Sun™ HPC Software, Linux Edition 1.0

Installation Guide

Sun Microsystems, Inc. www.sun.com

Part No. 820-5451-10

July 2008

Table of Contents

Overview

Sun™ HPC Software, Linux Edition, is an integrated open-source software solution for Linux-based HPC clusters running on Sun hardware. It provides a framework of software components to simplify the process of deploying and managing large-scale Linux HPC clusters.

Software Components

Sun HPC Software, Linux Edition provides software to provision, manage, monitor and operate large scale Linux HPC clusters. It serves as a foundation for optional add-ons such as schedulers and other components not included with the solution. A verification suite provides tools to verify interoperability of software and hardware and help ensure a stable environment.

System Requirements

The table below shows the Sun HPC Software, Linux Edition, system requirements.

Platform	Sun x64
Operating System	CentOS 5.1 x86_64 Red Hat Enterprise Linux 5
Architecture	Sun x64
Networking	Ethernet, Infiniband

Installing the Sun HPC Software, Linux Edition

The steps below provide an overview for how to install and configure the Sun HPC Software, Linux Edition. Click on a step to view a more detailed procedure for that step.

1. Obtain Sun HPC Software, Linux Edition
2. Install Sun HPC Software, Linux Edition
3. Create an inventory of devices
4. Set up management nodes
5. Create configurations for all nodes in the cluster
6. Boot up the client nodes
7. Build SSH keys
8. Add rootfs images for additional node types
9. Test the operational commands
10. Perform additional site customization
11. Complete cluster-wide verification

Step 1: Obtain Sun HPC Software, Linux Edition

1. For instructions for building a ISO DVD image of the Sun HPC Software, Linux Edition:
   a. Go to www.sun.com/software/products/hpcsoftware.
   b. Click on the Get It tab and then click on the Download button.

A file with instructions is displayed.

Step 2: Install Sun HPC Software, Linux Edition

Use the procedure below to install the Sun HPC Software, Linux Edition:

1. Use one of these options:
   - Insert the DVD into the head node drive and reset the system.
   - Use virtual DVD redirection to copy the ISO image to the head node.

Warning: The current contents of the head node hard drives will be overwritten when the Sun HPC Software is installed.

A screen is displayed with the Giraffe Logo (Figure 1).

Figure 1. Sun HPC Software, Linux Edition, installation screen

2. Press <Enter> to install the stack.

   A series of messages are displayed indicating that drivers and other software are being loaded.

3. Select the “Skip” option to bypass testing the media, unless you want the media to be tested.
4. When the message “Welcome to SunHPC” appears, click “OK”.
5. When the Disclaimer screen appears, click “OK”.
6. Select language and keyboard options.

   The default language is English. The default keyboard option is US.

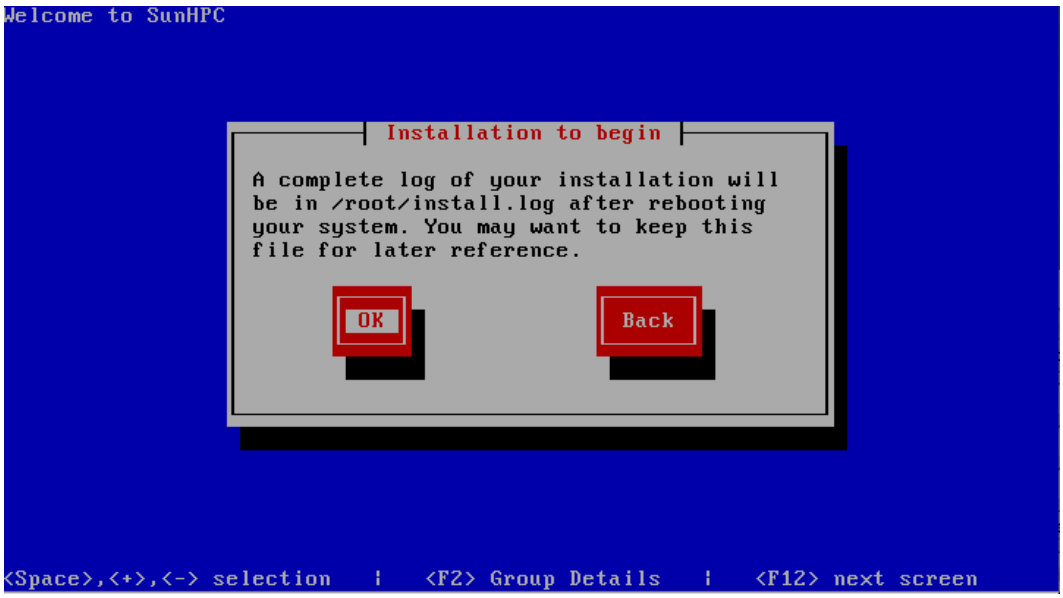
7. When the “Partitioning Type” screen appears, select either the default layout or create your own layout.

