CoCoNUT Over Windows Help

CoCoNUT over Windows Help

Version:

CoCoNUT Version 1.0.

Distribution:

The distributed CoCoNUT kit contains the following:

- Pre-compiled binaries of CoCoNUT.
- Pre-compiled binaries for genometools in the subdirectory /bin/genometools-current
- Source code for CoCoNUT
- Documentation in the folder "docs" containing manual and help documents as well as license agreement.

Installation and system requirements:

System Requirements:

- **Windows HPC Server 2008 RTM version.**
- **HPC component:** The prerequisites and the configuration of HPC component is explained in Appendix 1.
- **MS-MPI:** This already installed with the HPC component.
- **Cygwin:** Needed if you need to run your application based on cygwin. The installation is explained in Appendix 2.
- **Subsystem for UNIX-based Applications (SUA):** Installation is explained in Appendix 2.
- **Perl**: It is already installed with SUA and Cygwin.
- **Gnu plot** (at least version 3.7): optional for producing images of comparison results. The configuration is explained in Appendix 2.
- **Genometools:** We provide pre-compiled binaries. This package is open source and available at http://genometools.org/pub/genometools-1.3.0.tar.gz. We explain the installation of genometools in Appendix 2 in case you like to run the system on different Windows version.
- **Ghost script:** optional to display postscript files produced by **Gnu plot**. You can download it from *http://www.ghostscript.com/*

Running CoCoNUT

Download the package form the site. Be sure that cygwin is already installed on your system. If "genometools-current" not available or you want to updated, copy the compiled version of it to the "bin" directory in the CoCoNUT folder (see Appendix 2 for compiling genometools). We have three possibilities to run CoCoNUT. The last one can be used to assign a large number of jobs for comparing multi- chromosomal genomes.

1. Running CoCoNUT through Cygwin

The following steps show you how to run CoCoNUT based on Cygwin using command prompt.

1. Open a console/terminal by for example using the Run interface in windows 'Click start -> Run -> cmd'

2. Use cd command and move to the CoCoNut directory like this

```
> cd c:\cygwin\home\Microsoftcoconut
```

Now, you can test your installation. Do the following steps:

1. Write this command in the terminal

```
> C:\cygwin\home\MicroSoftCoCoNUT>perl coconut.pl –pairwise
```

This command will print a help of the system.

2. Download the testdata (if not already included in your distribution) and put it in CoCoCNUT directory. Then write this command in the terminal

```
> C:\cygwin\home\MicroSoftCoCoNUT>perl coconut.pl –pairwise\
testdata/chlamd/AE001273.fasta testdata/chlamd/AE001363.fasta -v –
plot
```

Running multiple comparisons: In comparing multi-chromosomal genomes, where each genome is a set of chromosomes, one needs to compare each chromosome of the first genome to each chromosome in the second genome. For example, the human genome is composed of 24 non-homologous chromosomes, and the mouse genome is composed of 21 non-homologous chromosomes. You can write a shell script with multiple command lines like the one above for different pairs of chromosomes.

Run CoCoNUT by SUA from bash file.

The command "**call ksh.bat**" is used to call the SUA and submit the command to it.

```
call ksh.bat -l -c "perl home/Microsoftcoconut/coconut.pl –pairwise
home/Microsoftcoconut/testdata/chlamd/chr22-human.fa.masked
home/Microsoftcoconut/testdata/chlamd/chr16-mouse.fa.masked -v -plot"
```

Running multiple comparisons: Create a batch file with extension ".bat", and write multiple command lines like the one above for comparing different pairs of chromosomes.

To run it from job scheduler

You can submit jobs through GUI of HPC component or through the command line. Here we use the command line interface with the job scheduler. For using GUI of job scheduler and more details, see Appendix 3.

To run CoCoNUT to compare two chromosomes, use the following command line:

```
>job submit /numofnodes:1 /workdir:\\H-Node\Users\Hishama\Desktop\Run\
/stdout: out\out.txt /stderr: errors\error.txt perl
home/Microsoftcoconut/coconut.pl –pairwise
Microsoftcoconut/testdata/chr22-human.fa.masked
Microsoftcoconut/testdata/chr16-mouse.fa.masked -v -plot
```

The following explains the arguments of this command line:

Job: specifies a cluster job

Submit: used to submit a job to the cluster

/numofnodes: defines number of nodes for this job. In our case we have 1 run to run on one node (here it will run on the head node). But if you assign a larger number of nodes, which is recommended, then the job scheduler will use the most available node of the cluster to run the job.

/workdir filepath: specifies a shared directory between all nodes, which is a requirement of the Windows HPC. This directory should be accessible by all nodes of the cluster. See Appendix 3 for more details.

/stdout filepath/filename: Specifies the folder and output file to store the standard output of the application (the output of the program output to the standard output). In this example, we have the file "out.txt" in the directory "out". The folder should be pre-created in the shared directors so that all nodes can write on it, but the file should not, as it can be automatically created.

/stderr filepath/filename: The same as stdout but for error messages.

Perl Microsoftcoconut/coconut.pl "arguments": This is the application to be run on the job scheduler, which is the CoCoNUT application with its arguments.

Running multiple comparisons:

Assume we have 5 human chromosomes (hschr1.fa, hschr2.fa, hschr3.fa, and hschr4.fa, hschr5.fa) to be compared to 4 mouse chromosomes (mmchr1.fa, mmchr2, mmchr3, and mmchr4). Assume all the chromosomes are in the testdata directory. To run all-against-all comparison, we have to submit 5 "parametric jobs" to the job scheduler. Each parameteric job compares one human chromosome to 4 mouse chromosomes. We make use of the "prefix" command (see CoCoNUT manual document) with the wild character of the parametric job to define different output directories. The submission command for line comparing human chromosome Z (Z ϵ [1..5]) is as follows (do not break lines):

```
>job submit /parametric: 4 /numofnodes:4 /workdir:\\H-
Node\Users\Hishama\Desktop\Run\ /stdout: out\out*.txt /stderr:
errors\error*.txt perl Microsoftcoconut/coconut.pl –pairwise
Microsoftcoconut/testdata/hschrZ.fa.masked
Microsoftcoconut/testdata/mmchr*.fa.masked -v –plot –prefix
Microsoftcoconut/testdataZx*/hschrZmmchr*
```

The extra argument of this command line are as follows:

`/Parametric:` defines different inputs and output. In our case, we have four jobs running on 4 different input files (chromosomes) .

Note that here we use the maximal number of nodes we have. Note that the output directory will be created for each comparison and with different prefix. For example, when Z =1 and we are running with respect to mouse chromosome 2, the output directory will be `Microsoftcoconut/testdata1x2/.` The prefix for all reported files with this directory will be `hschr2mmchr1.`

To run for each Z you have, you might write a batch file for this task.

Appendix 1

Installation of HPC Server 2008

Hardware requirements:

Processor (x64-based):
• Minimum: 1.4 GHz
• Recommended: 2 GHz or faster
RAM:
• Minimum: 512 MB
• Recommended: 2 GB or more
Available disk space:
• Minimum: 50 GB
• Recommended: 80 GB or more
Drive:
• DVD-ROM drive
Network adapters:
• The number of network adapters on the head node and on the compute nodes depends on the network topology that you choose for your cluster.

Decide which node is the head node and which are the compute Nodes:

Choose the Head Node which has the best specification cause it will be responsible for manage tasks between nodes in the cluster. Head node also can be a compute node at the same time by taking part of the task. To add a compute node there are three ways to add compute nodes:

- **From bare metal**. The operating system and all the necessary HPC cluster components are automatically installed on each compute node as it is added to the cluster. No manual installation of the operating system or other software is required.

- **Add preconfigured compute nodes.** The compute nodes are already running one of the 64-bit editions of the Windows Server2008 operating system, and Microsoft HPC Pack 2008 is manually installed on each node.

- **Import a node XML file.** An XML file that contains a list of all the nodes that will be deployed is used. This XML file can be used to add nodes from bare metal or from preconfigured nodes.

Active Directory Domain for CLUSTER

The head node and the compute nodes in your HPC cluster must be members of an Active Directory domain. Before deploying your cluster, you must choose the Active Directory domain that you will use for your HPC cluster. You can install Active Directory Domain Services on the head node. For more information about installing Active Directory Domain Services on a computer running Windows Server2008, see http://go.microsoft.com/fwlink/?LinkID=119580.

Choose a User Account for Installation and Diagnostics

During the configuration process of your HPC cluster, you must provide credentials for a domain user account that will be used for installation and diagnostics. You must choose an existing account or create a new account, before starting your cluster deployment. The user account that you choose:

- Must be a domain account.
- I f you want to restart nodes remotely from the cluster administration console (HPC Cluster Manager), the account must be a member of the local Administrators group on the head node.

Network Topology for Cluster

Windows HPC Server2008 RTM supports five cluster topologies. The five supported cluster topologies are:

- Topology No. 1: Compute Nodes Isolated on a Private Network
- Topology No. 2: All Nodes on Enterprise and Private Networks
- Topology No. 3: Compute Nodes Isolated on Private and Application Networks
- Topology No. 4: All Nodes on Enterprise, Private, and Application Networks
- Topology No. 5: All Nodes on an Enterprise Network We used Topology 5 for our Cluster.

Deploying Cluster Nodes

You must start by installing one of the 64-bit editions of the Windows Server2008 operating system on the computer that will act as the head node.

Install Microsoft HPC Pack 2008 on the Head Node:

1. To start the Microsoft HPC Pack2008 installation wizard on the computer that will act as the head node, run setup.exe from the HPC Pack 2008 installation media or from a network location.
2. On the Getting Started page, click **Next**.
3. On the Microsoft Software License Terms page, read or print the software license terms in the license agreement, and accept or reject the terms of that agreement. If you accept the terms, click **Next**.
4. On the Select Installation Type page, click Create New Compute Cluster, and then click **Next.** See Figure (1).
5. On the Basic New Compute Cluster page, click Create new instance using SQLExpress, and then click **Next.** See Figure (2).
6. On the Select Installation Location page, click **Next**. See figure (3).
7. On the Install Required Components page, click **Install**. See figure (4).
8. On the Installation Complete page, click **Close**.

