NOVASCALE UNIVERSAL

R440, R460
R421
R422, R422-INF

HOW TO: Create a Windows Compute Cluster (WCCS)

REFERENCE
Rev02v01

NOVASCALE UNIVERSAL

R440, R460
R421
R422, R422-INF

HOW TO: Create a Windows Compute Cluster (WCCS)

Hardware

July 2007

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

REFERENCE
Rev02v01

Printed in France

Suggestions and criticisms concerning the form, content, and presentation of this book are invited. A form is provided at the end of this book for this purpose.

To order additional copies of this book or other Bull Technical Publications, you are invited to use the Ordering Form also provided at the end of this book.

Trademarks and Acknowledgements

Table of Contents

Preface

This HOW TO explains how to configure a **Windows Computer Cluster Server 2003**

Intended readers

This HOW TO is written for very qualified user's or Bull SAS support technicians.

Chapter 1. Overview of the Installation Process

This HOW TO provides the information and steps necessary to create a Microsoft Compute Cluster (WCCS).

You have ordered and received the computers with Microsoft Compute Cluster Server 2003 pre-installed, pre-activated.

This means:

- Microsoft Windows Server 2003, Computer Cluster Edition is already installed on each computer.
- The Microsoft Compute Cluster Pack (CCP) and its installer have been copied to the hard disk with a link to the installer (Setup.exe) created on the Desktop.

Familiarize yourself with CCS terms and concepts found in the *Microsoft Compute Cluster Server 2003 Reviewers Guide,* found at http://www.microsoft.com/hpc

This HOW TO describes the next stage, specifically, the steps necessary to complete the installation of Compute Cluster Server 2003.

It describes, in particular, how to configure the cluster nodes for the following network topologies:

- All nodes on public, private and MPI networks
- Compute nodes isolated on private and MPI networks
- All nodes on public and private networks

Process summary:

1. Choose the network topology for your cluster
2. Name your node
3. Setup the physical network connections between the nodes and the public network
4. Start up the '*head node*', configure its networks and connect it to the active directory
5. Install and Configure the DHCP server on the head node for private and MPI networks
6. Start up each Compute Node, configure its networks and connect to the active directory
7. Check the network configuration before installing the **Compute Cluster Pack**: check the network connection status, verify routes to other nodes
8. Install the **CCP** on each node
9. On the head node, in the Compute Cluster Pack, select the network topology and add the Compute Nodes.

Chapter 2. How to Configure the Cluster Network

2.1 Overview of network topology A - All nodes on public, private and MPI networks

Windows Compute Cluster Server 2003 supports different network topologies with one to three NICs on each node.

In topology A, each node has three NICs and is connected to all networks including:

- Public networks to access corporate resources such as data and services.
- Private networks for communication between nodes including deployment and management.
- MPI networks for MPI traffic that is latency sensitive.

