
# Implementation Guide:
# Sun™ Cluster 3.0 Series: Guide to Installation—Part I

*Chris Dotson, Enterprise Engineering*

*Sun BluePrints™ OnLine—April 2003*


**http://www.sun.com/blueprints**

**Sun Microsystems, Inc.**
4150 Network Circle
Santa Clara, CA 95054 USA
650 960-1300

Part No.: 817-2015-10
Revision 2.0 3/25/03
Edition: April 2003

Please
Recycle

# Sun™ Cluster 3.0 Series: Guide to Installation—Part I

The purpose of this module is to guide you through the tasks you must perform before you install the Sun™ Cluster (SC) 3.0 software. These tasks include setting up the administrative workstation and configuring the Sun Cluster 3.0 hardware components.

The exercises in this module explain how to install and configure a workstation to perform Sun Cluster 3.0 software administrative functions in a cluster environment. Additionally, we provide instructions for configuring the cluster, implementing best practices, and performing design verifications, as well as administering a two-node cluster.

For information about managing a cluster, refer to the documents listed in the References section, specifically:

- *System Administration Guide* (“Preparing to Administer the Cluster,” “Beginning to Administer the Cluster,” “Administering the Cluster,” and “Administering Sun Cluster with the Graphical User Interface”)
- *SC3.0 Cluster Concepts* (“Cluster Administration and Application Development”)

Many of the steps in this guide refer to “manual (local) procedures,” which you should perform only if you have local (physical) access to the SunPlex™ platform. An example is, resetting the terminal concentrator (TC) in order to activate specific settings.

# Objectives

After completing this module, you will have successfully verified the installation and software configuration for each cluster component. These tasks must be performed before you can install the Sun Cluster 3.0 software on each cluster node. Additionally, you will have implemented the associated key practices during each task, including:

- Configuring the hardware and cabling.
- Configuring the Solaris™ Operating Environment (Solaris OE) and installing patches on each cluster node. This task is site specific.
- Verifying the management server setup on the administrative workstation.
- Verifying the terminal concentrator configuration.
- Installing and configuring the Cluster Console utility.
- Post-Installation and preparation for Sun Cluster software installation.

Many of the steps for configuring the cluster require you to have physical access to the equipment. We have made special note of specific configuration procedures that require local physical access to the SunPlex platform. For example, placing the TC into Monitor mode requires manual procedures. For this reason, these hands-on exercises specify steps that are to be performed as part of local (or manual) installation procedures.

These manual steps are not performed if you are accessing the Sun Cluster 3.0 hardware (SunPlex platform) remotely. Instead, these manual steps are included as documentation of the procedures required to complete the configuration of each cluster component referenced in Figure 1. Additionally, it is assumed that it is not necessary to provide the detailed procedures for installing the Solaris OE and Solaris OE patches on each SunPlex node. Instead, we define the parameters required for installation.

**Note –** In this module, you must perform all of the associated instructions to verify that the correct configuration has been achieved as specified. It is very important that you do this prior to configuring the additional cluster components to ensure consistent, reliable cluster services and SunPlex platform operations.

# Prerequisites

This article assumes you have experience and qualification as a Solaris network administrator. For installation queries, shell usage questions, patches, and packages, refer to the Sun Educational Services manuals for the *Solaris System Administration 1* and *Solaris System Administration 2* courses. It is also recommended to complete the web-based training for Enterprise Installation Standards (EIS). Additional prerequisites include ES-333 or equivalent.

The primary intended audience for this BPLAB guide is system support engineers, professional service consultants, and system and network administrators.

# Introduction

The administrative workstation (`clustadm`) is required for setting up and configuring the cluster, and can be used for ongoing cluster management and operations.

The administrative workstation is a remote workstation, which can be used to administer each cluster node by connecting through the terminal concentrator. The administrative workstation can also be used to administer and monitor the cluster through the cluster control panel, the command-line interface, or the Sun™ Management Center console (monitoring only).

**Note –** When configuring the administrative workstation, consider combining systems management functions, such as JumpStart server software, shared applications software, and patches. These can be combined, along with Sun Cluster 3.0 software administrative functions appropriate to your implementation.

In this module, we describe the hardware configuration and the procedures used to install the administrative workstation `clustadm`. We explain how to confirm that all requirements are met, verify that all cluster components (including patches) are installed correctly, and ensure that the shell environment is configured correctly.

# Enterprise Installation Standards: Standard Installation Practices

Sun Cluster installations must conform to the EIS installation standards that are current at the time of the installation. For this module, we have previously gathered all configuration information and ‘data-layout’ criteria, as required to successfully complete the installation.

For local (manual) installations, you must obtain and implement all current EIS standards when installing each node in the SunPlex platform.

This guide, including all steps and procedures, represents a specific Sun Cluster implementation, which conforms to EIS standards at the time of this writing. The EIS process is an efficient method for ensuring standard installations. Ensuring standard installation practices is important, and enables configuration auditing (as when performing root-cause analysis).

The EIS process includes a setup script that prepares a standard root shell environment. This script installs additional utilities which are used to capture important configuration data for each cluster node and the Administration Workstation (`clustadm`). For example, this script sets `PATH` and `MANPATH` variables, according to software installed on the system. Later, when software (such as Veritas) is installed, the root shell environment will dynamically contain the updated `PATH` information and so on.

Following the EIS methodology, the EIS setup script is run in the beginning of each installation, and will install the EIS components of which ACT and Explorer software are included. Then, the installation proceeds, using the EIS installation environment.

The EIS process is defined by checklists and instructions, which are appropriate to the node and applications being installed. For each checklist, line items on the checklist ensure certain features and configurations are configured on the SunPlex platform according to the best practices they represent. Furthermore, EIS instructions are followed, and the general installation sequence is documented for building the cluster. For example, EIS instructions begin by installing Solaris software (and patches), and ends with configuring the NFS data service.

Upon completion of the installation, the EIS process indicates to run the Explorer software, which gathers configuration data about the nodes that were installed. The configuration information is saved, and the installer is able to audit the site-specific documents to verify the accuracy of the installed configuration.

## Installation and Planning Considerations

New installations that are well planned and well executed are critical to ensuring reliability and, ultimately, availability. Reducing system outages involves using proven methods (that is, well-documented techniques, applications, components, configurations, operating policies and procedures) when configuring highly available platforms. This includes minimizing all single points of failure (SPOFs), and documenting any SPOFs that could occur, along with documenting any associated best practices.

The following points can contribute to successful configurations, assist in sustaining daily operations, and help maximize platform availability and performance:

- Ensure SC 3.0 administrators are highly trained and able to successfully test and conduct cluster failover operations for each highly available (HA) application and associated systems and subsystems, including fault isolation / troubleshooting, and disaster recovery procedures using all available utilities and interfaces.
- Carefully document site-specific and application-specific configurations and procedures, as part of implementing best practices and simplifying data center operations and maintenance.
- Plan for Change Management: Register and record all standard and non-standard configurations, implementing change management procedures for all systems and subsystems for auditing and tracking key systems and components throughout the life cycle of the data center).
- When designing SunPlex solutions, implement well-known and established configurations and techniques that minimize platform complexity and the number of active components, to simplify operations and maintenance.

- Prior to installing the SunPlex platform, ensure all hardware components are fully operational and set to known state (that is, all settings to factory defaults and all disk media ‘scrubbed’). This includes all system units and subsystems (disk arrays, controllers, firmware, and all shared disk media).
- Provide diagnostics, utilities, and interfaces that are easy to use and interpret, with clearly documented error messages and procedures for resolving potential problems.
- Refer to the *Sun Cluster 3.0 Configuration Guide* for current restrictions on hardware, software, and applications running in SC 3.0 environment.

## Customer Requirements

**Caution –** SC 3.0 applications are implemented across many layers, each layer has several aspects that must be considered when designing a solution. Only a comprehensive analysis of your applications requirements combined with thorough knowledge of Sun’s products and services can ensure that the cluster solution will meet the service level agreement (SLA).

To obtain SC 3.0 licenses for bundled HA agents, contact your local Sun Service provider. For HA agents developed by Sun Microsystems or third party vendors requiring licenses, contact your local Sun Microsystems representative for professional services. Additionally, SC 3.0 does not include a VERITAS Volume Manager (VxVM) or CVM license, which must be purchased. In some cases, Sun StorEdge™ arrays may include a VxVM license. Ensure all licenses, information and keys are available during installation, and stored safely (that is, available for use, if needed, during disaster / recovery operations).

# Hardware Configuration

Figure 1 and Table 1 through Table 5 represent the Sun Cluster hardware configuration used for this module, which specifies two or more Sun servers that are connected by means of a private network. Each server can access the same application data using multi-ported (shared) disk storage and shared network resources, thereby enabling either cluster node to inherit an application after its primary server becomes unable to provide those services.

Refer to Figure 1, which describes the SC 3.0 lab hardware implementation and Table 1 through Table 5, which define each connection.

Final verification of the cluster hardware configuration will be confirmed only after the required software has been installed and configured, and failover operations have been tested successfully.

