

Scyld ClusterWare HPC

Evaluation Installation Instructions

Thank you for your interest in evaluating *Scyld ClusterWare*. This guide describes how to install an evaluation copy of Scyld ClusterWare using Penguin's installation repository. You should perform any necessary backups of the system before installing this software, and should pay particular attention to keeping a copy of any local configuration files.

To proceed with the evaluation, you will only need this document, together with two files which you should have received by email: **clusterware.repo** and **scyld.lic**. You can also download the files from Scyld MasterLink at <http://www.penguincomputing.com/support>.

Scyld ClusterWare System Overview

System Components and Layout

Scyld ClusterWare streamlines the processes of configuring, running, and maintaining a Linux cluster using a group of commodity off-the-shelf (COTS) computers connected through a private network.

The front-end "master node" in the cluster is configured with a full Linux installation, distributing computing tasks to the other "compute nodes" in the cluster. Nodes communicate across a private network and share a common process execution space with common, cluster-wide process ID values.

A compute node is commonly diskless, as its kernel image is downloaded from the master node at node startup time using the Preboot eXecution Environment (*PXE*), and libraries and executable binaries are transparently transferred from the master node as needed. A compute node may access data files on locally attached storage or across NFS from an NFS server managed by the master node or some other accessible server.

In order for the master node to communicate with an outside network, it needs two network interface controllers (*NICs*): one for the private internal cluster network, and the other for the outside network. It is suggested that the master node be connected to an outside network so multiple users can access the cluster from remote locations.

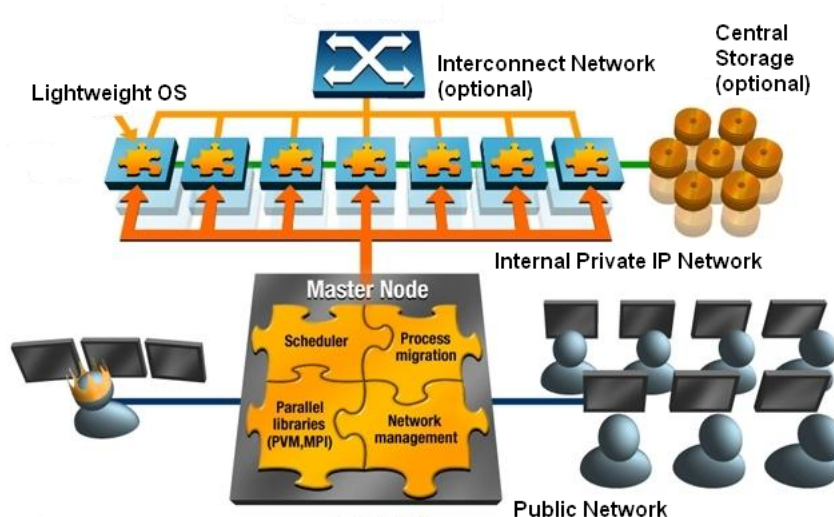


Figure 1. Cluster Configuration

Hardware Requirements

64-bit Intel® or AMD™ x86_64 processor architecture.

1024 MBytes (1 GByte) main memory, with 2048 MBytes (2 GBytes) or more preferred.

At least one Gigabit Ethernet Network Interface Controller on each compute node.

Preferably two Gigabit Ethernet Network Interface Controllers on the master node.

A Gigabit network switch for the private cluster network.

Optional: An Infiniband network infrastructure for compute nodes. Infiniband controllers must use the Mellanox™ chipset.

BIOS Requirements

Compute nodes must support the PXE network boot protocol. If an operating system has already been installed on a compute node's local disk, the node's BIOS must be configured to prioritize PXE network booting ahead of booting from the local disk.

Software Requirements

Scyld ClusterWare should be installed on a system running a base distribution of Red Hat Enterprise Linux 4 (RHEL4) Update 8 or CentOS 4.8. If you are running an earlier RHEL4 or CentOS 4 distribution, then the appropriate update procedure is described below.

Because Scyld ClusterWare depends on the presence of certain RHEL4 (or CentOS 4) packages, the ClusterWare installation will fail if any of these required packages have not been installed. Therefore, for simplicity, Penguin Computing recommends that you install all packages from the base distribution. This is accomplished by checking the *Everything* box in the Package Group Selection screen of the Red Hat installer GUI. Alternatively, a more labor-intensive approach is to attempt to install Scyld ClusterWare (as described below), note what dependencies are unmet and cause the install to fail, then manually install those specific required RHEL4 (or CentOS 4) packages and retry the Scyld ClusterWare installation, repeating as necessary.

The Infiniband interconnects (if present) must be supported by the RHEL4 *mtbca* driver. When in doubt, you should contact your Infiniband hardware vendor to determine if your hardware is supported by this driver.

Scyld ClusterWare includes a customized version of the RHEL4 kernel that can co-exist with the kernel(s) currently installed on your master node.