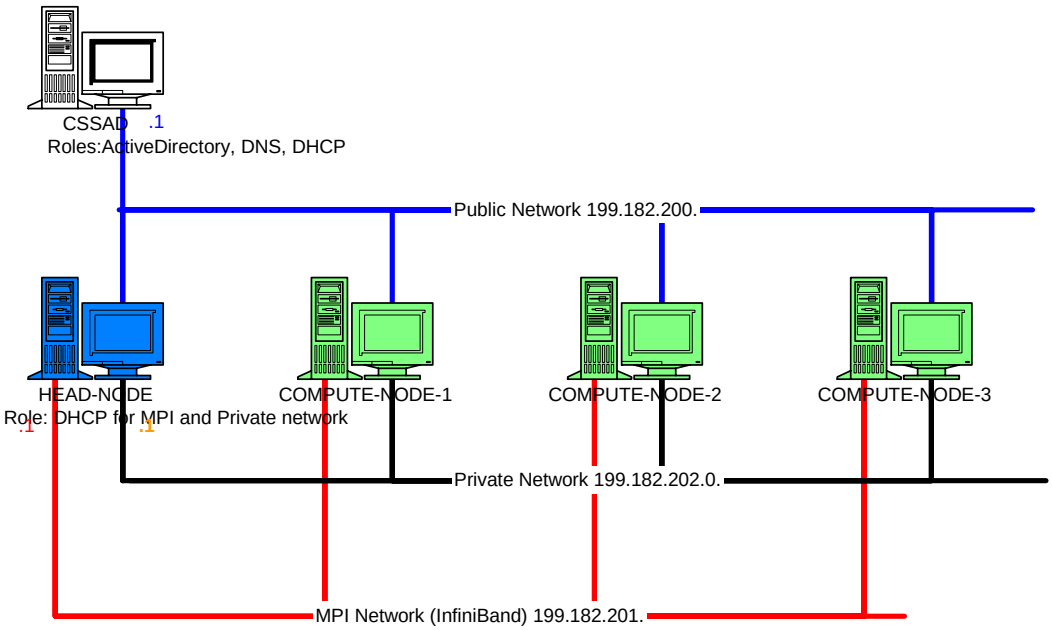


Figure 2-1. **Network Topology A** - All nodes on public, private and MPI networks

In this topology, each node of the compute cluster obtains the public IP address from the DHCP server on the public network and must connect to the same Active Directory domain as the head node.

If the public network does not have a DHCP server nor an Active Directory server, you can install the Active Directory on the head node and configure the DHCP server to provide public IP address for the public network.

The IP addresses of the head node on the private and MPI network are static.

On the head node, the DHCP server is configured to provide:

- Private IP address to compute node on the Private network for intra-cluster deployment and management traffic

- MPI IP address to compute node on the MPI network for parallel application traffic

The default gateway for each node is the public network one.

2.2 Overview of network topology B - Compute nodes isolated on private and MPI networks

In this network topology, accessibility from the public network is limited. Only the head node is on the public network and all nodes are on the private and MPI networks.

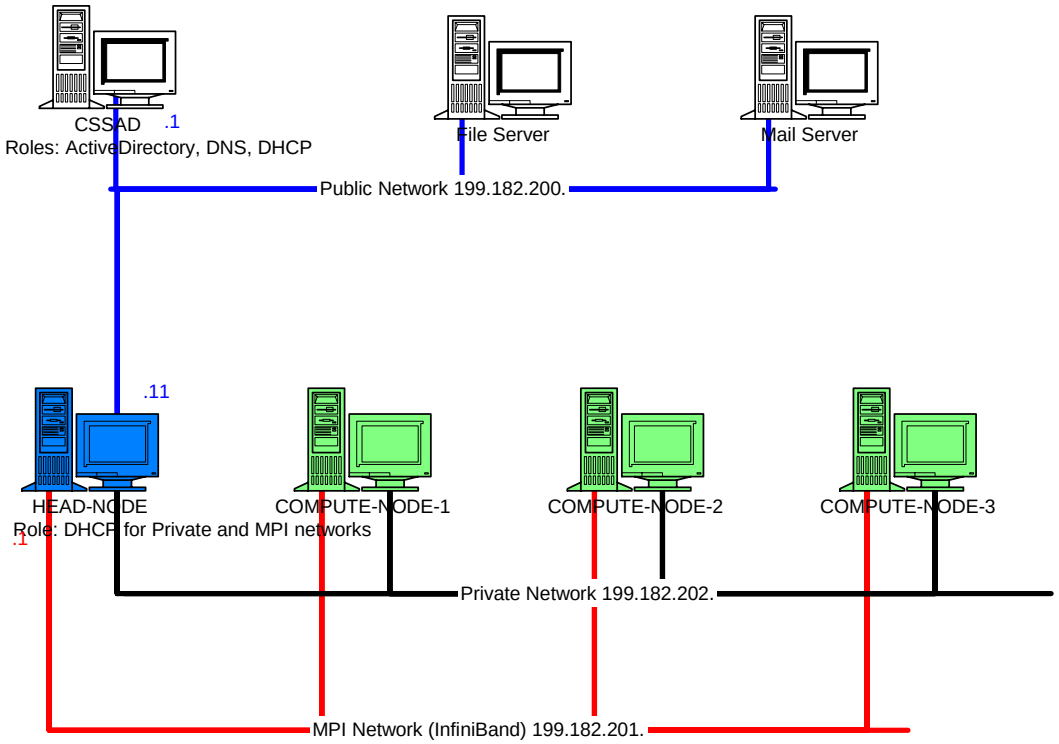


Figure 2-2. Network Topology B - Compute nodes isolated on private and MPI networks

The head node obtains its public IP address from the DHCP server on the public network. Its IP addresses on the private and MPI networks are static and assigned by the administrator.