   A series of screens are displayed with further instructions depending on the option you selected.

8. When the series of "Network Definition" screens appear, define the IP address for the head node, including the network address of the head node, gateway, and Domain Name Services (DNS) server, and the hostname.
9. When the "Time Zone Selection" screen appears, select the time zone in which you are located.
10. When the "Root Password" screen appears, select a root password for the head node and enter it twice.
11. When the "Package Selection" screen appears, select options to be added and click "OK".

    Installation of the SUN HPC Software on the head node begins.

12. When the "Installation to begin" screen shown below appears, click "OK".



    A series of screens are displayed as the installation progresses.

13. When the "Complete" screen appears to indicate the end of the installation process, remove the DVD or discontinue virtual DVD redirection of the ISO image.
14. Press "REBOOT" to reboot the system.

Step 3: Create an inventory of devices

1. Create an inventory of systems in your cluster including hostnames and MAC and IP addresses. This information will be used to configure the systems in a later step. Sample inventories are shown in Table 1 and Table 2.

Table 1. Management System Network

Testbed		MGMT			
platform		MAC	mgmt ip		mgmt name
c10 blade 7		00:14:4f:e6:41:9f	10.1.80	177	comp3-7sp
c10 blade 8		00:14:4f:e6:41:06	10.1.80	178	comp3-8sp
c10 blade 9		00:14:4f:d6:b3:71	10.1.80	179	comp3-9sp
G4		00:14:4f:a6:6e:89	10.1.80	243	mdsn#4-sp
G4					
thumper		00:14:4f:d3:5e:d8	10.1.80	147	ossn#18sp

Table 2. Inventory of Data-Facing Servers

Private					ibo		
MAC	eth0		hostname	GUID	ib0 ip		ib0 name
00:14:4f:80:14:a0	10.1.80	57	comp3-7		10.13.80	57	comp3-7ib
00:14:4f:82:31:5e	10.1.80	58	comp3-8		10.13.80	58	comp3-8ib
00:14:4f:9e:a0:ce	10.1.80	59	comp3-9		10.13.80	59	comp3-9ib
00:14:4f:45:26:e2	10.1.80	129	mdsn#4		10.13.80	129	mdsn#4ib

Step 4: Set up management nodes

1. Install Ethernet cabling for the management system if not already in place.
2. Set up the management interfaces, such as the Sun Integrated Lights Out Manager (ILOM) service processors (SPs), to prepare to provision the systems.

Step 5: Create configurations for all nodes in the cluster

In this step, oneSis and Dynamic Host Configuration Protocol (DHCP) will be used to configure the nodes in the cluster for a Preboot eXecution Environment (PXE) boot.

oneSIS is an open-source software tool for administering systems in a large-scale, Linux-based cluster environment. The default oneSIS configuration that results from building the head node can be used to begin provisioning the other nodes in the cluster as diskless clients.