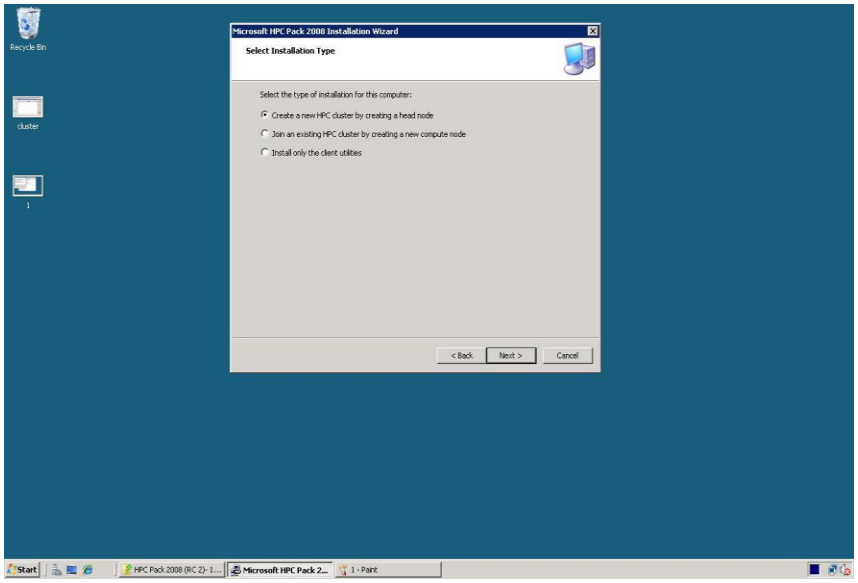

Figure 1: Create a new HPC

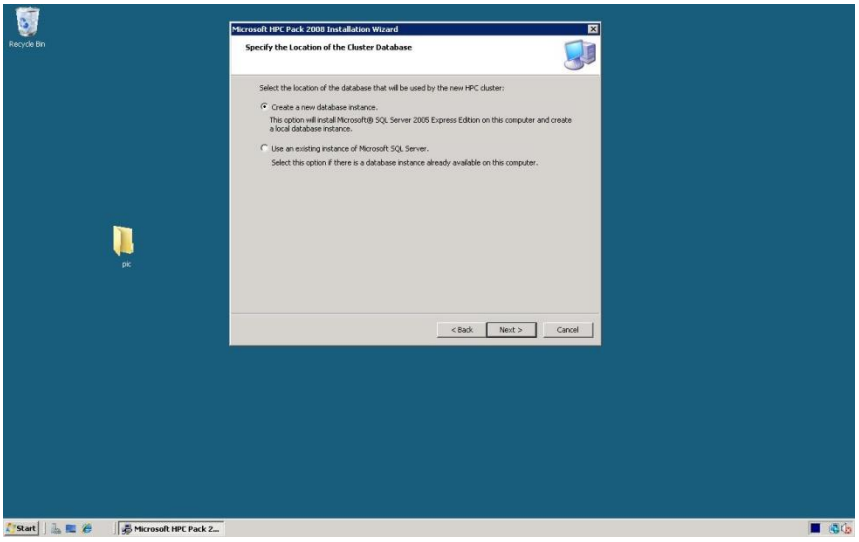

Figure 2 Create New Database

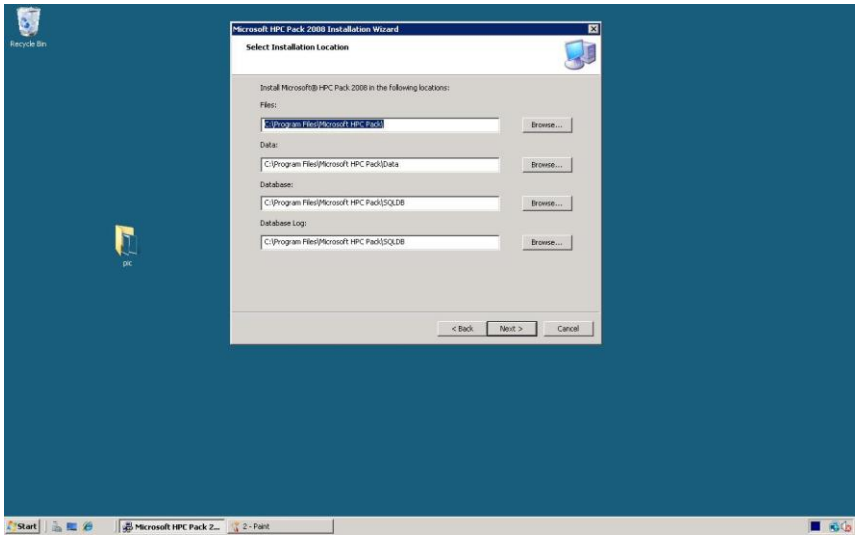

Figure 3 Installation Location

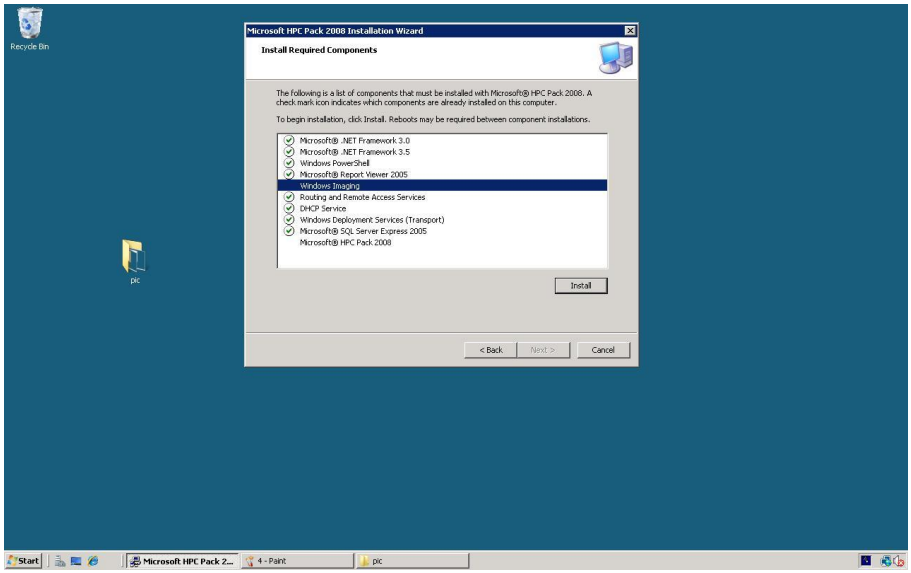

Figure 4 Required components

Configure the Head Node:

The HPC cluster network configuration is the first step in the configuration process of your head node. The HPC cluster network is configured by following the Network Configuration wizard in HPC Cluster Manager.

1. If HPC Cluster Manager is not already open on the head node, open it. Click Start, point to All Programs, click Microsoft HPC Pack, and then click HPC Cluster Manager.
2. To start the Network Configuration wizard, in the **To do list**, click **Configure your network.** See figure (5).
3. On the Network Topology Selection page, click the topology that you have chosen for your cluster, and then click **Next**. We chose topology number 5 for our cluster see figure(6).
4. On the Enterprise Network Adapter Selection page, in the Network adapter list, click the name of the network adapter that is physically connected to your enterprise network, and then click **Next.** See figure (7).
5. On the Firewall Setup page, select the firewall setting for the cluster. We disabled the firewall on the network see figure (8).
6. On the Review page, verify your settings and click **Configure**.
7. After the network configuration process is completed, on the Configuration Summary page, review the list of configuration items.
8. To close the wizard, click **Finish.**