On the head node, the DHCP server is configured to provide:

- Private IP addresses to Compute Nodes on the private network for intra-cluster deployment and traffic management.
- MPI IP addresses for Compute Nodes on the MPI network for parallel application traffic.
- The head node private IP address, used as a default gateway for all nodes.
- Public DNS IP addresses used as a DNS for all nodes.

2.3 Overview of network topology C - All nodes on public and private networks

In this topology scenario, each node has two NICs and is connected to all networks including:

- Public networks to access corporate resources such as data and services
- Private networks for communication between nodes and for MPI traffic

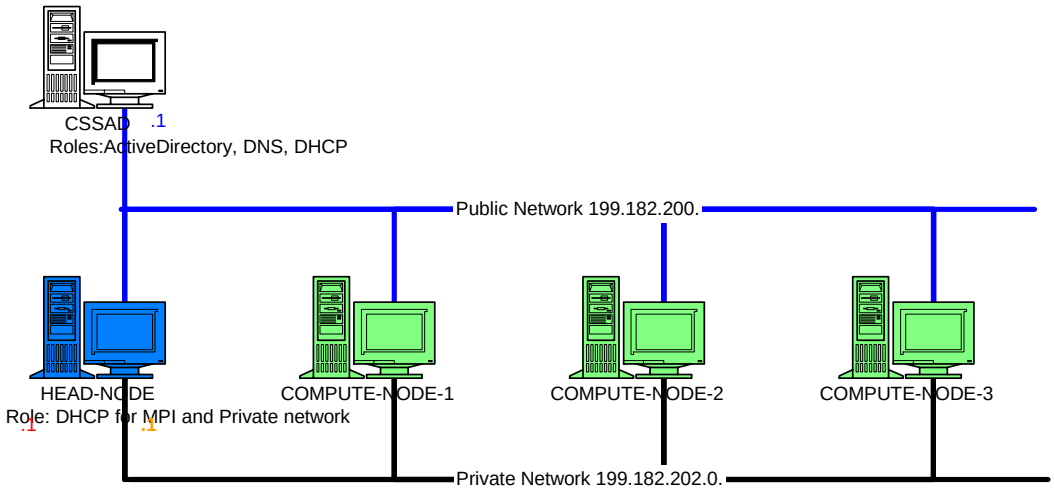


Figure 2-3. **Network Topology C** - All nodes on public, private and MPI networks

In this topology, each node of computing cluster obtains the public IP address from the DHCP server on the public network and must connect to the same Active Directory domain as the head node.

If the public network does not have a DHCP server nor an Active Directory server, you can install Active Directory on the head node and configure the DHCP server to provide public IP addresses for the public network.

The IP addresses of the head node on the private network are static.

On the head node, the DHCP server is configured to provide:

- Private IP address for Compute Nodes on the private network for intra-cluster deployment, traffic management and MPI traffic

The default gateway for each node is the public network one.

2.4 Configuring the network

2.4.1 Before starting

It is useful to create a bloc diagram of your cluster with information for each element (node, network, subnet, hub, switch, etc.): name, role (DHCP, Active Directory, etc.), IP address (or range).

2.4.2 Setup the head node.

Microsoft Windows Server 2003, Computer Cluster Edition is already pre-installed on the head node.

To setup the head node:

1. Setup the hardware and make any necessary connections to the existing network infrastructure.

2. Boot the head node, specify a name for the computer and join it to the Active Directory domain during mini-setup. If there is no Active Directory in your network infrastructure, continue the mini-setup but Active Directory will have to be installed on the head node later.

3. On the head node, open **Device Manager** (from the **Start** menu, right click **My Computer**, select **Manage** -> **System Tools** -> **Device Manager**) to check that all devices have been correctly installed. If not, follow the instructions to resolve the problem. Generally, you just have to re-install the correct driver(s) for the corresponding device(s).