NOTE: The Scyld ClusterWare kernel packages contain only those kernel modules that are included in the base distribution. This means that if your non-ClusterWare kernel is using a 3rd-party kernel module, e.g., for Panasas storage, or an Infiniband controller not supported by the *mtbca* driver found in the base distribution, then that 3rd-party module (and whatever hardware it controls) is unavailable in the Scyld ClusterWare kernel environment.

Install Yum Command

Scyld ClusterWare uses the **yum** command to install and update software. CentOS 4 includes **yum**, but RHEL4 does not. If **yum** is not installed on your master node, then log into MasterLink at <http://www.penguincomputing.com/support/masterlink> using the login name and password that was previously

communicated to you from Penguin Computing. Click on *Yum For RedHat*, and download the following RPMs to your system:

- yum-2.4.3-1.noarch.rpm
- python-elementtree-1.2.6-4.2.1.x86_64.rpm
- python-sqlite-1.1.7-1.2.x86_64.rpm
- python-urlgrabber-2.9.8-2.noarch.rpm
- sqlite-3.3.3-1.2.x86_64.rpm

Install the rpms with the following command:

```
[root@scyld ~]# rpm -ivh yum-2.4.3-1.noarch.rpm \
> python-elementtree-1.2.6-4.2.1.x86_64.rpm \
> python-sqlite-1.1.7-1.2.x86_64.rpm \
> python-urlgrabber-2.9.8-2.x86_64.rpm \
> sqlite-3.3.3-1.2.x86_64.rpm
```

Install Yum Configuration File and License File

You will have received two files by email that need to be installed in the appropriate places on your master node:

1. Install the **clusterware.repo** yum configuration file as `/etc/yum.repos.d/clusterware.repo`. This contains the credentials to identify your master node to Penguin Computing and to access Scyld ClusterWare software for download.
2. Install the **scyld.lic** evaluation license file as `/etc/scyld.lic`. This grants free use of ClusterWare for a month beginning from the date the license file was generated.

Update Base Distribution

1. Update the base distribution to the latest RHEL4 using:

```
[root@scyld ~]# up2date -u
```

or to CentOS 4 using:

```
[root@scyld ~]# yum update
```

as appropriate for your cluster.

2. If Infiniband hardware is present, ensure you have installed the entire Infiniband software stack from the base distribution, from RHEL4:

```
[root@scyld ~]# up2date openib* libib* *dap1* libmthca -i
```

or from CentOS 4:

```
[root@scyld ~]# yum install openib* libib* *dap1* libmthca
```

3. Finally, remove base distribution packages that conflict with ClusterWare 4:

```
[root@scyld ~]# yum remove openmpi* lam*
```

Install Scyld ClusterWare

- Clean the **yum** cache to a known state, and install Scyld ClusterWare:

```
[root@scyld ~]# yum clean all  
[root@scyld ~]# yum groupinstall Scyld-ClusterWare
```

yum displays the list of rpms to be installed and prompts you to accept the list before actually updating the system. If the command completes successfully, a *complete!* message is displayed to the terminal.

- However, if **yum** fails with a *Transaction Check Error* that complains that a base distribution rpm is newer than the Scyld ClusterWare rpm that is attempting to replace it, then you must manually install the downlevel Scyld ClusterWare rpm(s).

For example, if the offending rpm is a **kernel** rpm, then redo the *groupinstall*:

```
[root@scyld ~]# rpm -i --oldpackage /var/cache/yum/cw-eval/packages/kernel-  
[root@scyld ~]# yum groupinstall Scyld-ClusterWare
```

This should now *complete!* successfully. Examine `/etc/grub.conf` to verify that the `default=` entry selects a Scyld ClusterWare kernel.

Configure the Private and Public Networks

1. Configure the master node's IP address on the private cluster network:

```
[root@scyld ~]# system-config-network
```

presents you with a graphical configuration window that lists the available Ethernet devices. Typically, the public interface is `eth0` and the private cluster network is `eth1`. Select `eth1` and click on the *Edit* icon.

2. Now you are presented with a device-specific configuration window. Here you should select *Activate device when computer starts* and select *Statically set IP addresses:*. Choose an IP address and Subnet mask. We recommend using a non-routable address, such as 192.168.0.1 or 10.20.30.1. If the public interface also uses a static IP address, then choose a different subnet address. Then choose a subnet mask to accommodate all your compute nodes.

3. Configure the master node's network security level:

```
[root@scyld ~]# system-config-securitylevel
```

presents you with a graphical configuration window. If you select a *Security level:* of *Enable firewall*, then you should enable *Trusted services:* of at least *SSH*, and you must enable *Trusted devices:* of the private cluster interface device, e.g., `eth1`.

4. Configure the compute node IP addresses:

```
[root@scyld ~]# beonetconf
```

presents you with a graphical configuration window. There you must:

- Choose a *Cluster Interface:* that is the private cluster interface device, e.g., `eth1`.
- Enter a *Number of nodes:* value that is the maximum number of compute nodes in your cluster.

