Parallel System Support Programs for AIX

Diagnosis Guide

Version 3 Release 1.1

GA22-7350-01

Parallel System Support Programs for AIX

Diagnosis Guide

Version 3 Release 1.1

GA22-7350-01

Note!

Before using this information and the product it supports, read the information in “Notices” on page xiii.

Second Edition (October 1999)

This edition applies to version 3, release 1, modification 1 of the IBM Parallel System Support Programs for AIX (PSSP) Licensed Program (product number 5765-D51) and to all subsequent releases and modifications until otherwise indicated in new editions. This edition replaces GA22-7350-00. Significant changes or additions to the text and illustrations are indicated by a vertical line (|) to the left of the change.

Order publications through your IBM representative or the IBM branch office serving your locality. Publications are not stocked at the address below.

IBM welcomes your comments. A form for readers' comments may be provided at the back of this publication, or you may address your comments to the following address:

International Business Machines Corporation
Department 55JA, Mail Station P384
522 South Road
Poughkeepsie, NY 12601-5400
United States of America

FAX (United States & Canada): 1+914+432-9405
FAX (Other Countries):
Your International Access Code +1+914+432-9405

IBMLink (United States customers only): IBMUSM10(MHVRCFS)
IBM Mail Exchange: USIB6TC9 at IBMMAIL
Internet e-mail: mhvrcfs@us.ibm.com

If you would like a reply, be sure to include your name, address, telephone number, or FAX number.

Make sure to include the following in your comment or note:

- Title and order number of this book
- Page number or topic related to your comment

When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Tables

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Department LJEB/P905
522 South Road
Poughkeepsie, NY 12601-5400
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States or other countries or both:

AIX
AIX/6000
DATABASE 2
DB2
ES/9000
ESCON

HACMP/6000
IBM
IBMLink
LoadLeveler
Micro Channel
Netfinity
Netfinity Manager
NQS/MVS
POWERparallel
POWERserver
POWERstation
RS/6000
RS/6000 Scalable POWERparallel Systems
Scalable POWERparallel Systems
SP
System/370
System/390
TURBOWAYS

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, BackOffice, MS-DOS, and the Windows logo are trademarks or registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Tivoli Enterprise Console is a trademark of Tivoli Systems Inc. in the United States, other countries, or both.

UNIX is a registered trademark in the United States, other countries, or both and is licensed exclusively through X/Open Company Limited.

Other company, product, and service names may be the trademarks or service marks of others.

Publicly Available Software

PSSP includes software that is publicly available:

expect Programmed dialogue with interactive programs

Kerberos Provides authentication of the execution of remote commands

NTP Network Time Protocol

Perl Practical Extraction and Report Language

SUP Software Update Protocol

Tcl Tool Command Language

TclX Tool Command Language Extended

Tk Tcl-based Tool Kit for X-windows

This book discusses the use of these products only as they apply specifically to the RS/6000 SP system. The distribution for these products includes the source code and associated documentation. (Kerberos does not ship source code.)

/usr/lpp/ssp/public contains the compressed **tar** files of the publicly available software. (IBM has made minor modifications to the versions of Tcl and Tk used in the SP system to improve their security characteristics. Therefore, the IBM-supplied versions do not match exactly the versions you may build from the compressed **tar** files.) All copyright notices in the documentation must be respected. You can find version and distribution information for each of these products that are part of your selected install options in the **/usr/lpp/ssp/README/ssp.public.README** file.

About This Book

This book contains information to help you diagnose and resolve problems for IBM RS/6000 SP systems and Parallel System Support Programs for AIX (PSSP). It does not contain the following:

- Information about diagnosing other SP products, such as Parallel Environment (PE) and LoadLeveler. These products have their own publications.
- Information about other SP system management issues. For this information, see *PSSP: Administration Guide*.

For a list of related books and information about accessing online information, see the Bibliography in the back of the book.

This book applies to PSSP Version 3 Release 1 Modification 1. To find out what version of PSSP is running on your control workstation (node 0), enter the following:

```
splst_versions -t -n0
```

In response, the system displays something similar to:

```
0 PSSP-3.1.1
```

If the response indicates **PSSP-3.1.1**, this book applies to the version of PSSP that is running on your system.

To find out what version of PSSP is running on the nodes of your system, enter the following from your control workstation:

```
splst_versions -t -G
```

In response, the system displays something similar to:

```
1 PSSP-3.1.1
2 PSSP-3.1.1
7 PSSP-2.4
8 PSSP-2.2
```

If the response indicates **PSSP-3.1.1**, this book applies to the version of PSSP that is running on those nodes.

If you are running mixed levels of PSSP, be sure to maintain and refer to the appropriate documentation for whatever versions of PSSP you are running.

Who Should Use This Book

This book is intended for system administrators, who are responsible for setting up and maintaining the SP system. This book can also be used by system operators and others, who are responsible for monitoring the status of the SP system and interacting with the hardware.

It is assumed that the reader has a working knowledge of AIX or UNIX and experience with network systems.

Typographic Conventions

This book uses the following typographic conventions:

Typographic	Usage
Bold	• **Bold** words or characters represent system elements that you must use literally, such as commands, flags, and path names.
Italic	• *Italic* words or characters represent variable values that you must supply. • *Italics* are also used for book titles and for general emphasis in text.
`Constant width`	Examples and information that the system displays appear in `constant width` typeface.
[ ]	Brackets enclose optional items in format and syntax descriptions.
{ }	Braces enclose a list from which you must choose an item in format and syntax descriptions.
|	A vertical bar separates items in a list of choices. (In other words, it means “or.”)
< >	Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard. For example, <**Enter>** refers to the key on your terminal or workstation that is labeled with the word Enter.
...	An ellipsis indicates that you can repeat the preceding item one or more times.
<**Ctrl-*x*>**	The notation <**Ctrl-*x*>** indicates a control character sequence. For example, <**Ctrl-c>** means that you hold down the control key while pressing <**c>**.
****	The continuation character is used in coding examples in this book for formatting purposes.

Part 1. Detecting and Investigating PSSP Problems

Chapter 1. Diagnosing SP Problems Overview

This chapter contains information to help you diagnose problems you may encounter installing or operating the SP system and PSSP. It helps you to identify whether a problem is related to the hardware or the software. It also shows you the procedure to follow if you require assistance from the IBM Support Center.

How To Use This Book

This section discusses what level of is PSSP supported by this manual, when to use this manual, and what other manuals are needed to diagnose PSSP problems.

SP Systems and PSSP Software Supported by This Book

The book applies only to PSSP Version 3 Release 1.1 (PSSP 3.1.1 This book does not supersede previous versions. To display the levels of PSSP installed on all nodes of your SP system, see “About This Book” on page xvii. If your SP system has mixed levels of PSSP, each version of PSSP has its own *PSSP: Diagnosis Guide*. Use the proper version of the manual to diagnose problems with a particular node. In order to diagnose problems on a node running a particular level of PSSP, you must use the manual that applies to that level.

When To Use This Book

Consult this manual for assistance during these administrative efforts:

- Before contacting the IBM Support Center to report a problem

  This manual lists the basic information that you should have available before contacting the IBM Support Center, and how to obtain that information. In addition to the basic information, specific PSSP software subsystems may require you to provide additional data that is specific to the failing subsystem or to the particular problem that you are experiencing. To understand what information is required and how to obtain it, consult both the basic instructions for preparing information for the IBM Support Center, and the diagnostic instructions for the failing PSSP subsystems in Part 2, “Diagnosing PSSP Subsystems” on page 89.

- When you encounter problems while operating the SP hardware or the PSSP software

  This manual contains diagnostic procedures provided for PSSP software problems and some SP hardware problems. This manual also contains descriptions of available error information, how to retrieve it, and what to look for to analyze the problem. There are also solutions to some commonly encountered problems.

- When preparing to troubleshoot the SP hardware or PSSP software

  This manual provides instructions for executing diagnostic procedures for PSSP software subsystems. These procedures are specific to the individual subsystem or SP hardware device being examined. Consult Part 2, “Diagnosing PSSP Subsystems” on page 89 for these instructions.

- To become familiar with existing services that monitor system status

PSSP provides graphical and command-line facilities to display the current status of system resources and to monitor changes in these resources. This manual introduces these tools and demonstrates how they can be used to assess the current status of the SP hardware and software. Using these utilities, you can detect problems at an early stage and react to them before they propagate and magnify.

Essential Documentation - Other Manuals to Accompany this Book

The “Bibliography” on page 273 lists manuals of general interest for the SP system and PSSP. This section lists those manuals specific to problem detection and problem solving for the SP system and for AIX. This section contains references to manuals not listed in the Bibliography.

1. For all levels of PSSP running on your SP system, *PSSP: Diagnosis and Messages Guide*

   These manuals are needed because SP systems allow different levels of PSSP to run on different nodes. Information for previous versions of PSSP and the hardware that they support may be present only in previous versions of this manual. This manual (PSSP 3.1.1 version) applies only to software currently in PSSP 3.1.1 and the hardware that PSSP 3.1.1 supports.

   When to use these books:

   - When a problem occurs on a node running a level of PSSP other than PSSP 3.1.1
   - When failures occur in hardware that was introduced on SP systems prior to the release of PSSP 3.1.1, and this hardware is not supported on PSSP 3.1.1

     The information for that hardware will be included in *PSSP: Diagnosis and Messages Guide* for versions of PSSP that support the hardware.
   - When a problem occurs in a PSSP subsystem that runs on several different nodes within the SP and those nodes are at different levels of PSSP, including PSSP 3.1.1.

     In most cases, the subsystems that run on nodes with different levels of PSSP use a "backward compatibility" mode. This means that all nodes provide only that level of function available in the lowest version of PSSP on those nodes.

     For example, if Group Services (GS) is running on an SP system with four nodes, configured as follows:

     – two nodes running PSSP 3.1.1

     – one node running PSSP 2.4

     – one node running PSSP 2.3

     GS will essentially run as if PSSP 2.3 were installed on all nodes. Any additional function provided by GS for PSSP 2.4. 3.1 and 3.1.1 will not be available.

     This manual discusses diagnostic procedures for software provided in PSSP 3.1.1. Diagnostic procedures for software provided in PSSP 3.1 are documented in the PSSP 3.1 version of *PSSP: Diagnosis Guide*. Diagnostic procedures for software provided in versions of PSSP prior to PSSP 3.1 are documented in previous versions of *PSSP: Diagnosis and Messages*

Guide. These diagnostic procedures may still be valid when software is executing in a "backward compatibility" mode on PSSP 3.1.1 nodes.

2. *PSSP: Messages Reference*

   This manual is needed because it is the companion to this one (*PSSP: Diagnosis Guide*). The *PSSP: Messages Reference* provides a list of the specific error messages generated by PSSP 3.1.1, gives a detailed explanation of the error condition, and gives directions for responding to the error condition.

   When to use this manual:

   - When PSSP 3.1.1 software encounters problems or failures accompanied by error messages, to understand the nature of the failure and how to respond to it.
   - When you encounter a specific error message and wish to resolve the problem without having to diagnose the entire subsystem that issued the message.

3. *PSSP: Administration Guide*

   This manual is needed because it describes the supported configurations of the SP system. This manual also describes how to configure and administer the SP system.

   This manual, *PSSP: Diagnosis Guide*, has many references to sections in *PSSP: Administration Guide* when describing how to analyze, circumvent, or repair problems.

   When to use this manual:

   - When installing, customizing, and configuring the SP system
   - When adding nodes and other computer resources to the SP system
   - When diagnosing potential SP hardware and PSSP software problems, to verify that the configuration and customization of your SP system is correct
   - When diagnosing potential SP hardware and PSSP software problems, to verify that the configuration and customization of your SP system is supported
   - When responding to specific SP hardware or PSSP software error conditions, and you need more information on the commands and procedures you are instructed to use.

4. *RS/6000 SP: PSSP 2.2 Survival Guide (SG24-4928)*

   Although this manual is specific to PSSP 2.2, much of the information in this manual is still relevant to later releases of PSSP. This manual gives you insight to the SP hardware and PSSP structure, so that you can understand how problems in one component impact others.

   This manual provides specific error avoidance and recovery instructions for items that still exist (or are supported) in PSSP 3.1.1, such as:

   - Manual conditioning of a high node
   - Initial SP system setup
   - Node installation and Network Install Manager (NIM)
   - Switch topology and system partitioning

- Network Time Protocol (NTP), time-of-day synchronization, and the problems that can happen with these components
- Tips on using Problem Management for monitoring the SP and responding to error conditions
- Tips on using the SP Error Log Management Facility

When to use this manual:

- When problems are encountered on nodes running PSSP 2.2
- When nodes are encountered in distributed subsystems where one of the nodes involved is running PSSP 2.2
- When looking for assistance in installing and customizing any version of PSSP
- When trying to diagnose and recover from system partitioning problems

5. *RS/6000 SP: Problem Determination Guide (SG24-4778)*

   Although this manual is specific to PSSP 2.1, much of the information in this manual is still relevant to later releases of PSSP. This manual provides specific error detection and recovery instructions for items that are still used and supported in PSSP 3.1.1, such as:

   - Control workstation installation
   - Kerberos security subsystem
   - System partitioning
   - Changing IP addresses or hostnames of the nodes

   When to use this manual:

   - When troubleshooting Kerberos-related problems
   - When troubleshooting system partitioning problems
   - To use as a guide when changing node IP addresses or network hostnames

6. *RS/6000 SP Monitoring: Keeping It Alive (SG24-4873)*

   Although this manual is specific to PSSP 2.2, much of the information in this manual is still relevant to later releases of PSSP. This manual is needed because it describes the utilities built into the PSSP software that permit you to monitor what is going on in your SP system, how to be alerted when things go wrong, and how to automate the response to specific conditions. Topics include:

   - RS/6000 Cluster Technology (which is called "HA Infrastructure" in *RS/6000 SP Monitoring: Keeping It Alive*)
   - SP Perspectives
   - Problem Management

   When to use this manual:

   - To introduce yourself to the problem monitoring facilities available in PSSP, and how they can be of use to you
   - To understand what resources can be monitored

- When learning how to automate your response to certain problems that can be handled without human interaction. An example of this is the process of expanding file system space when there is a shortage.

7. *AIX Version 4 Problem Solving Guide and Reference*

   This manual provides assistance in investigating and resolving AIX operating system problems. Consult this manual when you suspect a problem with the AIX operating system, or when you suspect that an AIX problem is contributing to an SP system problem.

8. *AIX Version 4 Messages Guide and Reference*

   This manual contains the list of 3-digit LED/LCD display values for SP nodes. This manual is used when a node reports a 3-digit LED/LCD display value, so that the reader can understand its meaning and appropriate action to take in response. SP hardware and PSSP software also issue LED/LCD display values, which are documented in Chapter 37, "SP-Specific LED/LCD Values" on page 257. If the LED/LCD cannot be located there, consult *AIX Version 4 Messages Guide and Reference*.

Preparing For Your First Problem Before It Happens

This section explains how to obtain and record information about your SP system that you will need when you first problem occurs. You may not have the time or the means to obtain this information after a failure has occurred. The best strategy is to prepare this information before a failure occurs, and to have it handy before investigating possible problems.

Knowing Your SP Structure and Setup

Problem investigation efforts are streamlined considerably by knowing the characteristics of the SP system at the time that a problem occurs. This includes what node types are being used, what software is installed on these nodes, what level of software is installed, what software service is installed, and so forth.

Create a Log of Your SP Structure and Setup

PSSP: Planning Volume 1 provides guidance in planning your physical site and selecting your hardware. *PSSP: Planning Volume 2* provides guidance for logically laying out the SP system structure and the SP administrative network, and selecting your software. Examine your SP system, its structure and its software setup, and record this information in a log. Keep this log in a place where you will always have access to it, regardless of whatever failure occurs on your system. To avoid the possibility of losing this log to an online failure, it is best to keep this log in hardcopy format.

The list below is the minimum amount of information that should be recorded in the log:

1. Your customer information:
   - Your access code, which is your customer number
   - Names and phone numbers of people whom the IBM Support Center should contact to assist you with problem resolution
2. Control workstation Information
   - The level of AIX installed

- The PTF numbers for all fixes installed on AIX
- The product number of PSSP on the control workstation. You need to use this to report a problem.
- The PTF numbers for all fixes installed on PSSP

3. System Partitioning or Cluster information - what nodes are in the system partition or cluster. Also, what software is installed and active in that cluster.
4. Node information
   - The node's number, its frame number and associated slot number, which can be found by issuing the **/usr/lpp/ssp/bin/splstdata -n** command (note: this is a system-partition sensitive command.)
   - Node's hostname and IP address, which can be found using the **/usr/lpp/ssp/bin/splstdata -a** command (note: this is a system partition sensitive command.)
   - The level of AIX installed on the node
   - The PTF numbers for all fixes installed to AIX
   - The product number of PSSP installed, which may be different from that installed on the control workstation
   - The PTF numbers for all fixes installed to PSSP
   - Optional software installed, including version numbers and fixes
   - Special hardware characteristics: wide node, network attached node, twin-tailed DASD

Update the Log Whenever a Failure Occurs

Whenever an actual or suspected failure occurs on the SP system, make an update to this log. Record symptoms that are noticed at the time of the failure, and system conditions such as:

- The date and time that the problem was discovered
- The nature of the problem, such as a node halt, an abnormal program termination, request hang, poor response time, or hardware failure
- What nodes the problem was experienced on
- What software was running on the nodes at the time the problem was encountered
- What users were using the system at the time that the problem was encountered
- What actions that you or others took to repair or bypass the problem, and whether these actions were successful

Recording this information serves several purposes:

- It allows you to recognize recurring problems and to quickly find the steps needed to resolve or bypass the problem.
- It records details about the conditions that existed in the SP system at the time of the failure. This information is essential when contacting the IBM Support Center to report problems.
- It allows you to detect patterns in the occurrence of problems.

Perhaps these problems occur on a regular basis, or whenever a specific program executes, or when a specific system resource is unavailable or reaching maximum limits. These patterns are difficult to detect unless historical data on past failures is available. Having this information available can assist you and the IBM Support Center in detecting patterns in these failure conditions.

Update the Log Whenever System Conditions Change

Using outdated or incomplete information when investigating a failure leads to wasted time. The wrong information is obtained and analyzed, the wrong diagnostic procedures are performed, and in some cases an incorrect solution is applied. This causes problem conditions to remain the same or become worse. It also may introduce additional problems. To avoid this wasted effort, be sure to update this log whenever the SP system structure or setup changes.

Update the log whenever the following occurs:

- New software is installed
- Upgrades are made to new levels of AIX or PSSP
- AIX or PSSP PTFs are installed
- Node hostname or IP addresses change
- Hardware changes, Switch adapter changes
- Problems occur

Making Effective Use Of the IBM Support Center

There are several things you need to know in order to make effective use of the IBM Support Center. You need to know when to call IBM, how to contact IBM, and what information to collect before calling.

When To Contact the IBM Support Center

Contact the IBM Support Center for the following situations:

- A repeated or persistent halt of an SP node
- A repeated or persistent hang of an SP node
- A repeated or persistent failure or hang of specific SP software

  These failures may not always occur on the same node, given the distributed nature of this software.
- A failure in mission-critical PSSP software

A single node or infrequent software failure that is not mission-critical may not be a cause to contact the IBM Support Center immediately. These problems may be caused by conditions that can be remedied through administrative techniques. Investigate these failures, using this manual as a guide for conducting the investigation. Follow these steps:

- Determine what was active on the system at the time.
- See who was using the system.
- Record the date and time of the failure.

- Determine what hardware was in use.
- Determine what specific services were being used at the time that the failure was detected
- Use the information in this manual and “Essential Documentation - Other Manuals to Accompany this Book” on page 4 to analyze and correct the problem.

Log information about these failures that you discover in the course of your investigations. This information can be used for your own future reference, and by the IBM Support Center if this failure becomes frequent enough or critical enough to require their assistance, as follows:

- It permits you to respond more quickly to similar failures in the future, and helps you to remember how to resolve the problem.
- It can be used for pattern analysis.

Problems and failures may appear to be unrelated at first, but they may have some relationship that is not immediately evident. Examine the conditions that were recorded for previous infrequent failures to see if there may be a pattern to them, even if the failure seem to be unrelated. Consider the following items when looking at the historical data on problems and failures:

- Do they happen when similar programs, procedures, or jobs are run?
- Do they happen when certain people or groups use the system?
- Do they happen at specific times, days, or shifts (peak or off-peak hours)?
- Does the failure occur when specific hardware is used?
- Are node reboots the only way to resolve the problem?

Contact the IBM Support Center when you discover any patterns in infrequent failures because:

- The system configuration may need repair.
- The IBM Support Center may have information about the problem you are experiencing.
- You may be experiencing a problem that no one else has ever encountered or reported.

Information To Collect Before Contacting the IBM Support Center

> Read this section **in its entirety** and perform **ANY** of the instructions listed here before placing a call to IBM. Some of the required information must be captured **immediately** before system conditions change, the data is lost, or the data is overwritten.

1. Your Customer Information, which should be in the log discussed earlier. See “Create a Log of Your SP Structure and Setup” on page 7.
   - Your access code, which is your customer number

- Names and phone numbers (external numbers, complete with area codes) where you can be reached by IBM service representatives

2. Your Product Information:
   - The PSSP product number for the level of PSSP running on the Control Workstation

     If PSSP 3.1.1 is running, the product number is **5765-D51**. If a level of PSSP other than PSSP 3.1.1 is running, the product number can be obtained from *PSSP: Diagnosis and Messages Guide* for that level.
   - The PTF numbers for all PSSP fixes installed on the control workstation
   - The version number of AIX running on the control workstation
   - The PTF numbers for all AIX fixes installed on the control workstation
   - For all nodes involved in the problem, obtain this information:
     – The version number of AIX running on the node
     – The PTF numbers for all AIX fixes installed on the node
     – The PSSP version number on the node
     – The PTF numbers for all PSSP fixes installed on the control workstation
3. Information about your problem. Different information is needed for different kinds of problems, so you cannot collect the same set of information for all problems. Here are some general rules:
   - For a single node halt or crash, you will need the following information:
     a. A system dump from the halted node. The system may already have created this dump for you, but you must verify this. Examine the halted node's LED indicator using the **spmon -Led** command. If the display shows a flashing **888**, a dump was started on the node. Use the **spmon -reset** command to step through the LED values until the flashing **888** appears again, recording all these LED values. Use Table 6 on page 85 to determine if the dump has completed and verify its contents.
     b. The **/unix** file from the halted node. This file will be obtained automatically by the service tools.
     c. The error log from the halted node. This log will also be created by the service tools.
     d. System Data Repository (SDR) information from the control workstation, which is used to find environment or configuration problems.
        1) Login the control workstation as **root**.
        2) Issue the following commands, redirecting the output to a file:
           a) **splstdata -e >** *filename*
           b) **splstdata -n -G >>** *filename*
           c) **splstdata -s -G >>** *filename*
        3) Write the file to the media, and label it as "SDR information".

     To obtain this information:

a. Verify the contents of the dump. Use “Action 2. Verify the System Dump” on page 84.

b. Ensure that the **/tmp** file system has at least 8 MB of free space.

c. Make sure that the **/unix** file is the one that was used when the dump occurred (in the case where the **bosboot -k** command was used to select another **/unix**).

d. Login as **root** and issue the **snap -r** command to clear the current contents of the **/tmp/ibmsupt** directory.

e. Make sure that a tape drive is accessible to the node.

f. Issue the **snap -Dgo** *tape device* command.

g. Label the media with the node name, node number, its contents (dump, **/unix**, snap information), and the command used to create it: **snap -Dgo** *tape device*.

h. Enable write-protect on the media and put in a safe place.

i. For additional information on this process, see *RS/6000 SP: Problem Determination Guide*.

- For multiple node halts or crashes, you will need the same information as in single node halts and crashes. Examine each system using the **spmon -Led** command to determine if a system dump was taken on each halted node.