4. Remaining on the head node, explore the network connexions (right click on **Control Panel / Network Connections** and click **Explore**). For each connection, locate the physical port (this is done by plugging and unplugging the network cables) and rename each connection. It is recommended that you use one of these names: **public-network**, **private_network** or **mpi_network**.

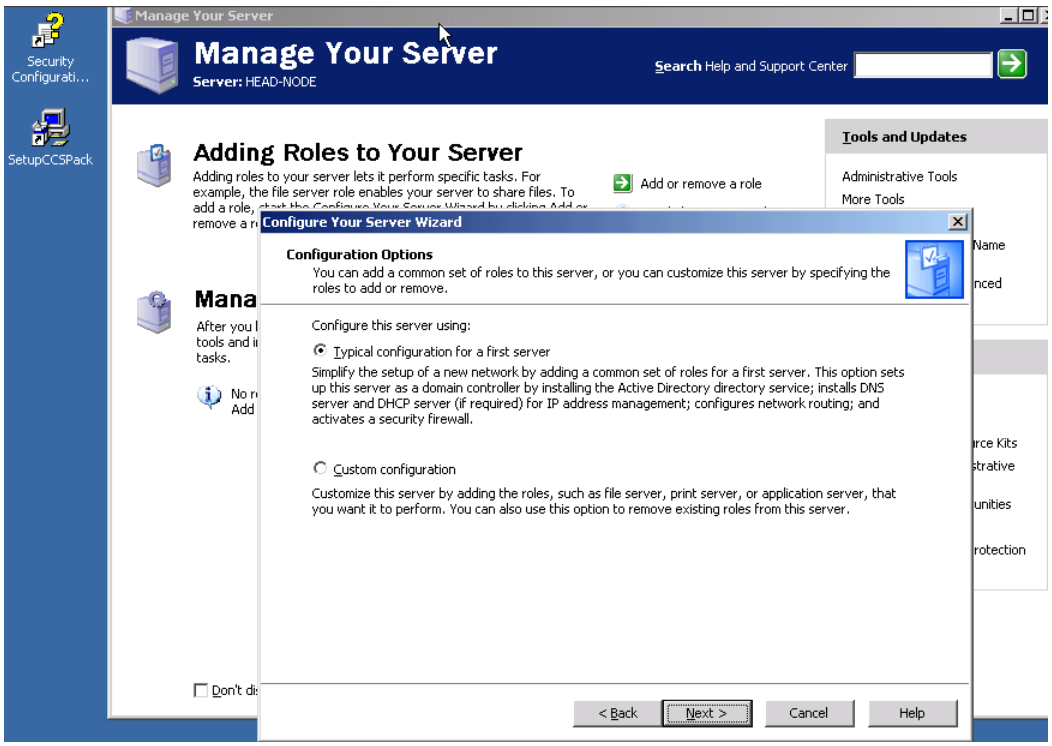
    a. If the public network does not have a DHCP server, initialize the static IP address for the public network connection.