# Cable Configuration

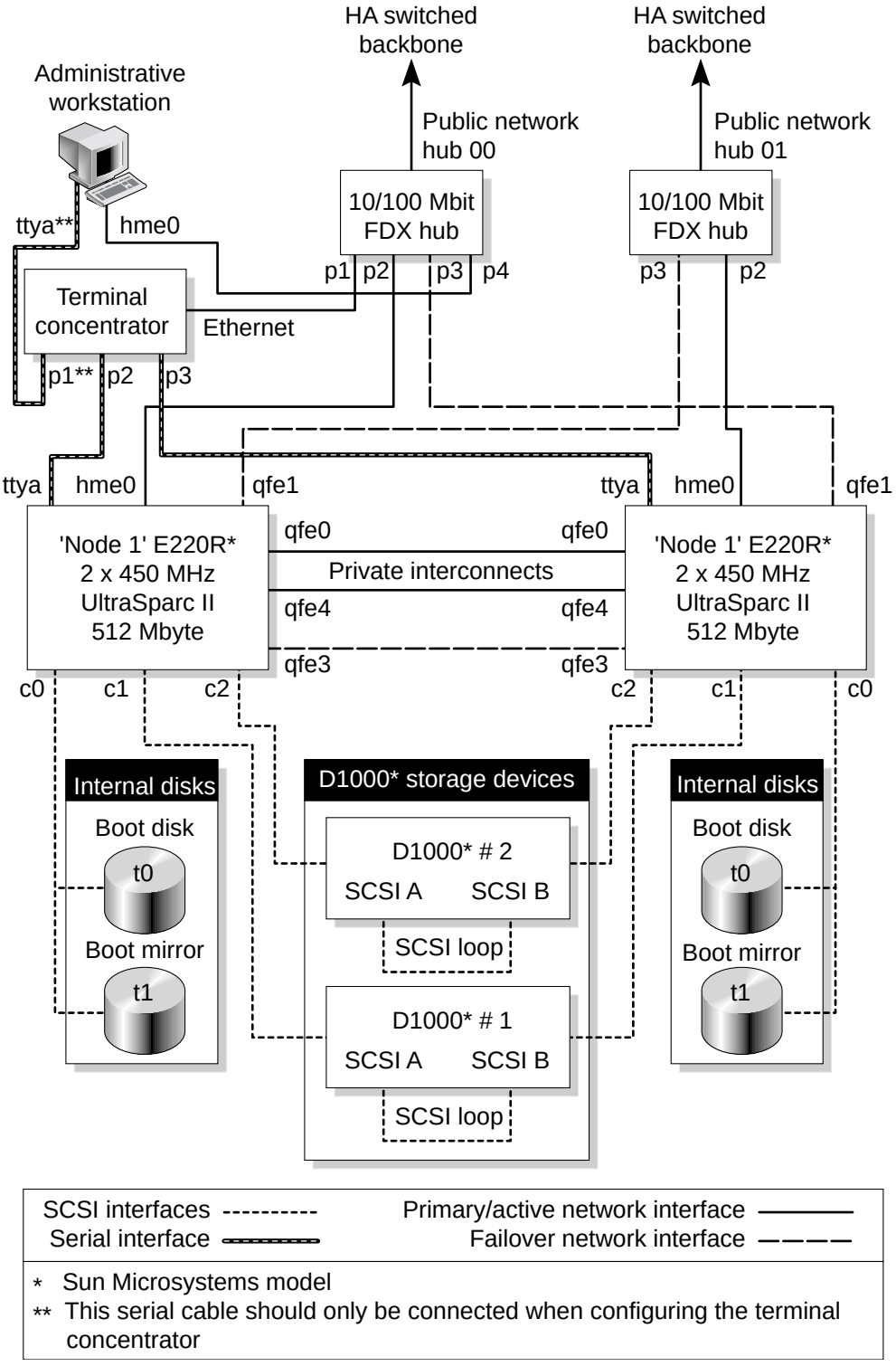


**Figure 1** Cable Configuration

**Note –** In the previous illustration: c1 = PCI3, c2 = PCI4; D1000s include t0, t1, t2, t8, t9, t10; Utilize spare Ethernet ports by configuring additional private interconnects (that is, use crossover cables between spare Ethernet port `qfe3` and `qfe7`, as indicated in Figure 1).

## Cable Connections

Table 1 through Table 5 list the required cable connections.

**Table 1** Server-to-Storage Connections

| From Device | From Location | To Device | To Location | Cable Label |
|---|---|---|---|---|
| E220R # 1 | SCSI A (PCI3) | D1000 #1 | SCSI A | C3/1 - C3/3A |
| E220R # 2 | SCSI A (PCI3) | D1000 #1 | SCSI B | C3/1 - C3/3B |
| E220R # 1 | SCSI A (PCI4) | D1000 #2 | SCSI A | C3/2 - C3/3A |
| E220R # 2 | SCSI A (PCI4) | D1000 #2 | SCSI B | C3/2 - C3/3B |

**Table 2** Private Network Connections

| From Device | From Location | To Device | To Location | Cable Label |
|---|---|---|---|---|
| E220R # 1 | `qfe0` | E220R # 2 | `qfe0` | C3/1 - C3/2A |
| E220R # 2 | `qfe4` | E220R # 2 | `qfe4` | C3/1 - C3/2B |

**Table 3** Public Network Connections

| From Device | From Location | To Device | To Location | Cable Label |
|---|---|---|---|---|
| E220R # 1 | `hme0` | Hub # 00 | Port #2 | C3/1 - C3/5A |
| E220R # 2 | `qfe1` | Hub # 01 | Port #3 | C3/1 - C3/6A |
| E220R # 1 | `hme0` | Hub # 01 | Port #2 | C3/1 - C3/6A |
| E220R # 2 | `qfe1` | Hub # 00 | Port #3 | C3/2 - C3/6A |

**Table 4** Terminal Concentrator Connections

| From Device | From Location | To Device | To Location | Cable Label |
|---|---|---|---|---|
| E220R # 1 | Serial Port A | Terminal Concentrator | Port #2 | C3/1 - C3/4A |
| E220R # 2 | Serial Port A | Terminal Concentrator | Port #3 | C3/2 - C3/4A |
| Terminal Concentrator | Ethernet Port | Hub # 00 | Port #1 | C3/4 - C3/5A |

**Table 5** Administrative Workstation Connections

| From Device | From Location | To Device | To Location | Cable Label |
|---|---|---|---|---|
| Administration Workstation | `hme0` | Hub # 00 | Port #4 | F2/1 - C3/5A |
| Administration Workstation | Serial Port A | Terminal Concentrator | Port #1 ** | F2/1 - C3/5B |

**Note –** The Cable Label column in Table 1 through Table 5 assumes the equipment is located in a specific grid location, for example C3. The number following the grid location identifies the stacking level for that piece of equipment with 1 being the lowest level. The letter at the end of the label tag indicates how many cables terminate at that level. For example, the letter A indicates one cable, B indicates two cables, and so on. Also, the label tag F2 is the grid location of the administrative workstation. The cable with "**" in the To Location column is only connected when configuring the terminal concentrator.

# Architectural Limitations

The Sun Cluster 3.0 architecture is able to provide highest levels of availability for hardware, the operating system, and applications without compromising data integrity. The Sun Cluster environment (that is, hardware, operating environment, Sun Cluster framework, and API applications) can be customized to create highly available applications.

## No Single Points of Failure

Multiple faults occurring within the same cluster platform (environment) can result in unplanned downtime. A SPOF can exist within, say, the software applications architecture. For the E220R, a SPOF for the single cluster node might be the embedded boot controller, or even a memory module.

- The basic Sun Cluster configuration based on Sun Enterprise Server Model 220R can be configured as an entry-level platform, providing no SPOFs for the cluster pair.

## Configuring Clusters for HA: Planning Considerations

The primary configuration / planning considerations for highly available applications and databases include identifying requirements for: software versions and features, boot environment, shared storage, and data services (and their agents).

Designing a production cluster environment for mission critical applications is a complex task, involving the careful selection of optimum components amid numerous options. We recommend that you work closely with a qualified consulting practice, such as Sun Professional Services, in making these selections.

An example of these choices are determining the optimum number and mix of database instances (services) per node, or ensuring no potential agent conflicts exists and that any service level conflicts are resolved.

Different cluster topologies require carefully prescribed setup procedures in relation to the following cluster components:

- Data center requirements (hardened to environmental and power-loss conditions)
- Number of logical hosts per node (including their agents, agent interoperability, and service level requirements)
- Type of volume manager
- Disk striping and layout
- File systems versus raw device database storage
- Performance (local storage versus GFS considerations)
- Network infrastructure requirements and redundancy
- Client failover strategy
- Logical host failover method (manual vs. automatic)
- Naming conventions such as host ID, disk label, disk groups, meta sets, and mount points.
- Normal (sustaining) operations policies and procedures
- Backup and recovery procedures for the SunPlex platform

# Section 1.1: Solaris Configuration (clustadm)

This section describes the steps necessary to install Solaris 8 (plus patches) on the Sun Cluster administrative workstation (clustadm). The same version of Solaris OE must run on both the clustadm workstation and on each of the cluster nodes. The workstation is used for the installation and for basic cluster operations.

---

**Note –** At this time, it is assumed that all systems and subsystems have been powered on, and the SunPlex platform is fully configured (as per Figure 1, and Table 1 through Table 5). All hardware components are operational.

---

**Key Practice:** Ensure all firmware is installed with the most recent (supported) versions for all systems and subsystems, including all servers, disk arrays, controllers, and terminal concentrators. For example, on each node in the SunPlex platform, ensure the system EEPROM contains the most recent (supported) version of OpenBoot PROM (OBP) such as, OpenBoot 3.15.

## Step 1.1.1

For local (manual) installations, it is a good idea to ensure that the clustadm workstation is configured with the most recent version of the OBP. Information about downloading can be obtained from SunSolve℠, at http://sunsolve.sun.com. After upgrading the OBP, manually turn system power off, and then on, before continuing with the next step.

## Step 1.1.2

For local (manual) installations, install Solaris OE on the clustadm workstation.

- Recommendations include installing the Entire Distribution, choosing the proper locale (English), and configuring the root disk layout. Choose Entire Distribution+OEM, only if required.
- Don’t forget to install the Solaris man pages.

**Key Practice:** Implement Solaris JumpStart to maintain consistency and fully automate the installation of the Solaris Operating Environment and additional software packages for each node in the SunPlex platform. Solaris JumpStart can improve installations and minimize operator errors that occur during a manual process. Combining Solaris JumpStart and Flash archives enables quick and consistent disaster / recovery operations.

## Step 1.1.3

For local (manual) installations, reboot the `clustadm` workstation after installing Solaris.

## Step 1.1.4

Verify the correct Solaris OE configuration. At this time, log in as the `root` user and ensure Solaris has been configured correctly, on the `clustadm` workstation.