  To obtain the necessary information from the halted nodes:

  a. With the node still in its crashed or halted state, make sure that a system dump has been taken. If a dump does exist, do not re-create it. If a dump does not exist, one needs to be created. DO NOT create a dump if the node's LED shows flashing **888**. Use “Action 1. Produce a System Dump” on page 83 to create the dump. Use the Primary Dump Device, unless this is impossible, because other tools will assume that the dump is located there.

  b. Verify the contents of the dump. Use “Action 2. Verify the System Dump” on page 84.

  c. Once the node is rebooted and the dump verified, ensure that the **/tmp** file system has at least 8 MB of free space available.

  d. Make sure that the **/unix** file is the one that was used when the dump occurred (in the case where the **bosboot -k** command was used to select another **/unix**).

  e. On the control workstation, ensure that the **/tmp** file system has at least 8 MB of free space for each node that has halted. For example, if four nodes have halted, make sure that 32 MB of space are available.

  f. On the control workstation, build a file containing the hostnames of the nodes that have halted or crashed. The hostnames should be on one line, separated by commas with no intervening white space characters. For example: node1a,node1b,node5d. Save the name of this file for use in the next two steps.

  g. On the control workstation, issue:

     ```
     splm -a service -t filename -r
     ```

to clear the current contents of the **/tmp/ibmsupt** directories on these nodes. *filename* is the name of the file from Step 3f.

h. On the control workstation, issue:

```
splm -a service -t filename -c -p Dg
```

to start the **snap -Dg** command on these nodes. *filename* is the name of the file from Step 3f on page 12.

i. Ensure that a tape drive is available on the control workstation.

j. On the control workstation, issue:

```
splm -a gather -k service -t \
filename –l /tmp/servcol -o tape_device_name
```

to retrieve the service information. The command retrieves the information from the nodes listed in the file, writes this information temporarily to the **/tmp/servcol** file, then archives the data to the tape device in **tar** format. *filename* is the name of the file from Step 3f on page 12. *tape_device_name* is the name of the tape drive.

k. If the **/tmp/servcol** file remains on the control workstation, remove it.

l. Label the media with the names of the nodes involved, their node numbers, the contents of the tape (system dumps, **/unix** files, **snap** information), and the command used to create the tape from Step 3j.

m. Enable the write protection on the media and put it in a safe place.

- Node hangs or experiences response problems

  a. A dump from the hung node is preferred. Go through the steps given above for manually creating and verifying a dump of the hung nodes.

  b. The **/unix** file is also needed, as in the above procedure.

  c. Reboot the node and run the **snap** command given in the previous section to collect the data and create the media.

- Failures in specific PSSP software subsystems, including denial of service problems and performance problems

  a. Consult Part 2, “Diagnosing PSSP Subsystems” on page 89 and find the diagnosis chapter for the failing PSSP subsystem. This chapter may specifically request that you collect certain information for the node, including information from remote nodes that do not seem involved in the problem.

  b. If error log information is needed from multiple nodes, the **splm** command can be used to consolidate these logs in one location.

  c. Do not generate a system dump unless the subsystem's diagnosis chapter instructs you to do so.

  d. Write all information requested to the media and clearly label it.

- Failure in other SP software which is supplied by IBM

  a. Consult the diagnosis documentation for the failing product. This information may specifically request that you collect certain information for the node, including information from remote nodes that do not seem involved in the problem

b. Do not generate a system dump unless the product's diagnosis instructions instruct you to do so.

c. Write all information requested to the media and clearly label it.

- Failure in non–IBM software

  a. Consult the diagnostic documentation for the failing product. This information may specifically request that you collect certain information for the node, including information from remote nodes that do not seem involved in the problem.

  b. Follow problem reporting procedures for that product.

- SP hardware failures

  a. Perform hardware diagnostic procedures associated with the hardware and record any information requested by these instructions. For details, see “RS/6000 SP Hardware Publications” on page 274.

How To Contact the IBM Support Center

In the United States:

The number for IBM software support is **1-800-237-5511**.
The number for IBM hardware support is **1-800-IBM-SERV**.

Outside the United States, contact your local IBM Service Center.

Contact the IBM Support Center using the phone number above, for these problems:

- Node halt or crash not related to a hardware failure
- Node hang or response problems
- Failure in specific PSSP software subsystems
- Failure in other SP software which is supplied by IBM

The person with whom you speak will ask for the information from “Information To Collect Before Contacting the IBM Support Center” on page 10 and give you a time period during which an IBM representative will return your call.

For failures in non-IBM software, follow the problem reporting procedures documented for that product.

For SP hardware failures, contact IBM Hardware Support at the number above.

For any problems reported to IBM Software Support, a Problem Management Record (PMR) is created. A PMR is an online software record used to keep track of software problems reported by customers.

- The IBM Support Center representative will create the PMR and give you its number.
- Have the SDR information you collected earlier handy because it may be needed for inclusion in the PMR.

- Record the PMR number. YOU WILL NEED IT to send data to the IBM Support Center. YOU WILL ALSO NEED IT on subsequent phone calls to the IBM Support Center to discuss this problem.
- Write the PMR number on ALL media you created in the previous steps, even if you are not going to send this data to the IBM Support Center at this time. The Support Center may request the data at a later time, so you want to ensure that neither the media nor the PMR number corresponding to it is lost.
- To send the media to the IBM Support Center, use this address:

IBM RS/6000 Scalable POWERparallel Systems
Dept. 39KA, M/S P961, Bldg. 415
522 South Road
Poughkeepsie, N.Y. 12601-5400

ATTN: APAR Processing

If you're using multiple packages or envelopes to send the media, be sure to label them in a series, such as "1 of 5", "2 of 5", and so forth.

Be sure that the person you identified as your contact can be reached at the phone number you provided in the PMR.

Chapter 2. Detecting SP Problems and Keeping Informed

The best way of streamlining your problem resolution is to prevent problems from occurring. To minimize the frequency and impact of problems, follow the configuration recommendations in *IBM RS/6000 SP: Planning, Volume 1, Hardware and Physical Environment* and *IBM RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, and use the tools documented in *PSSP: Administration Guide*. You should also follow the recommendations documented for any software you install.

However, problems may still occur. When they do, the best way to resolve these problems is to detect them as soon as they occur, and correct or bypass them before they impact the ability of other subsystems, causing secondary and tertiary failures. Several methods exist for detecting problems on the SP system.

The SP system provides the capability to detect problem situations in a runtime fashion when the administrator is actively monitoring system conditions. The SP system also has asynchronous notification methods for use when the administrator is not directly monitoring system conditions.

Runtime Notification Methods

PSSP provides tools to monitor system status and conditions in a runtime fashion, when the system administrator is actively monitoring the current status of the system. These tools are used when the administrator wants to know immediately the current status of system resources, or to be notified immediately of problems and potential trouble situations.

Two sets of runtime tools are available. The choice of the tools depends on the capabilities of the system administrator's terminal and the system administrator's preferences. PSSP provides **graphical tools** for use on the control workstation or network-attached terminals. PSSP also provides **command-line tools** for those situations when only modem or **s1term** access is available.

Graphical Tools - SP Perspectives

PSSP provides graphical tools for system administration and monitoring through the SP Perspectives tool suite. Perspectives is engineered for ease-of-use by the administrator, but in order to be used effectively, it requires graphics–capable terminals or workstations and high-speed connections. Use Perspectives when monitoring the SP from the control workstation or from a network-attached workstation.

The basic concepts of Perspectives and examples of its use are include in the SP Perspectives chapter of *PSSP: Administration Guide*. Perspectives also provides extensive online help information. To understand how to accomplish the tasks presented below, consult the Perspectives online help, using this section as a guide to the online help topics.

To use Perspectives, the **sysctld** daemon must be active on the node where the Perspectives suite is launched. The user must also have a Kerberos principal defined for that user in the system partition where Perspectives is launched. The user must also have write permission to the SDR.

Perspectives is launched with the **perspectives** command, which resides in the **/usr/lpp/ssp/bin** directory. Be sure to define the terminal's display by setting and exporting the **DISPLAY** environment variable, and make sure that the terminal will permit the remote host to create windows on the display with the **xhost** command.

Two Perspectives tools are useful for monitoring the system status and detecting problem situations:

The SP Event Perspective

This tool allows the user to specify system conditions that are of concern or importance, and to indicate what actions are to be taken when the condition exists. The Perspective interfaces with the Event Management software subsystem to monitor these conditions and alert the Perspective to the presence of the condition. To effectively use this Perspective you must understand certain terminology.

Condition. The circumstances within the system that are of interest to the system administrator. Conditions can be created, viewed, and modified through the **Conditions** pane in the SP Event Perspective. To specify a condition, the system administrator must provide the necessary components to form the condition, including an event expression and, optionally, a rearm expression. See their definitions below.

The rearm expression indicates when the SP Event Perspective should consider the event to have "stopped". For example, a file system is considered "almost full" when the available space is less than 10% of its capacity. The system administrator may want to consider the condition to exist until the available space reaches 13% of the file system's capacity. The event expression would then be set to 10% and the rearm expression to 13%. As with the event expression, the system administrator can indicate an action to take when the rearm expression occurs, such as deactivating reserve resources that had been activated when the event occurred.

Event Expression. A relational expression that specifies the circumstances under which an event is generated.

Rearm Expression. A relational expression that specifies that the condition that triggered the event is no longer true. It is usually the inverse of the event expression.

Event Definition. An association made by the system administrator between a condition and a response to the presence of that condition.

Registration. The activation of an event definition. By registering an event definition, the system administrator instructs the Perspective to begin monitoring for the condition and to take the associated action if the condition should occur.

Once the user registers the event definition, the action will be executed whenever the rearm expression occurs, whether or not the Event Perspective is active at the time the rearm event occurs.

Event. A change in the state of a system resource. For the purposes of this discussion, an event is more narrowly defined as the presence of the condition within the system.

To start the SP Event Perspective, double click on the **Event Perspective** icon in the main Perspectives launch pad window.

Users can create conditions for situations that are important to them through the **Conditions** pane of the SP Event Perspective. A number of default conditions have been provided through the SP Event Perspective, but you may wish to add more or to tailor the predefined conditions to meet the specific needs of your particular SP installation. The Perspectives online help provides assistance on how to create conditions and how to modify existing conditions. To access this help, click on the **Help** button from the SP Event Management Perspectives display, and select the **Tasks...** option. Assistance in handling conditions is available through the **Working with Conditions** topic.

Once a condition has been defined through the SP Event Perspective, an action can then be associate with it. The action may be as simple as a visual notification that the event has occurred, or the action can be more sophisticated, including automatically invoking a command in response to the event. To associate the appropriate action with the presence of the condition (or to the absence of the condition), an event definition must be created. You can create these definitions and examine default definitions through the **Event Definitions** pane of the SP Event Perspective. The Perspective online help provides assistance on how to create event definitions and how to modify existing definitions. To access this help, click on the **Help** button from the SP Event Management Perspective display and select the **Tasks...** option. Assistance in handling event definitions is available through the **Working with Event Definitions** topic.

Only after both the condition and its associated event definition have been defined to the Perspective can you begin the monitoring of the condition. This is done by registering the event definition through the **Event Definitions** pane in the SP Event Perspective. To find how this is done, consult the Perspective's **Working with Event Definitions** online help topic.

Other basic SP Event Perspective tasks are described in the online help. To access this information, click on the **Help** button from the SP Event Management Perspective display, select the **Tasks...** option, and click on the **How Do I ...?** topic.

Depending on how the event definition was constructed, the SP Event Perspective will react in one or more of the following ways when you register the event definition:

- The icon representing the event definition within the SP Event Perspective's Event Definitions pane changes to an envelope. This notification can only be detected if the SP Event Perspective is running. If the SP Event Perspective is shut down after you register for the condition, this visual notification is not presented.
- The action associated with the event definition is started. This action is specified when you create the event definition. Through this action, you can automate the response to the condition, such as sending e-mail to a system administrator, issuing a command to activate a pager, or issue an administrative command to allocate reserved resources to address the condition.

  Once you register the event definition, the action will be executed whenever the event occurs, whether or not the Event Perspective is active at the time the event occurs.

The actions for the event or the rearm event can be one of the following:

- A system command - This is an executable that is performed when the condition exists. The command can perform controls, enable or disable resources, or notify you by other means (like mail or online messages).
- An SNMP Trap - This transmits a notification to the network using SNMP protocols, indicating that an event has occurred. This trap can be configured so that certain SNMP applications can receive the notification, or all can receive it. NetView is an example of such an application. Use an SNMP trap when you are using SNMP-based monitoring tools (such as NetView), and you want these tools to detect when events occur on the SP.
- An entry in the AIX Error Log and the BSD System log - This is used to record a persistent record of the event, or the event's rearm condition. The AIX Error Log template **HA_PMAN_EVENT_ON** is used when the event condition occurs. The template **HA_PMAN_EVENT_OFF** is used when the rearm condition occurs. These templates can be viewed by issuing:

  ```
  errpt -at -J HA_PMAN_EVENT_ON -J HA_PMAN_EVENT_OFF
  ```

  Notification can be sent to the administrator whenever these templates are logged to the AIX Error Log. For instructions on setting up this notification, consult “Using the AIX Error Log Notification Facility” on page 72.

The SP Event Perspective is designed to be a multi-user tool. Multiple users can invoke the SP Event Perspective in parallel and monitor different conditions. Notifications are routed to those users that registered the associated event definition. The Perspective also stores any conditions and event definitions created by each user in the user's **$HOME/.$USER:Events** file. By storing these definitions in different files, the Perspective allows each user to tailor conditions and event definitions best suited to the user's needs. This also prevents users from accidentally modifying conditions or event definitions created or used by other Perspectives users.

The SP Hardware Perspective

This tool allows you to examine the current status of the SP system hardware. Through this tool, you can display a graphical representation of the system's overall structure, assess the current status of system hardware, and issue hardware control commands.

- **To examine the current status of a hardware device**, select the hardware device by single-clicking on the device's icon in the particular pane. Looking at the top row of icons, a notebook icon should now appear on the left. Open the notebook by single-clicking on the notebook icon. This creates a new display that contains the device's current status, settings, and monitored conditions. This is useful for examining a node's LED values, its responsiveness to the network and the switch, its network configuration, and other information. For further assistance in using the notebook to view hardware status, consult the Perspective's online help. To access this help, click on the **Help** button from the SP Hardware Perspective display and select the **Tasks...** option. Assistance in viewing hardware status is available in the **Viewing Hardware Attributes** topic.

  If you want to view the same hardware information from multiple entities, such as the responsiveness to the switch for a series of nodes, opening a notebook for each entity can be time-consuming. The SP Hardware Perspective offers an alternative method for obtaining this information. Select the multiple objects

from the same pane. The notebook icon is no longer available, but another icon towards the right remains available. The icon shows both a table and an icon. When you point at this icon, the descriptive text reads "Show objects in the table view or the icon view". By clicking on this icon, the pane changes from the icon representation to a table representation. Immediately, the SP Hardware Perspective presents a new display with a selection list, requesting those hardware characteristics that you wish to view. After selecting the characteristics, the table is populated with the current status for each characteristic from each hardware entity. The table entities are color coded to indicate "good", "bad", and "caution" status.

For further assistance on using the table view to examine hardware status, consult the Perspective online help. To access this help, click on the **Help** button from the SP Hardware Perspective display and select the **Overview..** option. From the new window that appears, select the **Starting and Customizing SP Perspective** topic, then select the **Customizing SP Perspective** subtopic, and finally select the **Using Table View** item. The Help option from the selection list window provides a fast path to the help topic.

- **To monitor the status of hardware devices**, select the pane where the devices are contained. Looking at the top row of icons, a graph icon should be visible on the right. When you point at this icon, the descriptive text reads "Set up and begin monitoring". Click on this icon to bring up a window of items that can be monitored. Select the items to be monitored from this list. All objects in the pane will now be monitored for these conditions.

  When monitoring is active, the icons of the entities use a visual indication of the status. If all conditions being monitored are not indicating any problems, the icons will be presented in a green color. If any of the conditions being monitored indicate a problem, the icon will appear to have a red X drawn through it. Note that this will occur even if only one condition indicates a problem. For example, if five nodes are being monitored for five conditions, and one of these five conditions appears on node1a, the icon for node1a will appear with a red X through it, while the remaining nodes will be represented with green icons.

  To determine what condition may exist on a marked entity, select the object by single-clicking on its icon or its table entry in the pane. Then open the object's notebook by clicking on the notebook icon in the upper left corner. When the notebook display comes up, page forward to the "Monitored Conditions" page. This page lists conditions being monitored for that object, along with the condition's current state. Any state listed as "Triggered" indicates that the condition is present.

  If any object in a pane is presented in gray with a question mark (**?**) drawn over it for longer than a few seconds, a communication problem exists between the SP Hardware Perspective and the Event Management software subsystem. For assistance in resolving the problem, consult Chapter 15, “Diagnosing SP Perspectives Problems” on page 151.

  For further assistance in setting up and starting the hardware monitor, consult the Perspectives online help facility. To access this help, click on the **Help** button from the SP Hardware Perspective display and select the **Tasks...** option. From the new window that appears, select the **Monitoring Hardware Objects** topic.

There are some characteristics of the SP Hardware Perspective that the user should keep in mind when using the tool. Unlike the SP Event Perspective, the SP Hardware Perspective does not permit the user to associate an action with the presence of a condition. For users that wish to automate a response to a specific system condition, the SP Event Perspective should be used. Also, the SP Hardware Perspective only monitors conditions while it is active. If the Perspective is shut down, any monitoring of hardware status is also shut down. Finally, the SP Hardware Perspective does not remember any prior monitor setting when it is restarted. The user must reissue the monitoring command to begin monitoring hardware status when the Perspective is restarted.

Previous versions of PSSP offered a graphical user interface as part of the System Monitor (**spmon**) command. PSSP Version 3.1 has incorporated this hardware control capability into the SP Hardware Perspective. While the capability of the **spmon** command is available through the Perspective, the "look and feel" of the control is somewhat different. The SP Hardware Perspective offers a special online help facility to acclimate former **spmon** graphical interface users to the new controls. To access this help, click on the **Help** menu bar item in the SP Hardware Perspective, select the **Tasks...** option, and then select the **Transforming System Monitor Experience into Hardware Perspectives Skills** topic from the help menu.

Each Perspective provides its own unique capabilities. For the purposes of problem monitoring and determination, this manual recommends that the SP Event Perspective be used to monitor conditions of interest for the SP system. When the SP Event Perspective indicates that a hardware failure condition exists, the SP Hardware Perspective should be used to examine the current status of the system hardware and obtain more detailed information about the hardware problem.

Command Line Tools

PSSP provides command-oriented tools for system administration in addition to graphical tools for system administration and monitoring. These tools require no special terminal capability or high-speed connection, making them usable by almost any terminal type in any mode of access. Use these tools when examining system status through a modem connection or through a node's S1 serial port. The tools discussed in this section are documented in greater detail in *PSSP: Command and Technical Reference*, *PSSP: Administration Guide*, and *AIX Version 4 Commands Reference*. The primary drawback to these tools is that they do not possess the same ease-of-use characteristics as their Perspectives based counterparts, although they do provide the same basic function.

Several commands are useful for monitoring the system status and detecting problem situations:

- spmon
- hmmon
- df
- lsps
- lssrc

The **spmon** command permits the user to control and monitor SP hardware resources through a command-line interface without requiring a graphics-capable terminal or high-speed connection. The **spmon** command does not provide the

capability to examine software status (such as paging space, file system space, or software subsystem activity). The **spmon** command provides access to more node-specific information than the **hmmon** command, which is introduced next. The **spmon** command provides a predefined system query to check the most basic problem conditions within the SP system.

The **hmmon** command provides hardware monitoring functions similar to the **spmon** command, and gives you access to more SP hardware information for frames and switches than the **spmon** command does. The **hmmon** command provides the capability to monitor frame and switch status as well as node status. The **hmmon** command is intended as a general-purpose SP hardware monitor. Although it has access to more SP information than the **spmon** command, it does not have access to some of the node-specific information that the **spmon** command does. The **hmmon** command does not provide a predefined system query, which the **spmon** command does.

The **df** command is an AIX command that examines the current status of file systems, such as current file system size and current available space within these file systems. While this command is designed to examine the AIX system on which it is executed, it can be invoked remotely with the **dsh** and **rsh** commands to acquire this information for all nodes. Three file systems are of particular importance for all SP nodes:

- **/spdata**

  This directory contains configuration information for PSSP software and also contains copies of information from the SDR. By default, this directory resides in the **/** (or **root**) file system. Insufficient space in this file system can result in failures in PSSP software, especially those dependent on the SDR for proper execution. As a rule of thumb, ensure that this file system has at least 5% of its capacity available at any time.

  One method for avoiding space problems for the **/spdata** directory is to create a separate file system for this directory. Follow the same rule of thumb for spotting potential trouble with this file system.

- **/var**

  This file system contains AIX system logs, such as the error log and user access logs. It also contains logs maintained by PSSP software for serviceability purposes. Some of these logs are never cleared except by explicit administrator actions. If left unattended, they can grow to consume all available space. As a rule of thumb, ensure that 10 MB of space is available within this file system at all times. If the file system reaches this threshold, consider either extending the file system's capacity with the **chfs** command, or examine the file system to determine where the space is being consumed and remove unneeded files.

  If **/var** is continually reaching the suggested threshold, this condition may indicate a chronic problem with some PSSP software or with specific hardware devices. Examine the logs listed in Chapter 4, “Error Logging Overview” on page 69 to determine if any show increased or extended activity, and perform any associated problem determination procedures if necessary.

- **/tmp**

  This file system is used by various user level applications, software products, and PSSP programs for temporary storage. Some legacy PSSP applications

use this file system to store trace logs used for serviceability purposes. Some applications may inadvertently leave temporary files in the **/tmp** file system, or these applications may terminate before they can remove these files. Insufficient space in **/tmp** can cause PSSP software to fail. As a rule of thumb, ensure that at least 8 MB of space is available in this file system at any time. Eight MB is the amount of space a **snap** command will require if the system has to produce a dump to be sent to the IBM Support Center.

These space capacities can be verified using the **dsh** or **rsh** command to invoke the **df** command on all nodes in the SP system.

The **lsps** command provides an instant assessment of the currently available paging space for an AIX system. As with the **df** command, the **lsps** command provides information for the AIX system on which it executes. Using the **lsps** command with the **dsh** or **rsh** command, you can obtain the assessment for all nodes in the SP system. Paging space availability by itself does not necessarily indicate a problem. Having only ten percent of 2 gigabytes of paging space available is not as significant a condition as having only ten percent of 100 MB available. Also, one system 's critical situation may be a tolerable situation for another system. Because of this discrepancy, this manual will not suggest any default figure for a critical paging space situation. Use your knowledge of the system setup, system workload, and any past paging space problems to determine this value.

The **lssrc** command provides information for software services currently installed on an AIX system. Using **lssrc**, you can determine if a software service is active or inactive. Use this command in cases where a software service does not appear to be responding to requests for service on a specific node. To check software service status on multiple nodes, use this command through the **dsh** or **rsh** commands.

The following scenarios demonstrate how these tools are used to query and monitor the status of the SP system.

Assessing the Current Status of the SP System

This task can be accomplished through the following series of steps:

1. **Preparing to Perform the System Check.** Prepare for this task by retrieving the log of the SP system structure. This log is discussed in “Create a Log of Your SP Structure and Setup” on page 7. This information is required to use the **hmmon** command effectively. The **hmmon** command obtains hardware information about nodes and switch devices using the frame number and slot number of the device, not the network name or IP address assigned to the device.

   This check should be performed by users authorized to invoke the **spmon** and **dsh** commands. By default, the **root** user has this authorization. To learn how an alternative user can acquire this authorization, see the section entitled "Using the SP System Monitor" in *PSSP: Administration Guide*.

2. **Perform a Preliminary Check of the SP System.** To perform a basic diagnostic check of the entire SP system, issue the following command from the control workstation:

   ```
   /usr/lpp/ssp/bin/spmon -G -d | more
   ```

   This test verifies several items in the monitor program itself to make sure it is running. Once the monitor verification completes, the **spmon** command checks

the status of the SP frames and obtains information about the SP nodes. The **spmon** command performs these tests in a dependent order, so that if one of the early checks fails, subsequent checks are not performed. For example, if a frame cannot be queried, the frame and the nodes within that frame are not checked.