    b. Initialize the static IP address for the private nework connection

    c. Initialize the static IP address for the mpi nework connection

2.4.3 Install Active Directory on the head node.

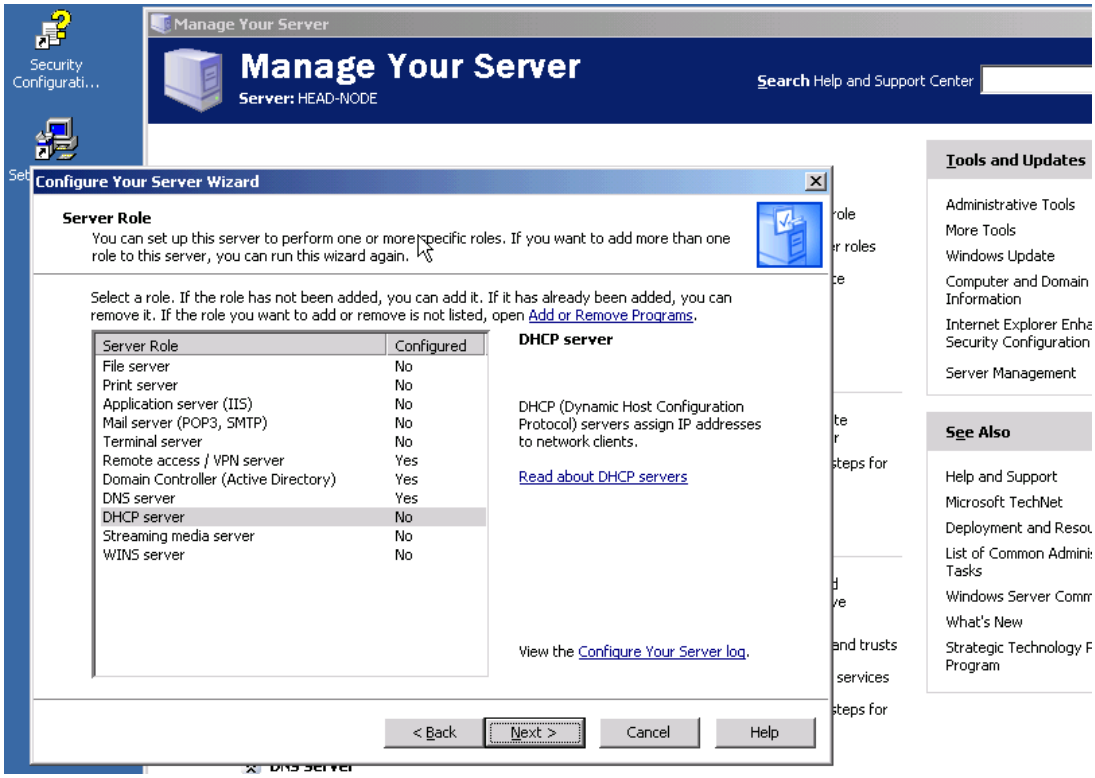
If your existing network infrastructure has no Active Directory, you must install it for the Compute Cluster (this is a requirement of the Microsoft Compute Cluster Solution), and the best place to install it is on the head node. Click on **Start** -> **Manage you server** -> **Add or Remove Role** to activate the **Configure Server Wizard** , then select **Typical configuration for the first server** to install Active Directory / DNS / DHCP.



Once the Active Directory installation has been done and the head node has been restarted, log in under a domain account: enter the user account and password with the new domain name (this has been done above) in the domain field of the login dialog box, so that the domain authentication is checked.

2.4.4 Configuring the DHCP server on the head node

Click on **Start** -> **Manage you server** -> **Add or Remove Role** to check if a DHCP server has been installed.



If not, select the DHCP server item and click **Next** to install it.

5. If a DHCP server for the public network is on the head node (case of network topology **A** or **C**, and no DHCP sever is in service on the public network), create a DHCP scope for the public network with **DHCP Scope Wizard**(A DHCP scope consists of a pool of IP addresses on a given subnet, that the DHCP server can lease to clients). The Scope Wizard is invoked when you:

   a. Install DHCP server on the head node

   b. Open **DNS Start** -> **Manage your server** -> **DNS** , in the console tree, right-click on the head node DHCP server and then select **New Scope.**

   Set the following public network scope options:

   c. Option **003 router** (default gateway)=head node IP address,

   d. Option **005 DNS servers** =head node IP address.

6. Create a DHCP scope for the private network. No scope option is necessary for network topology **A** or **C**. For topology **B**, set the following private network scope options:

   a. Option **003 router** (default gateway)=head node private IP address,

   b. Option **005 DNS servers** =public DNS IP address.

7. Create a DHCP scope for a MPI network without scope option.

2.4.5 Configuring the Compute Nodes

Microsoft Windows Server 2003, Computer Cluster Edition is already pre-installed on the compute node. The compute nodes may be configured in parallel.

1. Setup the hardware and make any connections necessary for the existing network infrastructure.

2. Boot the Compute Node, Windows automatically starts a mini setup. Specify a name for the computer, choose default configuration for network connection and connect it to the Active Directory domain. The Active Directory domain is the public one that exists already or the one you have created, above, on the head node.

3. Open **Device Manager** (on the **Start** menu, right click **My Computer**, select **Manage** -> **System Tools** -> **Device Manager** to check that all the devices have been correctly installed. If not, follow the instructions to resolve the problem. Generally, you have just to re-install the correct driver(s) for the corresponding device(s).