- Enter a *Beginning IP address for compute nodes*: that is consistent with the IP address that you assigned the master node on that private cluster network. For example, if you assigned 10.20.30.1 to the master node, then you may assign 10.20.30.100 to the first compute node.
- Click *OK* to save the entries and close the window. This updates the `/etc/beowulf/config` configuration file, which can also be manually edited (with care).

Configure Ganglia (optional)

Edit `/etc/xinetd.d/beostat` to change the 2nd-to-last line from `disable = yes` to `disable = no`.

Enable the **httpd** and **gmetad** services to start at master node boot time:

```
[root@scyld ~]# /sbin/chkconfig httpd on
[root@scyld ~]# /sbin/chkconfig gmetad on
```

After the installation has completed and the master node has been rebooted, visit `http://localhost/ganglia` with your browser. If accessing from a remote node, use the real hostname of the master node instead of `localhost`.

Configure Scyld IMF (optional webinterface)

Enable the `httpd` service, if it is not already enabled:

```
[root@scyld ~]# /sbin/chkconfig httpd on
```

Initialize the `admin` account by assigning it a unique password:

```
[root@scyld ~]# /usr/bin/htpasswd /etc/httpd/scyld-imf/htpasswd-users admin
```

After the installation has completed and the master node has been rebooted, point your Web browser to your master node with a URL of the form: `http://MasterNodeName/admin` and log in using the username `admin` and the password that was just assigned.

Start Cluster Operations

1. Reboot the master node:

```
[root@scyld ~]# reboot
```

After rebooting, run:

```
[root@scyld ~]# uname -r
```

and confirm that the master node is running a Scyld ClusterWare kernel.

2. Normally, Scyld ClusterWare services automatically start whenever the master node reboots. However, when starting Scyld ClusterWare the first time, you are asked to read and accept an End User License Agreement (EULA). Start Scyld ClusterWare with:

```
[root@scyld ~]# service beowulf start
```

which prompts you to press *Enter* if you want to review the EULA. If you proceed with an *Enter*, and after you accept the EULA with an affirmative *yes*, then Scyld ClusterWare **beowulf** services continue to initialize, and the master node is ready to respond to compute node DHCP and PXE requests. After acceptance, you will not be asked again about the EULA.

If, however, after 20 seconds you have not responded with an *Enter*, then the cluster services will not start and your master node will function simply as a non-ClusterWare Linux system. You can revisit the EULA review process at any time by again attempting to restart the **beowulf** service.

3. Once the **beowulf** service is up and running, the master node can PXE boot as many compute nodes into the cluster as were defined when you executed the **beonetconf** command. You can monitor the cluster status with the graphical **beostatus** or **ganglia** tools, or with the text-based **beostatus -c** or **bpstat -U**.

Note: Depending upon BIOS settings, the compute nodes' DHCP requests may timeout because the master node hadn't been ready to respond, and compute nodes would then revert to a BIOS prompt waiting for human input. If **bpstat** continues to show that compute nodes are *down*, then physically powercycle each compute node, either manually or using an already configured **ipmitool**.

Documentation and Support

- For a complete reference, the Scyld ClusterWare documentation set consists of:
 - The full *Installation Guide* containing broader, more detailed information for installing and configuring the cluster.
 - The *Administrator's Guide* describing how to configure, maintain, and update the cluster.
 - The *User's Guide*, *Programmer's Guide* and *Reference Guide* describing the commands, architecture, and programming interface for the cluster, including sample programs.
 - The *Release Notes* containing release-specific details, including information about known issues and workarounds.

These product guides are available in two formats, HTML and PDF.

- Visit MasterLink at <http://www.penguincomputing.com/support> using the username and password you received for this evaluation, and click on *Application Notes* for information about running specific applications.
- Note: If you **ssh -X** from a remote system that executes a more recent version of X11 to the master node, some graphical programs may fail with an error of the form:

```
Gdk-ERROR **: BadMatch (invalid parameter attributes)
  serial 798 error_code 8 request_code 72 minor_code 0
```

For a workaround, try setting the ssh client host X11 depth to 8:

```
[root@scyld ~]# export XLIB_SKIP_ARGB_VISUALS=1
```

on the master node before running the failing program. If that is successful, then consider adding that export to `/etc/bashrc` or to an individual's `~/.bashrc`.

- For additional support, contact Customer Support at scyldEval@penguincomputing.com.

Purchasing Scyld ClusterWare

To license Scyld ClusterWare, please email scyld@penguincomputing.com

Uninstalling Scyld ClusterWare

1. Edit `/etc/grub.conf` and change the `default=` entry to select a non-ClusterWare kernel. Then reboot the master node:

```
[root@scyld ~]# reboot
```

and confirm the system is running the desired kernel.

2. Use `yum groupermove` to uninstall ClusterWare:

```
[root@scyld ~]# yum groupermove Scyld-ClusterWare
```

3. Restore any packages that were explicitly removed when you installed Scyld ClusterWare, e.g., `openmpi*` and/or `lam*`.

