

Myrinet Roll: User's Guide

Myrinet



Myrinet Roll: User's Guide :

Version 5.2 Edition

Published Aug 2009

Copyright © 2009 Myricom, Inc.

Table of Contents

Preface	v
1. Requirements.....	1
Rocks Version.....	1
Other Rolls.....	1
2. Installing the Myrinet Roll	3
Adding the Roll	3
Adding the Roll to a running Frontend	4
3. Using the Myrinet Roll	7
Myrinet interfaces and IP Addresses	7
How to assign and IP address to a Myrinet interface	7
How to remove an IP address from a Myrinet interface	7
Using Myrinet with MPI	7
How to run mpich over myrinet	8
How to run mpich2 over myrinet	8
How to run openmpi over myrinet.....	8

Preface

The function of the Myrinet Roll is to install and configure the Myrinet drivers and an MPICH environment that uses the underlying Myrinet hardware.

The Myrinet Roll contains the following pre-built MPICH environments:

Table 1. MPICH Environments

Name	Version
MPICH over Myrinet	1.2.7..7
MPICH2 over Myrinet	1.1.1p1..4
OpenMPI	1.3.3

Preface

Chapter 1. Requirements

Rocks Version

The Myrinet Roll is for use with Rocks version 5.2 (Chimichanga).

Other Rolls

The Myrinet Roll requires the Base, Kernel and HPC to be installed on the Frontend. Compatibility has been verified with the following Rolls:

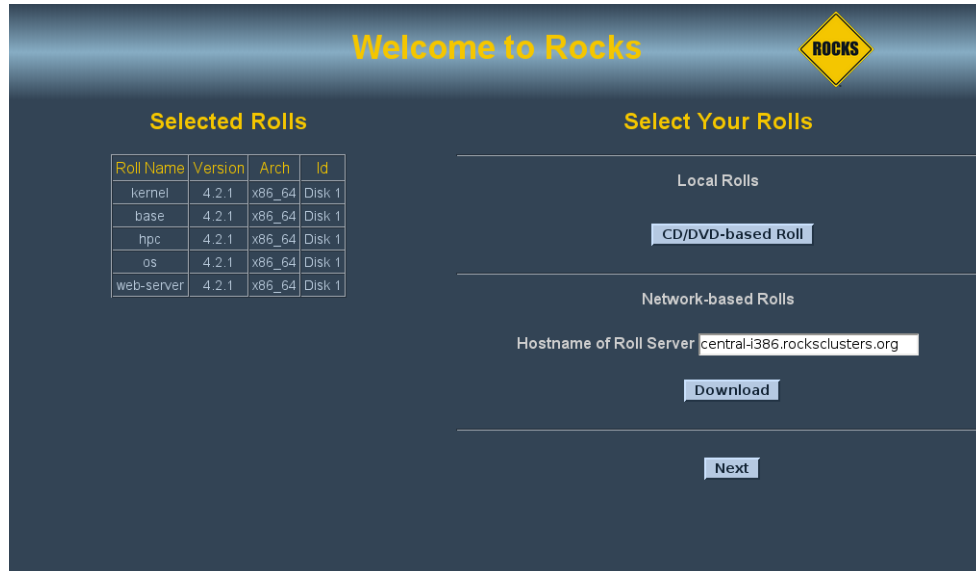
- HPC
- Kernel
- SGE
- Grid

Chapter 2. Installing the Myrinet Roll

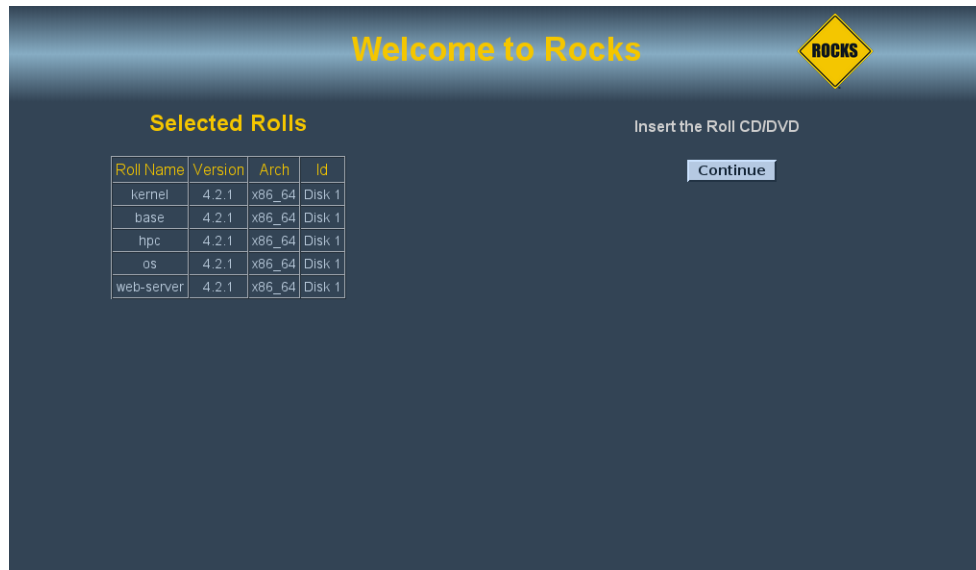
Adding the Roll

The Myrinet Roll should be installed during the Frontend installation step of your cluster (refer to section 1.2 of the Rocks usersguide).

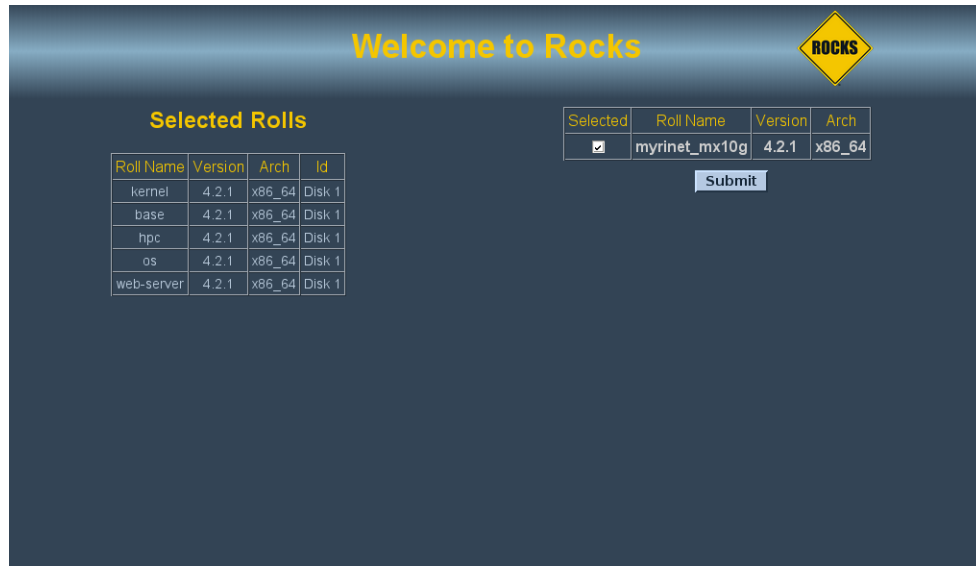
The Myrinet Roll is added to a Frontend installation in exactly the same manner as any other roll. For example, if you were only installing the minimum required rolls your screen should look something like this:



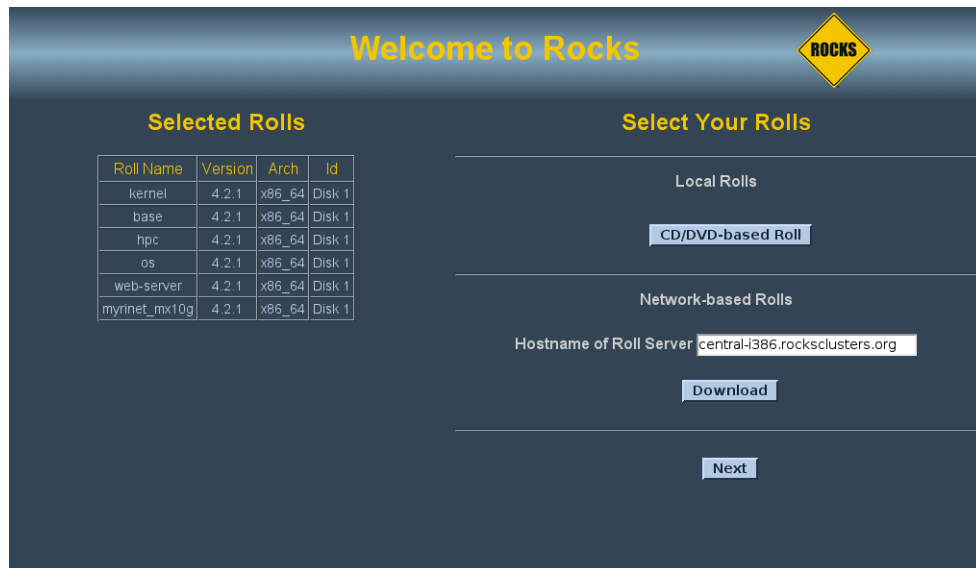
Just as with the other rolls, again click 'CD/DVD-based Roll' button.