Example output from the **spmon -G -d** command:

```
1.  Checking server process
Process 10512 has accumulated 192 minutes and 53 seconds.
Check ok

2.  Opening connection to server
Connection opened
Check ok

3.  Querying frames (s)
1 frames (s)
Check ok

4.  Checking frames

       Controller  Slot 17  Switch    Switch    Power supplies
Frame   Responds   Switch   Power    Clocking   A   B   C   D
-------------------------------------------------------------
   1       yes       yes      on        0       on  on  on  on

5.  Checking nodes
------------------------------ Frame 1 ----------------------
Frame  Node  Node         Host/Switch   Key   Env   Front Panel   LCD/LED is
Slot  Number Type  Power   Responds    Switch Fail    LCD/LED     Flashing
  1     1    wide   on     yes  yes    normal  no  LEDs are blank   no
  3     3    thin   on     yes  yes    normal  no  LEDs are blank   no
  4     4    thin   on     yes  yes    normal  no  LEDs are blank   no
  5     5    thin   on     yes  yes    normal  no  LEDs are blank   no
  6     6    thin   on     yes  yes    normal  no  LEDs are blank   no
  7     7    wide   on     yes  yes    normal  no  LEDs are blank   no
  9     9    wide   on     yes  yes    normal  no  LEDs are blank   no
  11    11   wide   on     yes  yes    normal  no  LEDs are blank   no
  13    13   wide   on     yes  yes    N/A     no  LCDs are blank   no
```

Note that these tests are numbered. This makes it easy to detect if a test was omitted. The results of this command indicate potential problems if any of these conditions exist:

- The command does not execute.
- The command does not perform all five verification checks.
- The fourth test indicates that the frame's controller is not responding, the switch power is not on, or any of the power supplies are listed as off.
- The fifth test indicates any abnormal conditions: a node's power is off, the **host responds** does not read **yes**, an environment failure is indicated, or the LCD or LED of the node is not blank (but not flashing).
- The fifth test indicates that the node's LCDs or LEDs are flashing.

  This indicates that a system dump was attempted.
- The fifth test indicates that the node is not responding to the switch device.

3. **Obtaining More Information**. If the **spmon** command mentioned previously indicates a potential problem situation, obtain more information in order to resolve the problem.

- If either the first or second test of **spmon -G -d** failed, consult Chapter 14, “Diagnosing System Monitor Problems” on page 147.
- If the third or fourth test failed, use the **hmmon** command to detect if there are problems with the frame itself. Issue this command to obtain this information:

  ```
  hmmon -G -q -s -v frPowerOff*,controllerResponds,\
  controllerIDMismatch,nodefail* range_of_frame_nums:0
  ```

  Example output from the above command:

```
1 0 nodefail1            FALSE   0x8802  node 01 I2C not responding
1 0 nodefail2            TRUE    0x8803  node 02 I2C not responding
1 0 nodefail3            FALSE   0x8804  node 03 I2C not responding
1 0 nodefail4            TRUE    0x8805  node 04 I2C not responding
1 0 nodefail5            FALSE   0x8806  node 05 I2C not responding
1 0 nodefail6            FALSE   0x8807  node 06 I2C not responding
1 0 nodefail7            FALSE   0x8808  node 07 I2C not responding
1 0 nodefail8            FALSE   0x8809  node 08 I2C not responding
1 0 nodefail9            FALSE   0x880a  node 09 I2C not responding
1 0 nodefail10           FALSE   0x880b  node 10 I2C not responding
1 0 nodefail11           FALSE   0x880c  node 11 I2C not responding
1 0 nodefail12           FALSE   0x880d  node 12 I2C not responding
1 0 nodefail13           FALSE   0x880e  node 13 I2C not responding
1 0 nodefail14           TRUE    0x880f  node 14 I2C not responding
1 0 nodefail15           FALSE   0x8810  node 15 I2C not responding
1 0 nodefail16           TRUE    0x8811  node 16 I2C not responding
1 0 nodefail17           FALSE   0x8812  switch  I2C not responding
1 0 frPowerOff           FALSE   0x8846  SEPBU   frame power off
1 0 controllerIDMismatch FALSE   0x8871  frame ID mismatch
1 0 controllerResponds   TRUE    0x88a8  frame responding to polls
```

  This command tests if any of the frame's power supplies are off, if the frame controller is experiencing problems, or any of the node slot connections are bad. Keep in mind the warning made earlier, since wide and high nodes occupy more than one node slot in a frame, node failures will be detected for node slots that cannot be used because a wide or high node occupies that space.

  Such a situation is demonstrated in the example output listed above. In this example, the nodes occupying slots 1 and 3 are wide nodes, as are the nodes occupying slots 13 and 15. Node slots 2, 4, 14, and 16 are therefore unusable, but the **hmmon** command indicates that nodes in these unavailable slots have failed. The log of the SP structure and setup is needed to understand which slots are "supposed" to indicate node failures, and which slots are not.

  Check for any of these conditions in the **hmmon** command output:

  – controllerResponds reads FALSE

  – controllerIDMismatch reads TRUE

  – nodefail17 reads TRUE (indicating a failure in the SP switch)

  – Other nodefails show TRUE, and these node failure cannot be attributed to a wide or high node occupying that slot

  If a controller ID mismatch is shown, consult the Managing a HACWS Configuration chapter in *PSSP: Administration Guide*. For controller

responsiveness problems, perform hardware diagnostics on the frame controller. For nodefail17 failures, perform hardware diagnostics on the switch device. For other node failures, perform hardware diagnostics on the node occupying that slot.

- If the fourth test of the **spmon -G -d** command indicated that the switch power was off, issue the following **hmmon** command to determine if this was caused by a hardware condition:

  ```
  hmmon -G -Q -s -v nodePower,powerLED,envLED,shutdownTemp frame_num:17
  ```

  Example output of the **hmmon** command, showing switch information for frame 1

```
1  17  powerLED                  1  0x8c47  node/switch LED 1 (green)
1  17  envLED                    0  0x8c48  node/switch LED 2 (yellow)
1  17  nodePower           TRUE     0x8c4a  DC-DC power on
1  17  shutdownTemp         FALSE    0x8c59  temperature shutdown
```

  This **hmmon** command will indicate if the switch has power, if power is available for the switch, if the switch's power was shut down automatically, and if the switch power was shut down due to high temperature. If the switch cannot obtain power, verify that the switch is correctly cabled to its power source. For other conditions, perform hardware diagnostics on the switch device.

- If the fifth test of the **spmon -G -d** command indicates that a node does not have power, and the node's power was not shut off manually, issue the following **hmmon** command to determine if the power was disabled because of a hardware condition:

  ```
  hmmon -G -Q -s -v nodePower,powerLED,envLED frame_num:node_num
  ```

  Example output of the hmmon command for a single node in a single frame

```
1   1  nodePower           TRUE     0x904a  DC-DC power on
1   1  powerLED                  1  0x9047  node/switch LED 1 (green)
1   1  envLED                    0  0x9048  node/switch LED 2 (yellow)
```

  This **hmmon** command will indicate if the node has power, if power is available for the node, and if the node's power was shut down automatically. If the node cannot get power, verify that the node is correctly cabled to its power source. For other conditions, perform hardware diagnostics on the node.

- If the fifth test of the **spmon -G d** command indicates that a node's LED/LCD display is not blank, a hardware or operating system error has occurred. The LED/LCD code contains important failure information. When

this condition exists, examine the node's LED/LCD value and record the value displayed. Use the following command to examine this value:

```
/usr/lpp/ssp/bin/spmon -L framenumber/nodenumber
```

To determine the explanation and action for the error, look up this code in Chapter 37, “SP-Specific LED/LCD Values” on page 257. If a three-digit LED/LCD code is not listed in this table, the code is generated by the AIX operating system; consult *AIX Version 4 Messages Guide and Reference*.

- If the fifth test of the **spmon -G -d** command indicated that the node's LED/LCD value was flashing, and the **spmon -L** command in the previous bullet indicates that the LED/LCD value is **888**, a system dump was initiated on this node. The flashing **888** LED/LCD value indicates that a series of values are stored in the LED/LCD display.

  Step through this list of codes and record each value shown using the following sequence of steps:

  a. Issue the command

     ```
     /usr/lpp/ssp/bin/spmon -reset -t framenumber/nodenumber
     ```

     to step to the next stored LED/LCD value.

  b. Issue the command

     ```
     /usr/lpp/ssp/bin/spmon -L framenumber/nodenumber
     ```

     to retrieve the new LED/LCD value.

  c. Record this LED/LCD value

  Repeat these steps until the **spmon -L** command displays a value of **888** again. Retain this list of codes; they will be required by IBM Support Center. To determine the explanation and action for these error codes, look up the codes in Chapter 37, “SP-Specific LED/LCD Values” on page 257. If a three-digit LED/LCD code is not listed in this table, the code is generated by the AIX operating system; consult *AIX Version 4 Messages Guide and Reference*. Finally, save and verify the system dump, following the instructions provided in Chapter 5, “Producing a System Dump” on page 83.

4. **Checking Basic Software Information.** Once hardware failures have been eliminated, it is time to perform some basic software verifications for the SP system. These checks will use the **dsh** command to invoke AIX commands on multiple nodes in parallel. To verify this, issue the following command from the control workstation:

   ```
   dsh -a -f32 hostname
   ```

   Example output of the **dsh -a -f32 hostname** command on a small SP system configuration:

   ```
   k21n01.ppd.pok.ibm.com: k21n01.ppd.pok.ibm.com
   k21n03.ppd.pok.ibm.com: k21n03.ppd.pok.ibm.com
   k21n04.ppd.pok.ibm.com: k21n04.ppd.pok.ibm.com
   k21n05.ppd.pok.ibm.com: k21n05.ppd.pok.ibm.com
   k21n06.ppd.pok.ibm.com: k21n06.ppd.pok.ibm.com
   k21n07.ppd.pok.ibm.com: k21n07.ppd.pok.ibm.com
   k21n09.ppd.pok.ibm.com: k21n09.ppd.pok.ibm.com
   k21n11.ppd.pok.ibm.com: k21n11.ppd.pok.ibm.com
   k21n13.ppd.pok.ibm.com: k21n13.ppd.pok.ibm.com
   ```

This test will verify that the **dsh** command can reach the nodes within the SP system. Only nodes that were previously detected as being offline in the earlier tests should fail to respond to this command. If any other nodes within the SP system fail to respond, check for Kerberos problems by referring to Chapter 10, “Diagnosing Remote Command Problems on the SP System” on page 105.

- Check the paging space consumed on all nodes by using the **lsps** command:

  ```
  dsh -av -f32 lsps -s | more
  ```

  Example output of the **dsh -av -f32 lsps -s** command on a small SP system configuration:

```
k21n01.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n01.ppd.pok.ibm.com:       768MB              8%
k21n03.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n03.ppd.pok.ibm.com:       768MB             17%
k21n04.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n04.ppd.pok.ibm.com:       768MB              8%
k21n05.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n05.ppd.pok.ibm.com:       768MB             13%
k21n06.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n06.ppd.pok.ibm.com:       768MB             12%
k21n07.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n07.ppd.pok.ibm.com:       768MB             11%
k21n09.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n09.ppd.pok.ibm.com:       768MB              9%
k21n11.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n11.ppd.pok.ibm.com:       768MB              9%
k21n13.ppd.pok.ibm.com: Total Paging Space   Percent Used
k21n13.ppd.pok.ibm.com:       768MB             15%
```

  Lack of available paging space can lead to thrashing conditions on a node. If these nodes are executing parallel applications, the entire application will be slowed to the rate of the slowest responding node. The extent to which low paging space and thrashing can be tolerated differs from one customer environment to the next, but as a general rule of thumb, investigate any nodes indicating that 80% or more of its paging space is currently in use.

- Check for file systems that are close to their capacity, concentrating on the file systems mentioned earlier in this section, by using the **df** command:

  ```
  dsh -av -f32 df /spdata /var /tmp | more
  ```

  Example output from the **dsh -av -f32 df /spdata /var /tmp** command on a small SP system configuration:

```
k21n01: Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
k21n01: /dev/hd4           32768       432   99%     1403    18%  /
k21n01: /dev/hd9var       147456     45480   70%      610     4%  /var
k21n01: /dev/hd3           98304     38632   61%       85     1%  /tmp
k21n03: Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
k21n03: /dev/hd4           32768     16960   49%     1431    18%  /
k21n03: /dev/hd9var       704512     99968   86%      595     1%  /var
k21n03: /dev/hd3           98304     50424   49%      278     3%  /tmp
k21n04: Filesystem    512-blocks      Free %Used    Iused %Iused Mounted on
k21n04: /dev/hd4           32768     16584   50%     1512    19%  /
k21n04: /dev/hd9var       147456    107312   28%      644     4%  /var
k21n04: /dev/hd3           98304     91232    8%       74     1%  /tmp
```

An obvious warning sign is if any of these file systems should appear to be more than 90% utilized. If any file systems appear over 90% utilized, examine the file systems for large files that can be removed or compressed, or consider extending the file system size. Attempt to keep 10 MB available in the **/var** file system and 8 MB available in the **/tmp** file system, to ensure that PSSP software and service software function correctly.

Keeping Informed of Status Changes

The previous discussion centered on obtaining the current status of SP system hardware and software. Such efforts are necessary if a problem is suspected and being actively investigated, but repeatedly issuing these commands periodically to examine the current status of the SP system can become tedious. To make the task of monitoring system status easier, PSSP provides monitoring capabilities within the **hmmon** and **spmon** commands as well. This avoids the necessity of reissuing the previously discussed commands over and over again to keep informed of the system status. This section describes some of the more common monitor commands.

To set up a monitor to check for frame hardware failures, issue the following background command:

```
hmmon -G -q -s -v frPowerOff*,controllerResponds,controllerIDMismatch,\
nodefail* range_of_frame_nums:0 &
```

Example initial output from the **hmmon** command:

```
1   0  nodefail1            FALSE    0x8802  node 01 I2C not responding
1   0  nodefail2            TRUE     0x8803  node 02 I2C not responding
1   0  nodefail3            FALSE    0x8804  node 03 I2C not responding
1   0  nodefail4            FALSE    0x8805  node 04 I2C not responding
1   0  nodefail5            FALSE    0x8806  node 05 I2C not responding
1   0  nodefail6            FALSE    0x8807  node 06 I2C not responding
1   0  nodefail7            FALSE    0x8808  node 07 I2C not responding
1   0  nodefail8            TRUE     0x8809  node 08 I2C not responding
1   0  nodefail9            FALSE    0x880a  node 09 I2C not responding
1   0  nodefail10           TRUE     0x880b  node 10 I2C not responding
1   0  nodefail11           FALSE    0x880c  node 11 I2C not responding
1   0  nodefail12           TRUE     0x880d  node 12 I2C not responding
1   0  nodefail13           FALSE    0x880e  node 13 I2C not responding
1   0  nodefail14           TRUE     0x880f  node 14 I2C not responding
1   0  nodefail15           TRUE     0x8810  node 15 I2C not responding
1   0  nodefail16           TRUE     0x8811  node 16 I2C not responding
1   0  nodefail17           FALSE    0x8812  switch I2C not responding
1   0  frPowerOff           FALSE    0x8846  SEPBU frame power off
1   0  controllerIDMismatch FALSE    0x8871  frame ID mismatch
1   0  controllerResponds   TRUE     0x88a8  frame responding to polls
```

This is a similar command to one presented previously, except this version continually monitors the frame condition and generates a message to the terminal if any of the status should change. To stop monitoring this information, terminate the background process.

To set up a monitor to check for SP switch hardware status changes, issue the following background command:

```
hmmon -G -q -s -v nodePower,powerLED,envLED,\
shutdownTemp range_of_frame_nums:17 &
```

Example initial output from the **hmmon** command:

```
1  17  powerLED                1  0x8c47  node/switch LED 1 (green)
1  17  envLED                 0  0x8c48  node/switch LED 2 (yellow)
1  17  nodePower          TRUE     0x8c4a  DC-DC power on
1  17  shutdownTemp       FALSE    0x8c59  temperature shutdown
```

This is a similar command to one presented previously, except this version continually monitors the frame condition and generates a message to the terminal if any of the status should change. To stop monitoring this information, terminate the background process.

To set up a monitor to check for changes in a node's LCD or LED status, issue the following background command:

```
hmmon -G -q -s -v LED7Seg* range_of_frame_nums:1-16 &
```

Example initial output from the **hmmon** command:

```
 1   1  LED7SegA                255  0x909f  7 segment LED A
 1   1  LED7SegB                255  0x90a0  7 segment LED B
 1   1  LED7SegC                255  0x90a1  7 segment LED C
 1   3  LED7SegA                255  0x949f  7 segment LED A
 1   3  LED7SegB                255  0x94a0  7 segment LED B
 1   3  LED7SegC                255  0x94a1  7 segment LED C
 1   4  LED7SegA                255  0x949f  7 segment LED A
 1   4  LED7SegB                255  0x94a0  7 segment LED B
 1   4  LED7SegC                255  0x94a1  7 segment LED C
 1   5  LED7SegA                255  0x949f  7 segment LED A
 1   5  LED7SegB                255  0x94a0  7 segment LED B
 1   5  LED7SegC                255  0x94a1  7 segment LED C
 1   6  LED7SegA                255  0x949f  7 segment LED A
 1   6  LED7SegB                255  0x94a0  7 segment LED B
 1   6  LED7SegC                255  0x94a1  7 segment LED C
 1   7  LED7SegA                255  0x909f  7 segment LED A
 1   7  LED7SegB                255  0x90a0  7 segment LED B
 1   7  LED7SegC                255  0x90a1  7 segment LED C
 1   9  LED7SegA                255  0x909f  7 segment LED A
 1   9  LED7SegB                255  0x90a0  7 segment LED B
 1   9  LED7SegC                255  0x90a1  7 segment LED C
 1  11  LED7SegA                255  0x909f  7 segment LED A
 1  11  LED7SegB                255  0x90a0  7 segment LED B
 1  11  LED7SegC                255  0x90a1  7 segment LED C
```

This command shows the initial status of these resources, and displays any status changes in these resources when they occur. All values should display a value of **255**, indicating that the associated readout element is blank. If any nodes indicate that a segment is not blank, issue the **spmon -L** command mentioned previously to obtain the current LCD or LED readout of the node.

To set up a monitor to check for nodes suddenly losing contact with the SP Switch, issue the following command:

```
spmon -q -M -l -t frame*/node*/switchResponds/value
```

Example initial output from the **spmon** command:

```
/SP/frame/frame1/node1/switchResponds/value/1
/SP/frame/frame1/node3/switchResponds/value/1
/SP/frame/frame1/node4/switchResponds/value/1
/SP/frame/frame1/node5/switchResponds/value/1
/SP/frame/frame1/node6/switchResponds/value/1
/SP/frame/frame1/node7/switchResponds/value/1
/SP/frame/frame1/node9/switchResponds/value/0
/SP/frame/frame1/node11/switchResponds/value/0
/SP/frame/frame1/node13/switchResponds/value/1
```

The **spmon** command also displays the current status, and a message to the terminal if any of these values change. All values should be 1. A value of 0 indicates that the node is not responding to the SP Switch. Note that this is the case with two of the nodes in this example, and these nodes should be investigated.

Other conditions can also be monitored using the **hmmon** and **spmon** commands; these suggestions offer the most basic of tests. To learn what other conditions can

be monitored with these commands and to tailor these commands to best suit your needs, refer to the **hmmon** and **spmon** sections of *PSSP: Command and Technical Reference*.

All commands can be executed from the same terminal session, but this can lead to confusing output when conditions change, or initial values can scroll off the terminal screen. To keep the monitoring manageable, consider issuing these commands from separate terminals, or from separate terminal windows from a XWindows capable terminal. Issue one monitoring command per terminal or terminal window. This will associate a terminal with each condition being monitored, and simplify the understanding of the monitor output.

Asynchronous (Batch) Notification Methods

As system administrator, you cannot always devote your entire attention to monitoring the current status of a system, trying to detect problem conditions before they occur. For even moderately sized SP configurations, this task can be time consuming and tedious. Other tasks require your attention, so actively monitoring the SP system for potential problem indications cannot become a task that consumes all your time and effort.

Fortunately, PSSP provides tools to monitor system status and conditions on your behalf, as well as the tools discussed previously to assess the current status of the system. Using these tools, you can indicate conditions of particular interest, request asynchronous notification of these events, and cause actions to be initiated when these conditions occur. In essence, the SP system monitors itself, takes action itself, and notifies you of the condition after it occurs. These monitoring tools can be used when you are not immediately available, such as during off-peak hours, or can be used to remove most of your monitoring burden.

Two sets of monitoring tools are available. As with the runtime notification tools mentioned previously, the choice of tool depends on the capabilities of your terminal and your preferences. PSSP provides **graphical monitoring tools** for use on the control workstation or a network attached terminals, and also provides **command-line monitoring tools** for those situations where only modem access or **s1term** access is available.

Graphics Tools - SP Event Perspective

This tool was introduced in “The SP Event Perspective” on page 18. This Perspective is designed for ease-of-use for the system administrator, but requires graphics-capable terminals or workstations, connected through high-speed network connections to be used effectively. Use this Perspective when setting up asynchronous notifications from the control workstation or a network-attached terminal. The basic concepts of the Perspectives tool suite and examples of its use are included in the chapter on SP Perspectives of *PSSP: Administration Guide*. As mentioned in the previous discussion, the SP Event Perspective provides extensive online help information to assist you in accomplishing Perspectives tasks. Use this source as a companion guide to this section when using the SP Event Perspective to set up asynchronous notification of conditions.

To use the SP Event Perspective, the **sysctld** daemon must be active on the node where the Perspective is to be executed. The Perspective user must also have a

Kerberos principal defined for the system partition where the Perspective is to be executed, and the user must have write permission to the System Data repository.

To use the SP Event Perspective effectively, you must understand certain terminology. These terms were introduced in “The SP Event Perspective” on page 18. Please refer to that section to become familiar with these terms.

Users of the SP Event Perspective can set up the Perspective to send a notification to the system administrator when conditions of interest exist on the system (or, to use the Perspectives terminology, when an event occurs). This is done by associating an action with the event in the event definition. This action can be any command or script that can be executed from the AIX command line, including the creation of an electronic mail message, starting a process that can place a telephone call to the system administrator's pager, send a message to a specific user at a specific terminal, or any other notification command. The action invoked when the event occurs is called the command.

When creating or modifying the event definition, the user can specify a command to be executed when the condition exists. The following AIX command can be used to have the SP Event Perspective send an electronic mail message to a specific user when the event occurs:

```
/usr/bin/echo \
"event_condition has occurred `/usr/bin/date` - Location Info: $PMAN_IVECTOR" | \
/usr/bin/mail -s "event_condition Notification" \
username@address
```

To understand the mechanics of setting up a command within an event definition, consult the Perspectives online help. Click on the **Help** menu button on the SP Event Perspective display, and select the **Tasks...** option. Assistance on specifying event definitions is available through the **Working with Event Definitions** topic.

Whenever the user registers for the event definition through the SP Event Perspective, the command will be executed if the condition exists in the system. The SP Event Perspective does not have to be currently active in order for this command to be executed; provided the user has registered for the event definition, the system will continue to monitor itself for this condition, and execute the command if the condition exists. In other words, the user can use the SP Event Perspective to set up event definitions, register for events, then shut down the SP Event Perspective, and the system will still execute the notification command when the condition occurs. The SP system continually monitors itself for the condition and issues the notification command until the user cancels the event registration.

The command associated with an event definition can also be used to automate a response to this condition, instead of merely notifying the system administrator or another user of the condition. This topic will be discussed in “Automating Your Response to Problems” on page 41.

Command Line Tools - Problem Management

Problem Management is a software subsystem used in command line and script oriented environments to specify conditions that should be monitored by the SP system, and to specify actions to take when these conditions exist on the SP system. This is the same software subsystem invoked internally by the SP Event Perspective discussed in the previous section. For users attempting to connect to the SP nodes through low-speed modems or using non-graphical terminals, Problem Management provides a command line interface that can be used in place of the SP Event Perspective. As with other command line oriented tools, the Problem Management command line interface is not as intuitive to use or designed for ease of use as is its graphical counterpart.