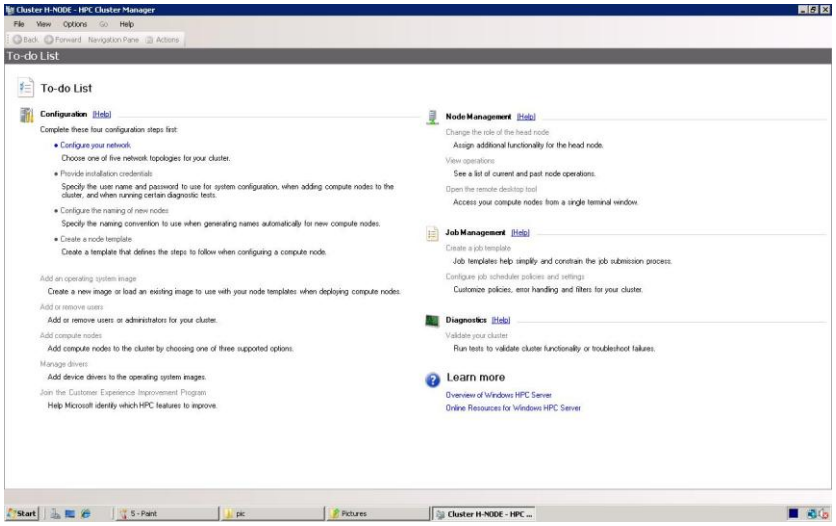

Figure 5 Configure Network

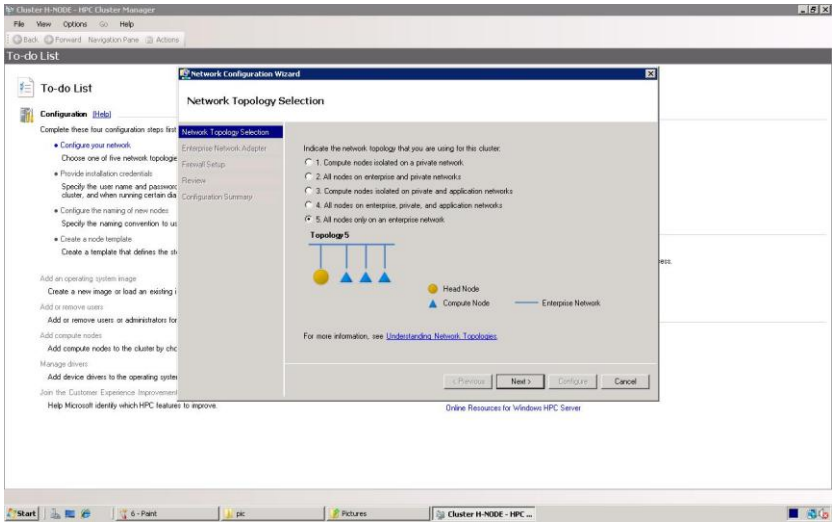

Figure 6 Select Topology

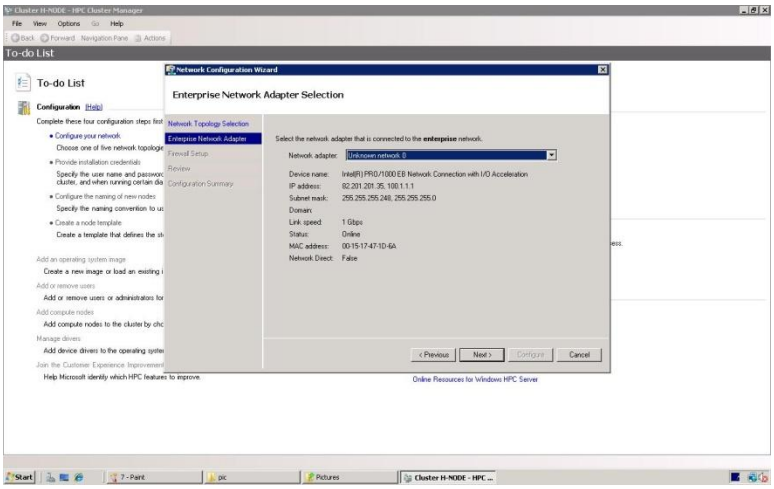

Figure 7 Network device

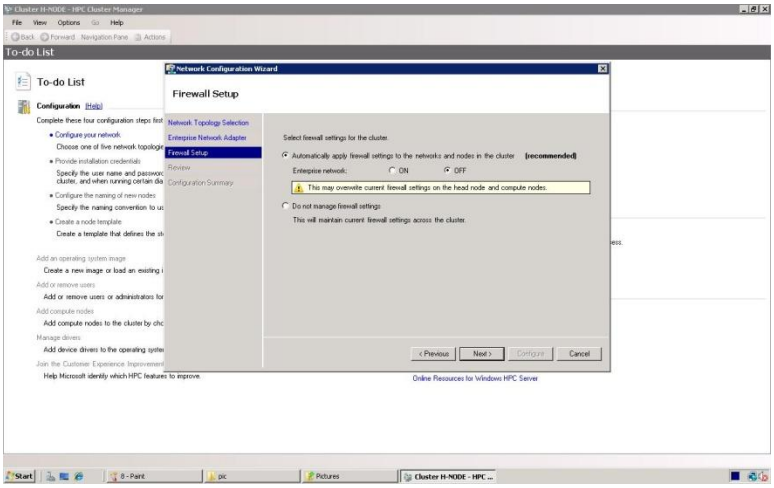

Figure 8 Firewall

Provide Installation Credentials

1. In the **To do list**, click Provide installation credentials see figure (9).
2. Type the user name, including the domain (DOMAIN User), and then the password for the domain user account see figure (10).
3. To save the specified credentials, click **OK**

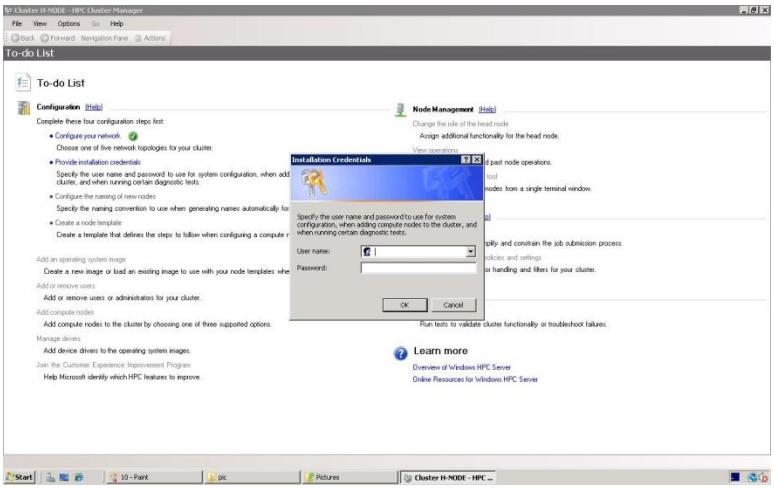

Figure 9 Installation credentials

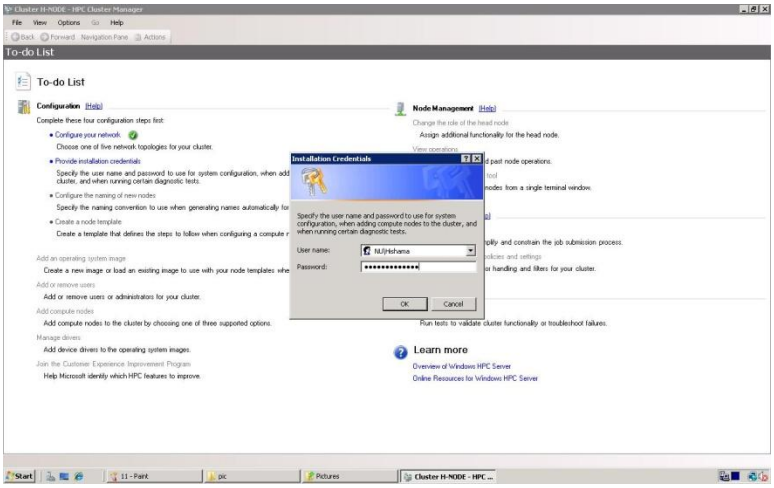

Figure 10 Network device

Configure the Naming of Nodes

You need to specify how those names will be generated, by defining a naming series. The naming series is defined by selecting a root name and the starting number that will accompany that name. The starting number is enclosed in percentage signs

1. In the **To do list**, click **Configure** the naming of new nodes.
2. Type the naming series that you want to use. We used 1000 but it is not important for us because we will add preconfigured nodes see figure (11).
3. Click **OK**.

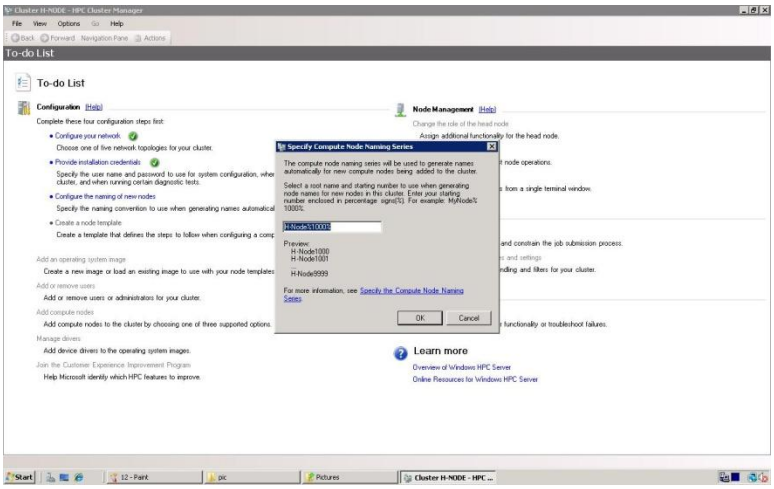

Figure 11 Naming of Nodes

Create a Node Template

Node templates are new in WindowsHPC Server2008. They define the necessary tasks for configuring and adding compute nodes to your cluster.

With a node template, you can deploy an operating system image, add specific drivers and software to compute nodes, or simply add a preconfigured node to your cluster. Because you might have more than one type of compute node, or you may be adding compute nodes to your cluster in different ways, you can create different templates that apply to different nodes or situations.

You can create two types of node templates:

• With an operating system image. This type of template includes a step to deploy an operating system on the compute nodes. Use this type of template when adding compute nodes from bare metal.

• Without an operating system image. This type of template is used to add preconfigured compute nodes to the cluster, or to update existing nodes.

We created a template without an operating system image.

1. In the **To do list,** click **Create a node template**.
2. On the Specify Template Name page, type a descriptive name for the template, and then click **Next** see figure (12).
3. Select **Deployment Type page**, click **without operating system**, and then click **Next**.
4. On the Specify Windows Updates page, specify if you want to add a step in the template to download and install updates using Microsoft Update or the enterprise Windows Server Update Services (WSUS).Click **Next** to continue see figure (13).
5. On the Review page, click Create.

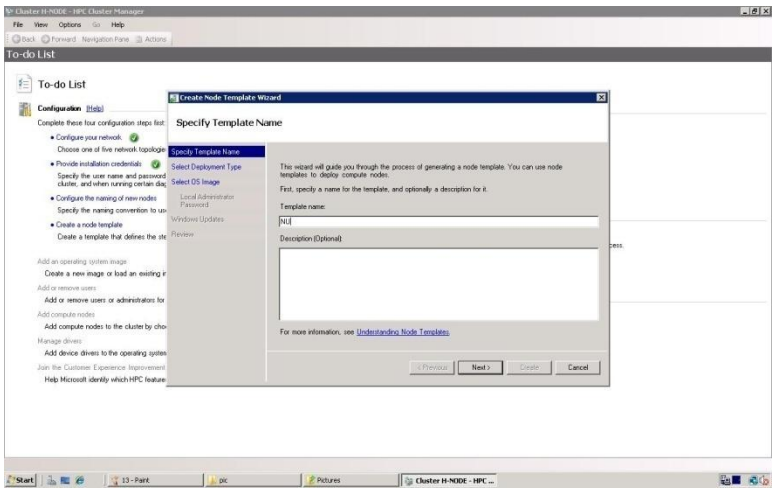

Figure 12 Template Name

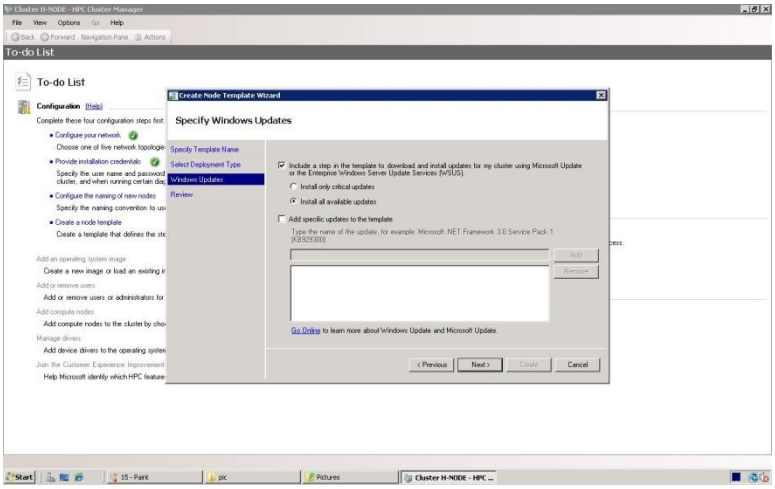

Figure 13 Update

Add Compute Nodes

After creating a node template, you can use the **Add Node** wizard to add compute nodes to your HPC cluster. There are three ways by which you can add compute nodes to your cluster:

1. Deploy Compute Nodes from Bare Metal
2. Add Compute Nodes by Importing a Node XML File
3. Add Preconfigured Compute Nodes

We used third one.

Add Preconfigured Compute Nodes

The following procedure describes how to add preconfigured compute nodes to your HPC cluster. The following steps will be done on all compute nodes.

1. Run setup.exe from the HPC Pack 2008 installation media or from a network location.
2. On the Getting Started page, click **Next**.