For these examples, the root password is **`abc`**

Example Solaris OE configuration – `clustadm` workstation

| | |
|---|---|
| **Host name:** | `clustadm` |
| **IP Address:** | `129.153.xx.xxx` |
| **Name Service:** | None (local files) |
| **Set Subnet:** | Yes |
| **Subnet Mask:** | `255.255.255.0` |
| **Default Gateway:** | None |

**Note –** The values quoted in the previous table are sample values. For local (manual) installations, substitute the appropriate site-specific values.

## Step 1.1.5

Verify that the following disk partitioning guidelines have been implemented for the `clustadm` workstation, which reserves space for use by a volume manager, and allocates 2 Gbytes (slice 1) for swap space, then assigns unallocated space to `/` (`root`, on slice 0).

**Note –** It is often easier to use the Solaris `format` command to calculate the exact number of even cylinders to be configured, as when determining the size of the `root` file system.

Configure boot disk slices using the following guidelines:

Slice 0 = assigned to "/" (all unallocated space)

Slice 1 = assigned to `swap` (2 GB)

Slice 2 = assigned as `backup` (full extent)

**Note –** These procedures, DO NOT include the steps necessary to configure a volume manager on the `clustadm` workstation; however, it is a good idea to configure one. For example, implement Solaris™ Volume Manager (SVM) and configure a highly available boot environment (that is, alternate boot disk and controller). At a minimum, data should be protected against failure of any single diskspindle.

**Key Practice:** Configure a standard, flexible disk partitioning scheme. Partitioning each disk spindle identically can save time, provide flexibility, and maintain consistency across nodes. In this case, use a consistent partitioning scheme that is flexible, allows for using either SVM or VxVM. Implement the following standard partitioning for boot disks:

Reserve the first few cylinders for use by a volume manager. For SVM, the reserved space (20 Mbytes) should be assigned to slice 7. For VxVM, a minimum of two cylinders are required, and slices 3 and 4 must be unassigned. Some customer applications require use of a slice. The alternate boot environment, Live Upgrade requires one slice.

**Note –** Installing the entire Solaris media kit can require nearly 13 Gbytes. Assuming 36 Gbyte disk drives, this should be more than sufficient for the boot disk layout being proposed here.

**Key Practice:** Verify that the Solaris OE installation was successful and that any errors reported are resolved before proceeding. Review the `/var/sadm/README` file to determine the location of the most recent installation logs (for example, `/var/sadm/system/logs`). Examine the current Solaris OE installation log files `begin.log`, `sysidtool.log`, or `install_log` for potential errors. Confirm the cause of any installation error messages which may occur, resolving failures before proceeding further.

## Step 1.1.6

Examine the installation logs, ensuring that any Solaris OE installation errors do not go undetected, or unresolved. On the `clustadm` workstation, type the following commands:

```
clustadm# cd /var/sadm/system/logs
clustadm# pwd
/var/sadm/system/logs

clustadm# ls
begin.log              finish.log              install.log
begin.log_2000_04_13   finish.log_2000_04_13   sysidtool.log
{{sample dates only}}
clustadm# more *
{{It is important to resolve any installation error messages noted in the logs.

Example:
pkgadd: ERROR: postinstall script did not complete successfully}}
```

## Step 1.1.7

Activate the root shell environment, by invoking a new shell, or simply executing '**`su -`** ', at this time. If you are connected to the `/dev/console` port, you should log off and then on again.

Verify that all settings and environment variables are configured, as required.

Example: Environment Variables (settings) - `clustadm`

| Variable | Label |
|---|---|
| TERM | `ansi or vt100` |
| stty | `istrip` |
| set prompt | `root@<hostname>#` |
| Ensure the following PATH variable settings | `/usr/bin:/usr/ucb:/etc:/sbin:/usr/sbin:/usr/cluster/bin:/usr/cluster/lib/sc:/opt/VRTSvmsa/bin:/opt/VRTSvxvm/bin:/etc/vx/bin:/usr/cluster/dtk/bin:/opt/CTEact/bin:/opt/SUNWexplo/bin:/opt/sun/bin:/usr/ccs/bin:/usr/openwin/bin:/usr/dt/bin` |
| Ensure the following MANPATH variable settings | `/usr/dt/man:/usr/openwin/share/man:/opt/VRTS/man:/usr/cluster/dtk/man:/opt/CTEact/man:/opt/SUNWexplo/man:/usr/cluster/man:/usr/man:/usr/share/man` |

### Step 1.1.8

Prior to using a text editor (such as vi) to view or modify files, ensure your terminal environment variable is correct, as appropriate. We recommend **TERM=ansi; export TERM** for proper video display.

## Section 1.2: Install SUNWccon package on SC 3.0 Admin Workstation

### Step 1.2.1

Install the SC 3.0 SUNWccon package (containing the SC 3.0 administrative and console software) on the clustadm workstation.

For example, change to the Sun Cluster 3.0 Packages directory containing the SC 3.0 SUNWccon package by entering the following commands:

```
root@clustadm# cd  /cdrom/suncluster_3_0/SunCluster_3.0/Packages

root@clustadm# pkgadd -d . SUNWccon
```

Verify the package installs, successfully.

**Note –** The installation of SC 3.0 client software is now complete. The cluster software provides tools needed to build and manage the cluster.

# Section 1.3: Patch Installation - Administration Workstation

In this section, install all recommended Solaris patches on the `clustadm` workstation.

## Step 1.3.1

For local (manual) installations only, obtain the latest recommended Solaris OE patches from Sun Solve Online. Go to `http://sunsolve.Sun.com` and click the Patches option on the left side column.

---

**Note –** SunSolve is a contract service from Sun Enterprise Services. It is a good idea to subscribe to this service, especially if you are running a production server.

---

**Key Practice:** Create a `/PATCHES` directory on a dedicated Management Server to store all patches. This enables centralized patch management. For example, the Sun BluePrints™ BPLAB hardware has been configured with a ‘master’ JumpStart server, which will serve all software binaries and patches, and act as the repository.

**Key Practice:** Refer to the individual patch `README` files to review any installation prerequisites before installing patches. Using this practice could possibly prevent conflicts with other patches, software, `bootprom` variables, or other unknowns.

## Step 1.3.2

For local (manual) installations only, install all recommended patches for Solaris 8.

## Step 1.3.3

On the `clustadm` workstation, ensure patches were installed successfully (and applied).

**Key Practice:** Review the `/var/sadm/README` file to identify important log files to be examined, as well as the `/var/sadm/install_data/Solaris_2.8_Recommended_log` file. Resolve any patch installation errors which may occur.

### Step 1.3.4

For local (manual) installations, reboot the `clustadm` workstation after all patches have been successfully installed and applied.

**Key Practice:** Reboot the system after patches have been installed and applied. It is often a good idea to reboot the system after changes are made to system software and configuration files. For example, at this time, reboot the `clustadm` workstation after patches have been installed, ensuring changes are applied, and that a consistent state has been achieved.

## Section 1.4: Configure Management Server for Administering Cluster Nodes

Configure the `clustadm` workstation to administer each node in the SunPlex platform. Implement basic cluster management services, which can help ensure consistent, efficient cluster installations and operations.

### Step 1.4.1

Verify the `/etc/inet/hosts` file is configured to support the SunPlex platform. Add the IP address for each Solaris install client (cluster node) to be serviced by JumpStart, as well as the additional host name entries required to support the SunPlex platform.

At this time, ensure that the corresponding entries are created in support of the SunPlex platform, including entries for each cluster node to be managed, the `clustadm` workstation, the terminal concentrator (`tc`), plus any additional site-specific entries, such as the logical host entry required for our HA-NFS data service.

```
root@clustadm# more /etc/inet/hosts

... {output omitted}

xxx.xxx.xx.xxx  clustadm loghost
xxx.xxx.xx.xxx  clustnode1
xxx.xxx.xx.xxx  clustnode2
xxx.xxx.xx.xxx  lh-hanfs
xxx.xxx.xx.xxx  tc
```

## Step 1.4.2

Create the `/etc/clusters` file with an entry for this cluster, with the following format:

```
<Name of Cluster> <Node1 Name> <Node2 Name> ...
```

where:

`<Name of Cluster> = nhl`

`<Node1 Name> = clustnode1`

`<Node2 Name> = clustnode2`

On the `clustadm` workstation, create the `/etc/clusters` file, configuring `clustnode1` and `clustnode2`. Verify that the `/etc/clusters` file is correct:

```
root@clustadm# cat /etc/clusters

nhl clustnode1 clustnode2
```

## Step 1.4.3

Next, configure the `/etc/serialports` file on the `clustadm` workstation, enabling a connection to the terminal concentrator through the `ttya` (serial) port. Create the `/etc/serialports` file, as shown in the following code box.

At the `clustadm` workstation, verify `/etc/serialports` file is correct, as follows:

```
root@clustadm# cat /etc/serialports

clustnode1 tc 5002

clustnode2 tc 5003
```

**Note –** In this example, the terminal concentrator (`tc`) has 8 ports, numbered 1 – 8. The `500`*2* entry refers to the terminal concentrator physical port 2, and `500`*3* refers to physical port 3. The `tc` entry must correspond to the host name entry in `/etc/inet/hosts`.

# Section 1.5: Configure the Terminal Concentrator

Some steps for configuring the terminal concentrator require that you are able to gain physical access to the terminal concentrator and equipment. For example, forcing the terminal concentrator into the 'Monitor' mode. The following steps are included only when performing these steps locally, as during a manual installation.

## Step 1.5.1

For local (manual) installations, prior to connecting the serial cable:

- Ensure that the terminal concentrator power is off.
- Connect the cable, noting the serial port `ttyb` is the default serial port you will be using on the `clustadm` workstation.

- As described in Figure 1 and Tables 1 through Table 5, a serial cable must be connected from a serial (`ttya` or `b`) port on the `clustadm` workstation to port 1 on the terminal concentrator. Port 1 (configuration port) of the terminal concentrator is required for performing all local (manual) steps.

**Note –** Step 1.5.2 is not implemented for these BPLABS. Also, Figure 1 and Tables 1 through Table 5 indicate that serial port "A" of the `clustadm` workstation is connected to the terminal concentrator (port 1) instead of the default `ttyb`.