1. **Preparing to Monitor The System.** To use Problem Management, the user must have a Kerberos principal defined for that user. The Problem Management subsystem must also be active on the node where the user will be executing Problem Management.

   To simplify later instructions for the use of Problem Management, ensure that Problem Management is active on the control workstation and on the nodes where conditions are to be monitored. Verify that the Problem Management subsystem is active on a node through the **lssrc -s pman** command. Example output from the **lssrc -s pman** command, executed on a node of an SP system:

   ```
   Subsystem         Group            PID     Status
    pman             pman             22886   active
   ```

   From the control workstation, verify that the Problem Management subsystem is active for the system partition that will be monitored, using the **lssrc -s pman.***system_partion_name* command. Example output from the **lssrc -s pman.***system_partion_name* command, executed from the control workstation of an SP system with multiple system partitions:

   ```
   Subsystem         Group            PID     Status
    pman.k5sp1       pman             34554   active
   ```

   Ensure that the pman subsystem for the system partition is active. If the subsystem is not active, activate it using the **startsrc -g pman** command from any node.

2. **Understand What You Want to Monitor.** Problem Management expects the user to know the conditions that are to be monitored. Unlike the SP Event Perspective, Problem Management does not provide an interactive method to query for the list of available conditions and a means to select from these conditions. The user must identify the conditions to be monitored, and provide them as a list to Problem Management. These conditions are identified by naming the associated resource variables, the internal mechanism that contains the current status of the associated resource.

   PSSP provides over 300 default resource variables. Chapter 3, “Conditions to Monitor on the SP System” on page 45 provides a suggested list of resource variables to monitor, but specific SP systems may require that additional resources also be monitored. The full list of resource variables is maintained by the Event Management subsystem, and can be retrieved using the **haemqvar** command. This command generates large amounts of information, so it is best to start with a brief report from this command to identify those resources to be monitored:

```
haemqvar -d | more
```

The command provides the resource variable names available, and a short description of the resource variable:

```
IBM.PSSP.aixos.Proc.swpque   Average count of processes
                             waiting to be paged in.
IBM.PSSP.aixos.Proc.runque   Average count of processes that are
                             waiting for the cpu.
IBM.PSSP.aixos.pagsp.size   Size of paging space (4K pages).
IBM.PSSP.aixos.pagsp.%free   Free portion of this paging space (percent).
IBM.PSSP.aixos.PagSp.totalsize Total active paging space size (4K pages).
IBM.PSSP.aixos.PagSp.totalfree Total free disk paging space (4K pages).
IBM.PSSP.aixos.PagSp.%totalused Total used disk paging space (percent).
IBM.PSSP.aixos.PagSp.%totalfree   Total free disk space (percent).
IBM.PSSP.aixos.Mem.Virt.pgspgout 4K pages written to paging space by VMM.
IBM.PSSP.aixos.Mem.Virt.pgspgin 4K pages read from paging space by VMM.
IBM.PSSP.aixos.Mem.Virt.pagexct   Total page faults.
IBM.PSSP.aixos.Mem.Virt.pageout   4K pages written by VMM.
IBM.PSSP.aixos.Mem.Virt.pagein    4K pages read by VMM.
IBM.PSSP.aixos.Mem.Real.size   Size of physical memory (4K pages).
IBM.PSSP.aixos.Mem.Real.numfrb   Number of pages on free list.
IBM.PSSP.aixos.Mem.Real.%pinned   Percent memory which is pinned.
IBM.PSSP.aixos.Mem.Real.%free   Percent memory which is free.
                          :
                          :
```

The **haemqvar** command lists resources available only on the node where the command is executed. Keep in mind that resources may exist on some nodes and not on others. *PSSP: Command and Technical Reference* gives a detailed description of the **haemqvar** command, and how it can be used to locate any resource variable available within the SP system.

Once the resource variable to be monitored have been identified, the value type and locator for each resource variable must be identified. The locator informs Problem Management where to monitor the resource. For example, Problem Management needs to know the name of the file system and the node on which a file system resides if it is to monitor that file system for the amount of space it has available; this information is conveyed to Problem Management through the locator value. To obtain the locator for a resource variable, issue the following **haemqvar** command:

```
haemqvar "" resource_variable_name "*"
```

This command provides details on the resource variable, including the locator keyword needed for Problem Management. The additional information can be helpful in constructing an effective Problem Management definition for the condition.

For example, to obtain the locator field for the **IBM.PSSP.CSS.ipackets_drop** variable, and to understand more about the variable issue:

```
haemqvar "" IBM.PSSP.CSS.ipackets_drop "*"
```

which produces this output:

```
Variable Name:  IBM.PSSP.CSS.ipackets_drop
Value Type:     Quantity
Data Type:      long
Initial Value:  0
Class:          IBM.PSSP.CSS
Locator:        NodeNum
Variable Description:
    Number of packets not passed up.

    A message received by a node from the switch of the
    Communication SubSystem (CSS) is comprised of packets.
    IBM.PSSP.CSS.ipackets_drop is the count of the number of good
    incoming packets at the subject node's CSS interface which
    were dropped by the adapter microcode, since that interface
    was last initialized.

    If a node has too heavy a general workload, it may not service its
    CSS interface often enough, causing its messages to linger in the
    switch network.  If this is allowed to continue, the switch can
    become backed up causing other nodes to encounter poor switch
    performance; in fact, this condition can cause the entire
    switch to clog. Instead, the adapter microcode drops any "excess"
    packet -- a reliable protocol will eventually retry the message.

    For performance reasons, counts such as this are only updated
    approximately once every 2 minutes.

    This variable is supplied by the "IBM.PSSP.harmld" resource monitor.

    Example expression:

    To be notified when IBM.PSSP.CSS.ipackets_drop exceeds 100 on any
    node, register for the following event:

          Resource variable:  IBM.PSSP.CSS.ipackets_drop
          Resource ID:        NodeNum=*
          Expression:         X>100
          Re-arm expression:  X<100

    Resource ID wildcarding:

    The resource variable's resource ID is used to specify the number
    of the node (NodeNum) to be monitored. The NodeNum resource ID
    element value may be wildcarded in order to apply a query or event
    registration to all nodes in the domain.

    Related Resource Variables:

      IBM.PSSP.CSS.ibadpackets    Number of bad packets received
                                  by the adapter.
      IBM.PSSP.CSS.ipackets_lsw   Packets received on interface
                                  (lsw bits 30-0).
      IBM.PSSP.CSS.ipackets_msw   Packets received on interface
                                  (msw bits 61-31).
Resource ID:    NodeNum=int
  NodeNum: The number of the node for which the information applies.
```

The **Locator**: field indicates the keyword to be used with Problem Management to identify where the resource should be monitored. Note that the **haemqvar** command offers advice on how to use the locator field in the output.

3. **Identify the Conditions of Interest.** Problem Management is informed of the conditions to be monitored through the **pmandef** command. One **pmandef** command is needed for each condition to be monitored. This command is used to subscribe to the event, which is similar in concept to the SP Event Perspective's registration of an event definition. To create the subscription, the following information is needed:

   - The resource variable name, obtained in Step 2 on page 35.
   - The resource variable locator, also obtained in Step 2 on page 35.
   - The event expression, which indicates the condition of interest in this resource variable
   - The rearm expression, which indicates when the condition is no longer of interest
   - An event handle, which is a symbolic name that the system administrator will use to refer to this definition.

   The event expression indicates the value that the resource variable will have when notice should be given. This value is assigned by the system administrator. The rearm expression indicates the value of the resource variable that indicates that the condition of interest is no longer present. How these expression is coded depends on the value type of the resource variable. The event handle is a name assigned by the system administrator, which should be descriptive of the condition being monitored.

   For example, consider a case where the system administrator is interested in paging space on any SP node. If paging space reaches 90% capacity, the system administrator considers the node to be "thrashing" and wants to be notified. The administrator considers the node to be "thrashing" once this threshold is reached, even if a little paging space frees up. The administrator doesn't consider the "thrashing" problem to be resolved unless 40% of the paging space becomes available again. Using this scenario and the **haemqvar** commands from Step 2 on page 35, the system administrator identifies these conditions of interest:

   - The resource variable name is **IBM.PSSP.aixos.PagSp.%totalused**, which contains the percentage of used paging space on a node.
   - The resource variable locator is **NodeNum**, meaning that a node number is needed to indicate where the resource is to be monitored. The administrator wants to monitor the condition on all nodes, so the locator expression is **NodeNum=***.
   - The event expression must indicate the 90% capacity condition, so the expression **X>90** is used.
   - The rearm expression must indicate the condition that "turns off" the event, which is when 40% of paging space becomes available. The expression **X< 60** is used.
   - The system administrator assigns the name **Node_Thrash_Monitor** to this event definition.

Identify these conditions for all resource variables to be monitored. Chapter 3, “Conditions to Monitor on the SP System” on page 45 lists some basic resource variables to monitor and the associated event expressions and rearm expressions.

4. **Decide How to Notify the System Administrator.** Problem Management associates an action with the event definition. When the condition exists within the system (or, to use the correct terminology, when the event occurs), Problem Management performs the action associated with the event. This action can be any command or script that can be executed from the AIX command line, including the creation of an electronic mail message, starting a process that can place a telephone call to the system administrator's pager, send a message to a specific user at a specific terminal, or any other notification command. This action is termed the command.

   The user can specify a command to be executed when the condition exists. The following AIX command can be used to have Problem Management send an electronic mail message to a specific user when the event occurs:

```
/usr/bin/echo \
"event_handle has occurred `/usr/bin/date` - Location Info: $PMAN_IVECTOR" | \
/usr/bin/mail -s "event_handle Notification" \
username@address
```

   An action can also be executed when the condition no longer exists: when the rearm expression has been met. This action, called the rearm command, can inform the system administrator that the condition no longer exists, so that the administrator knows that the condition no longer needs attention. For example:

```
/usr/bin/echo \
"event_handle condition ended `/usr/bin/date` - Location Info: $PMAN_IVECTOR" | \
/usr/bin/mail -s "event_handle Condition Ended"\
username@address
```

5. **Create an Event Definition File.** For every resource variable to be monitored, one **pmandef** command must be issued. If more than a handful of resources variables are to be monitored, this can result in a lot of typing. For convenience, create a file containing the **pmandef** commands to define these events to Problem Management. This will simplify the procedure for instructing Problem Management of the resources to monitor, and makes it easier to reissue these same commands at a later time.

   The **pmandef** command informs Problem Management of the conditions to be monitored by subscribing for events. This concept is almost exactly the same as the SP Event Perspective's concept of registering event definitions. To subscribe for events on the chosen resource variables, create a file to contain **pmandef** commands in the following format:

```
pmandef -s event_handle \
-e 'resource_variable_name:locator:event_expression' \
-r "rearm_expression" \
-c event_command \
-C rearm_command \
-n 0
```

Substitute the following information for the keywords in the above command format:

- *event_handle* is the event handle assigned by the system administrator in Step 3 on page 38.
- *resource_variable_name* is the name of the resource variable, obtained in Step 2 on page 35.
- *locator* is the locator expression, indicating where the resource is to be monitored, determined in Steps 2 on page 35 and 3 on page 38.
- *event_expression* indicates the value the resource variable will have when the condition of interest exists, determined in Step 3 on page 38.
- *rearm_expression* indicates the "shut off" value the resource variable will have when the condition no longer exists, determined in Step 3 on page 38.
- *event_command* indicates the notification command to use for informing the system administrator that the condition exists, created in Step 4 on page 39.
- *rearm_command* indicates the notification command to use for informing the system administrator that the condition no longer exists, created in Step 4 on page 39.

Continuing with the previous example, the **pmandef** command to subscribe for the node thrashing condition would be:

```
pmandef -s Node_Thrash_Monitor \
-e 'IBM.PSSP.aixos.PagSp.%totalused:Nodenum=*:X>90' \
-r "X<60" \
-c '/usr/bin/echo "Node_Thrash_Monitor Alert `/usr/bin/date` - Location Info: $PMAN_IVECTOR" |\
/usr/bin/mail -s "Node_Thrash_Monitor Alert" root@adminnode.ibm.com' \
-C '/usr/bin/echo "Node_Thrash_Monitor Cancellation `/usr/bin/date` - Location Info: $PMAN_IVECTOR" |\
/usr/bin/mail -s "Node_Thrash_Monitor Cancellation" root@adminnode.ibm.com' \
-n 0
```

One **pmandef** command is required for each condition being monitored. Save this file and note its name for future reference.

6. **Subscribe to the Events Through Problem Management.** To record these event definitions to Problem Management, execute the **pmandef** commands recorded in the file created in Step 5 on page 39 by issuing the **ksh** *filename* command, where *filename* is the name of the file created in Step 5 on page 39. Immediately after issuing the **ksh** command, issue the following command:

   ```
   pmandef -d all
   ```

   Problem Management not only subscribes to events with the **pmandef -s** command, but it also begins monitoring the resources as well. The **pmandef -d all** command disables the monitoring of these resources.

7. **Begin Monitoring.** Begin monitoring these resources when you are ready. To begin monitoring the events that were subscribed in Step 6, issue the following command:

   ```
   pmandef -a all
   ```

   This instructs Problem Management to begin monitoring all the conditions that were defined in Step 6. Should any of these events occur, Problem

Management will issue the associated event command to inform the system administrator of the event.

8. **Tailor the Monitoring.** At times, certain conditions should not be checked on certain nodes. For example, Problem Management may be monitoring the space available in the **/tmp** file system on all nodes, but the system administrator expects **/tmp** to exceed that limit on a specific node (for example: node 5 on a 32-node SP system) for a certain period of time. If the monitoring is not tailored or modified to compensate for this expected event, the system administrator will be notified that the event occurred just as if it were an unexpected event.

   The system administrator can modify the subscribed event to Problem Management. To do this, the administrator needs to know the following:

   - The event handle used for the condition, assigned in Step 3 on page 38.
   - The new locator that excludes the location where the condition is not to be monitored.

   Modification of the subscription is done in two steps:

   a. The event subscription is disabled, using the **pmandef** command:

      ```
      pmandef -d event_handle
      ```

      This deactivates the monitoring for this condition.

   b. A new event expression is created, using a locator that excludes the location where the condition is not to be monitored. Using the example of **/tmp** monitoring, where node 5 is not to be monitored, the event expression would appear as:

      ```
      IBM.PSSP.aixos.FS/%totused:NodeNum=1-4,6-32;VG=rootvg;
      LV=hd3:X>90
      ```

   c. Issue the **pmandef -s** command, using the structure provided in Step 5 on page 39 and the new event expression.

9. **Stop Monitoring.** To stop monitoring of all events previously enabled in Step 7 on page 40, issue the following command:

   ```
   pmandef -d all
   ```

These steps provide an overview of how Problem Management can be used to monitor system events and notify the system administrator when events occur. This is not a complete tutorial on the use of Problem Management. For greater detail on the capabilities and uses of Problem Management, consult the Problem Management chapter in *PSSP: Administration Guide*, the **pmandef** command, and the **haemqvar** command in *PSSP: Command and Technical Reference*.

Automating Your Response to Problems

Detecting potential problem conditions before they become critical situations is the best way to resolve SP system problems. The condition is brought to the attention of the system administrator, allowing the system administrator to respond before the condition impacts other hardware and software components. But what if the procedure for rectifying the situation is always the same? What if the system administrator will always execute the same set of commands to address the condition whenever the condition occurs? Is it really necessary to require system

administrator intervention, when the administrator is going to perform the same action in all cases? The answer is **no**.

PSSP provides the capability to set an automated response to a system condition. This capability is provided as part of the SP Event Perspective and the Problem Management subsystem, both described in the previous sections. Through this capability, the administrator can assign a specific action to be executed automatically by these tools when the system condition exists, and also when the condition "goes away" (when the rearm event occurs). In the previous discussions of this chapter, this action was kept rather simple: the action caused a notification to be sent to the system administrator. The associated action does not need to be this simple: the action can be any AIX command or script. When the event occurs, these tools will execute the command or script in response to the event.

Important - When Actions Are Performed

A response to a particular system condition may not always be the same, despite initial appearances. For example, when a file system is close to reaching its capacity, the appropriate response in most cases is to increase the file system's capacity using the **chfs** command. However, disk space is not a limitless resource, and eventually all disk space will be consumed if this approach is used whenever the file system reaches its capacity. Although this is a convenient solution, it is not always the correct solution. The file system should be checked for large obsolete files that can be deleted, users that exceed their quotas, directories that can be mounted on other file systems to save space in this file system, and other solutions.

When an administrator associates an action with an event through Problem Management or the SP Event Perspective, this action is performed **each time the event occurs** (and Problem Management or Perspectives is monitoring the condition). Neither Problem Management nor the SP Event Perspective can decide that the action should not be performed for this specific occurrence of the event, but rather the administrator needs to do some more analysis. This decision making process has to either be incorporated into the AIX command or script that will run in response to the event, or it has to be left to the system administrator's discretion.

Two strategies are offered:

- Specify an action for the event that performs two actions: notifies the system administrator of the event and also issues a command to address the event. This strategy allows the system to respond automatically to the condition and attempt to resolve the condition before it becomes a critical situation, and also alerts the system administrator. The system administrator can then assess if the action should have been applied in this case.

  If the action was appropriate, the system administrator does not need to take action. However, if recent history indicates that the event has been recurring at an unusual rate, or history indicates that the action really does not resolve the condition and the event continues to occur, the system administrator still receives the notification and can respond to the condition.

- Build a response command or script that incorporates a decision-making process within it. This response command can attempt to determine if a particular action is appropriate for the condition based on other information, such as other events or the recent history of actions taken in response to this event.

The second alternative can involve some rather complicated logic, making it more difficult to implement. For this reason, the first strategy is recommended.

Important - Where Actions Are Performed

Actions associated with events are performed **on the node that requested that the event be monitored**, which is **not necessarily the node where the condition exists**. For example, if an administrator used Problem Management from the control workstation to monitor conditions on all nodes in the SP system and that condition suddenly exists on Node 42, the action is executed on the control workstation, not Node 42. If the administrator had associated a **chfs** command with the event, the **chfs** command would execute on the control workstation and modify the control workstation's file system space, not the file system on Node 42.

When associating actions with events, keep in mind that the action will be performed by default on the node that asked for the event to be monitored. If action is to be taken on the node where the condition actually exists, the command invoked must determine where the condition occurred from the event information, and attempt to execute remote processes on that node.

Both the SP Event Perspective and Problem Management make available several environmental values to the commands associated with the event. These variables are described in the chapter on using the Problem Management subsystem in *PSSP: Administration Guide*. Any command or script executed by Problem Management or the SP Event Perspective has access to these variables. The variable **PMAN_IVECTOR** indicates where the condition exists. The command or script can parse the value of this environment variable, extract the node location information, and use that information to construct the appropriate remote command.

For example, consider the case where the **/var** file system is being monitored to ensure that it does not reach its capacity. When the file system does reach its capacity, the **chfs** command is to be invoked on the node where the condition exists to extend the size of the **/var** file system. To perform this action, a Korn Shell script is created. This script examines the contents of the **PMAN_IVECTOR** value, which has the following components to identify where the condition exists:

```
VG=rootvg;LV=hd9var;NodeNum=node_number_where_condition_exists
```

Once the node number has been found in the **PMAN_IVECTOR** value, the script will then find the host name for that node in the SDR. The script then uses the **dsh** command to execute the **chfs** command on the remote node to extend the size of the **/var** file system:

```
#!/bin/ksh
OLDIFS=$IFS
IFS=';'
set $PMAN_IVECTOR
for TOKEN in $*
do
if [[ $TOKEN = NodeNum* ]]
        then
              IFS='='
              print "$TOKEN" | read JUNK NODENUM
              HOST=$(SDRGetObjects -x Node node_number\=\=$NODENUM | \
                      awk '{print $11}')
        fi
done
IFS=$OLDIFS
if [[ "$HOST" != "" ]]
then
        dsh -w $HOST /usr/sbin/chfs -a size\=+1 /var
fi
```

This script is saved to an AIX file on the node where the monitoring request is made, and the execute permission is set on the file. The full path name of the file can then be provided to either Problem Management or the SP Event Perspective as a command to be issued when the event occurs.

Graphical Tools - SP Event Perspective

The “Asynchronous (Batch) Notification Methods” on page 33 introduced the concept of associating an action with a specific condition through the SP Event Perspective. This was done by providing a command within an event definition. In the previous section, this command was relatively simple: it issued an electronic mail message to a specific user to report the occurrence of the event.

To create an automatic response to the condition, provide an AIX command or script in the command field in addition to (or in place of) the notification command that was used in the previous discussion. Be sure to understand where the SP Event Perspective will attempt to execute the command before assigning a command to this event, by referring to “Important - Where Actions Are Performed” on page 43.

Command Line Tools - Problem Management

“Asynchronous (Batch) Notification Methods” on page 33 introduced the concept of associating an action with a specific condition through the **pmandef** command of Problem Management. This was done by specifying an argument to the **-c** and **-C** options of the **pmandef** command. In the previous section, this argument was relatively simple: it issued an electronic mail message to a specific user to report the occurrence of the event.

To create an automatic response to the condition, provide an AIX command or script as an argument to **-c** or **-C** options of the **pmandef** command, in addition to (or in place of) the notification command that was used in the previous discussion. Be sure to understand where Problem Management will attempt to execute the command before assigning a command to this event by referring to “Important - Where Actions Are Performed” on page 43.

Chapter 3. Conditions to Monitor on the SP System

Conditions to Monitor Using Perspectives or Problem Management

Whether you decide to monitor system condition using Perspectives or Problem Management, the sections below provide a list of the **minimum** hardware and software conditions to monitor. See “Descriptions of Each Condition” on page 46 for a detailed description of each condition.

Monitor These Hardware Conditions

For each frame, switch or node, monitor the following hardware conditions:

- For a frame:
  - Power
  - Controller responding
  - Controller ID mismatch (only applies to HACWS)
  - Temperature
  - Node slot failures
- For a switch:
  - Power and power availability
  - Environment light
  - Temperature
- For a node:
  - Power and power availability
  - Environment light
  - Temperature
  - Keymode switch
  - LED/LCD contents
  - LED/LCD flashing
  - Node responding
- For the control workstation:

  Same as for a node

Monitor These Software Conditions

For each node, monitor the following software conditions:

- On each node:
  - Node can be reached by RSCT
  - **/tmp** becoming full
  - **/var** becoming full
  - **/** becoming full

  - Paging space low
  - Rising mbuf failures
  - Switch I/O errors
  - **inetd** daemon activity
  - **srcmstr** deamon activity
  - **ftpd** daemon activity
  - **biod** daemon activity - applies only to NFS systems
  - **portmap** daemon activity - used by RPC
  - **xntpd** daemon activity (NTP time synch)
  - **httpd** daemon activity (applies only to HTTP servers)
  - **kerberos** daemon activity (if the Kerberos database resided on this node)
  - **hatsd** daemon activity (cannot check using Event Management)
  - **hadsd** daemon activity (cannot check using Event Management)
  - **haemd** daemon activity (cannot check using Event Management)
- On the control workstation:
  - Same as on each node
  - **sdrd** daemon active

Descriptions of Each Condition

Table 1 (Page 1 of 10). Details About Each Condition to Monitor

Condition	Details
Frame Power	**Description:** Whether the frame has its power on or off. When the frame power is off, nodes and switches in the frame cannot receive power. **Resource Variables:** • SP_HW.Frame.frPowerOff (overall power) • SP_HW.Frame.frPowerOff_A (power supply A) • SP_HW.Frame.frPowerOff_B (power supply B) • SP_HW.Frame.frPowerOff_C (power supply C) • SP_HW.Frame.frPowerOff_D (power supply D) • SP_HW.Frame.frACLED (AC power OK) • SP_HW.Frame.frDCLED (DC power OK) **Notes:** With the **frPowerOff_*** variables, only one needs to be **on** for the frame to receive power. If all of them are **off**, then the frame has no power. If power was not explicitly shut down by the system administrator, perform hardware diagnostics on the frame.

