>
<td>Compute Cluster Edition</td>
<td>0003</td>
</tr>
</tbody>
</table>

IBM Global Financing

IBM Global Financing offers competitive financing to credit-qualified customers to assist them in acquiring IT solutions. Offerings include financing for IT acquisition, including hardware, software, and services, from both IBM and other manufacturers or vendors. Offerings (for all customer segments: small, medium, and large enterprise), rates, terms, and availability can vary by country. Contact your local IBM Global Financing organization or visit

http://www.ibm.com/financing

IBM Global Financing offerings are provided through IBM Credit LLC in the United States, and other IBM subsidiaries and divisions worldwide to qualified commercial and government customers. Rates are based on a customer's credit rating, financing terms, offering type, equipment type, and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension, or withdrawal without notice.

Trademarks

System x is a trademark of International Business Machines Corporation in the United States or other countries or both.
eServer, ServerProven, ServicePac, and BladeCenter are registered trademarks of International Business Machines Corporation in the United States or other countries or both.

Microsoft and Windows are registered trademarks of Microsoft Corporation.

Linux is a trademark of Linus Torvalds in the United States, other countries or both.

Other company, product, and service names may be trademarks or service marks of others.

This announcement is provided for your information only. For additional information, contact your IBM representative, call 800-IBM-4YOU, or visit the IBM home page at: http://www.ibm.com