## Step 1.5.2

For local (manual) installations, use the UNIX `tip` command to communicate with the terminal concentrator during configuration.

**Note –** Before the `tip` command will work, ensure that the `/etc/remote` file includes the following lines (appended to the end of the file):

```
annexterm:\
        :dv=/dev/term/n:br#9600:el=^C^S^Q^U^D:ie=%$:oe=^D

{{In the above line, substitute the serial port letter you
are using for "n" - for example, if using ttyb, replace "n"
with "b", as: ".....dev/term/b....."}}
```

- An easy way create this entry is to simply copy the lines from the `hardwire` command, then change the entry from *hardwire* to *annexterm*, ensuring that the port letter is correct.
- The `tip`(1) command can be used to connect the `clustadm` workstation console I/O with the terminal concentrator I/O using the parameters specified in this file.

## Step 1.5.3

For local (manual) installations, on the `clustadm` workstation enter the following command:

```
root@clustadm# tip annexterm
connected
```

**Note –** At this point, for local (manual) installations, you will not receive a prompt back from the terminal concentrator until you physically power off the terminal concentrator, power the terminal concentrator back on and, within 30 seconds, and press and hold the TEST button until the power LED starts flashing. This is described in the next step.

## Step 1.5.4

For local (manual) installations, configure the terminal concentrator device in the following prescribed sequence:

- Power on the terminal concentrator while viewing the front panel and status LED indicators.
- After powering on the unit and *within 30 seconds* press and hold the TEST button until the power LED starts flashing.
- When the POWER LED starts to flash, release the TEST button for *at least* six seconds. Then briefly push the TEST button again to initiate diagnostics.

Before proceeding to the next step, ensure that the terminal concentrator successfully completes (the diagnostic tests take approximately 60 seconds to complete).

**Key Practice:** Verify that the terminal concentrator settings are correct. Terminal concentrator firmware options and settings can vary from between different terminal concentrator revision levels (that is, actual options will likely differ from those specified in this BPLAB). When configuring the terminal concentrator, refer to the manufacturer's documentation to ensure that settings are established correctly. Specifically, ensure that settings for the terminal concentrator Internet address, subnet mask, and broadcast address are as indicated in the following steps.

**Note –** For the next few steps to configure the terminal concentrator, the settings should be configurable 'as-listed.' However, terminal concentrator firmware settings vary from unit-revision to unit-revision. Your actual options may differ. When configuring the terminal concentrator, refer to the manufacturer's documentation to ensure that settings are established correctly. Specifically, ensure settings for the terminal concentrator, IP address, subnet mask, and broadcast address, as indicated in the following steps.

## Step 1.5.5

For local (manual) installations, when the diagnostic tests are completed, the `tip` window of the `clustadm` workstation should display:

```
System Reset - Entering Monitor Mode
monitor:
```

## Step 1.5.6

For local (manual) installations, we will use the `addr` command to set the Internet address, subnet mask, and broadcast address for the terminal concentrator.

Enter the following commands:

```
System Reset - Entering Monitor Mode
monitor:addr
Enter Internet address [192.40.85.60]:: 192.9.200.4
Enter Subnet mask [255.255.255.0]:: <CR>
Enter Preferred load host Internet address [<any host>]:: <CR>
Enter Broadcast address [129.153.49.255]:: 192.9.200.255
Enter Preferred dump address [0.0.0.0]:: <CR>
Select type of IP packet encapsulation(ieee802/ethernet)
[<ethernet>]:: <CR>
Type of IP packet encapsulation: <Ethernet>
Load Broadcast Y/N:: [N] <CR>
monitor:: sequence

At this point you need to enter a list of 1 to 4 interfaces to
attempt to use for downloading or upline dumping. Enter them in
the order they should be tried, separated by commas or spaces.

Possible interfaces are:
Ethernet: net
SELF: self

Enter interface sequence [net]:: self
Interface sequence: self
monitor:: ~.     {{this command ends the tip session}}
```

The `addr` command displays and sets several Annex terminal concentrator operating parameters:

- IP address
- Subnet mask
- Preferred load host IP address

- Load/dump gateway IP addresses
- Broadcast address
- IP encapsulation type

The sequence command edits the load/dump interface list. This list determines the order of the network interfaces, and either the local area network (LAN) or the SLIP interface to be used by the Annex terminal concentrator for loading and dumping. The default Ethernet selection is "net" which uses the LAN interface.

The `self` selection specifies the terminal concentrator is self-boot configured.

## Step 1.5.7

For local (manual) installations, the terminal concentrator must be power cycled for the previous (manual) configuration changes to take effect:

- Power off the terminal concentrator.
- Power on the terminal concentrator and wait approximately 90 seconds for it to configure.

**Key Practice:** Because port 1 of the terminal concentrator is the configuration port, minimize security vulnerability by disconnecting port 1 of the terminal concentrator after configuration. This will prevent unauthorized access to the terminal concentrator's configuration port.

## Step 1.5.8

For local (manual) installations, if the terminal concentrator requires access from an adjacent network, the `defaultrouter` configuration must be performed on each cluster node. This would be performed later, after the Solaris OE installation has completed on each cluster node. At that time, configure the default router information on the cluster nodes by performing the following steps.

- Create the file `/etc/defaultrouter` and insert the IP address of your gateway.

Example:

```
192.9.200.254 {{sample gateway address}}
```

**Note –** In this example, we use the gateway IP address instead of the host name (see following key practice).

**Key Practice:** For each node being managed, maintain all IP addresses in one location. Edit the file `/etc/defaultrouter` and add the host name of the gateway system. Next, edit the `/etc/inet/hosts` file and reference the gateway (host name, IP address, plus an appropriate comment to identify the entry—for example, default gateway). In this manner, you can easily record and maintain a complete list of all nodes in the SunPlex.

- If you want to enable a default route without rebooting, and verify your new router setting, enter the following commands:

```
root@clustadm# ping tc
tc is alive

root@clustadm# route add default 192.9.200.254

. . . {{this is only a sample gateway address}} . . .

root@clustadm# netstat -rn

. . . {{look for 192.9.200.254 entry with "ug" flag}}. . .
```

## Step 1.5.9

For local (manual) installations, complete the configuration of the terminal concentrator by entering the following commands. Enter the data as shown where prompted:

```
root@clustadm# telnet tc
Trying 192.2.200.4...
Connected to tc.
Escape character is '^]'.
cli
Annex Command Line Interpreter * Copyright 1991 Xylogics, Inc.
annex: su
Password: 192.9.200.4 {{the password defaults to the assigned IP address, but does not echo to the screen}}
annex# edit config.annex
Ctrl-W: save and exit Ctrl-X: exit Ctrl-F: page down Ctrl-B: page up
       # The following are definitions of the gateway entries
       #
       %gateway
       #
       # The following are definitions of the macro entries
       #
       %macros
       %include macros
       #
       # The following are definitions of the rotary entries
       #
       %rotary
       %include rotaries
       <Cntrl>+W #to exit and save
annex# admin
admin: port all
admin: set port mode slave
You may need to reset the appropriate port, Annex subsystem or
reboot the Annex for changes to take effect.
admin: reset
reset default port set [y] ? y
admin: quit
annex# hangup {{ends the login session}}

root@clustadm#
```

If the terminal concentrator requires access from an adjacent network, the `defaultrouter` should be included in the `config.annex` file of the terminal concentrator.

**Key Practice:** To avoid security vulnerability, change the default password for the terminal concentrator. After changing the password, maintain your Telnet session with the terminal concentrator and use another window to connect to the terminal concentrator by Telnet to verify that your new password works. A typographical error when entering a password change will render the terminal concentrator inaccessible. See SunSolve SRDB ID 24989 for resetting a lost root password.

## Step 1.5.10

For local (manual) installations, change the terminal concentrator password, which is similar to changing a UNIX password, as follows:

- Use the `telnet` command to connect to the terminal concentrator. Enter the `cli` command at the port prompt to enable the command interpreter.
- Enter `su`, then the default password (the terminal concentrator IP address).
- Enter the `passwd` command at the prompt and change the default password. Record any changes, and the new password setting.

**Note –** While this is important in the field, for the purpose of these manual procedures, we use the default password. If the default password is lost, you must enter the `erase` command to clear the EEPROM contents and then re-enter all data once again. See SunSolve article SRDB ID 24989

# Section 1.6: Configure Cluster Control Panel

During this section you will start the Cluster Control Panel by entering the `ccp` command for the cluster named "**nhl**". Please read this entire step before entering any commands. After starting the cluster control panel, you will double-click on the `cconsole` icon.

- At this time, verify that each cluster node is accessible to the `clustadm` workstation by starting the Cluster Console Panel and accessing the cluster consoles for each cluster node.

## Step 1.6.1

If you are accessing the clustadm workstation from a remote system, execute the **xhost +** command, enabling remote display from the clustadm workstation to your local system.

When accessing the clustadm workstation remotely you must, also, set the DISPLAY environment variable on the clustadm workstation to point to your local system. For example, for csh users, **setenv DISPLAY** *yoursystem*:**0.0**.

**Note –** This step can be performed when accessing the SunPlex platform from a remote workstation. It is often useful to access the Cluster Control Panel (ccp) remotely, as appropriate, or when configuring (administering) the Sun Cluster.

At this time, you must set the DISPLAY variable before invoking the CCP. First, on your local workstation (example only):

```
yoursystem# /usr/openwin/bin/xhost +clustadm
```

Next, on clustadm (note: replace yoursystem with local system name):

```
root@clustadm# setenv DISPLAY yoursystem:0.0
```

## Step 1.6.2

Enter the following commands, on the clustadm workstation:

```
root@clustadm# which ccp
/opt/SUNWcluster/bin/ccp

root@clustadm# ccp nhl &
```

When the ccp command is executed, the Cluster Control Panel window will appear. Verify that a menu bar and icon panel display all of the available tools, as listed:

- Cluster Console, console mode
- Cluster Console, rlogin mode
- Cluster Console, telnet mode.