3. On the Microsoft Software License Terms page, read or print the software license terms in the license agreement, and accept or reject the terms of that agreement. If you accept the terms, click **Next**.
4. On the Select Installation Type page, click **Join Existing Compute Cluster**, and then click **Next**.
5. On the **Join Cluster page**, type the computer name of the **head node** on your cluster, and then click **Next**.
6. On the Select Installation Location page, click **Next**.
7. On the Install Required Components page, click **Install**.
8. On the Installation Complete page, click **Close.**

Do these steps on all Compute Nodes then when node is ready it will appear in the HPC Cluster Manager on the Head node.

It will appear with status **Unknown** to bring them online:

1. In the **To do list**, click **Add compute nodes**.
2. Right click and then select **Add node** see figure (14).
3. On the **Select New Nodes page**, in the Node template list, click the name of a node template that does not include a step to deploy an operating system image.
4. Select the preconfigured compute nodes that you want to add to your cluster.
5. Click **Add**.
6. On the Completing the **Add Node Wizard page**, click **Respond only to PXE requests** that come from existing compute nodes.
7. Click Finish.
8. During the deployment process of a compute node, its state is set to **Provisioning.** When the deployment process is complete, the state changes to **Offline** see figure (15).
9. In Node Management, under **By State**, click **Offline**.
10. Select Nodes and then right click on them, select **Bring Online,** see figure (16).
11. Results in figure (17).

Now the cluster is ready to run the jobs.

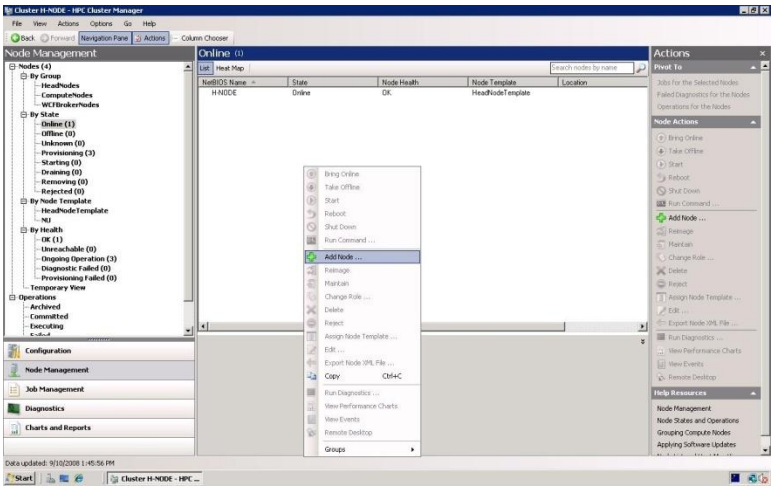

Figure 14 Add Node

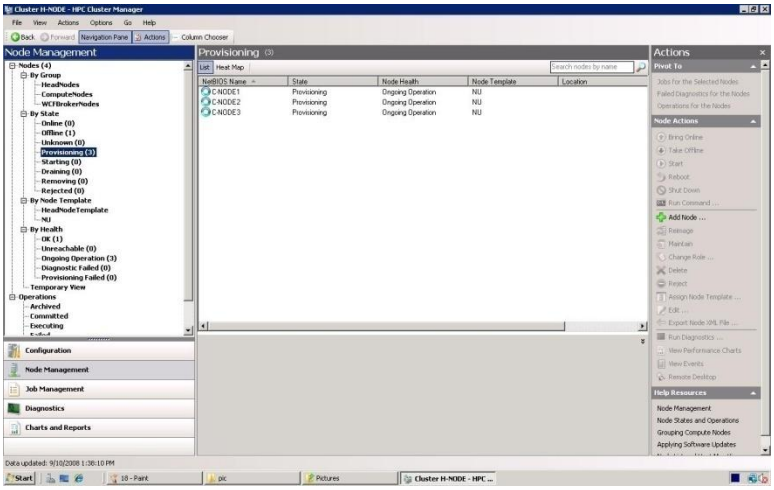

Figure 15 Add Node

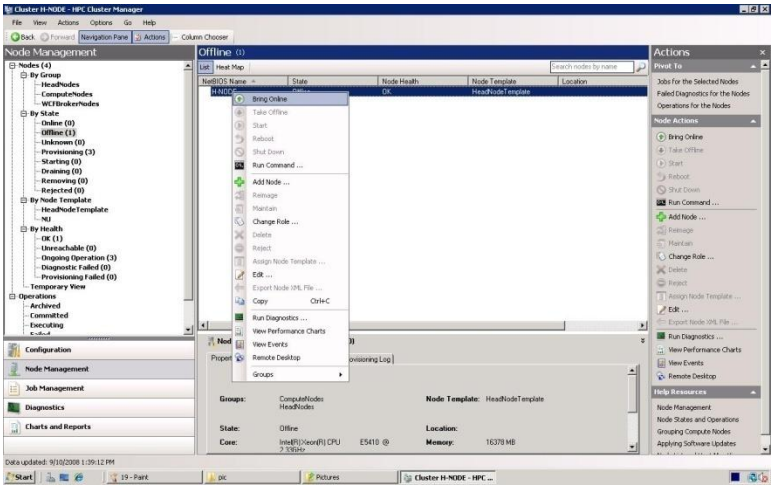

Figure 16 Bring online

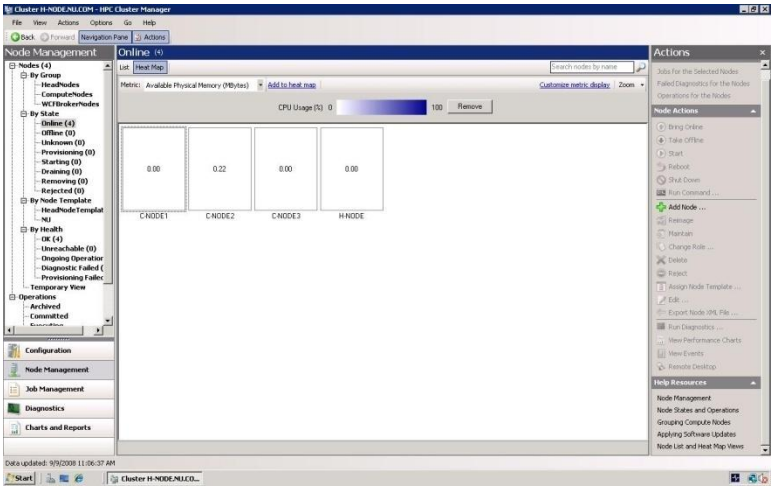

Figure 17 Results

Appendix 2

Installing Cygwin

Cygwin is a Linux-like environment for Windows. It consists of two parts:

1. A DLL (cygwin1.dll) which acts as a Linux API emulation layer providing substantial Linux API functionality.
2. A collection of tools which provide Linux look and feel.

The Cygwin DLL currently works with all recent, commercially released x86 32 bit and 64 bit versions of Windows, with the exception of Windows CE. There was a problem in running Cygwin under windows server 2008 RC version but this problem was solved in the RTM version

You can download Cygwin from this Link: http://www.cygwin.com.

Installation steps

1. Choose install from Internet if you do not have source as shown in figure (18).
2. Select root directory, choose Default Text file type (Unix/binary (Recommended Option)) and choose the local package directory as shown in figure (19).
3. Select the direct connection as shown in figure (20).
4. Choose the desired mirror site to download the program and its utilities as shown in figure (21).
5. From the select package install all the Devel Hierarchy to download and setup the G++ also choose ”XFree86” (UPDATE: ”X11” entry) (X Windows) entry (about second from last), and click on the text ”Default” which should change to say ”install” as shown in figure(22).