Then insert your Myrinet Roll cdrom.



Select the myrinet roll and submit the Myrinet roll.



And continue on with the installation by hitting the 'Next' button.

Note: Note: the exact name of the Myrinet roll may vary depending on the version.

Adding the Roll to a running Frontend

If you have an already installed Frontend, you can use the following steps to install roll. These instructions are from the Rocks mailing list, and are a little more hands-on than installing the roll during frontend installation. These instructions may also change in the future. They are provided here for your convenience.

```
cd /export/rocks/install
rocks add roll {path_to_iso}
rocks enable roll {roll name}
rocks create distro
```

```
rocks run roll {roll name} | /bin/sh
```

Where {path_to_iso} is the location of the Myrinet iso file, and {roll name} corresponds to the name of the roll, i.e., myrinet_gm, myrinet_mx, myrinet_mx2g. You may now shoot all your nodes and the myrinet roll will be installed on them.

Chapter 3. Using the Myrinet Roll

Myrinet interfaces and IP Addresses

How to assign and IP address to a Myrinet interface

After you bring up a compute node, then you can assign an IP address to the Myrinet adapter. First you have to add a network definition. Here is an example creating a network:

```
# rocks add network myrieth subnet=192.168.0.0 netmask=255.255.0.0
```

Now you can add the myri0 interface by using 'rocks add host interface'. Below is an example of assigning an address to the myri0 interface on compute-0-0:

```
# rocks add host interface compute-0-0 myri0 ip=192.168.0.1 \  
  subnet=myrieth name=compute-0-0-myri0
```

After adding all the interfaces, it is a good idea to sync up the frontend's configuration with the changes you just made. Run the following:

```
# rocks sync config
```

You can use `rocks list host interface compute-0-0` to list your interfaces on compute-0-0. Please refer to your Rocks documentation for more information on how to manage ethernet interfaces in the Rocks environment.

After you run the above command, you must reinstall the compute node in order to correctly apply the IP address. To reinstall the compute node, execute:

```
# shoot-node compute-0-0
```

How to remove an IP address from a Myrinet interface

If you wish to remove a previously added IP address, execute:

```
# rocks remove host interface compute-0-0 myri0
```

The above removes the entry for the Myrinet interface `myri0` associated with `compute-0-0`.

Using Myrinet with MPI

Note: You must be a non-root user to run an mpi job.

First, create a `machines` file in your home directory. Below is a sample:

```
compute-0-0  
compute-0-1
```

We will be compiling and running the `mpi` program, a common example program. Copy the source code to your home directory:

```
$ cd ~/
```

```
$ cp /var/www/html/roll-documentation/myrinet_mx/5.2/examples/cpi.c ~/
```

How to run mpich over myrinet

First compile cpi with mpich:

```
$ /opt/mpich/myrinet_mx/gnu/bin/mpicc -o cpi.mpich cpi.c
```

To run cpi, execute:

```
$ /opt/mpich/myrinet_mx/gnu/bin/mpirun -np 2 -machinefile machines cpi.mpich
```

How to run mpich2 over myrinet

First compile cpi with mpich2:

```
$ /opt/mpich2/myrinet_mx/gnu/bin/mpicc -o cpi.mpich2 cpi.c
```

Next, you need to start the mpds on each node. Before you can do that, you have to create an `.mpd.conf` file with a secret word.

```
$ echo "secretword=SecretWord" > ~/.mpd.conf  
$ chmod 600 ~/.mpd.conf
```

Note: Mpich2 behaves a little differently than the other mpi implementations when running jobs from a frontend that does not participate in the job. Notice that in the line below 3 mpds are started, one on the frontend and one on each of the compute nodes.

```
$ /opt/mpich2/myrinet_mx/gnu/bin/mpdboot -n 3 -f machines
```

You can run `mpdtrace` to see check that everything was started correctly.

```
$ /opt/mpich2/myrinet_mx/gnu/bin/mpdtrace
```

To run cpi, execute:

Note: The `'-1'` argument tells `mpiexec` to skip the first node, the frontend, when spawning the job.

```
$ /opt/mpich2/myrinet_mx/gnu/bin/mpiexec -1 -n 2 ./cpi.mpich2
```

How to run openmpi over myrinet

First compile cpi with openmpi:

```
$ /opt/openmpi-myrinet_mx/bin/mpicc -o cpi.ompi cpi.c
```

To run cpi , execute:

```
$ /opt/openmpi-myrinet_mx/bin/mpirun -np 2 -machinefile machines -x MX_RCACHE=2 --mca p
```