Example: Cluster Control Panel Window

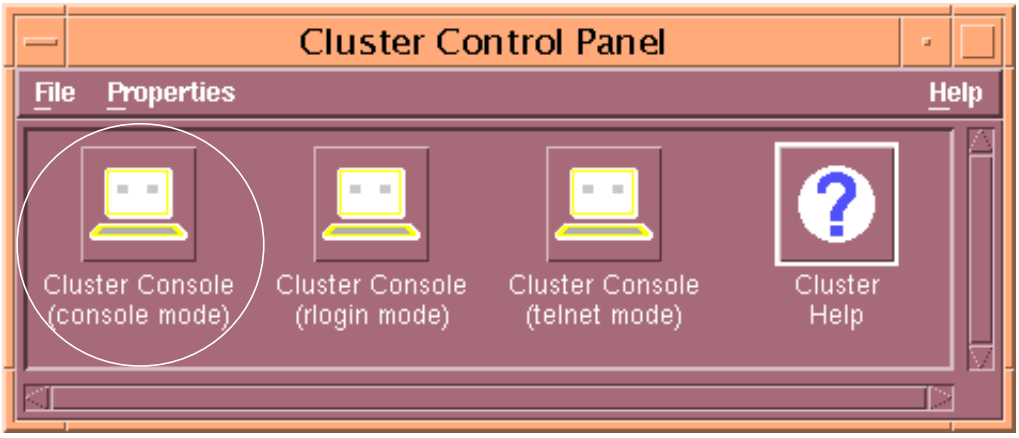


**Figure 2** Cluster Control Panel Window

## Step 1.6.3

Refer to the preceeding figure. Double-click the “Cluster Console (console mode)” icon (circled) to display the cluster console. An example cluster console shown in the following figure.

In this example, three windows are displayed: one small *Cluster Console* window, and two larger *cconsole: host [name]* windows. Note that each of the larger windows is associated with a specific host, or cluster node.

.Example: Cluster Console (console mode) and `cconsole` Windows

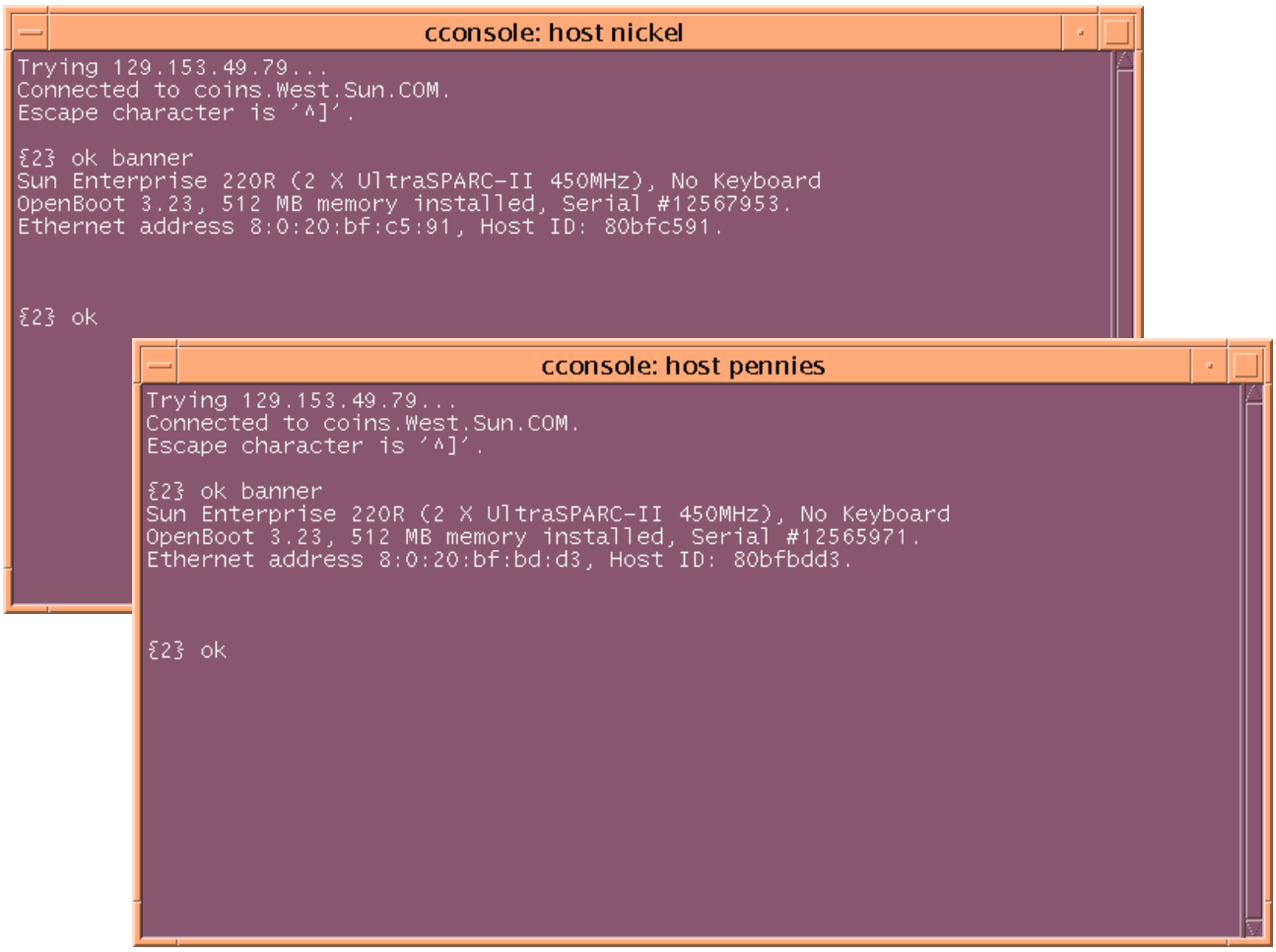


**Figure 3** Cluster Console and cconsole Windows

**Caution –** The Cluster Console utility provides a method of entering commands into multiple cluster nodes simultaneously (or individually, as required). ***Always be aware of which window is active prior to entering commands***. If a `cconsole` window does NOT appear for a cluster node, verify the following: From the *Cluster Console* window (console mode), select *Hosts*, followed by *Select Hosts*. Next, verify (insert) an entry for each cluster node (for example, `clustnode1`, `clustnode2`).

At this time, arrange each window for your own personal viewing preferences. Ungroup the *Cluster Console* window from the *cconsole: host [name]* windows. Select *Options* from the menu (Cluster Console window), and uncheck *Group Term Windows*.

For example, arrange the cconsole windows to be able to see each window clearly and at the same time by moving the Cluster Console window away from the other cluster node windows. This is done to ensure that commands are entered correctly into one, or both nodes, as required during these exercises (and to prevent entering commands into the wrong window).

It is NOT necessary to do so at this time, but when you wish to close the Cluster Console window, select *Exit* from the Cluster Console window "Hosts" menu.

**Note –** It is NOT necessary to do so at this time**,** but if you need to issue a "Stop-A" command to each cluster node, simultaneously placing them in the OBP mode, use the following procedure, for the Annex Terminal Server. First, activate the *Cluster Console* window, then press the ^] ( Ctrl + ] ) keys. This will display a telnet> prompt, for each cluster node. At the telnet> prompt, enter the send brk command, which will issue a Stop-A to each cluster node (placing them at the OBP ok prompt).

## Step 1.6.4

Verify operations using the Cluster Console Panel (CCP) utility by logging in to each cluster node. Begin configuring each system which will operate as a cluster node.

Log in as root from the *cconsole:host [hostname]* window, on each cluster node:

```
clustnode1 console login:root
Password: abc
```

```
clustnode2 console login:root
Password: abc
```

**Note –** During many of the following steps, you will be required to enter commands simultaneously into the console window for each cluster node. Use the Cluster Control Panel windows for this purpose. Double-click the Cluster Console (console mode) icon, as described in the previous section, "Configuring the Cluster Control Panel."

# Section 1.7: Configure Solaris OE (Each Cluster Node)

This section describes the steps necessary to configure the Solaris OE on each cluster node.

**Key Practice:** Ensure all firmware is installed with most recent (supported) versions, for all systems and subsystems, including all servers, disk arrays, controllers, terminal concentrators, etc. For example; On each node in the SunPlex platform, ensure the system EEPROM contains the most recent (supported) version of OpenBoot Prom (OBP). For example, `OpenBoot 3.15`

## Step 1.7.1

For local (manual) installations, it is a good idea to ensure that each cluster node is configured with the most recent version of the OBP. Information about downloading can be obtained from SunSolve Online, at `http://sunsolve.sun.com`.

**Note –** After upgrading the OBP successfully, you must manually turn system power off, then on. If you have just upgraded the OBP on each cluster node, you may skip the next step.

## Step 1.7.2

For local (manual) installations, it is a good idea to ensure that no previous EEPROM settings exist by setting the EEPROM to a known state (that is, factory defaults). It is recommended that this be performed only once, at this point in the procedure, prior to customizing the system EEPROM to meet cluster requirements (and BEFORE installing any software). Reset the system EEPROM to its factory default.

For local (manual) installations, enter the following OBP command on each cluster node:

```
ok set-defaults
```

Using the `set-defaults` command at this step establishes a consistent, known (default) state of all OBP variables prior to customizing the OBP environment.

**Caution –** Resetting the system EEPROM should only be performed at this time, during the initial preparation for the Solaris OE installation. This command resets all EEPROM (OBP) variables to their factory default values. All subsequent steps assume the EEPROM has been reset (at this point in the exercise). During the next few steps, the EEPROM will be modified (customized).

**Key Practice:** Ensure a consistent state on each cluster node before proceeding to configure site-specific (customized) OBP settings. Prior to implementing any configuration changes, and as part of initial Solaris OE installation preparations, reset the EEPROM to the factory defaults. This is done only once, and at this point in the procedure, and will easily and quickly ensure that a consistent state is achieved **before further customization** occurs.

**Note –** For local (manual) installations, prior to installing Solaris , we will re-configure the OBP settings for each cluster node. This is achieved by executing commands at the OBP 'ok' prompt (the OBP ok prompt should be viewable through the Cluster Control Panel windows).