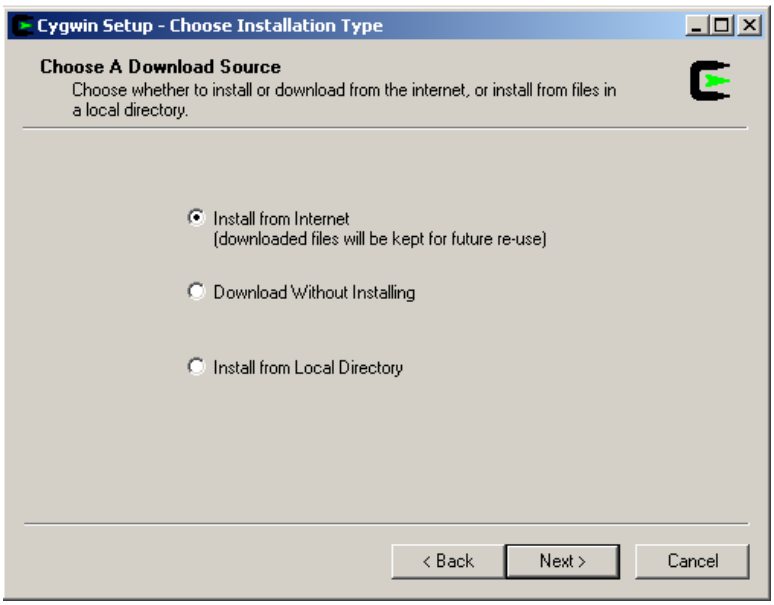

Figure 18 Choose install Source

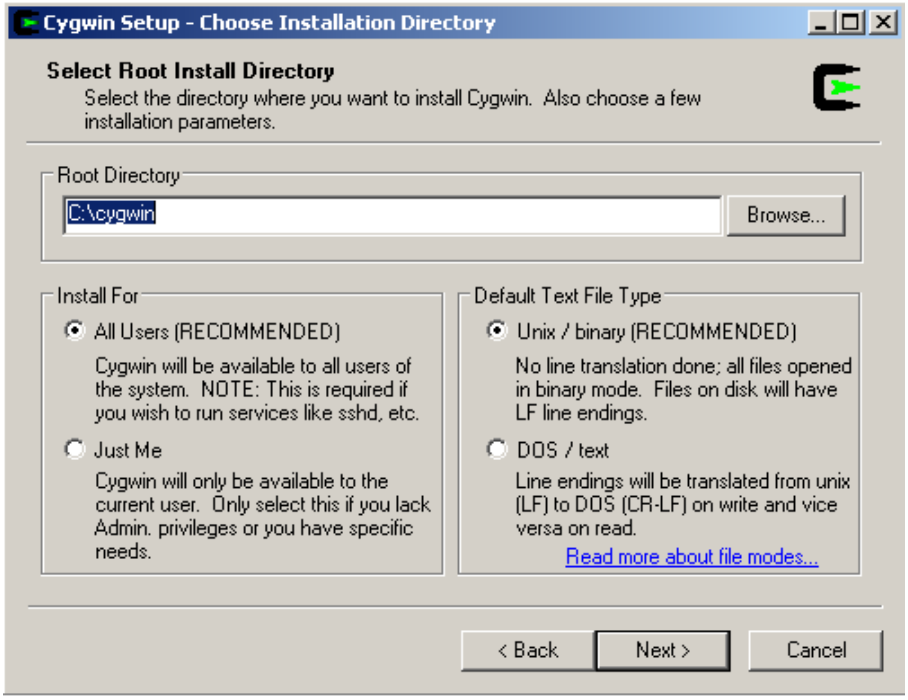

Figure 19 Choose Root Directory

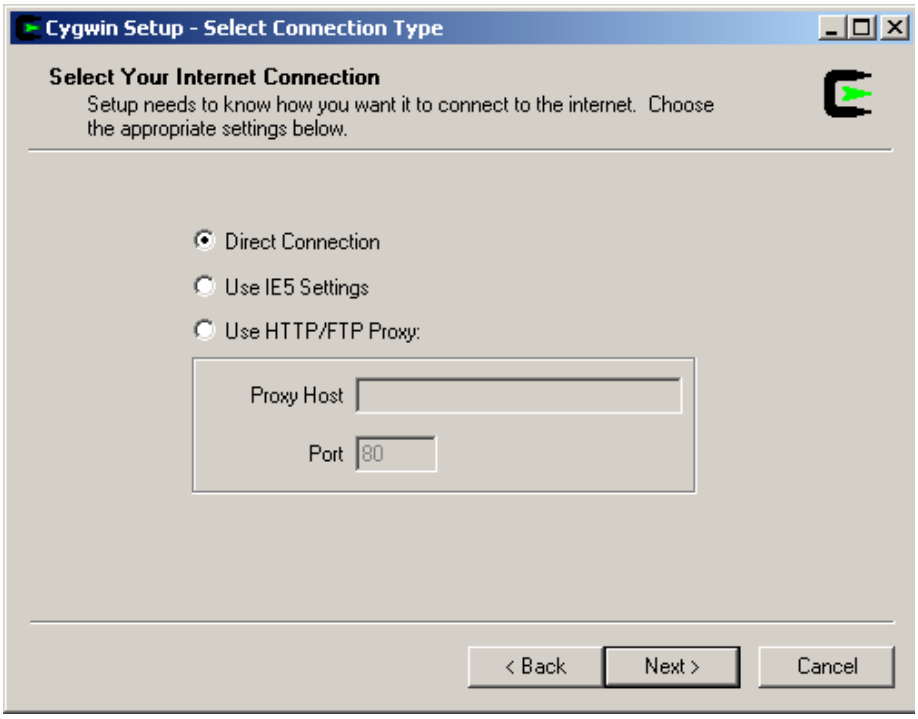

Figure 20 Connection

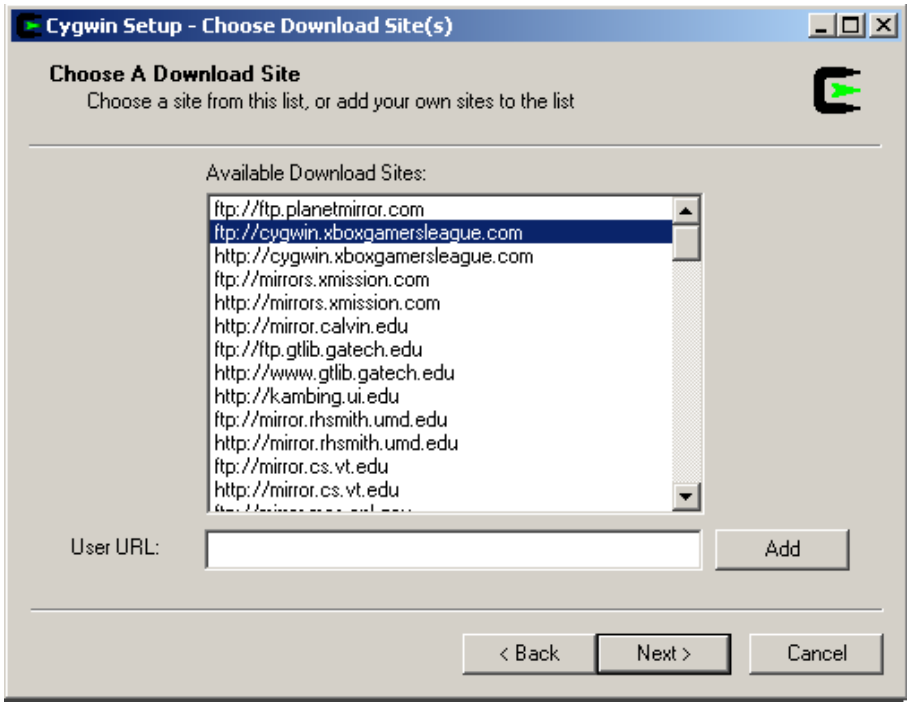

Figure 21 choose site

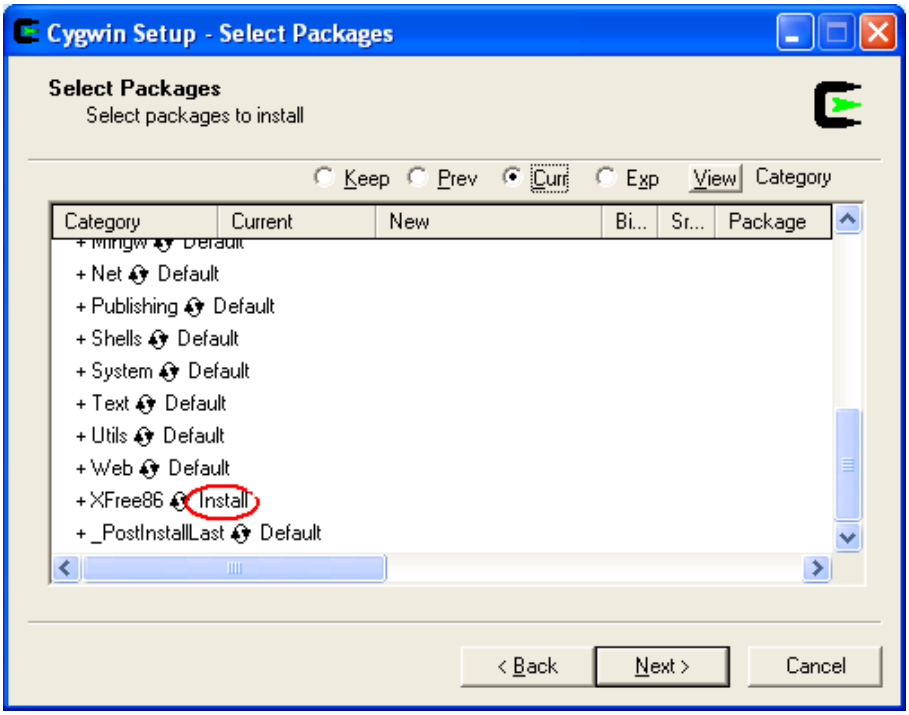

Figure 22 Packages to install

Setting up the environment variable

In order to make Cygwin works from command prompt do the following steps:

1. Right Click on My computer Icon on the desktop.
2. Choose properties and then Advanced from the tabs
3. Click on Environment variables, new windows will open as shown figure (23).
4. From variables choose Path and then edit as shown in figure (24).
5. In the Variable Value add this line after the semicolon.
6. Now you are ready to use Cygwin from command prompt.
7. C: \ cygwin\ bin ;C: \ cygwin\ us r \ sbin ;C: \ cygwin\ us r \X11R6\ bin ;

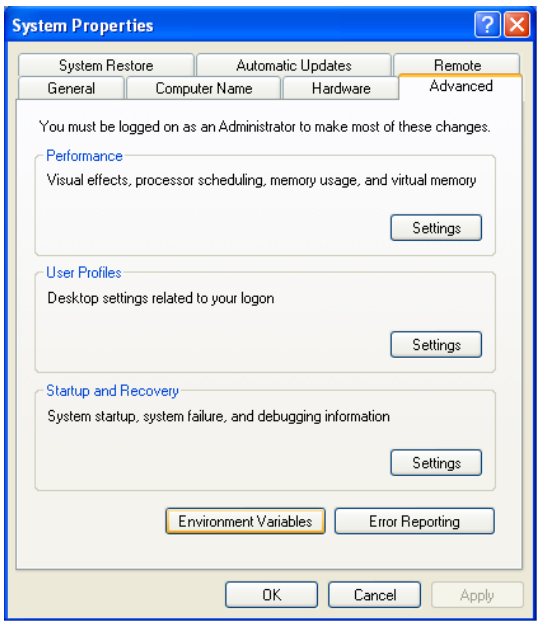

Figure 23 Environment variables

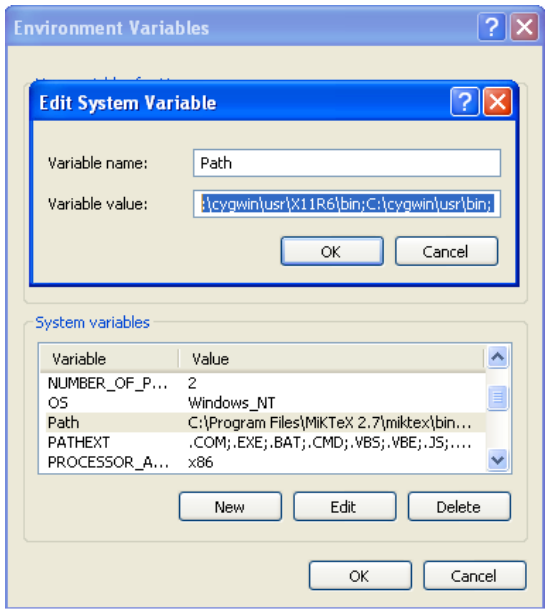

Figure 24 Environment variables

Install Subsystem for UNIX-based Applications (SUA)

1. Click Start -> Control Panel -> Programs -> Turn Windows features on or off, select "Subsystem for UNIX-based Applications". See figure (25)
2. Download SUA from http://www.microsoft.com/downloads/details.aspx?FamilyID=93ff2201-325e-487f-a398-efde5758c47f&displaylang=en
3. Run setup package, Select all packages and then click on install. See figure (26).