4. Check the network connexions (**Control Panel**, right click **Network Connections** and select **Explore**). For each connection, locate the physical port (this is done by plugging and unplugging the network cables) and rename each connection. It is recommended that you use one of these names: **public-network**, **private_network** or **mpi_network**.

2.4.6 Check the network configuration.

The cluster network has now been configured. However, the network configuration should be checked before proceeding to the next step (CCP installation):

1. On each Compute Node, run **ipconfig /all** and check
   a. The IP address of each connection: IP address is not auto configured, it is in the subnet IP address range (or DHCP scope),
   b. Default gateway is the public network default gateway for network topologies **A** & **C**, or head node private IP address in for network topology **B**.
   c. DNS is the public network DNS (it is the same as the head node public IP address if Active Directory has been used for the head node).

2. **Ping** each other using different IP addresses.

3. **Ping** each other using the node names.

4. Check all nodes are in the same domain: on each node, go to the **Start** menu, right click on **My Computer** and select the **Computer Name** tab to display the full compute name. If the domain name is not correct, click **Change** and enter the correct computer name and domain name.

Congratulations! Now the network installation is complete. The next step is to install Compute Cluster Pack (CCP).

Chapter 3. Install the Compute Cluster Pack on all Nodes

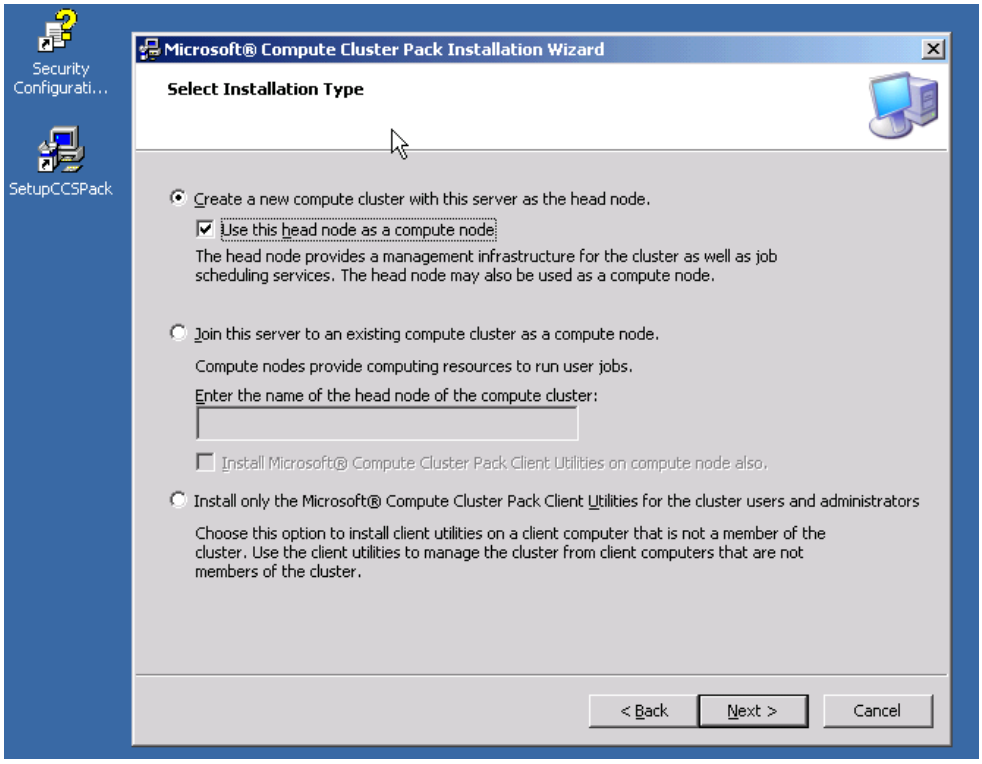
Now, Windows Server is installed on all nodes. The network is configured. It is time to install the Microsoft Compute Cluster Pack (CCP).

You have ordered a computer with Windows CCS pre-installed. This means CCP and its installer have been copied to the hard disk with a link to the installer (Setup.exe) created on the Desktop.