Dynamic Host Configuration Protocol (DHCP) is used by clients to obtain the parameters needed to operate in an Internet Protcol (IP) network, allowing devices to be added to the network with minimal or no manual configuration.

The Preboot eXecution Environment (PXE) is a client/server interface that allows networked computers without an operating system installed to be configured and booted remotely.

The contents of the default oneSis configuration file are shown below, where:

`DISTRO` is the Linux Centos version.
`ONESIS_IMAGE` is the rootfs image's PATH for diskless clients.
`DHCPD_IF` is the network interface on which dhcpd is listening.
`KERNEL_VERSION` is the kernel release (the value returned by `uname -r`).
`UNUSED_PKGS` is a list of packages that are not to be installed on the diskless clients.

```
# cat /etc/sunhpc.conf
DISTRO=centos-5.1
ONESIS_IMAGE=/var/lib/oneSIS/image/centos-5.1
DHCPD_IF=eth0
KERNEL_VERSION=2.6.18-53.1.21.el5.sunhpc1
UNUSED_PKGS="conman powerman ganglia-gmetad sunhpc-configuration"
```

1. To begin provisioning the client nodes, update the file /etc/hosts on the head node with the client hostnames and TCP/IP addresses.

Note: This step is not required if hostnames can be resolved using DNS services on the head node.

---

2. Modify the default contents of /etc/dhcp_hostlist to incorporate client names, IP addresses, and MAC addresses to be used by oneSIS to provision these nodes as shown below:

```
# cat /etc/dhcp_hostlist
host ezo02 {hardware ethernet 00:14:4F:28:45:4E; fixed-address
192.168.10.12;}
host ezo03 {hardware ethernet 00:14:4F:28:4A:00; fixed-address
192.168.10.13;}
```

3. Run the "sunhpc_setup" command to apply the oneSIS configuration to the client.

```
# /usr/sbin/sunhpc_setup -i
```

The message “Copying / to /var/lib/oneSIS/image/centos-5.1/... “ is displayed for about 10 minutes followed by several other messages.

4. Check the status of the iptables on the head node. If any are on, turn them off as described below.

   a. Check the iptables as shown below.

```
[root@giraffe ~]# chkconfig --list | grep tables
ip6tables       0:off   1:off   2:on    3:on    4:on    5:on    6:off
iptables        0:off   1:off   2:on    3:on    4:on    5:on    6:off
```

   b. Disable the iptables as shown below.

```
[root@giraffe ~]# /etc/init.d/iptables stop
[root@giraffe ~]# /etc/init.d/ip6tables stop
[root@giraffe ~]# chkconfig --level 2345 iptables off
[root@giraffe ~]# chkconfig --level 2345 ip6tables off
[root@giraffe ~]# chkconfig --list | grep tables
ip6tables       0:off   1:off   2:off   3:off   4:off   5:off   6:off
iptables        0:off   1:off   2:off   3:off   4:off   5:off   6:off
```

5. Check the status of the iptables on the oneSIS clients. If any are on, turn them off as shown below.

```
[root@giraffe ~]# chroot /var/lib/oneSIS/image/centos-5.1 /sbin/chkconfig
--level 2345 iptables off
[root@giraffe ~]# chroot /var/lib/oneSIS/image/centos-5.1 /sbin/chkconfig
--level 2345 ip6tables off
[root@giraffe ~]# chroot /var/lib/oneSIS/image/centos-5.1 /sbin/chkconfig
--list | grep table
ip6tables       0:off   1:off   2:off   3:off   4:off   5:off   6:off
iptables        0:off   1:off   2:off   3:off   4:off   5:off   6:off
```

6. Determine if any of your clients have specific hardware or driver requirements. If so, you may need to build a new `initramfs`.