4. Click Start ->Program files-> Subsystem for UNIX-based Applications->C Shell see figure (27).

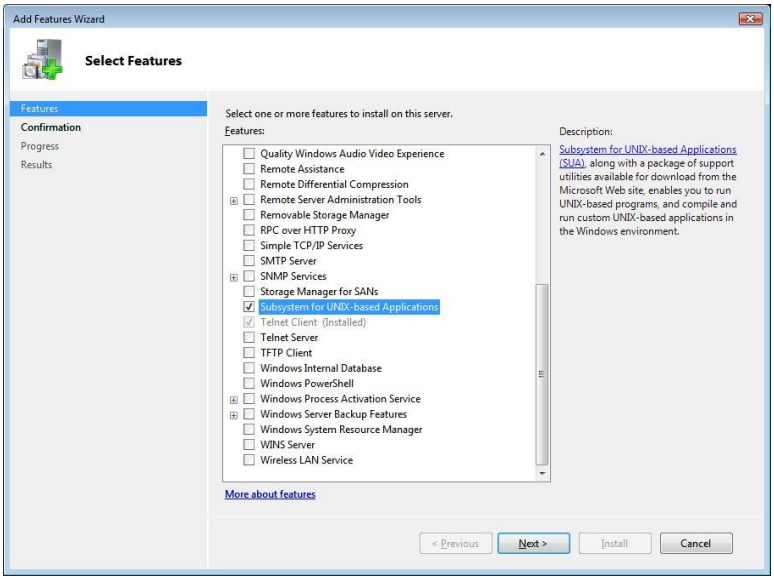

Figure 25 SUA

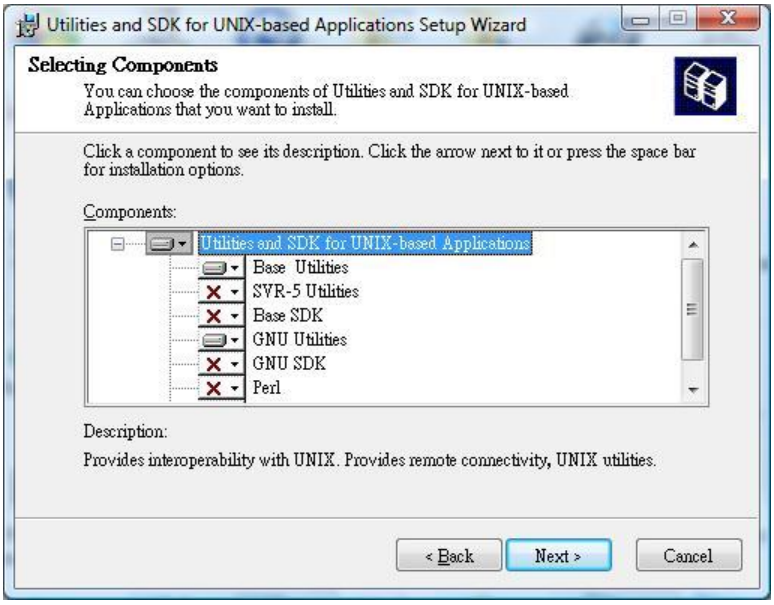

Figure 26 SUA

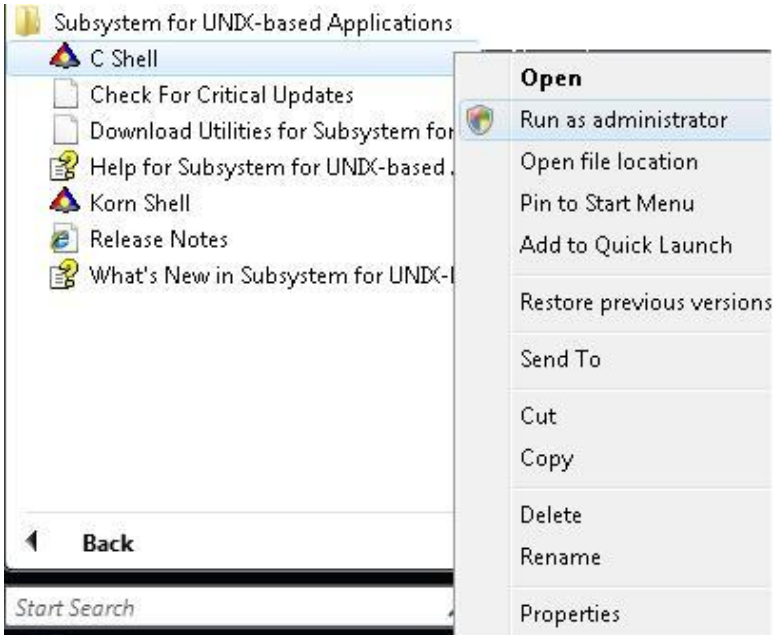

Figure 27 Start C Shell

Setting up and running GNU-Plot

1. Make a new shortcut on your desktop to this batch file:
   C: \ cygwin\usr\X11R6\bin\Startxwin.bat
2. Run your new shortcut.
3. A black window should appear briefly and then disappear.
4. A new icon should appear on the windows system tray as shown in figure (28).
5. A new window should appear as in figure(29)
6. On your windows desktop, right click on"My Computer" and select"Properties". Then click
7. On the "Advanced" tab. This should have a button at the bottom left called "Environment Variables", click on this.
8. Create a new System Variable by clicking on the new button as shown in figure (30).
9. Type in variable name of DISPLAY and Type in variable value of 127.0.0.1:0.0 .
10. Now you are ready to use the GNU-plot from the command prompt by writing gnuplot as shown in figure (31).

Figure 28 X Icon

Figure 29 Bash Window

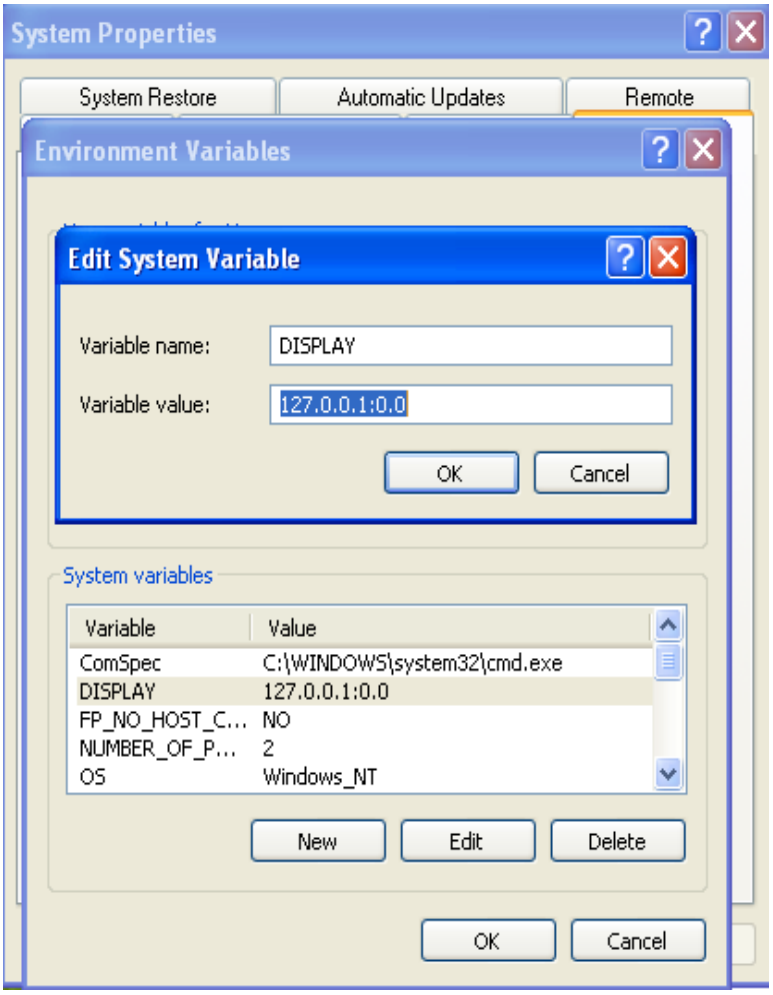

Figure 30 DISPLAY Variable

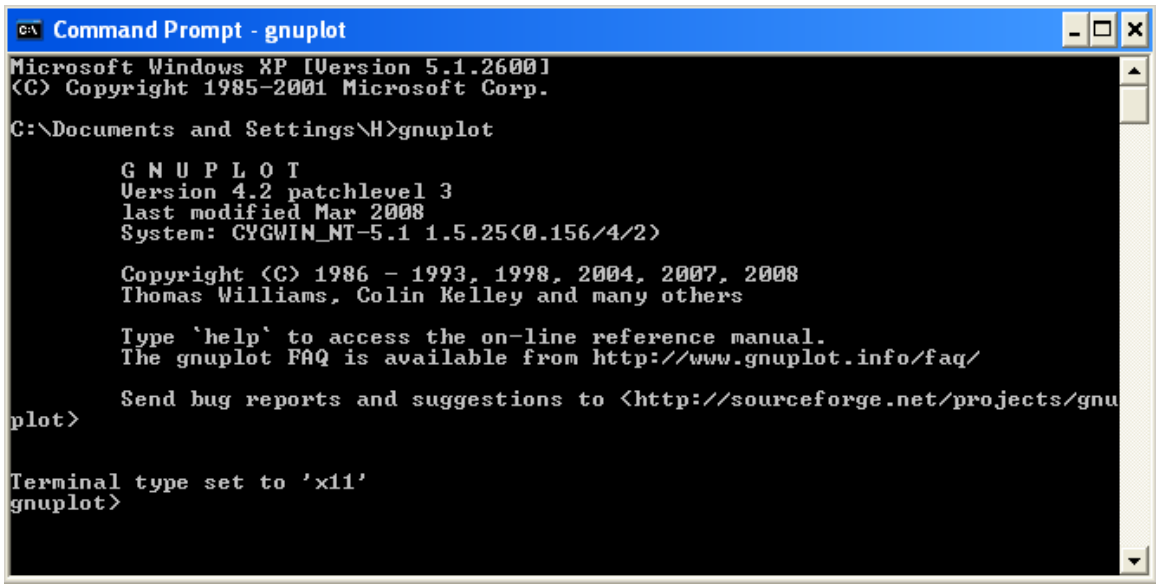

Figure 31 Gnuplot

Installing Genometools

We tried to install genome tools through SUA but we have faced some problems, which we will explain later.

Steps using Cygwin

After downloading the package. Put it in C:\cygwin\home

1. Extratct the packages

```
>tar -xvf genometools-1.3.0.tar.gz
```

2. Move to genometools directory

```
>cd genometools-1.3.0
```

3. Move to C:\cygwin\home\genometools-1.3.0 and then open "Makefile" using WordPad
4. Press "ctrl + f" write –FPIC then click **Find Next remove** this flag then save and close the file
5. Back to Cygwin write

```
>make prefix=/home/genometools-1.3.0 install
```

Where prefix define the installation directory

6. After finishing compiling write

```
>make install
```

7. Rename genometools-1.3.0 to genometools-current.
8. Now the genometools is installed we will need to test it, from Cygwin write

```
> /home/genometools-current/bin/gt.exe –help
```

The output will be something like that:

```
Usage: /home/genometools-current/bin/gt [option ...] [tool | script] [argument ...]
The GenomeTools (gt) genome analysis system (http://genometools.org).

-i          enter interactive mode after executing 'tool' or 'script'
-test       perform unit tests and exit
-help       display help and exit
-version    display version information and exit

Tools:

bed_to_gff3     extractfeat     matstat         sequniq         uniq
bioseq          extractseq      merge           shredder        uniquesub
cds             filter          mgth            sketch
chseqids        fingerprint     mkfmindex       splicesiteinfo
clean           gff3            mmapandread     splitfasta
congruence      gff3_to_gtf     mutate          stat
csa             gff3validator   packedindex     suffixerator
eval            gtf_to_gff3     prebwt          tagerator
exercise        ltrharvest      seqfilter       tallymer

Set the environment variable GT_ENV_OPTIONS=-spacepeak to show a spacepeak after program run
(GT_ENV_OPTIONS=-help shows possible ``environment options'').

Report bugs to <gt-users@genometools.org>.
```

Install genometools using SUA

We faced some problem and we have solved them to install genometools using SUA.