## Step 1.7.3

On each cluster node, execute the OBP banner command to verify system information, such as the system model number, OBP version, Ethernet address, hostid, and serial number.

```
ok banner
```

Each node will respond with configuration information.

- Document system information of each cluster node.

**Key Practice:** Until the EEPROM configuration has been completed, you should disable the auto-boot EEPROM feature on each cluster node. Disabling the auto-boot feature will alleviate any problems which could arise if both systems attempted to boot their Solaris OEs while, at the same time, both systems are set with the same, and therefore, conflicting SCSI-initiator ID settings.

We temporarily disable `auto-boot?` on each cluster node, during this phase of the installation. We do this because, as yet, the system has not been configured—if there is an accidental reboot of a node or nodes, and the system `auto-boot?` variable has been set to FALSE, the system will reset to the OBP prompt instead of attempting to boot from disk. At this phase, any attempts to boot from disk may require an administrator to manually put the system back to the OBP for further configuration changes.

---

**Note –** You will be instructed to re-enable `auto-boot?` at the end of this procedure.

---

## Step 1.7.4

Disable `auto-boot?` by entering the following command into each cluster node:

```
ok setenv auto-boot? false
auto-boot? =          false
```

## Step 1.7.5

On each cluster node, set the following OBP variables, as indicated:

```
ok setenv local-mac-address? false
local-mac-address? = false

ok setenv diag-level min
diag-level = min

ok setenv diag-switch? false
diag-switch? = false
```

## Step 1.7.6

For this two-node cluster, set the global SCSI-initiator ID on the second cluster node only. On `clustnode2`, set `scsi-initiator-id` to a value of ‘6’. By changing the SCSI-initiator-ID on `clustnode2`, we are making a global variable change which will impact other devices attached to `clustnode2`’s internal SCSI controller. Specifically, setting the global SCSI-initiator-ID to 6 will create a conflict between the internal SCSI controller and the internal CD-ROM.

**Note –** To solve this conflict, in the NEXT step we will explicitly set the SCSI-initiator-ID of `clustnode2`'s internal SCSI controller to a value of '7' , by entering a simple script into the `clustnode2's` non-volatile RAM, or `'nvramrc`'.

At this time, enter the following command into the *cconsole: host clustnode2* window:

```
ok setenv scsi-initiator-id 6
scsi-initiator-id = 6
```

**Note –** SCSI-initiator-ID modification. Refer to Figure 1 and Table 1 through Table 5, specifically noting the disk subsystem cabling and configuration. Because two cluster nodes (both Sun Enterprise Model 220R servers) are connected to the same pair of Sun StorEdge D1000s, the OBP settings require modification. We will set the SCSI-initiator-ID on one of the cluster nodes (clustnode2 in this exercise) to a value of 6 and insert a script into `clustnode2's nvramrc` (non-volatile memory) to maintain a SCSI-initiator-id of 7 on the `clustnode2` internal SCSI controller. Setting the `clustnode2` global SCSI-initiator-id to 6 will prevent a conflict on the shared SCSI bus that connects both Sun Enterprise 220Rs to the Sun StorEdge D1000s.

**Note –** Use the `OBP nvedit` command in the following procedure. **The `nvram` editor is always set to `insert` mode.** Use the following keystrokes when editing (refer to the following).

Using `nvedit`: Keystrokes

| Keystroke | Action |
|---|---|
| `Ctrl+B` | Move backward one character |
| `Ctrl+C` | Exit the `nvramrc` editor, returning to the Open Boot PROM command interpreter. The temporary buffer is preserved, but is not written back to the `nvramrc` editor. (Use `nvstore` afterwards to write it back.) |
| `Delete` | Delete previous character. |
| `Ctrl+F` | Move forward one character. |
| `Ctrl+K` | From current position in a line, deletes all text after cursor and joins the next line to the current line (that is, deletes the new line). |

| Keystroke | Action |
|---|---|
| Ctrl+L | List all lines. |
| Ctrl+N | Move to the next line of the nvramrc editing buffer. |
| Ctrl+O | Insert a new line at the cursor position and stays on the current line. |
| Ctrl+P | Move to the previous line of the nvramrc editing buffer. |
| <CR> | Insert a new line at the cursor position and advance to the next line. |

**Note – Using nvedit can be tricky because there is no *delete* command.** If you want to delete a line you must delete all the characters for the line. Use Ctrl+K to join the empty line with the subsequent line.

## Step 1.7.7

On clustnode2, set the internal SCSI controller (example: /pci@1f,4000/scsi@3) SCSI-initiator-ID value to 7 by using the nvedit command. Enter the following commands into the *cconsole: host clustnode2* window:

```
ok printenv nvramrc

nvramrc =

{{ensure that no previous commands / entries exist in nvram, before proceeding}}

ok nvedit {{invoke the nvram editor}}
   0: probe-all
   1: cd /pci@1f,4000/scsi@3
   2: 7 " scsi-initiator-id" integer-property {{note the space before AND after the first quotation prior to
the word 'scsi' in this line}}
   3: device-end
   4: install-console
   5: banner
   6: {{at this point, use Ctrl + C to exit nvedit}}

ok nvstore
ok printenv nvramrc {{verify/compare this exactly matches with your screens output}}

nvramrc =             probe-all
                      cd /pci@1f,4000/scsi@3
                      7 " scsi-initiator-id" integer-property
                      device-end
                      install-console
                      banner
```

## Step 1.7.8

Enter the following command into the *cconsole: host clustnode2* window to enable the nvramrc:

```
ok setenv use-nvramrc? true
use-nvramrc? =        true
```

## Step 1.7.9

Verify the nvramrc script works by performing the following steps on *clustnode2*:

1. On clustnode2, reset the system by entering the reset command into the OBP ok prompt.
2. After clustnode2 resets, enter the printenv scsi-initiator-id command into the OBP ok prompt to confirm that the global SCSI-initiator ID is set to 6:

```
ok reset-all
Resetting ...
ok printenv scsi-initiator-id
scsi-initiator-id =   6
```

3. On clustnode2, use the cd command to navigate to the node (that is, directory) that represents the internal SCSI controller /pci@1f,4000/scsi@3.
4. At the OBP ok prompt, enter the .properties command to verify that clustnode2's internal SCSI bus SCSI-initiator ID is set back to 7, as indicated in the next code box.

```
ok cd /pci@1f,4000/scsi@3
ok .properties
scsi-initiator-id        00000007
.
.
.
```

## Step 1.7.10

**For local (manual) installations**, when dual-hosted SCSI devices are configured (that is, cluster-pair configuration using dual-hosted D1000 arrays), verify that the probe-scsi-all command completes successfully on both cluster nodes. This is performed after the reset-all command succeeds.

For local (manual) installations, enter the following OBP command, simultaneously, on both cluster nodes, and verify the command completes successfully:

```
ok probe-scsi-all
```

## Step 1.7.11

After you have verified all settings are correct, reset auto-boot? to "true" on each cluster node:

```
ok setenv auto-boot? true
auto-boot? =          true
```

## Step 1.7.12

For local (manual) installations, install and configure Solaris software.

Checklist recommendations include installing the Entire Distribution (+OEM, as required by third party applications). Select the proper locale (English) and root disk layout.

**Key Practice:** Implement Solaris JumpStart to maintain consistency and fully automate the installation of the Solaris OE and additional software packages for each node in the SunPlex platform. Solaris JumpStart can improve installations and minimize operator errors that occur during a manual process. Combining Solaris JumpStart and Flash archives can be used to enable quick and consistent disaster / recovery operations.

## Step 1.7.13

For local (manual) installations, reboot each cluster node after installing the Solaris software.

## Step 1.7.14

At this phase of the installation, verify that all Solaris site specific data is correct on each cluster node:

Example Solaris OE Configuration - `clustnode1`

| | |
|---|---|
| **Host name:** | clustnode1 |
| **IP Address:** | 129.153.xx.xxx |
| **Name Service:** | None (local files) |
| **Set Subnet:** | Yes |
| **Subnet Mask:** | 255.255.255.0 |
| **Default Gateway:** | None |

**Note –** The values quoted in the previous table are sample values. In a live installation, substitute the appropriate site-specific values.

# Step 1.7.15

Verify the primary boot disk (`c0t0`) is partitioned correctly on each cluster node.

**Key Practice:** We recommend configuring a standard, flexible disk partitioning scheme. In this case, a consistent, flexible partitioning scheme that allows for using either SVM or VxVM. Implement the following standard partitioning for boot disks. Furthermore, partitioning each disk spindle identically can save time, provide flexibility, and maintain consistency across nodes.

**Note –** It is often easier to use the Solaris `format` command to calculate the exact number of even cylinders to be configured, as when determining the size of the `root` file system. Using these guidelines, the size of the `root` file system is dependent on the actual size of the disk.

On each cluster node, verify that the primary boot disk (`c0t0`) is partitioned for site-specific requirements.

| |
|---|
| Slice 0 = assigned to "`/`" (10 GB) |
| Slice 1 = assigned to `swap` (2 GB) |
| Slice 2 = assigned as "`backup`" (full extent) |
| Slice 6 = assigned to `/globaldevices` (100 MB) |

**Note –** Boot disk partitioning must adhere to site-specific requirements. For each cluster node, we recommend the following boot (root) disk partitioning guidelines: Combine `/`, `/usr`, and `/opt` (recommended), and add an additionial 100 Mbytes to the size of the root (`/`) file system. If `/usr` is a separate file system, include an extra 40 Mbytes. If `/var` is a separate file system, it should be sized appropriately, to ensure core dumps can be saved. 750 Mbytes swap is minimum, however, swap size should be 2 Gbytes, or 2x RAM whichever is greater. Configure `/globaldevices` (100 Mbytes); Leave 20 Mbytes for SVM, and assign to slice 7. For VxVM, reserve first two cylinders for private region and encapsulation area, and ensure slice 3 and 4 are *unassigned.* Customer applications may require one slice. Live Upgrade requires one slice.