   The example below shows how to incorporate the driver module needed to support the ck804 nVidia ethernet controller used on Sun Blade X6220 Server Modules.

   a. On the head node, enter the following, where `forcedeth` is the driver module for the ck804 nVidia Ethernet controller and `initrd-2.6.18-53.1.14.el5.lem.img` is the image name to be used to initially boot the head node:

```
[root@giraffe ~]# mk-initramfs-oneSIS --with=forcedeth -s 8192
initrd-2.6.18-53.1.14.el5.lem.img
```

   b. To enable support for different devices in the cluster, modify the `pxelinux.cfg` file as shown below. Use the mk-initramfs-oneSIS command with the “image name” variable.

Note: You do not need to modify the `pxelinux.cfg` file if all the devices in the cluster are the same type as the head node.

```
[root@giraffe ~]# cat /tftpboot/pxelinux.cfg
      PROMPT 1
      TIMEOUT 20
      IPAPPEND 2
      DEFAULT linux
      label linux
      kernel vmlinuz-2.6.18-53.1.14.el5 append root=/dev/nfs
      initrd=initrd-2.6.18-53.1.14.el5.lem.img selinux=0
      label rescue
      kernel vmlinuz-2.6.18-53.1.14.el5 append root=/dev/nfs
      initrd=initrd-2.6.18-53.1.14.el5.lem.img selinux=0
```

Step 6: Boot up the client nodes

1. Boot up the client nodes in the cluster for PXE provisioning.

   The following example shows how to boot up a client for PXE provisioning. In this example, three data server nodes are provisioned as diskless clients. FreeIPMI commands are used to power the nodes off and on. For more information, see the `ipmipower` man page.

```
[root@giraffe ~]# ipmipower -h 10.1.80.[177-179] -u root -p changeme -f -W
"endianseq"
10.1.80.178: ok
10.1.80.179: ok
10.1.80.177: ok
[root@giraffe ~]# ipmipower -h 10.1.80.[177-179] -u root -p changeme -n -W
"endianseq"
10.1.80.179: ok
10.1.80.177: ok
10.1.80.178: ok
```

Note: The `-W "endianseq"` option must be used for the current version of ILOM SP-based Sun Systems. It is not needed for older Sun Systems such as v20z/v40z.

Step 7: Build SSH keys

Some Sun HPC Software components, such as pdsh, require Secure Shell (SSH) public key authentication to access clients in the cluster. Follow the procedure below to set up ssh public keys.

1. On the head node, create the SSH public key.

```
# ssh-keygen -t rsa -N ''
```

2. Copy authorized keys to the to the client oneSys root filesystem.

```
# mkdir -p /var/lib/oneSIS/image/centos-5.1/root/.ssh
# chmod 700 /var/lib/oneSIS/image/centos-5.1/root/.ssh
# cat /root/.ssh/id_rsa.pub  >
/var/lib/oneSIS/image/centos-5.1/root/.ssh/authorized_keys
# chmod 600 /var/lib/oneSIS/image/centos-5.1/root/.ssh/authorized_keys
```

3. Define the nodes in the cluster in /etc/genders.

```
# cat /etc/genders
node[09-12] compute
```

Note: These nodes must either be in /etc/hosts or must be able to be resolved through DNS.

4. Generate host entries defined in /etc/genders

```
# ssh-keyscan -t rsa `nodeattr -s compute` > /root/.ssh/known_hosts
```

Step 8: Add rootfs images for additional node types

When the Lustre file system option is installed, additional rootfs images must be created for the metadata servers (MDSs) and some types of object storage servers (OSSs) used in the Lustre file system. These procedures are described below.

Object Storage Server (OSS) for Sun Fire X4500

A Sun Fire X4500 OSS, or any OSS requiring software RAID, must use cento-4.6. In the procedure below, a rootfs is created for the OSS in which the centos-4.6 packages are installed.