The steps of installation like Cygwin but you will need some set some things before installing:

1. To install genometools you will need to download “gmake” it will not work with “make”

2. When we were trying to compile genometools there was error at start of compilation

```
ar: Either the AR_LIBRARIAN environment variable must be set
ar: or you must include the directory containing 'lib.exe' in PATH
```

This is because there are problems in the libraries.

To solve this problem

Note: The **ar** in /bin is for 64bit cc compiles.

There is a 32bit lib ar in /opt/gcc.4.2/bin, so the easy way out is to put this path first.

We made these changes for gcc 4.2 in my .environ file, Write this commands in USA:

```
>export PATH="/usr/local/bin: /bin: /usr/contrib/bin:\
/usr/contrib/win32/bin: /usr/X11R5/bin :$(winpath2unix\
${SYSTEMROOT}/system32):/common"

>export TMPDIR="$(winpath2unix ${TEMP})"

>export LD_LIBRARY_PATH=/usr/lib/x86:/opt/gcc.4.2/lib:${__X11}/lib

     >export GCC_EXEC_PREFIX=/opt/gcc.4.2/lib

>export LIBRARY_PATH=/usr/lib/x86:/opt/gcc.4.2/lib
```

Further more:

1) Delete crt0.o and libc.a in /usr/lib - they are 32bit and the same files as in /usr/lib/x86.
2) Rename /bin/nm, /bin/strings, /bin/strip to e.g. /bin/nm.x86 ... …..
3) Rename the following files in /usr/lib/x86 to e.g. *.old and create new gcc.4.2 symbolic links. ld.so, ld.so.1, libdl.so, libdl.so.5.2, libg2c.so.5.2, libmath.so.5.2, libstdc++.so, libstdc++.so.5.2, libsupc++.so.5.2, libsupc++convenience.so.5.2

Then write:

```
> mv ld.so.1 ld.so.1.old
>ln -s /opt/gcc.4.2/lib/ld.so.1 ld.so.1 …
```

Do these two steps for all the files

So, with this it's easy to enable gcc 4.2 whenever you wish in the shell by simply enter:

```
>export PATH=/opt/gcc.4.2/bin:$PATH
```

3- SUA don't include some of Linux API like getpagesize(). And if I need to use it I have to include a header file <sys/mman.h> so I will need to open about 30 file and put this header in it. It took a lot of time to solve this problem and to know the solution of it. Cygwin didn't give me any errors for that.

4- int8_t,int16_t,int64_t,uint8_t,uint16_t,uint32_t ,uint64_t, PRId16, PRId32, PRId8.
All of these are exact-width integer definitions. In Cygwin it is already defined during compilation and it gave me no error, but with SUA it gives me errors and to solve that we have to open the file and but the header in it which is <inttypes.h> or to define these variables by hand.

Appendix 3

Job Submission

In order to submit a job to Cluster you need to make a share folder between all nodes of the cluster.

To create a share folder:

1. Make a new folder.
2. Right click on the folder select properties then click on **Sharing** tab then **share** button as shown in figure (32).
3. Select user which was defined in the domain as explained in Appendix 1.
4. Click **share** then click **done** as shown in figure (33).

Figure 32 Sharing

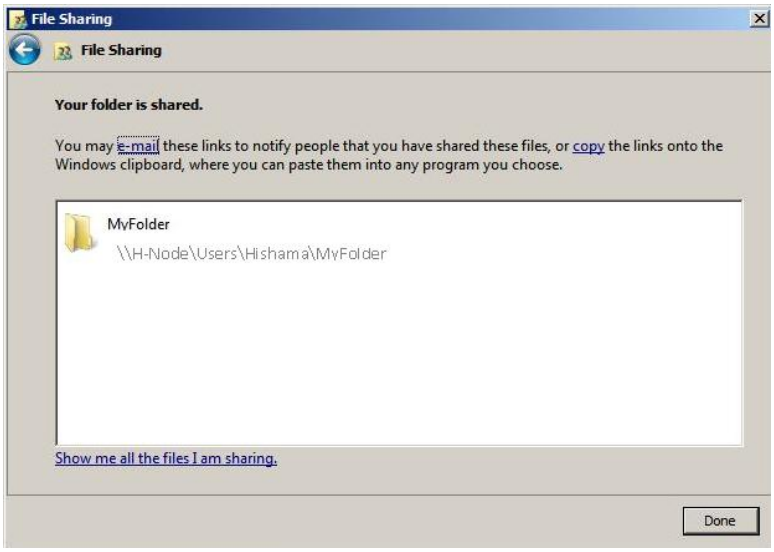

Figure 33 Sharing

Submit job to Cluster form GUI

Submit one job

In **Job Management**, in the **Actions** pane, click **New Job**. New windows will appear see figure (34).

1. In **Job Details**, specify the following job parameters:

1. In the **Job template** list, click **Test Template** (the template that you created).
2. In the **Job name** box, type **any name you want for the job**.
3. In the **Priority** list select it according to the task is it with high priority or medium or low.

2. In **Job run option**, you can specify these parameters or you can let it without specification.

1. To define duration for the job select **Do not run this job more than** and select the duration.
2. Run the job until it canceled by the user.
3. If the job contains more than one task make the job fail.

3. In **Job resources**, you define the resources that will be used in this job. You can select number of nodes or number of cores.

4. To add a task, click **Task List**, and then specify the following task parameters:

1. To add a new basic task to the job, click **Add**.
2. In the **Task name** box, type a name for the new task.

3. In the **Command line** box, type **the program name that you want to run**.
4. In the **Work directory** box, select the shared folder between all nodes which we have done before
5. If the program need inputs from user create a file called **input.txt** and writes the inputs in this file and put it in the shared folder. The **out file** show what the output of the program. The **errors file** show errors in running the program. You will need to create two file _out.txt and _err.txt and put them in the shared folder see figure (35).
6. To submit mpi task to the cluster write **mpiexec –machinefile hosts –n 4 program.exe.** Where **hosts** is a text file contains the names of the nodes of the cluster. **hosts** file need to be placed in the working directory which is the chared folder and **–n 4** define number of nodes on the cluster in our case there was only **4** nodes
7. Maximum and minimum define number of nodes

8. To add this task, click **Save**.

5. To limit the job so that it only runs on a specific compute node in your HPC cluster, click **Resource Selection**, and then specify the following resource parameters see figure (34):

1. Select the **Run this job only on nodes in the following list** check box.
2. Select the check box for one of the nodes in your HPC cluster.