Table 1 (Page 2 of 10). Details About Each Condition to Monitor	
Condition	**Details**
Frame Controller Responding	**Description:** This indicates whether the frame controller is responding to command requests. The SP system can function when the frame controller is not responding, but it will not be possible to obtain certain node hardware status (such as key switch position and LED readouts) or issue certain hardware commands (such as resetting the node). When the controller fails, perform hardware diagnostics and replace the frame controller, if this is called for. Replacing the frame controller requires you to schedule down time for all nodes in that frame. **Resource Variable:** SP_HW.Frame.controllerResponds
Frame Controller ID Mismatch	**Description:** This indicates whether the ID of the frame controller agrees with the ID stored for it in the HACWS supervisor card. If the IDs do not match, this indicates that the HACWS supervisor card is not properly wired to the frame (possibly using the wrong "tty" line). Have the wiring between the control workstations (primary and backup) and the frame controller checked, and if that does not solve the problem, perform hardware diagnostics on both the control workstations and the frame controller. Monitor this condition only when HACWS is installed on the SP system. **Resource Variable:** SP_HW.Frame.controllerIDMismatch
Frame Temperature	**Description:** This indicates whether the frame's temperature is within the normal operational range. If the temperature becomes out of range, hardware within the frame may fail. Make sure that all fans are working properly. There are resource variables that you can check with the SP Event Perspective or the **hmmon** command to determine this. Make sure that the frame has proper ventilation. **Resource Variable:** SP_HW.Frame.tempRange

Table 1 (Page 3 of 10). Details About Each Condition to Monitor

Condition	Details
Frame Node Slot Failures	**Description:** This indicates whether or not the frame supervisor can communicate with the node supervisor attached to the frame slot. It is possible to see a "failure" in this condition when no real failure exists. For example, since a wide node occupies two slots in the frame but only has one node supervisor, one of the slots associated with the wide node will always show a "failure". Any slots where nodes are not attached will show a "failure", but this is OK. This is why it is important to know the layout of the SP system. You should be concerned when the status changes to show a failure because it can indicate a failure in the node supervisor. The node may continue to function in this type of failure, but certain hardware status (LEDs, switch position) may not be available and commands (node reset) may not work. Run hardware diagnostics on the node connected to the frame slot showing a failure. **Resource Variables:** 1. SP_HW.Frame.nodefail1 2. SP_HW.Frame.nodefail2 3. SP_HW.Frame.nodefail3 4. SP_HW.Frame.nodefail4 5. SP_HW.Frame.nodefail5 6. SP_HW.Frame.nodefail6 7. SP_HW.Frame.nodefail7 8. SP_HW.Frame.nodefail8 9. SP_HW.Frame.nodefail9 10. SP_HW.Frame.nodefail10 11. SP_HW.Frame.nodefail11 12. SP_HW.Frame.nodefail12 13. SP_HW.Frame.nodefail13 14. SP_HW.Frame.nodefail14 15. SP_HW.Frame.nodefail15 16. SP_HW.Frame.nodefail16 17. SP_HW.Frame.nodefail17 **Note:** SP_HW.Frame.nodefail17 indicates a failure in the SP Switch supervisor. If the frame has no switch, this will always show a "failure".
Switch Power	**Description:** This indicates whether the switch power is on or off. If the frame has no power, the switch will not have power, so this should be checked first. If the frame has power but the switch does not, ensure that the switch was not manually shut down, and perform hardware diagnostics on the switch **Resource Variables:** • SP_HW.Switch.nodePower • SP_HW.Switch.powerLED • SP_HW.Switch.shutdownTemp **Notes:** SP_HW.Switch.nodePower indicates whether the power is on or off. SP_HW.Switch.powerLED indicates this as well, but also indicates whether the switch can receive power but is powered off. SP_HW.Switch.shutdownTemp indicates if the switch was powered off because of a high temperature condition.

Table 1 (Page 4 of 10). Details About Each Condition to Monitor	
Condition	**Details**
Switch Hardware Environment Indicator	**Description:** This indicates if the switch has detected any hardware anomalies that can cause or has caused a shut down of the switch. Such anomalies are: incorrect voltage, fan failure, temperature out of range, and internal hardware failure. This indicator shows whether all is well, whether a condition exists that should be investigated, or whether the switch was forced to shut down because of these errors. **Resource Variable:** SP_HW.Switch.envLED **Notes:** Any change in this indicator is worth investigating, even if the indicator shows that the problem is not yet critical. Check for fan failures in the SP Switch. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command. Perform hardware diagnostics on the switch. Schedule repair for any failing hardware components.
Switch Temperature	**Description:** This indicates if the temperature inside the switch hardware is out of the normal operational range. If the temperature becomes out of the normal range, the device may overheat and the hardware may fail. **Resource Variable:** SP_HW.Switch.tempRange **Notes:** Check for fan failures in the SP Switch. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command, and ensure that the frame has proper ventilation.
Node Power	**Description:** This indicates whether the node power is on or off. If the frame has no power, the node will not have power, so this should be checked first. If the frame has power but the node does not, ensure that the node was not manually shut down, and perform hardware diagnostics on the node. **Resource Variables:** SP_HW.Node.nodePower SP_HW.Node.powerLED **Notes:** SP_HW.Node.nodePower indicates whether the power is on or off. SP_HW.Node.powerLED indicates this as well, but also indicates whether the node can receive power but is powered off.
Node Controller Responding	**Description:** This indicates whether the node's controller is responding to the frame controller. The node controller allows you to set hardware controls on the node, such as the power, key mode switch and the reset button. It also queries hardware status of the node, such as the LCD or LED values. A failure in the node's controller does not necessarily indicate that the node is unusable, but it will make it difficult or impossible to query or control the node hardware using the SP Hardware Perspective or the **spmon** and **hmmon** commands. If the node's power is off, the node controller cannot respond, so this should be checked first. If the node's power is on, perform hardware diagnostics on the node. **Resource Variable:** SP_HW.Node.hostresponds

Table 1 (Page 5 of 10). Details About Each Condition to Monitor	
Condition	**Details**
Node Hardware Environment Indicator	**Description:** This indicates if the node has detected any hardware anomalies that can cause or have caused a shut down of the node. Such anomalies are: incorrect voltage, fan failure, temperature out of range, or internal hardware failure. This indicator shows whether all is well, whether a condition exists that should be investigated, or whether the node was forced to shut down because of these errors. **Resource Variable:** SP_HW.Node.envLED **Notes:** Any change in this indicator is worth investigating, even if the indicator shows that the problem is not yet critical. Check for fan failures in the node. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command. Perform hardware diagnostics on the node. Schedule repair for any failed hardware components.
Node Temperature	**Description:** This indicates if the temperature inside the node hardware is out of the normal operational range. If the temperature becomes out of the normal range, the node may overheat and the hardware may fail. **Resource Variable:** SP_HW.Node.tempRange **Notes:** Check for fan failures in the node. There are additional resource variables that you can use to check this with the SP Event Perspective and the **hmmon** command. Ensure that the frame has proper ventilation, and check that all air paths within the frame are not clogged.
Node Key Mode Switch Position	**Description:** This shows the current setting of the node's mode switch. During node boot, the key switch position controls whether the operating system is loaded, and whether service controls are activated. During system operation, the position controls whether the node can be reset and whether system dumps can be initiated. For everyday operation, the key should be in the "Normal" position and should not change. A command must be issued to change the key position. If this does occur, locate the person changing this control and ensure that this action was taken for a proper reason. **Resource Variable:** SP_HW.Node.keyModeSwitch **Note:**Not all nodes have a key switch.
Node LED or LCD Readout	**Description:** Each node has a 3-digit or 4-digit LCD or LED display. This display indicates the status of hardware and software testing during the node's boot process. This display is also used to display specific codes when node hardware or the operating system software fails. These codes indicate the nature of the failure, and whether any additional error data may be present. After a node has successfully booted, this display should be blank. If the display is not blank, use either the SP Hardware Perspective or the **spmon** command to determine what value is being displayed, and consult the AIX documentation, Chapter 37, “SP-Specific LED/LCD Values” on page 257, and Chapter 38, “Network Installation Progress” on page 261 to determine what the LED/LCD value means. **Resource Variable:** SP_HW.Node.LCDhasMessage

Table 1 (Page 6 of 10). Details About Each Condition to Monitor	
Condition	**Details**
Node Reachable by RSCT Group Services	**Description:** This indicates whether RSCT Group Services can reach the node through any of its network adapters and the switch. If this indicates that the node is not reachable, all of the node's network and switch adapters have either failed or been disabled. In this case, the only way to reach the node when it is powered on is through the node's serial link, using the **s1term** command. When this happens, check the network adapter status and issue the **/etc/ifconfig on** command to enable the adapter. Also, check the switch status and perform problem determination procedures for Group Services and Topology Services. **Resource Variable:** Membership.Node.state
/tmp file system becoming full	**Description:** Each node has its own locally available **/tmp** file system. This file system is used as temporary storage for many AIX and PSSP utilities. If this file system runs out of space, these utilities can fail, causing failures in those PSSP and LPP utilities that depend on them. When this file system nears its storage capacity, it should be checked for large files that can be removed, or the file system size should be increased. **Resource Variable:** aixos.FS.%totused **Resource Identifier:** VG = rootvg LV = hd3
/var file system becoming full	**Description:** Each node has its own locally available **/var** file system. This file system contains system logs, error logs, trace information, and other important node-specific files. If this file system runs out of space, log entries cannot be recorded, which can lead to loss of error information when critical errors occur, leaving you and IBM service personnel without an audit or debug trail. When the file system nears its storage capacity, it should be checked for old log information that can be removed or cleared, the file system size should be increased, or separate file systems should be made for subdirectories that consume large amounts of disk space. **Resource Variable:** aixos.FS.%totused **Resource Identifier:** VG = rootvg LV = hd9var
/ file system becoming full	**Description:** Each node has its own locally available **root** file system. This file system contains important node boot and configuration information, as well as LPP binaries and configuration files. If this file system runs out of space, it may not be possible to install products on the node, or update that node's configuration information (although the SMIT and AIX-based install procedures should attempt to acquire more space). When this file system nears its storage capacity, it should be checked for **core** files or any other large files that can be removed, or the file system's size should be increased. **Resource Variable:** aixos.FS.%totused **Resource Identifier:** VG = rootvg LV = hd4

Table 1 (Page 7 of 10). Details About Each Condition to Monitor

Condition	Details
Paging Space Low	**Description:** Each node has at least one locally available paging device. When all these paging devices near their capacity, the node begins to thrash, spending more time and resources to process paging requests than to process user requests. When operating as part of a parallel process, the thrashing node will delay all other parts of the parallel process that wait for this node to complete its processing. It can also cause timeouts for other network and distributed processes. A temporary fix is to terminate any non-critical processes that are using large amounts of memory. If this is a persistent problem, a more permanent fix is to restrict the node to specific processing only, or to add additional paging devices. **Resource Variable:** aixos.PagSp.%totalused
Kernel Memory Buffer Failures	**Description:** Kernel memory buffers, or "mbufs", are critical to network processing. These buffers are used by the kernel network protocol code to transfer network messages. If the kernel begins to encounter failures in acquiring these buffers, network information packets can be lost, and network applications will not run efficiently. An occasional failure can be tolerated, but numerous failures or a continuous stream of small failures indicates that not enough memory has been allocated to the kernel memory buffer pool. **Resource Variable:** aixos.Mem.Kmem.failures **Resource Identifier:** Type = mbuf
Switch Input and Output Errors	**Description:** The switch device driver tracks statistics on the device's use, including any errors detected by the driver. These errors are tracked as counters that are never reset, unless the node is rebooted. Consult Chapter 11, “Diagnosing Switch Problems” on page 109 for assistance in diagnosing any reported errors for the SP Switch. **Resource Variables:** • CSS.ibadpackets • CSS.ipackets_drop • CSS.ierrors • CSS.opackets_drop • CSS.oerrors • CSS.xmitque_ovf **Notes:** Any increment in the value of the CSS.ierrors or CSS.oerrors counters indicates that the switch adapter is about to go offline. Continual increments to the CSS.ibadpackets counter can indicate transmission problems or "noise" in the connection between the SP Switch adapter and the SP Switch, so the SP Switch cabling should be checked and hardware diagnostics performed. Continual increments to the CSS.ipackets_drop and CSS.opacktes_drop counters indicate that there is either too much input or too much output for the SP Switch device driver to handle, and packets are lost.

Table 1 (Page 8 of 10). Details About Each Condition to Monitor	
Condition	**Details**
inetd Daemon Activity	**Description:** The **inetd** master daemon is responsible for activating many AIX and PSSP service daemons when a client for that service connects to the node. If the daemon fails, these services cannot be started. Since many SP applications are network applications, this can cause widespread failure in all SP applications. If the daemon cannot be restarted manually, force a system dump of this node, collect information for the IBM Support Center, and restart the node. The reboot may temporarily resolve the problem. **Resource Variable:** Prog.xpcount **Resource Identifier:** ProgName=inetd UserName = root
srcmstr Daemon Activity	**Description:** The **srcmstr** daemon implements the System Resource Controller functions. If this daemon fails, services registered with the SRC cannot be controlled using SRC commands. If the daemon cannot be restarted manually, force a system dump of this node, collect information for the IBM Support Center, and restart the node. The reboot may temporarily repair the problem. **Resource Variable:** Prog.xpcount **Resource Identifier:** ProgName=srcmstr UserName = root
biod Daemon Activity	**Description:** The **biod** daemon handles block I/O requests for the NFS file system. In order for NFS to function on a node, at least one **biod** daemon must be active. For normal NFS activity, six to eight **biod** daemons are usually active on a node. For higher NFS activity, some nodes may have more. These daemons are started from node boot, run continuously, and should not shut down. If any daemons shut down, consult the NFS documentation for diagnostic procedures, and attempt to restart the daemon. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=biod UserName = root
portmap Daemon Activity	**Description:** This daemon knows all the registered ports on the node, and which programs are available on each of these ports. The daemon's task is to convert Remote Procedure Call (RPC) port numbers to Internet port numbers. RPC clients use this daemon to resolve their RPC port numbers. If the daemon fails, the daemon itself and all RPC servers on the node must be restarted. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=portmap UserName = root

Table 1 (Page 9 of 10). Details About Each Condition to Monitor	
Condition	**Details**
xntpd Daemon Activity	**Description:** This daemon is active on a node when the Network Time Protocol (NTP) time synchronization protocol is running. This daemon ensures that the node's time-of-day hardware is synchronized with the network's time server. A failure in the daemon does not necessarily mean that the time of day hardware on the node will no longer be synchronized with the network, although this danger does exist. A failure in the daemon does mean that time change updates from the network server will not be made on this node. Such problems can lead to failures in RSCT's Topology Services component, which may begin to see packets arriving out of chronological order, and may cause RSCT to falsely detect that one of its peer nodes has failed. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=xntpd UserName = root
kerberos Daemon Activity	**Description:** The **kerberos** daemon executes on the node where the Kerberos databases are stored. You need to know which node this is to properly check this condition. The daemon is responsible for accepting Kerberos client requests for principal information, service tickets, and Kerberos database maintenance. Failure in this daemon will cause failures in Kerberos clients to acquire or validate credentials, which will lead to denial of service for users of the Kerberos clients. If this daemon fails, consult Chapter 9, “Diagnosing Authentication Problems” on page 95 for Kerberos diagnostics, and attempt to restart the daemon. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=kerberos UserName = root
hatsd Daemon Activity	**Description:** This is the RSCT Topology Services daemon, which is responsible for maintaining an internal topology map of the SP system on this node. The daemon is under SRC control, and should restart automatically if it is accidentally terminated. If this daemon fails and does not restart, the node will be seen as "down" by all other nodes in this system partition. Other consequences of this daemon's failure to restart include the RSCT Group Services daemon on the node will fail and the RSCT Event Management daemon will fail. This daemon's status cannot be monitored by the SP Event Perspective or Problem Management, because these two facilities depend on the daemon for their own processing. To check this daemon's activity, you must use the **lssrc -g hats** command or the **ps -ef | grep hats** command.
hagsd Daemon Activity	**Description:** This is the RSCT Group Services daemon, which is responsible for handling Group Services functions for all Group Services clients on this node. The daemon is under SRC control, and should restart automatically if it is accidentally terminated. If this daemon fails and does not restart, all Group Services clients on this node will appear to have failed, as far as the Group Services group members are concerned. Those groups will begin their member failure processing for the Group Services clients on this node. The daemon's status cannot be monitored by the SP Event Perspective or Problem Management, because these two facilities depend on the daemon for their own processing. To check this daemon's activity, you must use the **lssrc -g hags** command or the **ps -ef | grep hags** command.

Table 1 (Page 10 of 10). Details About Each Condition to Monitor	
Condition	**Details**
haemd Daemon Activity	**Description:** This is the RSCT Event Management daemon, which is responsible for handling Event Management registrations on this node and communicating with other Event Management daemons in this system partition. The daemon is under SRC control, and should restart automatically if it is accidentally terminated. If this daemon fails and does not restart, none of the Event Management resource variables from this node will be available to Event Management applications for monitoring or event generation purposes. These affected applications include Problem Management and the SP Event Perspective. This daemon's status cannot be monitored by the SP Event Perspective or Problem Management, because these two facilities depend on the daemon for their own processing . To check this daemon's activity, you must use the **lssrc -g haem** command or the **ps -ef | grep haem** command.
sdrd Daemon Activity	**Description:** This daemon runs on the Control Workstation (and therefore must be checked only on that node), and services all requests made of the System Data Repository (SDR). Although a failure in this daemon may not have any immediate consequences, PSSP software services will not be able to access SDR information, and can fail at later times when this information is needed. Certain hardware monitoring capability can also be lost, and may result in widespread, falsely detected "node not responding" failures. **Resource Variable:** Prog.pcount **Resource Identifier:** ProgName=sdrd UserName = root

Preparing to Examine and Monitor this Information

This section describes how to define the above conditions to the SP Event Perspective and how to define event definitions associated with these conditions.

SP Event Perspective — Conditions That You Can Monitor Using the Default Event Definition

The following conditions have a default event definition which you can install using the SP Event Perspective.

Table 2 (Page 1 of 2). Conditions and Default Event Definitions	
Condition to Monitor	**Default Event Definition**
Frame Power	framePowerOff
Frame Controller Responding	frameControllerNotResponding
Switch Power	switchPowerLED
Node Power	nodePowerLED
Node controller Responding	hostResponds
Node H/W Environment	nodeEnvProblem
Node Key Switch	keyNotNormal
Node LED/LCD Readout	LCDhasMessage

Table 2 (Page 2 of 2). Conditions and Default Event Definitions	
Condition to Monitor	**Default Event Definition**
Node Reachable by RSCT	nodeNotReachable
/tmp File System Filling	tmpFull
/var File System Filling	varFull
Page Space Low	pageSpaceLow
Switch Input Errors	switchPacketInputDropped switchPacketInputErrors switchPacketNoBuffers
Switch Output Errors	switchPacketOutputDropped switchPacketOutputErrors switchPacketOverflow
inetd Daemon Activity	Monitor_inetd_daemon
sdrd Daemon Activity	sdrDown

Each of the above conditions is to be monitored in all locations, with the exception of **sdrd** Daemon Activity, which is to be monitored only on the control workstation. Each default event definition contains a definition for the condition, so you need only load and register the event definition. Use the following procedure for each of the event definitions above:

1. Bring up the SP Event Perspective.
2. Click in the Event Definitions pane.
3. Select Actions → Load Defaults from the menu bar.

   This opens the Load Default Event Definitions dialog box.
4. Select the default event definition of interest from the list.
5. If you wish to register the selected event definition, click **Register the selected event definitions**.
6. Click **OK** to load the selected default event definition and close the dialog box.

SP Event Perspective — Conditions That You Can Monitor That You Must Define to the Event Perspective

This section contains instructions for using the Event Perspective to create the conditions that do not have default definitions.

To start:

1. Bring up the Event Perspective
2. If the Conditions pane is hidden, select View → Add Pane from the menu bar, to add the Conditions pane.
3. Click in the Conditions pane.
4. Select Actions → Create from the menu bar.
5. This displays the Create Conditions Notebook.

Now use the Event Perspective to create a condition for each of the conditions below, as follows:

- Frame Controller ID Mismatch - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, frameControllerIDMismatch.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Frame.controllerIDMismatch.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Frame Temperature - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, frameTempOutOfRange.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select tempRange.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Frame Node Slot Failure - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, frameSlotFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Frame.nodefail.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
  8. Repeat the above steps for each IBM.PSSP.SP_HW.Frame.nodefail* resource variable.
- Switch Power Shutdown - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, switchShutdownTemp.
  2. Type a description of the condition.

3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.switch.shutdownTemp.
5. In the **Event expression** field, enter X==1
6. Leave the **Rearm expression** field blank.
7. Press the **Create** button.

- Switch Hardware Environment Indicator - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, switchEnvLED.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Switch.envLED.
  5. In the **Event expression** field, enter X>0
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Switch Temperature - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, switchTemp.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Switch.tempRange.
  5. In the **Event expression** field, enter X==1
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.
- Node Temperature - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, nodeTemp.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.SP_HW.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.SP_HW.Node.tempRange.
  5. In the **Event expression** field, enter X>0
  6. Leave the **Rearm expression** field blank.
  7. Press the **Create** button.

- **/** File system Filling - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, rootFull.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.aixos.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.%totused.
  5. In the **Event expression** field, enter X>90
  6. In the **Rearm expression** field, enter X<80
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter LV=hd4;VG=rootvg.
  8. Press the **Create** button.
- Kernel Memory Buffer Failures - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, mbufFailures.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.aixos.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.Mem.Kmem.failures.
  5. In the **Event expression** field, enter X>X@P
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter Type=mbuf.
  8. Press the **Create** button.
- **srcmstr** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, srcmstrFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.xpcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=srcmstr;UserName=root
  8. Press the **Create** button.
- **biod** Daemon Activity - create a condition by following these steps:

1. In the **Name** field, enter a name that describes the condition. For example, biodFailure.
2. Type a description of the condition.
3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
5. In the **Event expression** field, enter X@0<X@1
6. Leave the **Rearm expression** field blank.
7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=biod;UserName=root
8. Press the **Create** button.

- **portmap** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, portmapFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=portmap;UserName=root
  8. Press the **Create** button.
- **xntpd** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, xntpFailure.
  2. Type a description of the condition.
  3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
  4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
  5. In the **Event expression** field, enter X@0<X@1
  6. Leave the **Rearm expression** field blank.
  7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=xntpd;UserName=root
  8. Press the **Create** button.
- **kerberos** Daemon Activity - create a condition by following these steps:
  1. In the **Name** field, enter a name that describes the condition. For example, kerberosFailure.

2. Type a description of the condition.
3. Scroll down in the **Resource variable classes** list and select IBM.PSSP.Prog.
4. Scroll down in the **Resource variable names** list and select IBM.PSSP.aixos.FS.pcount.
5. In the **Event expression** field, enter X@0<X@1
6. Leave the **Rearm expression** field blank.
7. In the **Resource ID Elements to be Fixed for the Condition** field, enter ProgName=kerberos;UserName=root
8. Press the **Create** button.

SP Event Perspective — Creating the Event Definitions

Create one event definition for each of the conditions listed above. Most of these will specify that the conditions be checked in all locations, while some are on specific nodes. Those that will be checked in all locations are:

- Frame Controller ID Mismatch
- Frame Temperature
- Switch H/W Environment Indicator
- Switch Temperature
- Node Temperature
- **/** file system becoming full
- Kernel Memory Buffer Failures
- **srcmstr** Daemon Activity
- **portmap** Daemon Activity

Conditions to Monitor Only in Specific Locations

Conditions to be monitored only in specific locations are:

- Frame Node Slot Failures - The slots to be monitored are those slots where nodes are directly connected in the frame.

  Those slots that cannot be used because wide or high nodes occupy previous slots are not to be monitored. You will need to identify which slots in which frames are to be monitored, and you can create one event per slot.

  When creating the event definition, select the proper "nodefail" resource name for the slot, and select the range of frame numbers where this slot is in use.

- **biod** Daemon Activity - This daemon is monitored only for nodes that use the NFS file system.

  When creating the event definition, select those nodes where NFS is used.

- **xntpd** Daemon Activity - This daemon is monitored only for nodes that use NTS time service.