1. Use one of these options:
   - Mount the DVD created in Step 1.
   - Mount the ISO image created in Step 1 using the loop back option.
2. Use the centos4_chroot script to install the centos-4.6 packages to a chroot environment on the headnode on which centos-5.1 is running.
3. After installing the packages, run the "sunhpc_setup" command to apply the oneSIS configuration to the OSS.

```
# mount -o loop sunhpc-1.0.iso /mnt
# mkdir /var/lib/oneSIS/image/centos-4.6
# centos4_chroot /mnt/CentOS4 /var/lib/oneSIS/image/centos-4.6
Cannot open logfile /var/lib/oneSIS/image/centos-4.6/var/log/yum.log
Setting up Local Package Process
Examining /mnt/CentOS4/acl-2.2.23-5.3.el4.x86_64.rpm: acl -
2.2.23-5.3.el4.x86_64
Examining /mnt/CentOS4/acpid-1.0.3-2.x86_64.rpm: acpid - 1.0.3-2.x86_64
Examining /mnt/CentOS4/anacron-2.3-32.x86_64.rpm: anacron - 2.3-32.x86_64
```

```
- snip -

Transaction Summary
================================================================================
Install    407 Package(s)
Update       0 Package(s)
Remove       0 Package(s)

Total download size: 1.2 G
Downloading Packages:
Running Transaction Test
warning: gnupg-1.2.6-9: Header V3 DSA signature: NOKEY, key ID 443e1821
Finished Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing: libgcc                      ###################### [  1/407]
  Installing: libgcc                      ###################### [  2/407]
- snip -

# sunhpc_setup -u -f /etc/sunhpc_oss.conf
Using template: /usr/share/oneSIS/initrd-templates/initrd-dhclient-x86.gz
Added /lib/modules/2.6.9-67.0.7.EL_lustre.1.6.5smp/kernel/drivers/net/e1000/
e1000.ko
Added /lib/modules/2.6.9-67.0.7.EL_lustre.
1.6.5smp/kernel/drivers/net/forcedeth.ko
Added /lib/modules/2.6.9-67.0.7.EL_lustre.
1.6.5smp/kernel/drivers/scsi/scsi_mod.ko
Added /lib/modules/2.6.9-67.0.7.EL_lustre.
1.6.5smp/kernel/drivers/scsi/mv_sata.ko
- snip -

Starting NFS mountd:                                       [  OK  ]
Stopping xinetd:                                           [  OK  ]
Starting xinetd:                                           [  OK  ]
```

4. Add the MAC address and IP address of the OSS to /etc/dhcp_hostlist_oss.

```
# vi /etc/dhcp_hostlist_oss
# /etc/init.d/dhcpd restart
```

Metadata Storage Server (MDS)

Each MDS uses the same centos-5.1distribution as the client with a Lustre patch applied. In the procedure below, the Lustre patch is applied to the centos-5.1 distribution and a rootfs is created for the MDS.

1. Use one of these options:
   - Mount the DVD created in Step 1.
   - Mount the ISO image created in Step 1 using the loop back option.
2. Execute the rsync or mirror-image command to create a clone of the client rootfs to use as the MDS rootfs.

```
# rsync -av /var/lib/oneSIS/image/centos-5.1/*
/var/lib/oneSIS/image/centos-5.1-mds/
building file list ... done
created directory /var/lib/oneSIS/image/centos-5.1-mds
tmp -> /ram/tmp
bin/
bin/arch
bin/awk -> gawk
bin/basename
bin/bash
bin/cat
bin/chgrp
- snip -

sent 2712352797 bytes  received 1931396 bytes  11192924.51 bytes/sec
total size is 2705983227  speedup is 1.00
```

3. Issue the mk-sysimage command with the revert option to prepare the rootfs image for the MDS.

```
# mk-sysimage -r /var/lib/oneSIS/image/centos-5.1-mds
oneSIS: Restoring LINKDIR: /var/lib/oneSIS/image/centos-5.1-
mds/var/run.default
oneSIS: Restoring LINKDIR: /var/lib/oneSIS/image/centos-5.1-
mds/var/log.default
oneSIS: Restoring LINKDIR: /var/lib/oneSIS/image/centos-5.1-
mds/var/lock/subsys.default
- snip -
oneSIS: Restoring file: /var/lib/oneSIS/image/centos-5.1-
mds/etc/sysconfig/network
oneSIS: Reversing patch: /usr/share/oneSIS/distro-patches/centos-5.1.patch
```