The following operations are necessary to build a compute cluster:

On the head node

1. Double click on the CCS Setup icon on the desktop to install the Compute Cluster Pack.
2. Select the **Create a new compute cluster with this server as head node** option



3. Install the components required and the CCP
4. The CCS head node creation is complete. Go to the Compute Cluster Pack **To Do List** to declare the network topology that has been configured above.

On each compute node

1. Double click on the CCS Setup icon on the desktop to install the Computer Cluster Pack.

2. Select **Join this server to an existing compute cluster node** and enter the full computer name of head node.

3. Install the **.NET** framework and the **CCP**

On the head node

1. Click **Start** -> **All Programs** -> **Microsoft Compute Cluster Pack** -> **Compute Cluster Administrator** -> **To Do List** -> **Node Management** -> **Add nodes**; add all computer nodes of the cluster in the list box before validating by clicking on **Next**.

2. Under **Compute Cluster Administrator**, click on **Node Management** to display the list of compute nodes. These nodes are under **Pending Approval**. Select all, right click and select **Approve**.

3. All the above nodes will change to the **Paused** state

4. Select all these nodes again, right click and select **Resume**

5. All the above nodes will change to the **Ready** state

6. Complete the security section of the **To Do List** to add users and administrators to the cluster.

   a. If MPI network is based on InfiniBand, execute the following command on the head node:

```
cluscfg setenvs MPICH_SOCKET_SBUFFER_SIZE=0
```

   b. If **Voltaire HC410** is used for **InfiniBand**, add the following environment variables to the system (Click on **Control Panel** -> **System** -> **Advanced** -> **Environment Variables** -> **System Variables** -> **New**):

```
IBWSD_POLL=500
IBWSD_NO_IPOIB=1
```

   c. On Compute Nodes based on **R422** with InfinBand included, add

```
C:\Program Files\Mellanox\WinIB\IPOIB
```

   to the path environment variable (to do this, Click on **Control Panel** -> **System** -> **Advanced** ->**Environment Variables** -> **System Variables** ->**Path** -> **Edit**)

7. The cluster installation is complete.
   Execute the following command on the head node:

```
clusrun /all ipconfig /all
```

   to obtain the ouput of the **ipconfig /all** command for each node.

Appendix A. Windows HPC Links

- Windows HPC Home

http://www.microsoft.com/windowsserver2003/ccs

- Windows Compute Cluster Server 2003 Reviewers Guide

http://www.microsoft.com/windowsserver2003/ccs/reviewersguide.mspx

- Windows Compute Cluster Server 2003 System Requirements

http://www.microsoft.com/technet/ccs/sysreqs.mspx

- Windows HPC Community

http://windowshpc.net

Technical publication remarks form

Title:	R440, R460, R421, R422, R422-INF HOW TO: Create a Windows Compute Cluster (WCCS)

Reference:	Rev02v01

Date:	July 2007

ERRORS IN PUBLICATION

SUGGESTIONS FOR IMPROVEMENT TO PUBLICATION

Your comments will be promptly investigated by qualified technical personnel and action will be taken as required.
If you require a written reply, please include your complete mailing address below.

NAME: _____ DATE: _____

COMPANY: _____

ADDRESS: _____

_____

Please give this technical publication remarks form to your BULL representative or mail to:

Bull - Documentation D^ept.
1 Rue de Provence
BP 208
38432 ECHIROLLES CEDEX
FRANCE
info@frec.bull.fr

Technical publications ordering form

To order additional publications, please fill in a copy of this form and send it via mail to:

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

Phone: +33 (0) 2 41 73 72 66
FAX: +33 (0) 2 41 73 70 66
E-Mail: srv.Duplicopy@bull.net

Reference	Designation	Qty
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
_ _ _ _ _ _ _ _ _ [_ _]		
[_ _] : The latest revision will be provided if no revision number is given.		

NAME: ______________________________ DATE: __________

COMPANY: ______________________________

ADDRESS: ______________________________

____

PHONE: ____________________ FAX: ____________________

E-MAIL: ______________________________

For Bull Subsidiaries:
Identification: ______________________________

For Bull Affiliated Customers:
Customer Code: ______________________________

For Bull Internal Customers:
Budgetary Section: ______________________________

For Others: Please ask your Bull representative.

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

REFERENCE
Rev02v01