  When creating the event definition, select only those nodes where NTS is used.

- **kerberos** Daemon Activity - This daemon is monitored only on nodes where the Kerberos database resides.

When creating the event definition, specify only the nodes where the Kerberos database resides.

Creating an Event Definition Using the Create Event Definition Notebook

On the Definition Page:

1. Give the Event Definition a descriptive name. For example, you can use rootFullEvent for the rootFull condition.
2. Click the down-arrow button for **Name** in the **Condition** box, and find the condition you want to use.
3. For those conditions that are to be observed in all locations, click the **Wild Card Element** selector under **Specify remaining resource ID elements** box.
4. For those that are to be observed in specific locations, select those locations under **Element Values**.

On the **Notifications** page, make sure the **Notification** (Get notified of events during the Event Perspective session) button is selected.

On the **Actions** page:

1. Click the **Take these actions when the event occurs** button.
2. In the **Run command** field, enter the following:

```
'echo $PMAN_HANDLE Alert: Instance Vector $PMAN_IVECTOR |
mail -s "$PMAN_HANDLE Alert" username@address'
```

Then click the **Create** button at the bottom of the page.

SP Hardware Perspective

The SP Hardware Perspective can be used to investigate further when the SP Event Perspective detects a hardware problem. This Perspective can show you the node's LED/LCD values (the Event Perspective cannot), and can be used to control the hardware or reset nodes when necessary.

Problem Management

Problem Management can be used as an alternative to the SP Event Perspective, to define event definitions that monitor conditions on your SP system.

Ensuring the Conditions You Intend to Monitor Are Known to Problem Management

The SP Event Perspective uses Problem Management whenever the event definition indicates that a command should be executed when the event occurs. If you used the SP Event Perspective earlier to set up the event definitions and conditions, these definitions may already be known to Problem Management. Use the **pmanquery** command to check if the definition already exists. For example:

```
pmanquery -n varFullEvent
```

will check for the **varFullEvent** definition.

If the event definitions are not known, use the **pmandef** command to define them. Any event definitions made using the **pmandef** command are also usable by the SP Event Perspective later on. To create new event definitions, use the **pmandef**

command. It is best to write a shell script with all the **pmandef** commands in it, and then execute the shell script.

When the **pmandef** commands complete, the event definitions will be registered, which means that Problem Management will begin to check for these conditions immediately. If this is a problem, issue the **pmandef -d** command to disable the event until you are ready to monitor the condition, then use the **pmandef -a** command to activate it.

For more information about problem management, see the chapter on using the Problem Management subsystem in *PSSP: Administration Guide*. For more information about the **pmandef** command, see the manpage for the command in *PSSP: Command and Technical Reference*.

pmandef Commands for Specific Conditions

Use these **pmandef** commands to set up event definitions for the following conditions. The field *username* is the user ID to receive the notification, and the *address* is a hostname.

- Frame Power

```
pmandef -s framePowerEvent \
-e 'IBM.PSSP.SP_HW.Frame.frPowerOff:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Frame Controller Responding:

```
pmandef -s frameCtrlRespondingEvent \
-e 'IBM.PSSP.SP_HW.Frame.controllerResponds:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Frame Controller ID Mismatch:

```
pmandef -s frameCtrlIDMismatchEvent \
-e 'IBM.PSSP.SP_HW.Frame.controllerIDMismatch:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Frame Temperature:

```
pmandef -s frameTempEvent \
-e 'IBM.PSSP.SP_HW.Frame.tempRange:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Frame Node Slot Failure:

```
pmandef -s frameSlot1FailureEvent \
-e 'IBM.PSSP.SP_HW.Frame.nodefail1:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Frame Switch Slot Failure:

```
pmandef -s frameSwitchSlotFailEvent \
-e 'IBM.PSSP.SP_HW.Frame.nodefail17:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Power:

```
pmandef -s switchPowerEvent \
-e 'IBM.PSSP.SP_HW.Switch.powerLED:Frame=*:X!=1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Hardware Environment Indicator:

```
pmandef -s switchLEDEvent \
-e 'IBM.PSSP.SP_HW.Switch.envLED:Frame=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Shutdown Due To Extreme Temperature:

```
pmandef -s switchTempShutdownEvent \
-e 'IBM.PSSP.SP_HW.Switch.shutdownTemp:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Temperature:

```
pmandef -s switchTempEvent \
-e 'IBM.PSSP.SP_HW.Switch.tempRange:Frame=*:X==1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node Power:

```
pmandef -s nodePowerEvent \
-e 'IBM.PSSP.SP_HW.Node.powerLED:NodeNum=*:X!=1' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node Controller Responding:

```
pmandef -s hostRespondsEvent \
-e 'IBM.PSSP.SP_HW.Node.hostResponds:NodeNum=*:X==0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node Hardware Environment Indicator:

```
pmandef -s nodeLEDEvent \
-e 'IBM.PSSP.SP_HW.Node.envLED:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node Temperature:

```
pmandef -s nodeTempEvent \
-e 'IBM.PSSP.SP_HW.Node.tempRange:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node Key Mode Switch Position:

```
pmandef -s nodeKeySwitchNotNormalEvent \
-e 'IBM.PSSP.SP_HW.Node.keyModeSwitch:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node LCD or LED Readout:

```
pmandef -s nodeLCDMessageEvent \
-e 'IBM.PSSP.SP_HW.Node.LCDhasMessage:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Node Reachable by RSCT Group Services:

```
pmandef -s nodeNotReachableEvent \
-e 'IBM.PSSP.Membership.Node.state:NodeNum=*:X>0' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **/tmp** File system Filling:

```
pmandef -s tmpFullEvent \
-e 'IBM.PSSP.aixos.FS.%totused:Nodenum=*;\
VG=rootvg;LV=hd3:X>90'\
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **/var** File system Filling:

```
pmandef -s varFullEvent \
-e 'IBM.PSSP.aixos.FS.%totused:Nodenum=*;\
VG=rootvg;LV=hd9var:X>90" -r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **/** File system Filling:

```
pmandef -s rootFullEvent \
-e 'IBM.PSSP.aixos.FS.%totused:Nodenum=*;VG=rootvg;LV=hd4:X>90'\
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Paging Space Low:

```
pmandef -s pageSpaceLowEvent \
-e 'IBM.PSSP.aixos.PagSp.%totalused:NodeNum=*:X>90' \
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Kernel Memory Buffer Failures:

```
pmandef -s mbufFailureEvent \
-e 'IBM.PSSP.aixos.Mem.Kmem.failures:NodeNum=*;Type-mbuf:X>X@P'\
-r "X<80" \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Input Errors:

```
pmandef -s switchInputErrEvent \
-e 'IBM.PSSP.CSS.ierrors:NodeNum=*:X>X@P' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Output Errors:

```
pmandef -s switchOutputErrEvent \
-e 'IBM.PSSP.CSS.oerrors:NodeNum=*:X>X@P' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- Switch Transmit Queue Overflows:

```
pmandef -s switchInputErrEvent \
-e 'IBM.PSSP.CSS.xmitque_ovf:NodeNum=*:X>X@P' \
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **inetd** Daemon Activity:

```
pmandef -s inetdFailureEvent \
-e 'IBM.PSSP.Prog.xpcount:NodeNum=*;\
ProgName=inetd;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **srcmstr** Daemon Activity:

```
pmandef -s srcmstrFailureEvent \
-e 'IBM.PSSP.Prog.xpcount:NodeNum=*;\
ProgName=srcmstr;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **biod** Daemon Activity:

```
pmandef -s biodFailureEvent \
-e 'IBM.PSSP.Prog.pcount:NodeNum=*;\
ProgName=biod;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **portmap** Daemon Activity:

```
pmandef -s portmapFailureEvent \
-e 'IBM.PSSP.Prog.pcount:NodeNum=*;\
ProgName=portmap;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **xntpd** Daemon Activity:

```
pmandef -s xntpdFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=range of nodes where xntpd daemon runs;\
ProgName=xntpd;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

- **kerberos** Daemon Activity:

```
pmandef -s kerberosdFailureEvent \
-e 'IBM.PSSP.Prog.pcount:\
NodeNum=node where kerberos daemon runs;\
ProgName=kerberos;UserName=root:X@0<X@1'\
-c 'echo $PMAN_HANDLE Alert: Location Information $PMAN_IVECTOR\
| mail -s "$PMAN_HANDLE Alert" username@address' \
-h local
```

Chapter 4. Error Logging Overview

This section describes methods for managing the AIX Error Log on the SP system. Commands and SMIT interfaces for performing general log management, and for managing Syslog and the AIX Error Log, are also installed with the **sysman** option. Refer to the Error Log Management chapter in *PSSP: Administration Guide* for an overview.

Error logging is the writing of information to persistent storage to be used for debugging purposes. This type of logging is for subsystems that perform a service or function on behalf of an end user. The subsystem does not communicate directly with the end user and, therefore, needs to log events to some storage location. The events that are logged are primarily error events.

Error logging for the SP uses Berkley Software Distribution (BSD) syslog and AIX Error Log facilities to report events on a per node basis. The intent is to have the AIX Error Log be the starting point for diagnosing system problems. The AIX Error Log is named **/var/adm/ras/errlog**.

The BSD system log has a default name of **/var/adm/SPlogs/SPdaemon.log**. It is controlled by the **syslogd** daemon. This log must be created by issuing the **syslogd** command. For instructions on using **syslogd**, refer to the manpage for the **syslogd** command, or *AIX Version 4.3 Commands Reference*.

Error log entries include a "DETECTING MODULE" string that identifies the software component, module name, module level, and line of code or function that detected the event that was logged. The information is formatted depending on the logging facility the user is viewing. For example the AIX Error Log facility information appears:

```
DETECTING MODULE
LPP=LPP name Fn=file name SID_level_of_the_file L#=Line number
```

The BSD syslog facility information appears:

```
timestamp, hostname, ID, PID
LPP=LPP name Fn=file name SID_level_of_the_file L#=Line number
```

Classifying Error Log Events

The following table displays the mapping of SP error log label suffixes to syslog priorities and AIX Error Log error types.

Table 3 (Page 1 of 2). SP Error Log Label Suffixes Mapped to syslog Priorities and AIX Error Log Types

Error Label Suffix	syslog Priority Field	syslog Description	AIX Error Log Error Type	AIX Error Log Description
EM	LOG_EMERG	Emergency, system unstable	PEND	The loss of availability of a device is imminent.
ER	LOG_ERR	Error condition	PERM	No recovery from this condition. A permanent error occurred.

Table 3 (Page 2 of 2). SP Error Log Label Suffixes Mapped to syslog Priorities and AIX Error Log Types				
Error Label Suffix	**syslog Priority Field**	**syslog Description**	**AIX Error Log Error Type**	**AIX Error Log Description**
ST	LOG_NOTICE	Normal, but significant condition	UNKN	It is not possible to determine the severity of the error.
TR	LOG_INFO	Informational message	UNKN	It is not possible to determine the severity of the error.
RE	LOG_DEBUG	Debug message	TEMP	Condition was recovered after several unsuccessful attempts.
DE	LOG_DEBUG	Debug message	UNKN	It is not possible to determine the severity of the error.

Effect of Not Having a Battery on Error Logging

In a typical RS/6000, a battery is installed to maintain NVRAM. On an SP system there is no battery and NVRAM may be lost when the node is powered off. AIX writes the last error log entry to NVRAM. During system startup the last entry is read from NVRAM and placed in the error log when the errdemon is started. This last error log entry may be important in diagnosis of a system failure.

On SP wide nodes the NVRAM does have power to it as long as the node is plugged into the frame and the frame is plugged into a working power source. On SP thin nodes, NVRAM is lost whenever the node is powered down. If the last error log entry is desired, the thin nodes should not be powered off. They should be re-IPLed in the “normal” key mode switch position if at all possible.

Managing and Monitoring the Error Log

To manage and monitor the error log, you can do the following:

- View error log information in parallel.
- View SP switch error log reports.
- Use AIX error log notification.

Viewing Error Log Information in Parallel

It may be helpful when diagnosing a system problem to look at all of the error logs at once in parallel.

It is not a good idea to copy the **/var/adm/ras/errlog** files from the various nodes to a central place and then run **errpt** against the combined file. First, copying time is added to the sequential processing time of all the nodes and the total time required will be longer than viewing the logs in parallel. Second, error log analysis requires per node information from the ODM database (on each node).

Use the **dsh** command with the **errpt** command and its options to view the error log. Perform the following steps:

1. View the summary information for all nodes to determine which ones are to be examined more closely. For example:

   ```
   dsh -a errpt -s 0930020094 |pg
   ```

In this example, all error entries that occurred after September 30, 1994 at 2 a.m. for every node defined in the System Data Repository, are listed. The output is piped to **pg** in a one entry per line format.

2. Pick out the nodes that have error entries that require further examination.
3. View the selected nodes. For example:

   ```
   dsh -w host1,host2,host3 errpt -a -s 0930020094 > /tmp/930errors
   ```

   This example collects all the fully expanded error log reports after September 30, 1994 at 2 a.m. from nodes with a hostname of **host1, host2, host3**.

Summary Log for SP Switch and SP Switch Adapter Errors

For systems running PSSP 3.1 or higher, a centralized error log has been added to record information about SP Switch and SP Switch adapter errors. Logging of switch and adapter errors in the AIX error log on nodes and on the control workstation causes the generation of a summary record in the summary log. This log has the name: **/var/adm/SPlogs/css/summlog** and is located on the control workstation. The summary log provides a centralized location for monitoring system-wide error activity. It also improves the usability of log output collected from individual nodes.

The summary log contains one summary entry for each CSS error log entry recorded on the failing node or control workstation. Entries in the log have the following fields, which are separated by blanks:

- **Timestamp** - A timestamp in the form: MMDDhhmmYYYY.
- **Node** - The *reliable_hostname* as stored in the SDR, for the originating node, with the domain portion removed.
- **Snap indicator** - A value that indicates whether a snap dump was taken. **Y** indicates that a snap dump was taken, **N** indicates that no snap dump was taken.
- **Partition** - The name of the system partition to which the node belongs.

  For error log entries that do not pertain to a particular system partition, this field contains **global**.
- **Index** - The error log index for the entry being reported.
- **Label** - The error log entry label field for the entry being reported.

The summary log contains a record for each CSS error log entry produced on each node in the system. You can use this log to obtain a single image of error activity across the entire SP system. Using the log, you can identify situations involving multiple nodes and determine the nodes that are affected. You can use the timestamps to determine which node experienced a problem first, so that you can more easily identify the root cause of a problem.

Viewing SP Switch Error Log Reports

Enter the following command to view all the SP switch adapter error reports in parallel:

```
dsh -a errpt -a -N css
```

It sends to stdout all the fully-formatted error log entries for all unusual status detected for the switch adapter device drivers that are contained in the error log.

This may be for the past 90 days. AIX has a default **crontab** entry that removes all hardware error entries after 90 days and all software error entries after 30 days.

Enter the following command to view all the SP switch information in parallel:

```
dsh -a errpt -a -N Worm
```

It sends to stdout all the fully-formatted error log entries for the switch. This includes errors found during switch diagnostics.

Using the AIX Error Log Notification Facility

You can be notified of an SP error when it occurs by using the AIX Error Notification Facility.

IBM General Concepts and Procedures for RS/6000 (GC23-2202) explains how to use the AIX Error Notification Facility. *IBM RS/6000 Problem Solving Guide* (SC23-2204) explains the use of the AIX Error Log. This facility will perform an ODM method defined by the administrator when a particular error occurs or a particular process fails. The following classifications of errors can have notification objects defined by the administrator. Many of these messages will not occur often, so these notification objects can be defined even for large SP systems.

1. **PSSP AIX Error Log Labels that end in _EM.**

   The EM suffix signifies an emergency error and is usually used to tell the administrator information that would be needed to re-IPL a node. To find these messages, issue:

   ```
   errpt -t |grep "_EM "
   ```

2. **Any AIX Error Log entries that have an Error Type of PEND.**

   PEND signifies an impending loss of availability and that action will soon be required of the administrator.

3. **Any AIX Error Log entries for the boot device of the node.**

   The boot device of the node usually has a resource name of **hdisk0**, but the name may vary if the installation has been customized.

4. **The AIX Error Label EPOW_SUS.**

   The EPOW_SUS error log entry is generated prior to power down when an unexpected loss of electrical power is encountered.

5. **The AIX Error Labels KERN_PANIC and DOUBLE_PANIC.**

   KERNEL_PANIC or DOUBLE_PANIC error log entries are generated when a kernel panic occurs.

The examples on the following pages may help the administrator in adding Error Notification Objects on the SP. Adding a **dsh -a** command to the ODM commands will perform the action on all nodes of the SP system.

Example 1

Mail the error report to **root@controlworkstation** when a switch adapter fails online diagnostics.

- **Step 1**. Set up directories for the Error Notification objects and methods.

  ```
  mkdir /customerdefinedpath/errnotify/objects
  mkdir /customerdefinedpath/errnotify/methods
  ```

Keep the methods scripts on each node so you can run them if distributed file system problems occur. File Collections is an excellent way to keep these scripts updated. The object files may be in a distributed file system since they are not used unless changes to the object are required.

- **Step 2**. Create the Error Notification Method scripts.

  Create a script or program that will be run when the error occurs. For example:

```
#!/bin/ksh
#############################################################################
# Run errpt to get the fully expanded error report for the error
# that was just written and redirect to a unique tempfile with the PID
#of this script.
#############################################################################
errpt -a -l $1 > /tmp/tempfile.$$
#############################################################################
# Mail the fully expanded error report to root@controlworkstation
# This could be anywhere in the network.
# root@controlworkstation is the user and hostname that the
# administrator wants to be notified at.
#############################################################################
mail root@controlworkstation < /tmp/tempfile.$$
```

- **Step 3**. Create the Error Notification Object

  Create a file that contains the Error Notification Object to catch the switch diagnostic failed error.

```
errnotify:
          en_name = "tbx_diagerr.obj"
          en_persistenceflg = 1
          en_label = "SWT_DIAG_ERROR2_ER"
          en_method = "/customerdefinedpath/methods/errnot.
                      test.ksh$1"
```

  (The **en_name** value can be a maximum of 16 characters long.) Enter the **odmshow errnotify** command to view the Error Notification object.

  It is easy to modify an existing set of ODM errnotify stanzas. To do this, enter:

```
odmget errnotify > file
```

  and edit the file. Include only attributes that have values.

- **Step 4.** Add the Error Notification Object to the errnotify class.

```
odmadd  /customerdefinedpath/object/tbx_diagerr.obj
```

  (The file name is the name of the file with the Error Notification Object in it.)

  To delete this object, enter:

```
odmdelete -o errnotify -q "en_name = tbx_diagerr.obj"
```

  To view this object in the ODM database, enter:

```
odmget  -q "en_name = tbx_diagerr.obj" errnotify
```

- **Step 5.** The following mail will be sent to **root@controlworkstation** when an SP Switch MX adapter fails diagnostics:

```
From root@sp2n5.kgn.ibm.com Mon Oct  3 11:25:59 1994
Received: from sp2n5.kgn.ibm.com by ppsras.kgn.ibm.com
          (AIX 3.2/UCB 5.64/4.03)
          id AA24781; Mon, 7 May 1995 10:14:59 -0400
Date: Mon, 3 Oct 1994 11:25:59 -0400
From: root
Message-Id: <9410031525.AA24781@sp2n5.kgn.ibm.com>
To: root
Status: RO

 ---------------------------------------------------------------------------
ERROR LABEL: SWT_DIAG_ERROR2_ER
ERROR ID: 323C48A0

Date/Time:       Mon Oct  3 11:25:57
Sequence Number: 18282
Machine Id:      000004911800
Node Id:         sp2n5
Error Class:     H
Error Type:      PERM
Resource Name:   Worm
Resource Class:  NONE
Resource Type:   NONE
Location:        NONE

Error Description
Switch adapter failed On-Line diagnostics

Probable Causes
Switch clock signal missing
Switch adapter failure

User Causes
Switch cable loose or disconnected

Recommended Actions
Run adapter diagnostics

Failure Causes
Switch adapter hardware

Recommended Actions
Run adapter diagnostics

Detail Data
DETECTING MODULE
LPP=PSSP,Fn=dtb3mx,SID=1.35,L#=1303,
Service Request Number
763–942
```

Example 2

Error Notification when any Error Type of PEND occurs.

- **Steps 1 and 2** are the same as defined in the switch diagnostic failure example.
- **Step 3.** Create the Error Notification Object

Create a file that contains the Error Notification Object to catch the pending availability problems. For example:

```
errnotify:
        en_name = "errnot.PEND.obj"
        en_persistenceflg = 1
        en_type = "PEND"
        en_method = "/tmp/errnot.test.ksh $1"

errnotify:
        en_name = "errnot.pend.obj"
        en_persistenceflg = 1
        en_type = "pend"
        en_method = "/tmp/errnot.test.ksh $1"

errnotify:
        en_name = "errnot.Pend.obj"
        en_persistenceflg = 1
        en_type = "Pend"
        en_method = "/tmp/errnot.test.ksh $1"
```

(The variations of PEND are added because upper case is not strictly adhered to by all AIX LPPs and vendors.)

- **Step 4.** Add the Error Notification Objects to the errnotify class. For example:

  ```
  odmadd  /customerdefinedpath/object/errnot.pend.obj
  ```

  (The file name is the name of the file with the Error Notification Object in it.)

  To delete these objects enter:

  ```
  odmdelete -o errnotify -q "en_name = errnot.PEND.obj"
  odmdelete -o errnotify -q "en_name = errnot.pend.obj"
  odmdelete -o errnotify -q "en_name = errnot.Pend.obj"
  ```

  To view this object in the ODM database, enter:

  ```
  odmget  -q "en_name = errnot.PEND.obj" errnotify
  odmget  -q "en_name = errnot.pend.obj" errnotify
  odmget  -q "en_name = errnot.Pend.obj" errnotify
  ```

- **Step 5**. Mail is sent to the administrator when an error that has an Error Type of PEND occurs.

Example 3

Error Notification when any Error on the boot device of **hdisk0** occurs.

- **Step 1 and 2** are the same as defined in "Example 1" on page 72.
- **Step 3**. Create the Error Notification Object.

  Create a file that contains the Error Notification Object to catch the boot disk errors. (Assumes **hdisk0** is the boot device.)

  ```
  errnotify:
          en_name = "errnot.boot.obj"
          en_persistenceflg = 1
          en_resource = "hdisk0"
          en_method = "/tmp/errnot.test.ksh $1"
  ```

- **Step 4**. Add the Error Notification Object to the errnotify class.

  ```
  odmadd  /customerdefinedpath/object/errnot.boot.obj
  ```

To delete this object, enter:

```
odmdelete -o errnotify -q "en_name = errnot.boot.obj"
```

To view this object in the ODM database, enter:

```
odmget  -q "en_name = errnot.boot.obj" errnotify
```

- **Step 5**. Mail with the fully expanded error report will be sent to the administrator when an error on **hdisk0** occurs.

Example 4

Error Notification when unexpected power loss and kernel panics occur.

- **Steps 1 and 2** are the same as defined in "Example 1" on page 72.
- **Step 3**. Create the Error Notification Object

  Create a file that contains the Error Notification Object to catch the kernel panic and power loss Error Labels. For example:

```
errnotify:
                en_name = "power.obj"
                en_persistenceflg = 1
                en_label = "EPOW_SUS"
                en_method = "/customerdefinedpath/methods/
                            errnot.test.ksh $1"

errnotify:
                en_name = "panic.obj"
                en_persistenceflg = 1
                en_label = "KERNEL_PANIC"
                en_method = "/customerdefinedpath/methods/
                            errnot.test.ksh $1"

errnotify:
                en_name = "dbl_panic.obj"
                en_persistenceflg = 1
                en_label = "DOUBLE_PANIC"
                en_method = "/customerdefinedpath/methods
                            /errnot.test.ksh $1"
```

- **Step 4.** Add the Error Notification Object to the errnotify class. For example:

```
odmadd  /customerdefinedpath/object/power.panic.obj
```

  The file name is the name of the file with the Error Notification Object in it.