4. Remove the Lustre packages for the patchless client.

```
# rpm -e --root /var/lib/oneSIS/image/centos-5.1-mds lustre-
modules-1.6.5-2.6.18_53.1.21.el5.sunhpc1_lustre.1.6.5
# rpm -e --root /var/lib/oneSIS/image/centos-5.1-mds
lustre-1.6.5-2.6.18_53.1.21.el5.sunhpc1_lustre.1.6.5
# rpm -e --root /var/lib/oneSIS/image/centos-5.1-mds kernel-
ib-1.2.5.5-2.6.18_53.1.21.el5.sunhpc1
# rpm -e --root /var/lib/oneSIS/image/centos-5.1-mds kernel-ib-
devel-1.2.5.5-2.6.18_53.1.21.el5.sunhpc1
```

5. Install the Lustre-related packages for the MDS.

```
# rpm -ivh --root /var/lib/oneSIS/image/centos-5.1-mds /mnt/SunHPC/kernel-
lustre-smp-2.6.18-53.1.14.el5_lustre.1.6.5.x86_64.rpm
# rpm -ivh --root /var/lib/oneSIS/image/centos-5.1-mds /mnt/SunHPC/lustre-
ldiskfs-3.0.4-2.6.18_53.1.14.el5_lustre.1.6.5smp.x86_64.rpm
# rpm -ivh --root /var/lib/oneSIS/image/centos-5.1-mds /mnt/SunHPC/lustre-
modules-1.6.5-2.6.18_53.1.14.el5_lustre.1.6.5smp.x86_64.rpm
# rpm -ivh --root /var/lib/oneSIS/image/centos-5.1-mds
/mnt/SunHPC/lustre-1.6.5-2.6.18_53.1.14.el5_lustre.1.6.5smp.x86_64.rpm
# rpm -ivh --root /var/lib/oneSIS/image/centos-5.1-mds /mnt/SunHPC/kernel-
ib-1.2.5.5-2.6.18_53.1.14.el5_lustre.1.6.5smp.x86_64.rpm
# rpm -ivh --root /var/lib/oneSIS/image/centos-5.1-mds /mnt/SunHPC/kernel-
ib-devel-1.2.5.5-2.6.18_53.1.14.el5_lustre.1.6.5smp.x86_64.rpm
```

6. Run the "sunhpc_setup" command to apply the oneSIS configuration to the MDS.

```
# sunhpc_setup -u -f /etc/sunhpc_mds.conf
# rm /var/lib/oneSIS/image/centos-5.1-mds/etc/sysconfig/network-
scripts/ifcfg-eth*
```

7. Add the MAC address and IP address of the MDS to /etc/dhcp_hostlist_mds.

```
# vi /etc/dhcp_hostlist_mds
# /etc/init.d/dhcpd restart
```

Step 9: Test the operational commands

Several monitoring, provisioning, and management tools – pdsh, FreeIPMI, Powerman, ConMan, and Ganglia – are provided with Sun HPC Software, Linux Edition. Follow the instructions below to configure these tools as needed and test that operational commands are functional on your system.