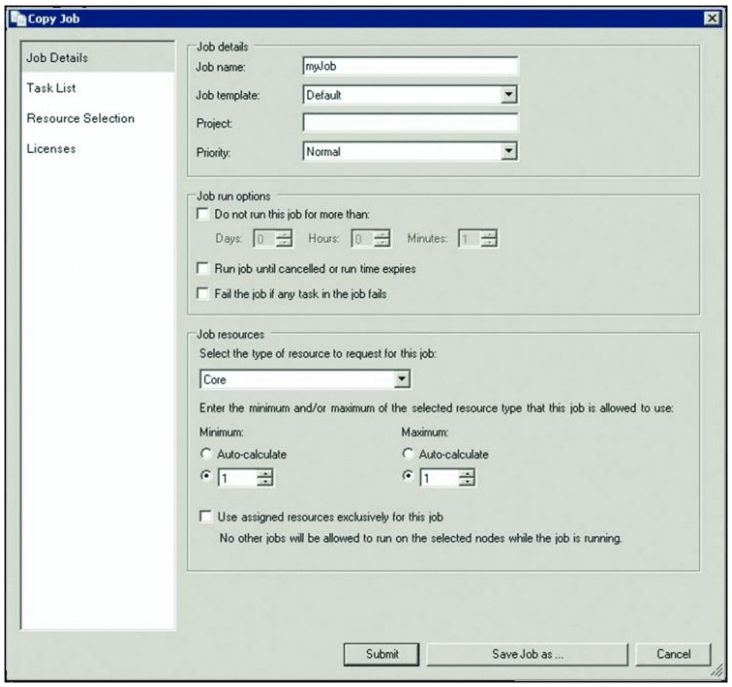

Figure 34 job submission

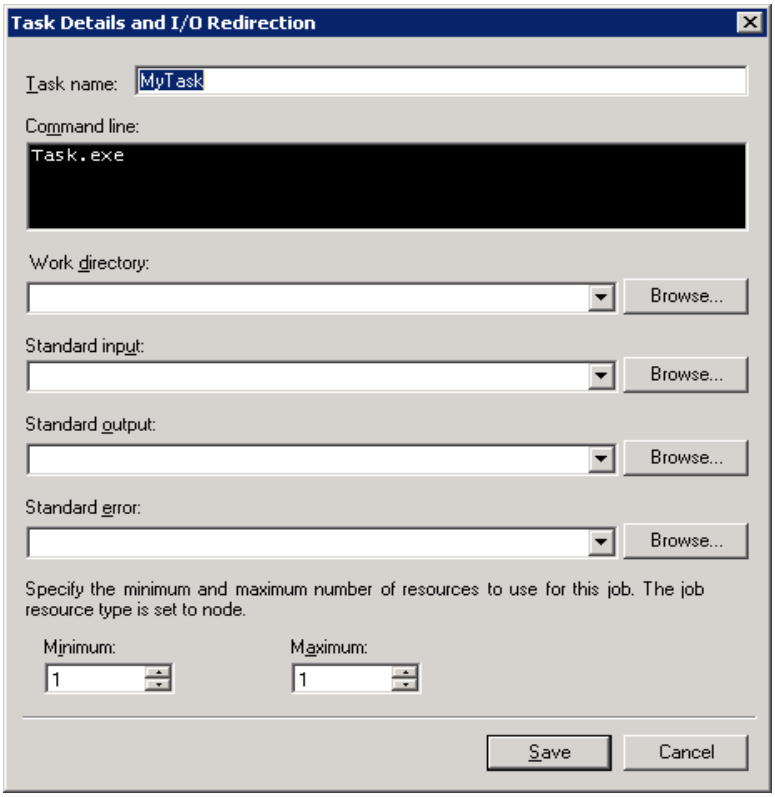

Figure 35 Add Task

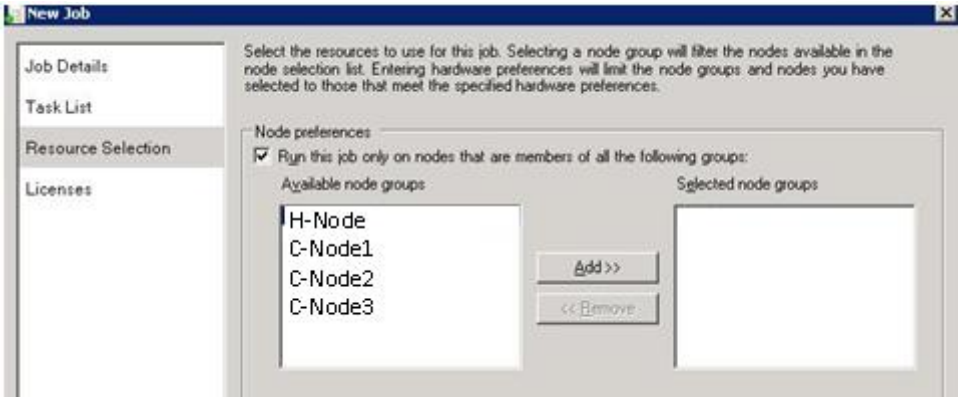

Figure 36 Select resources

6 .To submit the job, click **Submit**.

Now the job is submitted and the output will be in _out.txt in the shared folder and the error will be in _err.txt

Submit parametric job

Repeat steps from 1-3 like a single job then

1. To add a task, click **Task List**, and then click on Add button and choose parametric task see figure (37).
2. Type a name for your task.
3. The index for your parametric sweep applies to the instances of your application, and to your input, output, and error files. Define the index for your parametric sweep as follows:

    - In **Step 1** in the dialog box, set the start and end index values for your sweep.
    - In **Step 2**, choose the increment for the sweep index.

You can verify your index series under **Preview your sweep** at the bottom of the dialog box.

4. Type the task command, relative to the working directory, in the **Command line** entry box. For example, type **`myApp.exe *.`** This creates multiple indexed instances of you application.
5. Specify the **Working directory** for your task which is the shared directory. The working directory can include the wildcard character (*) if each step in the sweep needs a discrete working directory.
6. Specify the **Standard input**, **Standard output**, and **Standard error** file names relative to the working directory. Use the wildcard character (*) to represent a file number that will increment from one iteration to the next according to the index settings you chose in Step 5 of this procedure. For example: **myInput*.dat**.
7. Preview your parametric sweep job. If the job is not what you intended, make changes and preview again.
8. Click **Save,** see figure (38).
9. For **Resource manager** it is like single job.
10. Click **Submit**.

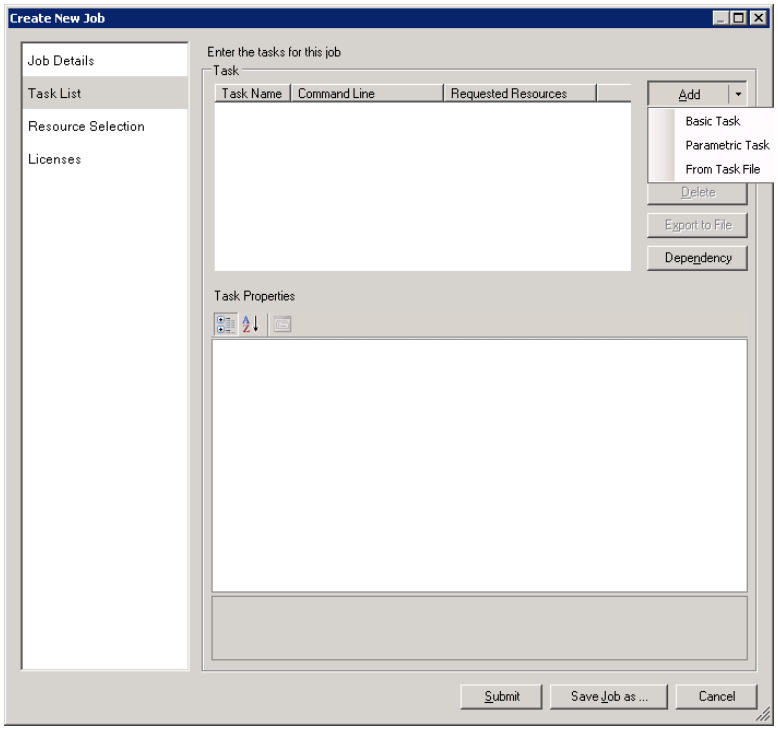

Figure 38 Parametric Task

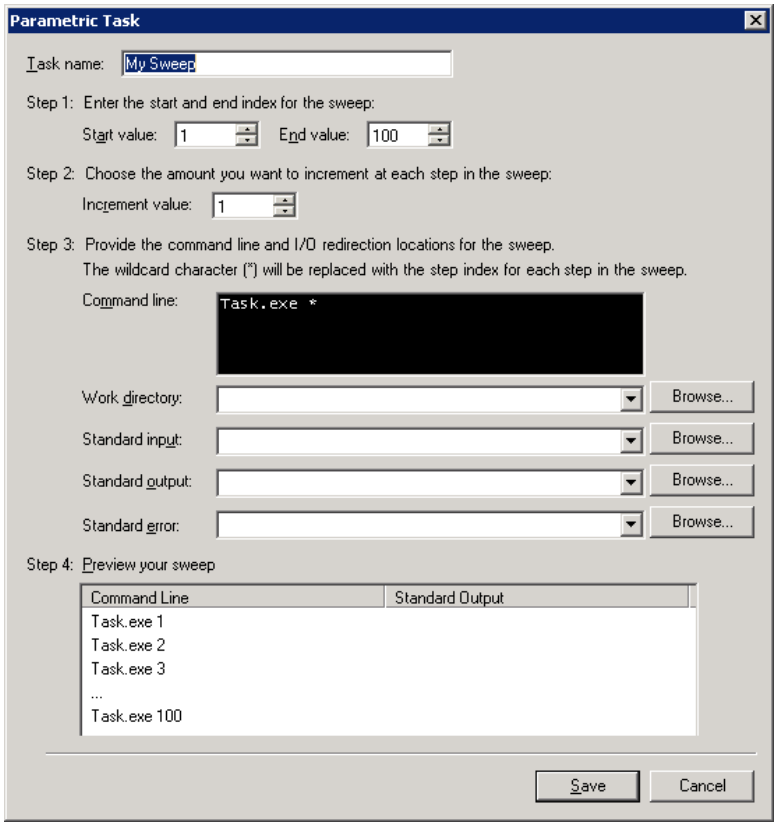

Figure 37 Parametric Task

Submit job to cluster from command line

Submit one job

To create and submit job from command line

```
 >job new /jobname:"job1" /priority:"Lowest" /RunTime:0:0:1
/requestednodes:”<ComputeNodeName>”

>job add <JobID> /numnodes:4 /workdir:\\H-
Node\users\Hishama\Myfolder\ /stdin:input.txt /stdout:out.txt
/stderr:err.txt  program.exe

>job submit /id:<JobID>
```

Or you can create the job and submit it in one step

```
>job submit /numnodes:4 /workdir:\\H-
Node\users\Hishama\Desktop\Run\ /stdin:input1.txt /stdout:out.txt
/stderr:err.txt  program.exe
```

In this case the job scheduler will set the other parameters.

To submit MPI job

```
> job submit /numnodes:4 /workdir:\\H-
Node\users\Hishama\Desktop\Run\ /stdin:input1.txt /stdout:out.txt
/stderr:err.txt mpiexec -machinefile hosts.txt -n 4 program.exe
```

Submit parametric job

```
> job submit /parametric:100 /numnodes:4 /workdir:\\H-
Node\users\Hishama\Desktop\Run\  /stdIn:"input*.txt"
/stdOut:"output*.txt" program.exe
```

To submit MPI job

```
> job submit /parametric:100 /numnodes:4 /workdir:\\H-
Node\users\Hishama\Desktop\Run\  /stdIn:"input*.txt"
/stdOut:"output*.txt" mpiexec –machinefile hosts.txt program.exe
```

Command Line Reference

Command	Operators	PowerShell Cmdlet	Description
job	job new [job_terms]	New-HpcJob	Create a job.
	job add jobID [task_terms]	Add-HpcTask	Add tasks to a job.
	job submit /id:jobid	Submit-HpcJob	Submit a job created using the **job new** command.
	job submit [job_terms][task_terms]	Submit-HpcJob	Submit a job.
	job cancel jobID	Stop-HpcJob	Cancel a job.
	job modify jobID [options]	Set-HpcJob	Modify a job.
	job requeue JobID	Submit-HpcJob	Requeue a job.
	job list	Get-HpcJob	List jobs in the cluster.
	job listtasks jobID	Get-HpcTask –JobID JobID	List the tasks of a job.
	job view JobID	Get-HpcJob JobID	View details of a job.
task	task view taskID	Get-HpcTask	View details of a task.
	task cancel taskID	.Cancel()	Cancel a task.
	task requeue taskID	.Requeue()	Requeue a task.
cluscfg	cluscfg view	Get-HpcClusterOverview	View details of a cluster.
	cluscfg listparams/setparams	Get-HpcClusterProperty Set-HpcClusterProperty	View or set configuration parameters.
	cluscfg listenvs/setenvs	Get-HpcClusterProperty Set-HpcClusterProperty	List or set a cluster-wide environment.
	cluscfg delcreds/setcreds	Remove-HpcJobCredential Set-HpcJobCredential	Set or delete user credentials.
node	node list	Get-HpcNode	List the nodes in the cluster.
	node listcores		List the cores in the cluster.

Command	Operators	PowerShell Cmdlet	Description
	node view nodename	Get-HpcNode nodename	View the properties of a node.
jobtemplate	jobtemplate add	New-HpcJobTemplate	Add a job template
	jobtemplate delete	Remove-HpcJobTemplate	Remove a job template
	jobtemplate grant/deny/remove	Set-HpcJobTemplateAcl	Set job template permissions.
	jobtemplate list	Get-HpcJobTemplate	List job templates.
	jobtemplate view	Get-HpcJobTemplate template	View template details.
clusrun	clusrun [options] command		Run a command across a set of compute nodes.

For more details about submitting jobs you can see this link

http://technet.microsoft.com/en-us/library/cc720168.aspx