- **Step 5**. Mail with the fully expanded error report will be sent to the administrator when any power loss or kernel panic occurs.

Using the SP Logs

The SP System uses the standard logs provided by both AIX and the public domain software it includes, as well as SP-specific logs. Some logs reside on the control workstation only, and some reside only on the SP nodes. Others reside on both. Table 4 on page 77 summarizes and shows the location of the logs to use when diagnosing SP problems. The abbreviation **CWS** stands for control workstation in this table.

Your IBM Support Center representative may ask you to provide information from these logs.

Table 4 (Page 1 of 5). SP Log Files

Type of Message	Log File Name	Location
Standard AIX error log entries including the SP Switch	**/var/adm/ras/errlog**	Nodes
Error messages and verbose messages from security programs.	**/var/adm/SPlogs/auth_install/log**	CWS, nodes
Automounter messages	**/var/adm/SPlogs/auto/auto.log**	CWS, nodes
Messages from the **cstartup** command	**/var/adm/SPlogs/cs/cstart.***timestamp.pid*	CWS
Messages from the **cshutdown** command	**/var/adm/SPlogs/cs/cshut.***timestamp.pid*	CWS
Details of recovery actions for the IBM RVSD function	**/var/adm/SPlogs/csd/vsd.debuglog**	Nodes
Summary of recovery actions for the IBM RVSD function	**/var/adm/SPlogs/csd/vsd.log**	Nodes
Trace output of **pssp_script**, which performs the post install customization	**/var/adm/SPlogs/css/$nim_client_shr.config.log**	Nodes
Switch cable miswire information	**/var/adm/SPlogs/css/cable_miswire**	Primary node
Messages from the **css.snap** script	**/var/adm/SPlogs/css/css.snap.log**	CWS, nodes
Switch admin daemon messages	**/var/adm/SPlogs/css/cssadm.debug**	CWS
Switch admin daemon messages (stdout)	**/var/adm/SPlogs/css/cssadm.stdout**	CWS
Switch admin daemon messages (stderr)	**/var/adm/SPlogs/css/cssadm.stderr**	CWS
Fault service daemon messages (stderr)	**/var/adm/SPlogs/css/daemon.stderr**	Nodes
Fault service daemon messages (stdout)	**/var/adm/SPlogs/css/daemon.stdout**	Nodes
System error messages that occurred while distributing the topology file to the nodes.	**/var/adm/SPlogs/css/dist_topology.log**	Primary node
Trace of adapter diagnostics failures	**/var/adm/SPlogs/css/dtbx_failed.trace**	Nodes
Adapter diagnostics trace information	**/var/adm/SPlogs/css/dtbx.trace**	Nodes
Messages from adapter diagnostics (stderr)	**/var/adm/SPlogs/css/dtbxworm.stderr**	Nodes
Log from the **Eclock** command	**/var/adm/SPlogs/css/Eclock.log**	CWS
Log from all **Ecommands** executed	**/var/adm/SPlogs/css/Ecommands.log**	CWS
Log from the **Emonitor** command	**/var/adm/SPlogs/css/Emonitor.log**	CWS

Table 4 (Page 2 of 5). SP Log Files

Type of Message	Log File Name	Location
Result of **Estart** commands issued by the **Emonitor** daemon	**/var/adm/SPlogs/css/Emonitor.Estart.log**	CWS
Trace of last **Eunpartition** operation	**/var/adm/SPlogs/css/Eunpart.file**	Primary node
Trace file of fault service daemon messages	**/var/adm/SPlogs/css/fs_daemon_print.file**	Nodes
Switch fault information	**/var/adm/SPlogs/css/flt**	Nodes
stdout and stderr of the Event Management Resource Monitor and Methods	**/var/adm/SPlogs/css/logevnt.out**	CWS
SP Switch advanced diagnostics Messages Daemon log	**/var/adm/SPlogs/css/msdg.log**	CWS
Description of problems that arise while switch is initializing	**/var/adm/SPlogs/css/out.top**	Primary node
Initialization messages from the SP Switch support code	**/var/adm/SPlogs/css/rc.switch.log**	Nodes
Log from switch router generation	**/var/adm/SPlogs/css/router.log**	Nodes
Output log from switch router generation when it detects a failure	**/var/adm/SPlogs/css/router_failed.log**	Nodes
SP Switch advanced diagnostics tests and architecture components log	**/var/adm/SPlogs/css/spd.trace**	CWS, nodes
SP Switch advanced diagnostics GUI log	**/var/adm/SPlogs/css/spd_gui.log**	CWS
Summary records for events logged to AIX error log on nodes.	**/var/adm/SPlogs/css/summlog**	CWS
stdout and stderr for the CSS logging daemon's Event Management client	**/var/adm/SPlogs/css/summlog.out**	CWS
Current state of the switch network, details about attached nodes and the topology file	**/var/adm/SPlogs/css/topology.data**	Primary node
Worm trace file from switch initialization	**/var/adm/SPlogs/css/worm.trace**	Primary node
Results of the last CSS verification test	**/var/adm/SPlogs/CSS_test.log**	CWS
Output of the **supper** command	**/var/adm/SPlogs/filec/sup***date.time*	Nodes
Actions **supper** performs when updating file collections	**/var/adm/SPlogs/filec/sup***date.time***r**	Nodes
Authentication database administration daemon	**/var/adm/SPlogs/kerberos/admin_server.syslog**	Primary authentication server

Table 4 (Page 3 of 5). SP Log Files

Type of Message	Log File Name	Location
Primary authentication server log	**/var/adm/SPlogs/kerberos/kerberos.log**	Primary authentication server
Secondary authentication server log	**/var/adm/SPlogs/kerberos/kerberos.slave_log**	Secondary authentication server
Authentication database propagation daemon	**/var/adm/SPlogs/kerberos/kpropd.log**	Secondary authentication server
Messages generated by the Problem Management daemon	**/var/adm/SPlogs/pman/pmand.log**	Nodes
Messages generated by the Problem Management daemon	**/var/adm/SPlogs/pman/pmand.***partition name***.log**	CWS
Hardware Monitor **s70d** daemon error messages	**/var/adm/SPlogs/spmon/s70d/s70d**.*frame*.**log**.*julian_date*	CWS
System Data Repository configuration messages	**/var/adm/SPlogs/sdr/SDR_config.log**	CWS
Output of the **SDR_test** command	**/var/adm/SPlogs/sdr/SDR_test.log**	CWS, nodes
System Data Repository error messages	**/var/adm/SPlogs/sdr/sdrdlog.***pid*	CWS
Login control messages	**/var/adm/SPlogs/spacs/spacs.log**	Nodes
SP configuration Vital Product Data	**/var/adm/SPlogs/SPconfig**	CWS, nodes
Vital Product Data output for the node	**/var/adm/SPlogs/SPconfig/***node number***.umcl**	CWS, nodes
lscfg -v command output	**/var/adm/SPlogs/SPconfig/***node number***.lscfg**	CWS, nodes
Messages generated by system daemons, including hardware errors	**/var/adm/SPlogs/SPdaemon.log**	CWS, nodes
SP extension node messages	**/var/adm/SPlogs/spmgr/spmgrd.log**	CWS
Hardware Monitor initialization and error messages	**/var/adm/SPlogs/spmon/hmlogfile.** *julian_date*	CWS
Node conditioning messages	**/var/adm/SPlogs/spmon/nc/nc.***frame.node*	CWS
Netfinity daemon (**nfd**) messages	**/var/adm/SPlogs/spmon/nfd/nfd.***frame***.log.***julian-date*	CWS
s70d messages	**/var/adm/SPlogs/spmon/s70d/s70d.***frame***.log.***julian-date*	CWS
Activity of the logging daemon (**splogd**).	**/var/adm/SPlogs/spmon/splogd.debug**	CWS
Contains the PID of the logging daemon, (**splogd**).	**/var/adm/SPlogs/spmon/splogd/splogd.pid**	CWS

Table 4 (Page 4 of 5). SP Log Files

Type of Message	Log File Name	Location
SP Logging Daemon state changes Note: this log is not shipped activated. To activate, see Action 10 of Chapter 14, "Diagnosing System Monitor Problems" on page 147.	**/var/adm/SPlogs/spmon/splogd.state_changes.** *timestamp*	CWS
Microcode download messages recorded when using smit (**smitty supervisor** command)	**/var/adm/SPlogs/spmon/ucode/ucode_log.***frame.node*	CWS
Output of the **spmon_ctest** command	**/var/adm/SPlogs/spmon_ctest.log**	CWS
Output of the **spmon_itest** command	**/var/adm/SPlogs/spmon_itest.log**	CWS
Job Switch Resource Table Services information and error messages	**/var/adm/SPlogs/st/st_log**	Nodes
Sysctl server log messages	**/var/adm/SPlogs/sysctl/sysctld.log**	CWS, nodes
AIX error messages from mirroring a root volume group using SP volume group commands	**/var/adm/SPlogs/sysman/mirror.out**	Nodes
System Management configuration messages	**/var/adm/SPlogs/sysman/***node*.**config.log**.*pid*	Nodes
System Management first boot configuration messages	**/var/adm/SPlogs/sysman/***node*.**configfb.log**.*pid*	Nodes
System Management console messages	**/var/adm/SPlogs/sysman/***node*.**console.log**	CWS, nodes
System Management configuration messages	**/var/adm/SPlogs/sysman/spfbcheck.log**	Nodes
AIX error messages from unmirroring a root volume group using SP volume group commands	**/var/adm/SPlogs/sysman/unmirror.out**	Nodes
Informational and error messages from the **SYSMAN_test** command	**/var/adm/SPlogs/SYSMAN_test.log**	CWS, nodes
hags internal trace and log file	**/var/ha/log/hags***	CWS, nodes
hagsglsm internal trace and log file	**/var/ha/log/hagsglsm***	CWS, nodes
Event Management activity log	**/var/ha/log/em.default.***partition-name*	CWS, nodes
Trace information for the topology services daemon	**/var/ha/log/hats.***dd.hhmmss.partition-name*	CWS, nodes
Information from the topology services startup script	**/var/ha/log/hats.***partition-name*	CWS, nodes
Hardmon resource monitor messages	**/var/ha/run/haem.***hostname***/IBM.PSSP.hmrmd/ IBM.PSSP.hmrmd_log.***julian-date*	CWS

Table 4 (Page 5 of 5). SP Log Files		
Type of Message	**Log File Name**	**Location**
filec_config command log, no longer in use	**/var/sysman/logs/***	Nodes
SP SNMP Agent messages	**/var/tmp/SPlogs/spmgr/spgrd.log**	CWS, nodes
SP TaskGuide activity messages (exit values, stdout and stderr from commands run by SP TaskGuides	*TaskGuide_name.timestamp***.tglog** View these logs through the SP TaskGuides themselves.	SDR system files

Log files are cleaned up on the nodes by the **cleanup.logs.nodes** command. Log files are cleaned up on the control workstation by the **cleanup.logs.ws** command. By default, continuously growing logs are trimmed to 400 lines every night and non-growing files are deleted after seven days. The exceptions are:

- The **supper** message log, which is deleted after two days.
- The SDR log, which is deleted after seven days, only if it is **not** the current log.
- The **SPdaemon.log**, which is trimmed.
- The SP extension node SNMP Manager log file size is controlled by the user when the daemon is started.
- The Event Management log, which is trimmed when it reaches 256KB.
- The **JSRT Services** log, which is truncated when it reaches 100KB.
- The Group Services logs, **/var/ha/log/hags*** and **/var/ha/log/hagsglsm/***, which are trimmed according to the LOGSIZE specified when starting Group Services.
- The Topology Services logs - three instances of the daemon logs are kept. Current logs are trimmed to a given number of lines. This number of lines is a tunable parameter stored in the **Log_Length** attribute of the **TS_Config** SDR class. The default value is 5000 lines. The startup script log is not trimmed, and the seven latest instances are kept.
- The following logs are not trimmed:
  - The css (switch) logs
  - The **sysctld** logs
  - The logging daemon log
  - The SP logging daemon state changes log
  - The Netfinity log
  - The **s70d** log
  - The hardmon resource monitor log
  - The **SYSMAN_test** log

Chapter 5. Producing a System Dump

Table 5. System Dump Information	
Symptom	**Recovery**
Nodes do not respond or system crashes	“Action 1. Produce a System Dump” “Action 2. Verify the System Dump” on page 84

Actions

Action 1. Produce a System Dump

When your nodes do not respond or when your system crashes, a system dump may help you determine the cause of the problem. A system dump contains a copy of the kernel data on the system at the time of the crash. This section explains how to produce a dump, verify it, copy the dump to tape, and send the tape to IBM.

In some cases the system produces a dump automatically. If the system senses a fatal condition, it usually dumps automatically to the primary dump device and puts flashing **888** in the node's three-digit display.

Attention

Do not initiate a system dump if the node's three-digit display is **888**. If you initiate a dump, you will overwrite the dump taken at the time of the problem.

Instead, proceed to “Action 2. Verify the System Dump” on page 84.

Dump Methods

There are several ways you can produce a system dump. Some of the methods work with all configurations and others do not. Each method explained here includes this configuration information.

Note

Use the Hardware Perspective to operate the system controls. You can reset the node or put it in service mode either from the Nodes Status page of the Node Notebook, or the Actions menu.

Dump to the Primary Dump Device

Choose one of these methods to produce a dump on the primary dump device.

Method 1: This method works for all systems that have a key switch.

Set the key mode switch to the Service position and press the Reset button once.

Method 2: This method can only be done from a directly-attached keyboard. It cannot be done from a **tty** connection. This method works only on the control workstation.

Set the key mode switch to the Service position and, while holding the Ctrl and Alt keys, press the 1 on the numeric key pad.

Method 3: This method works for all system configurations, if the system is responding to commands.

Login as **root** and enter:

```
sysdumpstart -p
```

Dump to the Secondary Dump Device

Choose one of these methods to produce a dump on the secondary dump device.

Note

If the secondary dump device is a removable media device, such as a tape or diskette drive, make sure that the medium is in the device.

Method 4: This method can only be done from a directly-attached keyboard. It cannot be done from a **tty** connection. This method works only on the control workstation.

Set the key mode switch to the Service position and, while holding down Ctrl and Alt keys, press the 2 on the numeric key pad.

Method 5: This method works for all system configurations, if the system is responding to commands.

Login as **root** and enter:

```
sysdumpstart -s
```

Action 2. Verify the System Dump

You may have a system dump because you initiated it yourself or because the system produced one automatically. In either case, follow these steps to verify that the system dump was successful and that the information it contains is usable.

1. Record the three-digit codes.
   - If the system dumped automatically, the three-digit display will show flashing **888**. Press Reset repeatedly until **888** displays again and write down each three-digit code that is displayed. The last code before **888** displays again indicates if the dump was successful. Check the dump code status in the next table for more information.
   - If you initiated the dump yourself, the three-digit code that is displayed indicates if the dump was successful. Check the dump code status in the next table for more information.

Table 6. System Dump Status Codes	
Three-digit code	**Meaning**
0c0	The dump completed successfully.
0c4	The dump device was too small but the dump may still be usable. If zero bytes are written and 0c4 is displayed, it means the dump device was large enough but the system was hung and not able to initiate dump.
0c5	The dump device was not accessible. No dump was taken.
0c6	Prompts you to make the secondary dump device available.
0c8	No dump device is defined.
0c9	The dump is in progress. Wait for completion.

2. Return the key mode switch to "normal" and re-IPL. This will allow the last error log entry stored in NVRAM to be placed in the error log. (See "Effect of Not Having a Battery on Error Logging" on page 70)
3. Log in as **root**.
4. Verify the dump and the dump device by entering:

   ```
   sysdumpdev
   ```

   This should return something like:

   ```
   primary            /dev/hd7
   secondary          /dev/sysdumpnull
   ```

   Note the primary dump device name, and substitute it for **/dev/hd#** in the following steps.
5. Verify the usability of the dump by entering:

   ```
   crash /dev/hd#
   ```

   This should return:

   ```
   Using /unix as the default namelist file.
   Reading in symbols.....................
   ```

   - If you get the message: "`ATTENTION: dumpfile does not appear to match namelist`", either the dump did not take place or the **/unix** file does not match the dump that was in the dump device.

     Enter **q** to quit the **crash** command.

     The dump file is not useful. Do not continue with these steps and do not send the dump to IBM.
   - If messages are not displayed, proceed with the next step.
6. Enter the **errdead** command to extract the error records from the **/dev/error** buffer and place them in the error log:

   ```
   /usr/lib/errdead /dev/hd#
   ```
7. When you see the **>** prompt, enter:

   ```
   stat
   ```

   It produces a status report similar to this sample:

```
sysname:   AIX
nodename:  journey
release:   2
version:   3
machine:   000052643100
time of crash:  Sun Jan 24 19:18:53 1993
age of system:  18 day, 1 hr., 29 min.
```

- If the **time of crash** in the output approximately matches the time the system crashed, the dump is sufficient for analysis. Continue with the next step.
- If the **time of crash** in the output does not approximately match the time the system crashed,

  Enter **q** to quit the **crash** command.

  The data is not useful. Do not continue with these steps and do not send the dump to IBM.

8. Enter:

   ```
   trace
   ```

   Look for a trace report similar to this sample:

   ```
   STACK TRACE:
      .m_freem ()
      .soreceive
      ._recv
      .recv
   ```

   Enter **q** to quit the **crash** command.

Gather the dump and other **snap** or log information for the IBM Support Center. Contact your local service representative or call the IBM Support Center to open a Problem Management Record as explained in “How To Contact the IBM Support Center” on page 14.

Chapter 6. Diagnosing Hardware and Software Problems

This section provides troubleshooting information for the SP hardware and software. It contains tables to help you isolate the cause of SP problems and recover from them. The first table describes the high level symptoms and gives you a course of action or directs you to other tables to further analyze the problem. Note that some of the tables in the diagnosing sections for the various components list recovery actions. Such actions are further described within the body of the section.

High-Level SP Symptoms

The following table lists the high-level symptoms you may experience and directs you to the corresponding chapter for each one.

Table 7 (Page 1 of 2). High-Level SP Symptoms

Symptoms	Condition	Action
controllerResponds indicator on the Frame Notebook in the Hardware Perspective is red, **SPdaemon.log** reports this variable as 0. or the Frame Notebook in the Hardware Perspective displays blank nodes.	Frame supervisor failure	Go to Chapter 7, “Diagnosing Frame Supervisor Communication Problems” on page 91
You get error messages when you start the switch or applications that use the switch fail or hang.	Switch Failure	Go to Chapter 11, “Diagnosing Switch Problems” on page 109
switchResponds indicator on the Nodes Noteboook in the Hardware Perspective is red, meaning that the node is not currently available to the switch network. When the indicator is yellow, the node is available to, but not part of, the network.	Switch or switch adapter problems	Go to Chapter 11, “Diagnosing Switch Problems” on page 109
The **spmon**, **hmmon**, or **hmcmds** commands fail.	System Monitor problem	Go to Chapter 14, “Diagnosing System Monitor Problems” on page 147
Perspectives is having trouble starting or running.	Perspectives problem	Go to Chapter 15, “Diagnosing SP Perspectives Problems” on page 151
Cannot monitor or operate hardware controls, cannot power frames, nodes, switch on or off.	Hardware problem or System Monitor problem	Go to Chapter 14, “Diagnosing System Monitor Problems” on page 147
The configuration indicated by the SDR and Perspectives do not match your system's configuration.	Frame Supervisor connectivity problem.	Go to Chapter 7, “Diagnosing Frame Supervisor Communication Problems” on page 91.
3DigitDisplay indicator in the Hardware Perspective displays three-digit codes. LED/LCD displays present for a given node.	Hardware or software problem	Check Chapter 37, “SP-Specific LED/LCD Values” on page 257 to see if the code is discussed there. If not, refer to *IBM RS/6000 Problem Solving Guide*, SC23-2204
Node hangs, cannot access system (**ping, rsh**) or node crashes with **888** in the *3DigitDisplay* indicator.	Hardware or software problem	Go to Chapter 5, “Producing a System Dump” on page 83

Table 7 (Page 2 of 2). High-Level SP Symptoms		
Symptoms	**Condition**	**Action**
You have an orange icon at any level of the System Monitor topology display.	System connectivity problem	Go to Chapter 13, “Diagnosing System Connectivity Problems” on page 145
Cannot **ping** on the external network, **rsh**, **telnet**, **rlogin** commands fail.	System connectivity problem	Go to Chapter 13, “Diagnosing System Connectivity Problems” on page 145
User access problems, cannot log in, password is invalid, cannot get to home directory.	Software problem	Go to Chapter 18, “Diagnosing User Access Problems” on page 167
You get an error message in response to an SP command on the control workstation.	Hardware or software problem	Look up the message in *PSSP: Messages Reference* and follow the action suggested.

Part 2. Diagnosing PSSP Subsystems

Chapter 7. Diagnosing Frame Supervisor Communication Problems

Table 8. Frame Supervisor Symptoms

Symptom	Recovery
The **controllerResponds** indicator on the Frame Status page of the Frame Notebook in the Hardware Perspective is red The **/var/adm/SPlogs/SPdaemon.log** reports that the Frame Controller is not responding	Check the RS-232 cable connection from the control workstation to the affected frames.
The **controllerResponds** indicator on the Frame status page of the Frame Notebook in the Hardware Perspective shows red.	"Action 1. Verify Parallel System Support Programs Installation" "Action 2. Verify RS-232 Connection" on page 92 "Action 3. Verify Serial Port Configuration" on page 92 "Action 4. Check the Frame Configuration" on page 92 "Action 5. Check the Log for Messages" on page 92 "Action 6. Check Frame Power" on page 92
The configuration indicated by the SDR and the Perspectives do not match your system's configuration.	"Action 2. Verify RS-232 Connection" on page 92 "Action 3. Verify Serial Port Configuration" on page 92 "Action 4. Check the Frame Configuration" on page 92
None of the other symptoms exist and the nodes have their power on	Call the IBM Support Center.

Actions

Action 1. Verify Parallel System Support Programs Installation

Run verification tests using SMIT or the command line to ensure installation is complete.

Using SMIT:

TYPE **smit SP_verify**

- The Installation/Configuration Menu appears.

SELECT The SSP System Management option.

PRESS Enter.

Using the command line, enter:

```
/usr/lpp/ssp/bin/SYSMAN_test
```

Action 2. Verify RS-232 Connection

Ensure that the RS-232 line is connected to the correct frame and the correct serial port on the control workstation.

Using SMIT:

TYPE **smit SP_verify**

- The Installation/Configuration Menu appears.

SELECT SSP System Monitor Configuration

PRESS Enter.

Using the command line, enter:

```
/usr/lpp/ssp/bin/spmon_ctest
```

If the RS-232 cables were not connected to the proper frames and you have already configured the frames using the **spframe** command, you must perform the following:

1. Use **spdelfram** to delete the affected frames *before* you re-cable the frames.
2. Recable the RS-232 to the proper frames.
3. Readd the frames using the **spframe** command.

Action 3. Verify Serial Port Configuration

Check the baud rate of the serial port. The System Monitor configures this for you when it starts. Use the **stty** command and redirect input from the correct serial port. For example:

```
stty < /dev/tty0
```

Action 4. Check the Frame Configuration

Use SMIT (**smit list_data**) or the **splstdata -f** command to verify that the proper serial port is specified for the frame. Check the values for your frames.

Action 5. Check the Log for Messages

Check the **/var/adm/SPlogs/SPdaemon.log** on the control workstation and respond to any messages. SP hardware problems will have a resource name of **sphwlog**. The same messages are found in **errpt**. To get full details of all SP hardware messages in **errpt**, enter:

```
errpt -aN sphwlog
```

(You should redirect the output of this command into a file; the output could be very large.)

Action 6. Check Frame Power

Visually check for power LEDs lit on any nodes, at frame power switch, and base power unit in the rear. If none are lit, you do not have power to the frame.