Using pdsh

The pdsh tool allows commands to be executed at multiple nodes at the same time. pdsh provides a module interface and supports rsh and ssh for communication. A command line example is shown below:

```
# pdsh -w node[09-12] uptime
node10:  21:10:46 up  5:28,  0 users,  load average: 0.00, 0.00, 0.00
node09:  21:11:49 up  5:29,  0 users,  load average: 0.00, 0.00, 0.00
node11:  01:08:52 up  5:28,  0 users,  load average: 0.00, 0.00, 0.00
node12:  00:10:49 up  5:28,  0 users,  load average: 0.00, 0.00, 0.00
```

To use a dsh (Dancer's shell) group file with pdsh, create an /etc/dsh/group directory. For each group, place a file in this directory that contains the hostname of each node in the group. In the example below, three groups are created named “client”, “oss”, and “mds”:

```
# mkdir -p /etc/dsh/group
# vi client
node09
node10
node11
node12

# vi oss
node10
node12

# vi mds
node09
node11
```

In the example below, pdsh is used with the `-g groupname` option:

```
# pdsh -g all uptime
node09:  21:17:25 up  5:34,  0 users,  load average: 0.03, 0.01, 0.00
node11:  01:14:29 up  5:34,  0 users,  load average: 0.08, 0.02, 0.01
node12:  00:16:25 up  5:34,  0 users,  load average: 0.04, 0.01, 0.00
node10:  21:16:23 up  5:34,  0 users,  load average: 0.00, 0.00, 0.00

# pdsh -g oss uptime
node10:  21:16:58 up  5:35,  0 users,  load average: 0.00, 0.00, 0.00
node12:  00:17:01 up  5:35,  0 users,  load average: 0.02, 0.01, 0.00
```

Using FreeIPMI

FreeIPMI provides support for Sun's ILOM with the `-u root` option and the `-W "endianseq"` option. You can enter these as command line options or incorporate them into a configuration file. A command line example is shown below.

```
# ipmipower -h 192.168.10.[2-3]  -p changeme -u root -s -W "endianseq"
192.168.10.2: off
192.168.10.3: on
# ipmi-chassis -h 192.168.10.[2-3]  -p changeme -u root -s -W "endianseq"
192.168.10.3: System Power               : on
192.168.10.3: System Power Overload      : false
192.168.10.3: Interlock switch           : Inactive
192.168.10.3: Power fault detected       : false
192.168.10.3: Power control fault        : false
- snip -
```

The example belows shows how ILOM options are defined in the configuration file `/etc/ipmipower.conf`.

```
hostname 192.168.10.[2-3]       # client's hostname
username root                   # ILOM's username
password changeme               # ILOM's password
workaround-flags "endianseq"    # work-around flags for ILOM
on-if-off enable
wait-until-on enable
wait-until-off enable
```

Once you've set up the configuration file as described above, you can use commands similar to the example below.

```
# ipmipower -s
192.168.10.2: off
192.168.10.3: on
```

Using PowerMan

PowerMan is used to operate remote power control (RPC) devices from a central location. You can use the list of the devices with MAC and IP addresses in the spreadsheet created in Step 3 to configure PowerMan.

1. Configure the PowerMan power management tool.

   a. Edit the PowerMan configuration file.

```
# vi /etc/powerman/powerman.conf
include "/etc/powerman/ipmipower.dev"
alias "all" "node[09-12]-sp"
device "node07" "ipmipower"  "/usr/sbin/ipmipower -h node[09-12]-sp |&"
node "node[09-12]-sp" "node07"  "node[09-12]-sp"
```

   b. Start Powerman.

```
# /etc/init.d/powerman start
```

2. Display the current power status of all clients:

```
# powerman -q all
on:      node09-sp,node10-sp,node11-sp,node12-sp
off:
unknown:
```

3. Shut down all nodes and display the power status again:

```
# pdsh -w node[09-12] shutdown -h now
# powerman -q all
on:
off:     node09-sp,node10-sp,node11-sp,node12-sp
unknown:
```

4. Turn on power for all nodes :

```
# pm -1 all
Command completed successfully
```

Note: The "pm" command is an alias for the “powerman” command.

Using ConMan

ConMan is a centralized console management tool capable of handling a large number of clients. It supports Sun ILOM. You can use the list of the devices with MAC and IP addresses in the spreadsheet created in Step 3 to configure ConMan.