**Caution –** Our disk data layout combines `/var` under the root file system (`/`). Consider an alternate approach, placing `/var` on a separate disk slice. When `/var` is combined with the root file system, as in our configuration, consider disabling `sendmail`. Alternatively, if `sendmail` is required by your applications you should limit the amount of free space available to the file system by explicitly setting `MinFreeBlocks` (`sendmail`) variable. Upon reaching this limit, this will cause `sendmail` to reject incoming messages, rather than causing the `/var` file system to run out of space; thus, preventing the basis for this type of denial of service attack.

**Caution –** By default, `/tmp` is installed as `tmpfs` file system, which can potentially result in total consumption of all system virtual memory, and ultimately cause the system to hang. Avoid this by explicitly setting the `size=` *option* for `mount_tmpfs(8)` to help prevent the basis for this type of denial of service attack by any login user. Alternatively, consider converting `/tmp` to use real disk space, though some applications would suffer.

## Step 1.7.16

Examine the installation logs, ensuring that any Solaris OE installation errors do not go undetected or unresolved. On each cluster node, enter the following commands:

```
# cd /var/sadm/system/logs
# pwd
/var/sadm/system/logs

# ls
begin.log              finish.log              install.log
begin.log_2000_04_13   finish.log_2000_04_13   sysidtool.log
{{sample dates only}}
# more *
{{It is important to resolve any installation error messages noted in the logs.

Example:
pkgadd: ERROR: postinstall script did not complete successfully}}
```

**Key Practice:** Verify that the Solaris OE installation was successful and that any errors reported are fully resolved before proceeding. Review the `/var/sadm/README` file to determine the location of the most recent installation logs (for example, `/var/sadm/system/logs`). Examine the current Solaris OE installation log files for potential errors (that is, `begin.log`, `sysidtool.log`, or `install_log`). Confirm the cause of any installation error messages which may occur, resolving failures before proceeding further.

## Step 1.7.17

On each node in the SunPlex platform, examine the `/etc/inet/hosts` file, verifying that IP addresses and host names are configured correctly.

**Note –** Prior to using a text editor (such as, vi) , set your terminal environment variable, as appropriate—for example, **`TERM=ansi or vt100`—**for proper video display.

On each cluster node, configure this file to include the entries for each cluster node (that is, `clustnode1`, `clustnode2`), the SC3.0 Admin. Workstation (`clustadm`), our single terminal concentrator (`tc`) and logical host (`lh-hanfs`). For each cluster node, append '`hostname.some.com`' to the IP Address entry as per the example, below, eliminating `sendmail` messages. Verify that each cluster node is configured correcly, as indicated in the code boxes below:

```
clustnode1# more /etc/inet/hosts

... {{output omitted}} . . .

xxx.xxx.xx.xxx  clustnode1 loghost clustnode1.some.com
xxx.xxx.xx.xxx  clustadm
xxx.xxx.xx.xxx  clustnode2
xxx.xxx.xx.xxx  tc tc-nhl

... {{output omitted}}

xxx.xxx.xx.xxx  lh-hanfs

... {{output omitted}}
```

```
clustnode2# more /etc/inet/hosts

... {{output omitted}} . . .

xxx.xxx.xx.xxx  clustnode2 loghost clustnode2.some.com
xxx.xxx.xx.xxx  clustadm
xxx.xxx.xx.xxx  clustnode1
xxx.xxx.xx.xxx  tc tc-nhl

... {{output omitted}}

xxx.xxx.xx.xxx  lh-hanfs

... {{output omitted}}
```

**Note –** Each example shows standard host name entries for the two-node cluster. On each cluster node, verify that your configuration matches the actual site specific installation requirements. This file will be modified further, during SC3.0U3 software installation procedures.

## Step 1.7.18

On each cluster node, ensure that all shell environment variables and path names are configured, as required.

Example: Environment Variables (settings) - `clustnode1`

| Variable | Label |
|---|---|
| TERM | ansi or vt220 |
| stty | istrip |
| set prompt | root@<hostname># {{e.g., root@clustnode1}} |
| Ensure the following PATH variable settings | /usr/bin:/usr/ucb:/etc:/sbin:/usr/sbin:/usr/cluster/bin:/usr/cluster/lib/sc:/opt/VRTSvmsa/bin :/opt/VRTSvxvm/bin:/etc/vx/bin:/usr/cluster/dtk/bin: /opt/CTEact/bin:/opt/SUNWexplo/bin:/opt/sun/bin:/usr /ccs/bin:/usr/openwin/bin:/usr/dt/bin |
| Ensure the following MANPATH variable settings | /usr/dt/man:/usr/openwin/share/man:/opt/VRTS/man:/usr/cluster/dtk/man:/opt/CTEact/man:/opt/SUNWexplo/man :/usr/cluster/man:/usr/man:/usr/share/man |

## Step 1.7.19

On each cluster node, configure the root shell installation environment. For example, you will modify `/.profile` to uncomment the DISPLAY settings that are located in the "`clusters`" section (end of file). For VxVM, you must, also, correct the default entry for LD_LIBRARY_PATH, on each cluster node.

**Note –** In some circumstances, setting LD_LIBRARY_PATH in this manner could result in performance penalties, and should be restricted.

At this time, edit the `/.profile` file on each cluster node, making the changes referenced in the following code box. After making changes, verify that the entries are correct, as indicated below.

```
# more /.profile
# Initial settings for user root
# Version . . .

{{..... output omitted....}}


# Modify the following entry, per EIS
LD_LIBRARY_PATH=/usr/lib:$LD_LIBRARY_PATH:/usr/openwin
/lib

{{..... output omitted....}}

# Uncomment next section for cluster node . . .

LOGINFROM=‘who am i | cut -f2 -d"(" | cut -f1 -d")"‘
DISPLAY=${LOGINFROM}:0.0
export LOGINFROM DISPLAY
echo ""
echo "DISPLAY=$DISPLAY"
echo "“


# Note: BPLAB’s recommendation is to use TERM=ansi,
# instead of vt100
if [“‘tty‘” = “/dev/console”]; then
TERM=vt100; export TERM
fi
```

## Step 1.7.20

If not already done, verify that remote root login is enabled during the installation. This change should only be temporary, and is useful during the installation.

Verify the `/etc/default/login` file has been modified, as indicated:

```
# more /etc/default/login

...{{output omitted}} . . .

# Uncomment the next line to allow remote root login

#CONSOLE=/dev/console
```

Remote root login should NOT be allowed (that is, it should be DISABLED) after the cluster installation has been completed successfully.

## Step 1.7.21

On each cluster node, verify all changes made during the previous step(s) are successful, and activated. For example, as the `root` user, log in remotely (`rlogin`) to the `clustadm` workstation, and log in remotely to each cluster node. For each node, verify environment variables are correct, after logging in remotely.

## Step 1.7.22

At this time, from the `cconsole` window of each cluster node, first log off then log in again as the `root` user. Next, from the Cluster Control Panel, choose `ctelnet`, which can be used during the installation, when entering commands to each cluster node.

---

**Note –** When rebooting cluster nodes, or examining the console windows for error messages, always refer back to the `cconsole` window.

---

## Step 1.7.23

Prior to installing VxVM 3.2, install the SUNWsan package on each cluster node, and add additional patches described during subsequent steps.

## Step 1.7.24

Add all Solaris 8 recommended patches available at `http://sunsolve.sun.com`.

**Note –** Verify that patches install correctly. **Do not reboot** at this time; instead, proceed immediately to the next step, and install any additional patches.

## Step 1.7.25

At this time, install all additional patches, as required for the configuration. For example, for Solaris 8/VxVM, patch `111413-xx` must be installed before installing VxVM 3.2 software. At this time, enter the following commands on each cluster node:

```
# cd /cdrom/PATCHES/VERITAS_3.2
# patchadd 111413-08
checking installed patches...

. . . {{ output omitted }} . . .

Patch packages installed:
SUNWluxop
SUNWluxox
#
```

**Note –** Ensure that installation errors do not go undetected, or unresolved, before continuing the installation. Refer to log files under `/var/sadm/patch/...` directory, for each patch installed. Note the `/var/sadm/README` file (when navigating the directory structure).

## Step 1.7.26

For local (manual) installations, verify that all patches are installed correctly. For example, to list all patches applied, enter the following command into each cluster node.

```
# patchadd -p | more
```

**Note –** Both the `/usr/sbin/patchadd -p` and `/usr/bin/showrev -p` commands will display a list of patches that have been added to the system. We will use `/usr/bin/patchadd -p | grep` <patch#>, where <patch#> is the number of the patch you are checking.

```
# patchadd -p | grep <patch#>
```

## Step 1.7.27

For local (manual) installation, modify this file, as indicated in the code box below, adding `[SUCCESS=return]`, after `files` (for the `hosts:` entry). After editing, verify that the changes are correct, as indicated in the following code box:

```
# more /etc/nsswitch.conf

. . . {{output omitted}} . . .

group: files
hosts: files [SUCCESS=return]
services: files
netmasks: files
```

**Key Practice:** The cluster environment requires that local (`/etc`) files supporting network services are searched ahead of any naming services. This increases availability by not having to rely on an outside agent. To do this, always put '`files`' first (ahead of `dns`, `nis`, etc.), for hosts, `netmasks`, `group`, and `services`.

## Step 1.7.28

Create the required `/etc/system` file entries, such as shared memory settings for Oracle, etc. After making changes, verify the settings are correct, as indicated.

**Caution –** Before editing, always take precautions, and reverify each entry is correct before proceeeding to the next step. Note that we DO NOT set the `rpcmod:svc_default_stksize` variable at this time. `SUNWscr` will create an entry later, setting this variable to `rpcmod:svc_default_stksize=0x6000`. We will modify this file later during subsequent Modules.