1. Configure ILOM in the ConMan serial console management tool.

   a. Add the nodes to the file `/etc/conman.conf` that you want to manage using ConMan.

```
# vi /etc/conman.conf
- snip -

console name="ezo02" dev="/usr/lib/conman/exec/sun-ilom.exp ezo02i root"
console name="ezo03" dev="/usr/lib/conman/exec/sun-ilom.exp ezo03i root"
console name="node09" dev="/usr/lib/conman/exec/sun-ilom.exp node09i root"
```

   b. Create a password file called `/etc/conman.pswd` with entries in the format
   `<host-regex> : <user> : <pswd>`

```
# vi /etc/conman.pswd
ezo[0-9]*i : root : changeme
node[0-9]*i : root : changeme
# chmod 400 /etc/conman.pswd
```

   c. Start the `conman` service.

```
  # /etc/init.d/conman start
```

   d. Use the command `conman<hostname>` to access remote consoles being managed by the "conmand" daemon.

```
# conman node09
<ConMan> Connection to console [node09] opened.
CentOS release 5 (Final)
Kernel 2.6.18-53.1.14.el5 on an x86_64
node09 login:&.
```

Using Ganglia

Ganglia is a monitoring tool for clusters. In the example below, the head node is set up to use the Ethernet 1 interface for ganglia monitoring.

```
# pdsh -w node07,node[09-12] route add -net 224.0.0.0 netmask 240.0.0.0 dev
eth1
# echo "any net 244.0.0.0 netmask 240.0.0.0 dev eth1" >
/var/lib/oneSIS/image/centos-5.1/etc/sysconfig/static-routes
```

You can use an advanced ganglia configuration for your cluster environment, but the simplest configuration assumes a single cluster name. This cluster name is defined by modifying `/etc/gmond.conf` on each client.

```
# vi /var/lib/oneSIS/image/centos-5.1/etc/gmond.conf
- snip -

 * NOT be wrapped inside of a <CLUSTER> tag. */
cluster {
  name = "giraffe_cluster"
  owner = "unspecified"
  latlong = "unspecified"
  url = "unspecified"
}
- snip -
```

Use pdsh to start the Ganglia daemon `gmond` on all nodes in the cluster.

```
# pdsh -w node[09-12] /etc/init.d/gmond start
```

On the head node, in `/etc/gmetad.conf`, add the client name to `data_source` and change the `gridname` to the name of the cluster.

```
# vi /etc/gmetad.conf
- snip -

# data_source "my cluster" 10 localhost  my.machine.edu:8649  1.2.3.5:8655
# data_source "my grid" 50 1.3.4.7:8655 grid.org:8651 grid-backup.org:8651
# data_source "another source" 1.3.4.7:8655  1.3.4.8

data_source "node09" node09
data_source "node10" node10
data_source "node11" node11
data_source "node12" node12
- snip -

# The name of this Grid. All the data sources above will be wrapped in a
GRID
# tag with this name.
# default: Unspecified
gridname "giraffe_cluster"
#
- snip -
```

Start the Ganglia daemon `gmetad` on the head node.

```
# /etc/init.d/gmetad start
```

Step 10: Perform additional site customization

Complete additional site customization such as configuring an Infiniband network or implementing and configuring a scheduler.

Note: The Sun Grid Engine scheduler is provided with the SUN HPC Software, Linux Edition 1.0. For more information about Sun Grid Engine, visit http://www.sun.com/software/gridware/. User documentation can be found at http://docs.sun.com/app/docs/coll/1017.4?l=en.

Step 11: Complete cluster-wide verification

Use tools in the verification suite to determine if the Sun HPC Software and third-party components have been installed and are inter-operating correctly.

GNU GENERAL PUBLIC LICENSE

Version 2, June 1991

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software--to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Lesser General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or

any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

c) If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or

distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and "any later version", you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO

MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

END OF TERMS AND CONDITIONS

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA *
Phone: 1-650-960-1300 or 1-800-555-9SUN (9786) Web: sun.com