Verify that the changes are correct, as indicated, on each cluster node:

```
# more /etc/system

{{..... output omitted....}}

* added per SC3.0 U3 installation recommendations
exclude: lofs
set ip:ip_enable_group_ifs=0
forceload: misc/obpsym
set nopanicdebug=1
set lwp_default_stksize=0x6000
```

## Step 1.7.29

For the root file system, add the '`logging`' option by editing the `/etc/vfstab` file. On each cluster node, verify that the file has been modified, as indicated:

```
# more /etc/vfstab

{{..... output omitted....}}

/dev/dsk/c1t0d0s0 /dev/rdsk/c1t0d0s0 / ufs 1 no logging

{{..... output omitted....}}
```

**Note –** This file will be further modified by SUNWscr, in a future module.

## Step 1.7.30

For local (manual) installations, it is a good idea to verify system diagnostics run successfully before installing Sun Cluster software.

For example, at this time, we recommend using SunVTS to run CPU/MEMORY stress tests, and verify that no failures occur. At this point in the installation, however, do NOT execute any tests which overwrite disk data, unless you plan to reload software again.

Prior to further customization, verify system diagnostics are executed successfully on each cluster node. For example, EIS installation procedures require running SunVTS for a minimum of two hours to verify that each Solaris node can successfully complete diagnostics.

**Note –** At this time, it is important to resolve any diagnostic errors that occur before performing the next step in the cluster installation.

# Section 1.8: Solaris OE —Post Installation and Configuration

Upon successful completion of Solaris on each cluster node, perform post-installation (and site-specific) procedures, verifying the operating environment is configured correctly, prior to building the Sun Cluster.

For example, we start by ensuring the date/time is set correctly, on all SunPlex nodes, prior to installing additional software.

# Step 1.8.1

In this step, prior to installing additional software, ensure the date/time are correct on each SunPlex node, using the Management Server (or, designated time host).

**Note –** This is **NOT** intended to synchronize the date/time on a running production cluster. For additional references, and when configuring NTP in a running cluster, see the References section, specifically *SC3.0 Installation Guide* "How to Update Network Time Protocol (NTP)," "SC3.0 Concepts," and "Cluster Time."

**Key Practice:** In preparation for installing additional software, set the system date/time on each SunPlex node. In the future, this can also help ensure that relevant timestamps (that is, log file entries and timestamps, error messages, and cluster events) are closely synchronized. On all nodes, verify date/time are set correctly, as per site-specific requirements for date/time, timezone, or timehost.

**Note –** It is often useful to track events closely and accurately over specific time periods, and can be important when auditing system statistics/logs for performance monitoring, capacity planning, configuration management, troubleshooting, as well as when ensuring scheduled, or sustaining operations are completed successful.

For a running cluster, the date and time between all cluster nodes must be synchronized. For this, the SunPlex employs the Network Time Protocol (NTP) to synchronize all nodes to designated timehosts. In normal cluster operation you should not need to reset the date and time unless it was set incorrectly, perhaps duing system installation.

At this time, ensure the system date and time is correct on each node. Synchronize the date and time to `clustadm`. On each cluster node, enter the following:

```
# ping clustadm
clustadm is alive

# rdate clustadm
```

**Note –** Verify that `rdate` completes successfully and the date and time are set correctly on each SunPlex node.

## Step 1.8.2

We recommend setting up the Solaris `syslog` facility by configuring the `/etc/syslog.conf` file on each cluster node. After configuring this facility, verify messages are logged properly.

**Key Practice:** On each cluster node, configure the Solaris `syslog` facility to forward errors and system messages to the `clustadm` workstation. Each logged message includes a message header and a message body. The message header consists of a facility indicator, a severity level indicator, a timestamp, a tag string, and optionally the process ID. See `man` pages for `syslogd(1M)` and `syslog(3C),` for additional information.

## Step 1.8.3

At this time, **ensure the alternate boot disk (`c0t1`) is partitioned correctly.** Execute the `format` command and examine the current settings *for each slice*.

- On each cluster node, execute the `format` command and verify that the primary boot disk (`c0t0`) is properly partitioned, as previously described in Step 1.7.15.
- Next, ensure the alternate boot disk (`c0t1`) is configured to match the primary boot disk (`c0t0`). Execute the `format` command, and select the alternate boot disk (`c0t1`), creating partitions to match the primary boot disk (`c0t0`).

**Note –** In this step, the partition information shown is for ‘typical’ disk drives, where disk0 (`c0t0`) is the primary boot disk and disk1 (`c0t1`) will be configured as the alternate boot disk (boot mirror). Both physical disks are of identical model and type, and must be formatted identically.

## Step 1.8.4

At this time, configure disk partitions for all shared storage (disk spindles). Ensure that each shared disk is configured correctly, as per site-specific requirements.

- Ensure that each shared disk (spindle) is configured and verified before proceeding to the next step.
- Examine the partition table for *each* shared disk (that is, each D1000 spindle), including: `c1t0, c1t1, c1t2, c1t8, c1t9, c1t10, c2t0, c2t1, c2t2, c2t8, c2t9,` and `c2t10`.

Configure the partitions on each shared disk spindle.

- Verify the configuration follows these guidelines, as indicated below:

| |
|---|
| Slice 0 approximately 2 GB (for shared data). |
| Slice 2 (backup) defined as the full extent of the disk. |
| Slice 7 (alternate) Reserve cylinders 1-6 (the first six cylinders). |

**Key Practice:** Disk data layout must meet site-specific requirements. We recommend configuring a standard, flexible disk partitioning scheme. In this case, a consistent partitioning scheme that is flexible, allows for using either SVM or VxVM. Implement the following standard partitioning for boot disks. Partitioning each disk spindle identically can save time, provide flexibility, and maintain consistency across nodes.

## Step1.8.5

The following is an example method for easily, and consistently replicating ‘standard’ disk drive partitioning (that is, as when configuring multiple spindles). This example uses the `prtvtoc` and `fmthard` commands to replicate an existing ‘valid’ VTOC.

**Caution –** In the following sequence, always save a copy of the original VTOC before modifying and rewriting to disk. Take extra care and precaution when using these commands to modify and rewrite disk information. Also, our example *temporarily* saves the original VTOC data in the `/tmp` directory. Note that `/tmp/vtoc.orig` would be erased during a system reboot.

This example assumes that both disk spindles are identical type, model, and geometry, before reading the existing VTOC from the primary boot disk (c0t0), and writing the new VTOC to the alternate boot disk (c0t1).

```
# prtvtoc /dev/rdsk/c0t0d0s2 > /tmp/D0.vtoc
{{create copy of a “valid” VTOC; here, disk0 was previously
verified and will be replicated}}

# fmthard -s /tmp/D0.vtoc /dev/rdsk/c0t1d0s2
{{write the correct disk0 VTOC to disk1}}

# prtvtoc /dev/rdsk/c0t1d0s2
{{verify disk1 VTOC matches disk0 VTOC}}
```

**Note –** The OpenBoot Prom (OBP) commands probe-scsi or probe-ide can be used to determine disk information. For the boot disk (and alternate boot disk), ensure both disk drives are the same size. If they are not, use the values for the smaller drive.

**Caution –** Prior to issuing either of these OBP commands (for example, probe-scsi) you should temporarily set auto-boot? false and perform a reset-all command to avoid hanging the system (which would require local, manual intervention in order to re-set the hardware).

**End of Module One**

Module 1 is now complete. You have successfully performed the following procedures:

- Verify Administrative Workstation Installation (OBP, Solaris OE, and patches).
- Perform SC3.0 Admin Workstation (Management Server) setup.
- Terminal Concentrator configuration.
- Install and configure Cluster Console utility.
- Verify cluster node installation (OBP, Solaris OE, and patches).
- Configure syslog facility.

# References

*Cluster Platform 220/100 Architecture* – A Product from SunTone Platforms Portfolio by Enrique Vargas available at:

http://www.sun.com/blueprints/browsesubject.html

Sun WebStart Flash documentation is available at:

http://www.sun.com/solaris/webstartflash/

The following Sun BluePrints OnLine Articles are available at:

http://www.sun.com/blueprints/browsesubject.html

- *Availability - What It Means, Why It's Important, and How to Improve It* by Richard McDougal
- *Disaster Recovery Requirements* by Stan Stringfellow
- *High Availability Best Practices* by Enrique Vargas
- *HA: Boot/Root/Swap* by Jeannie Johnstone Kobert
- *The Art of Production Environment Engineering* by Bill Walker
- *Datacenter Naming Scheme* by Mark Garner
- Robust Clustering: A Comparison of SC3.0 vs. SC 2.2
- *Architecting a Service Provider Infrastructure for Maximum Growth* by Stan Stringfellow
- *Building Secure N-Tier Environments* by Alex Noordergraaf
- *Planning for Large Configurations of Netra t1 Servers* by Stan Stringfellow
- *Planning to Fail* by John S. Howard
- *High Availability Fundamentals* by Enrique Vargas
- *Tales from the Trenches: The Case of the RAM Starved Cluster* by Richard McDougal
- *Building a JumpStart Infrastructure* by Alex Noordergraaf
- *Using NTP to Control and Synchronize System Clocks (Parts 1 - 3)*
- *Solaris 8 Additions to* `sysidcfg` by Rob Snevely
- *An Overview of Methodology* by Adrian Cockcroft
- *JumpStart™ Mechanics: Using JumpStart Application for Hands-Free Installation of Unbundled Software (Parts 1 - 2)* by John S. Howard
- *Setting Up A Solaris Operating Environment Install Server and the Solaris JumpStart Feature* by Rob Snevely
- *Scrubbing Disks using Solaris OE format Program*
- *Fast Oracle Exports* by Stan Stringfellow

- *IP Network Multipathing* by Mark Garner
- *MR System for Rapid Recovery* by John S. Howard
- *VxVM Private Regions: Mechanics and Internals of the VxVM Configuration Database* by Gene Trantham
- *Veritas VxVM Storage Management Software* by Gene Trantham
- *SSP Best Practices* by John S. Howard

The following additional references are available at:
`http://docs.sun.com`

- Sun Cluster 3.0 System Administration Guide
- Sun Cluster 3.0 Installation Guide
- Sun Cluster 3.0 Hardware Guide
- Sun Cluster 3.0 Data Services Installation and Configuration Guide
- Sun Cluster 3.0 (U3) Release Notes “Patches and Firmware Requirements,” “Sun Cluster 3.0 Concepts,” and “Sun Cluster 3.0 Error Messages”