Chapter 8. Diagnosing SDR Problems

Table 9. System Data Repository (SDR) Symptoms	
Symptom	**Recovery**
Nonzero return codes	"Action 1. Get the Return Code"
Cannot connect to server	"Action 2. Analyze System or Network Changes"
Class corrupted or nonexistent	"Action 3. Analyze Class Situation" on page 94

Actions

Action 1. Get the Return Code

If you cannot run SDR commands, or a program that uses the SDR is failing when running SDR commands, get the return code or the message number from the failing SDR routine. The return codes from SDR routines are imbedded in the message numbers. The first four numbers in the SDR cataloged message are always **0025**, followed by a hyphen and a three-digit number. The three digit number is the return code. For example, the following SDR message is issued with a return code of 80 from any SDR routine that cannot connect to the SDR server:

```
0025-080 The SDR routine could not connect to server.
```

Some programs report the return code from an SDR routine, but not the message. Use **0025** and the return code to find the appropriate message in *PSSP: Messages Reference*.

Action 2. Analyze System or Network Changes

System or network changes could affect the SDR. If an SDR command fails to connect to the server, do the following:

1. Type **spget_syspar** on the node showing failing SDR commands.
2. If the **spget_syspar** command fails, check the **/etc/SDR_dest_info** file on the same node. It should have at least two records in it. These records are the **primary** and the **default** records. They should look like:

   ```
   primary: syspar_ip_address
   default: default_syspar_ip_address
   ```

   where *syspar_ip_address* is the address of the system partition that this node is in, or the default system partition if this file is on the control workstation; *default_syspar_ip_address* is the address of the default system partition. Note that these two addresses might be the same.

   If this file is missing or does not have these two records, the node may not be properly installed or the file has been altered or corrupted. You can edit the file that contains the two records above or copy the file from a working node in the same system partition.

   The **spget_syspar** command may also fail if:

- The value of the **SP_NAME** environment variable is a hostname (not an IP address)

  AND
- The system name server is not functioning properly.

3. If the **spget_syspar** command is successful, check to make sure that the address is also the address of a valid system partition. If it is, try to **ping** that address. If the **ping** fails, contact your system administrator to investigate a network problem.
4. If the value returned by the **spget_syspar** command is not the same as the address in the **primary** record of the **/etc/SDR_dest** information file, the **SP_NAME** environment variable is directing SDR requests to a different address. Make sure that this address (the value of the **SP_NAME** environment variable) is a valid system partition.
5. If the value of the **SP_NAME** environment variable is a hostname, try setting it to the equivalent dotted decimal IP address. If SDR commands now work, the system name server is not functioning.
6. If the address returned by **spget_syspar** is a valid system partition address and **pings** to that address are successful, check for the existence of the SDR server process (**sdrd**) on the control workstation with:

   ```
   ps -ae | grep sdrd
   ```

   If the process is not running, do the following:

   a. Check the **/** directory for a core dump. If one exists, report it to IBM.

   b. Check the SDR server logs in **/var/adm/SPlogs/sdr/sdrdlog.***pid* where *pid* is a process ID.

   c. Issue **/usr/bin/startsrc -g sdr** to start the SDR daemon. Start checks again at Step 5. If the SDR daemon is now running and continues to run, check the **sdrd** entry in the file **/etc/inittab** on the control workstation. It should read:

   ```
   sdrd:2:once:/usr/bin/startsrc -g sdr
   ```

Action 3. Analyze Class Situation

If an SDR command ends with RC=102 (internal data format inconsistency) or 026 (class does not exist), first make sure the class name is spelled correctly and the case is correct (see the table of classes and attributes in “The System Data Repository” appendix in *PSSP: Administration Guide*). Then, follow the steps in “SDR Shadow Files” in the System Data Repository appendix in the *PSSP: Administration Guide*.

This condition could be caused by the **/var** file system filling up. If this is the case, either define more space for **/var** or remove unnecessary files.

If the problem persists, contact the IBM Support Center.

Chapter 9. Diagnosing Authentication Problems

Users of the Kerberos Version 4 authenticated services have only a few ways of interacting with the authentication services of the SP system. In general, the sequence of operations required to perform an administrative task using these services consists of:

1. Identifying yourself to the authentication service. This process obtains a ticket-granting ticket for the client principal based on either a user's Kerberos Version 4 password or a service's private key. The former is achieved interactively using the **k4init** command. Background processes running as root and executing shell scripts can use the **rcmdtgt** command to get a ticket as service principal **rcmd.***hostname*.
2. Invoking a client command for one of the authenticated services, such as **sysctl** or **spmon**. The client command uses Kerberos Version 4 facilities to obtain a service ticket, the credentials that it passes to the application server to identify the invoking principal.
3. Using the **k4destroy** command to terminate the authenticated state by destroying any tickets found in a ticket cache file belonging to the client. If you do not remove the ticket cache file, then tickets it contains can be reused until they expire and are automatically removed by the next successful **k4init** command.

Error messages on stderr are the principal diagnosis tool for the user who experiences problems using these facilities. Errors reported by the authentication services themselves generally have the message number prefix **2502**, **2503**, or **2504**.

Table 10. Authentication Symptoms

Symptom	Recovery
Problems establishing a user principal's identity	“Action 1: Analyze Error Messages” on page 96 “Action 2. Force the Propagation of Database Changes” on page 96
Problems establishing a service principal's identity	“Action 3. Analyze Error Messages” on page 96 “Action 4. Compare Service Key Versions” on page 97
Problems using authenticated services	“Action 5. Analyze Error Messages” on page 101
Problems with the Authentication Server Daemons	“Action 6. Check for Network Problems, Interface, or Routing Problems” on page 102
Daemons have Terminated	“Action 7. Check Authentication Daemon Log Files” on page 102

Actions

Action 1: Analyze Error Messages

When identifying yourself to the authentication service, these are the most common error messages that you might encounter:

Bad Kerberos name format

You probably entered *name*.admin in response to the prompt “Kerberos name:”. You are not allowed to specify an instance in addition to the name. If you want to enter the name and instance together, enter them as the command-line argument when you invoke **k4init**. To force **k4init** to prompt you separately for the instance, invoke it with the **-i** flag.

Kerberos principal unknown

You did not enter a principal name defined in the database. Perhaps you misspelled it, or the administrator did so when entering it into the database. This error can also occur when using a secondary authentication server, if the primary database has not been propagated since your principal was added. Check with the administrator responsible for maintaining the authentication service to determine if this is the case.

Incorrect Kerberos password

You entered the wrong password or your password was recently changed and has not yet been propagated to the secondary authentication server you are using. If you incorrectly entered your password, just try again. Otherwise, if you suspect an out-of-date database, contact the administrator of your authentication service. The recovery action in this situation is to force the propagation of database changes without waiting for the normal **cron** process.

Action 2. Force the Propagation of Database Changes

To force the propagation of database changes:

1. Issue the **k4init** command, specifying as your principal name any user principal listed in the root user's **.klogin** file on the primary server. The administrative principal name that was used to set up authentication on the primary server can be used, or any other user principal the administrator has subsequently added to the file.
2. Issue the following command to remotely perform the database propagation from the primary server to all secondary servers:

   ```
   rsh primary /usr/kerberos/etc/push-kprop
   ```

   where *primary* is the hostname of the primary server.

   Successful propagation is reported by a message for each secondary server hostname. If unsuccessful, review the **kpropd.log** file (see information on daemon log files below).

Action 3. Analyze Error Messages

The authentication service also provides commands that can be used to allow background processes to invoke authenticated services, when the **k4init** command cannot be issued, because no user is logged in to the client AIX system. Various SP installation and system management scripts use the **rcmdtgt** command to identify themselves to the authentication system and thereby obtain a ticket-granting ticket. This procedure should not fail; if it does, the most likely error message reported will be:

```
Incorrect Kerberos password
```

If this error occurs during installation or when performing administrative tasks requiring remote execution on SP nodes, it can indicate one of several error conditions:

- The process is not running as **root**. It cannot read the server key file, **/etc/krb-srvtab**, on the client system. You must, if authorized, log in as **root** and retry the failing task.
- The server key file is out-of-date with respect to the authentication database.
- The server key file does not exist.

Action 4. Compare Service Key Versions

An administrator running as **root** can compare the versions of the server keys in the server key file and the database using the **ksrvutil** or **k4list** command on the server system and by examining a dump of the authentication database on the primary authentication server, when AFS is not being used.

Use the **ksrvutil** or **k4list** command to show the version numbers of service keys in **/etc/krb-srvtab**. The following example shows the use of **ksrvutil** to display the key versions for service principals on a control workstation:

```
# ksrvutil list
Version    Principal
   2     rcmd.cwktr@XYZ.ABC.COM
   2     hardmon.cwktr@XYZ.ABC.COM
   2     rcmd.cwkfddi@XYZ.ABC.COM
   2     hardmon.cwkfddi@XYZ.ABC.COM
   2     rcmd.cwken1@XYZ.ABC.COM
   2     hardmon.cwken1@XYZ.ABC.COM
   2     rcmd.cwksta@XYZ.ABC.COM
   2     hardmon.cwksta@XYZ.ABC.COM
#
```

The following example shows the use of **k4list** to display the key versions for service principals on an SP node:

```
# klist -srvtab
Server key file:   /etc/krb-srvtab
Service         Instance        Realm      Key Version
-----------------------------------------------------
rcmd            node3fi         XYZ.ABC.COM       1
rcmd            node3tr         XYZ.ABC.COM       1
rcmd            node3sw         XYZ.ABC.COM       1
rcmd            node3en         XYZ.ABC.COM       1
#
```

You can determine the versions of service keys in the authentication database by locating the entry for the target service principal in a dump of the SP authentication database. If you have secondary authentication servers, or if you use the procedure for backing up your database that is recommended in *PSSP: Administration Guide*, the database dump can be found in file **/var/kerberos/database/slavesave** on the primary server host.

For example, if you encountered an authentication failure attempting to **rsh** to **node3sw**; you could inspect the database entry as follows:

```
# grep "^rcmd node3sw " /var/kerberos/database/slavesave
rcmd node3sw 255 1 2 0 a49bf286 d45c6560 200001010459 199503301502 root admin
#
```

The fifth field on the line is the key version number. In this example, the key version number is two (2).

When the key file, **/etc/krb-srvtab**, does not exist on the server system, or has the wrong key version, you must re-create the file. When the control workstation is an SP authentication server, customizing an SP node will automatically create a new server key file for the node. If customizing the node for this purpose is too disruptive, or if the system whose key file must be replaced is not an SP node, follow the procedures below.

Procedure: Re-creating Server Key Files

You can use authentication administration commands to re-create an erroneous or missing server key file. Each system with an SP authentication service installed has its own unique service key file, containing the encrypted keys for the service instances that are available on the system.

On the control workstation and other IBM RS/6000 workstations that have client services installed and initialized, the file contains entries for services **rcmd** and **hardmon**. Separate instances of these principals are defined for each network interface on the system, where the instance-name is the short form of the network name. So, for example, on a client system with token-ring and FDDI interfaces named **wksta5t.xyz.abc.com** and **wksta5f.def.abc.com**, the following principals are defined:

hardmon.wksta5f
hardmon.wksta5t
rcmd.wksta5f
rcmd.wksta5t

The service key files on these systems are created by the **setup_authent** command. On the SP nodes, only the **rcmd** entries are included, and the files are created by the **setup_server** command. During installation, they are kept in the **/tftpboot** directory on the control workstation, with file names of the form *hostname***-new-srvtab**, where *hostname* is the short form of the hostname for each node.

When the control workstation is an SP authentication server, these files are retained there only until copied to the node during network boot (or to the node's boot-install server, if the node boots from another node). A new server key file is generated any time the node is set up for a network boot.

When the control workstation is not configured as an authentication server, or when AFS authentication is used, the server key files for the SP nodes are not removed from the **/tftpboot** directory on the control workstation, once created.

If they are deleted or corrupted, or if you choose to change keys for any reason, follow the rest of the procedures to create new key files. In these procedures, *instance1 ...* are the network names (short form) of all the system's interfaces, and *hostname* is the short form of the system's hostname.

Procedure: Replacing an Authentication Server's Key File

To re-create the file for a workstation (control workstation or other) that is configured as an authentication server, the **root** user should use the following procedure.

```
# create new key files in the /tmp  directory for each instance
cd /tmp
/usr/kerberos/etc/ext_srvtab -n instance1 ...

# combine the key files into a single file
/bin/cat instance1-new-srvtab ... >/etc/krb-srvtab
/bin/rm -f instance1-new-srvtab  ...

# make sure the key file is readable only by root
/bin/chmod 400 /etc/krb-srvtab
```

Procedure: Replacing a Client Workstation's Key File

When a workstation is not an authentication server, the root user can use the remote commands to perform the same function on a server system and move the file to the local system. The principal name specified on the **k4init** command must be in the **root** user's **.klogin** file on the server:

```
# get a ticket-granting ticket to allow use of rsh/rcp
k4init principal

# create the new key files in /tmp on the server
rsh server cd /tmp\;
/usr/kerberos/etc/ext_srvtab -n instance1 ...

# copy the files we created to the local /tmp directory
rcp server:/tmp/instance1-new-srvtab ... /tmp

# delete the files on the server
rsh server /bin/rm -f /tmp/instance1-new-srvtab ...

# combine the local files into a single file
cd /tmp
/bin/cat instance1-new-srvtab ... >/etc/krb-srvtab
/bin/rm -f instance1-new-srvtab ...

# make sure the key file is readable only by root
/bin/chmod 400 /etc/krb-srvtab
```

Procedure: Replacing an SP Compute Node's Key File

The most straightforward way to replace a node's key file is to customize the node, using the **spbootins** command or SMIT. If you prefer, you can use procedures similar to the preceding example. If the control workstation is an SP authentication server, the **root** user, logged into the SP node whose server key file needs to be replaced, can use the procedure described previously for client systems. Specify the control workstation hostname as the *server*. When the control workstation is not an authentication server, the server key file must be re-created at the server and then placed in the **/tftpboot** directory on the control workstation. For this, the **root** user should be logged into the control workstation.

When the authentication server is another workstation running the SP server or another MIT Kerberos Version 4 implementation, use the following procedure:

```
# get a ticket-granting ticket to allow use of rsh/rcp
k4init principal

# create the new key files in /tmp on the server
rsh server cd /tmp\;
/usr/kerberos/etc/ext_srvtab -n instance1 ...

# copy the files we created to the control workstation
rcp server:/tmp/instance1-new-srvtab ...  /tmp

# delete the files on the server
rsh server /bin/rm -f /tmp/instance1-new-srvtab ...

# combine the local files into a single file in /tftpboot, whose
# name is hostname-new-srvtab.
cd /tmp
/bin/cat instance1-new-srvtab ...  >/tftpboot/hostname-new-srvtab
/bin/rm -f instance1-new-srvtab ...

# make sure the key file is readable only by root
/bin/chmod 400 /tftpboot/hostname-new-srvtab
```

The new key file can then be installed on the node by either:

- customizing the node.
- logging into the node as root, and using **rcp** to copy the file from **/tftpboot** on the control workstation to **/etc/krb-srvtab**.
- using **ftp** to copy the file (not a secure method).

Procedure: Replacing a Server Key File Using AFS Servers

When an AFS authentication server is being used, follow this procedure and be sure you are logged in to the control workstation as **root**.

```
# change the service password to a new known but random value
# repeat this step for each instance (i.e. short network name)
kas setpassword -name rcmd.instance1
  -new_password any-random-password \
  -kvno 1 -admin_username afs-admin-name
  -password_for_admin afs-admin-passwd

# use the same principals and passwords to create a srvtab file
/usr/kerberos/bin/ksrvutil -afs -f /tftpboot/hostname-new-srvtab add

# ksrvutil is an interactive program whose sequence of prompts and
# messages appear as follows:
Name: rcmd
Instance: instance1
Realm: <Enter>
Version number: 0
New principal: rcmd.instance1@realm; Version 0
Is this correct? y
Password:
Key successfully added.
Would you like to add another key?
(reply y until all instances have been entered)

# make sure the key file is readable only by root
/bin/chmod 400 /tftpboot/hostname-new-srvtab
```

Action 5. Analyze Error Messages

When you use any of the Kerberos Version 4 authenticated client/server applications to administer or control the SP system, the error messages you receive on authentication failures will vary according to the application. For example, if you are using the command-line interface, you might see error messages such as the following indicating that **spmon** is unable to obtain credentials from the authentication service:

```
0026-706 Cannot obtain service ticket for hardmon.cwksta
Kerberos error code is 76, Kerberos error message is:
     2504-076 Kerberos ticket file was not found.
spmon: 0026-001 Opening session failed.
```

The message states that you have no tickets, expired or unexpired, or your **KRBTKFILE** environment variable specifies a nonexistent file.

The following message states that you have tickets that have expired in the ticket cache file specified by your KRBTKFILE environment variable or defaulted.

```
0026-706 Cannot obtain service ticket for hardmon.cwksta
Kerberos error code is 32, Kerberos error message is:
     2504-032 Kerberos ticket expired.
spmon: 0026-001 Opening session failed.
```

If application error messages indicate probable authentication failure, use the **k4list** command to check your authentication status. The command always displays the current active ticket cache file, whether specified by the **KRBTKFILE** environment variable or the default file, **/tmp/tkt***uid*.

The following is a listing of the default ticket cache file for the **root** user (uid 0):

```
# k4list
Ticket file:    /tmp/tkt0
Principal:  root.admin@XYZ.ABC.COM

  Issued           Expires          Principal
Nov 12 16:26:11  Dec 12 16:26:11  krbtgt.XYZ.ABC.COM@XYZ.ABC.COM
Nov 12 16:26:46  Dec 12 16:26:46  hardmon.cwksta@XYZ.ABC.COM
Nov 12 16:45:15  Dec 12 16:45:15  rcmd.cwksta@XYZ.ABC.COM
#
```

The second line shows the Kerberos Version 4 principal acting as client, to whom the tickets belong. This is the user principal you supplied to the **k4init** command, or the **rcmd**.*instance* service principal used by **rcmdtgt**. The list of tickets always begins with the ticket-granting ticket. The others are service tickets; in this case for the System Monitor service on the control workstation (**hardmon**) and the **Sysctl** service also on the control workstation (**rcmd**).

Action 6. Check for Network Problems, Interface, or Routing Problems

Other more or less common errors may result from the inability to communicate with the Kerberos Version 4 authentication and administrative servers. If this is the case, check for network problems or interface and routing problems. You can also check the state of the server daemons themselves on systems running the SP authentication server. The primary server system should have **kerberos** and **kadmind** daemons running (from the **init** command).

The following example shows the **/etc/inittab** entries for the Kerberos daemons on a primary server:

```
> lsitab kerb
kerb:2:respawn:/usr/kerberos/etc/kerberos
> lsitab kadm
kadm:2:respawn:/usr/kerberos/etc/kadmind -n
>
```

A secondary server system should have **kerberos** and **kpropd** daemons running (from init).

The following example shows the **/etc/inittab** entries for the Kerberos daemons on a secondary server:

```
> lsitab kerb
kerb:2:respawn:/usr/kerberos/etc/kerberos -s
> lsitab kpropd
kpropd:2:respawn:/usr/kerberos/etc/run-kpropd
>
```

Action 7. Check Authentication Daemon Log Files

Each Kerberos daemon program records errors and some status in a log file in the **/var/adm/SPlogs/kerberos** directory. Check these files if you suspect one or more of the daemons have terminated. They are designed to hang indefinitely or for several minutes, depending in some cases on command-line options, to prevent problems caused by constant respawning of a failing daemon.

The following is an example of the log file created by the **kerberos** daemon:

```
# cat /var/adm/SPlogs/kerberos/kerberos.log
13-Oct-1994 07:53:07 Kerberos started, PID=8012

13-Oct-1994 07:53:07 kerberos: 2503-604 Cannot verify master key.
13-Oct-1994 07:53:07 Kerberos will pause so as not to loop init
13-Oct-1994 08:47:33 Kerberos started, PID=13243
#
```

The following is an example of the log file created by the **kadmind** daemon:

```
# cat /var/adm/SPlogs/kerberos/admin_server.syslog
13-Oct-1994 08:42:16 Kerberos admin server started, PID=9831
13-Oct-1994 08:42:16 kadmind: 2503-101 error: 2504-318
                     Could not verify master key

13-Oct-1994 08:42:16 Shutting down admin server
13-Oct-1994 08:47:33 Kerberos admin server started, PID=13759
#
```

The following is an example of the log file created by the **kpropd** daemon:

```
# cat /var/adm/SPlogs/kerberos/kpropd.log
13-Oct-1994 09:27:29

***** kpropd started *****
13-Oct-1994 09:27:29 Established socket
13-Oct-1994 09:27:31 Connection from cwksta.xyz.abc.com, 129.49.100.41
13-Oct-1994 09:27:31 kpropd: Connection from rcmd.cwksta@XYZ.ABC.COM
13-Oct-1994 09:27:31 File received.
13-Oct-1994 09:27:31 Temp file renamed to /tmp/sdump10560
13-Oct-1994 09:27:32

***** kpropd started *****
13-Oct-1994 09:27:32 Established socket
13-Oct-1994 10:19:09 Connection from cwksta.xyz.abc.com, 129.49.100.41
13-Oct-1994 10:19:09 kpropd: Connection from rcmd.cwksta@XYZ.ABC.COM
13-Oct-1994 10:19:09 File received.
13-Oct-1994 10:19:09 Temp file renamed to /var/kerberos/database/slavedb
14-Oct-1994 07:36:41 Connection from cwksta.xyz.abc.com, 129.49.100.41
14-Oct-1994 07:36:41 kpropd: Connection from rcmd.cwksta@XYZ.ABC.COM
14-Oct-1994 07:36:41 File received.
14-Oct-1994 07:36:41 Temp file renamed to /var/kerberos/database/slavedb
```

Chapter 10. Diagnosing Remote Command Problems on the SP System

In AIX 4.3.1, the AIX Remote Command suite was enhanced to support Kerberos Version 5 authentication through DCE. These commands include **rsh**, **rcp**, **rlogin**, **telnet**, and **ftp**. For SP migration purposes, the AIX remote commands **rsh** and **rcp** were enhanced to call an SP-supplied Kerberos Version 4 set of **rsh** and **rcp** routines. Therefore, the AIX **/usr/bin/rsh** and **/usr/bin/rcp** commands on the SP system support the following authentication methods: Kerberos Version 5 (through DCE), Kerberos Version 4, and standard AIX.

The previously-supplied remote commands are no longer shipped with PSSP. The **/usr/lpp/ssp/rcmd/bin/rsh** and **/usr/lpp/ssp/rcmd/bin/rcp** commands are now symbolic links to the AIX **/usr/bin/rsh** and **/usr/bin/rcp** commands, respectively.

You may see error messages for one type of authentication method before another enabled authentication method is attempted and is successful. Display of these messages depends on the authentication method enabled, the tickets owned by the user, and the order in which the authentication methods are attempted. In general, Kerberos Version 5 messages (through DCE) have a prefix of "Kerberos:".

Diagnosing Kerberos Version 5 problems is similar to the procedure below in that:

- The user must have tickets for authentication.
- The user must be authorized though an authorization file (**.k5login**).
- The authentication method must be installed, configured and enabled.

Refer to the DCE documentation and AIX documentation when diagnosing problems involving Kerberos Version 5 through DCE.

The SP-supplied Kerberos Version 4 set of **rsh** and **rcp** subroutines depend heavily on Kerberos authentication and the correct configuration of Kerberos on the SP system. For Kerberos problems, see Chapter 9, “Diagnosing Authentication Problems” on page 95 in this book, and the Kerberos commands in *PSSP: Command and Technical Reference*. For information on authentication methods, see *PSSP: Administration Guide*.

Use the following procedure to diagnose remote command problems.

1. Verify that the symbolic links are correct by issuing the **ls -al** command. The following should appear:

   ```
   /usr/lpp/ssp/rcmd/bin/rcp –> /usr/bin/rcp
   /usr/lpp/ssp/rcmd/bin/rsh –> /usr/bin/rsh
   ```

2. Check the authentication method on the target and source hosts by issuing the **lsauthent** command.
3. Check the authentication method enabled in the system partition where the source and target hosts reside, by issuing the **lsauthpar** command.

   If Kerberos Version 4 is not listed as a valid, enabled authentication method by the **lsauthpar** command do the following to enable Kerberos Version 4:

   a. Verify that Kerberos Version 4 is installed and configured in the system partition.

b. Enable the Kerberos Version 4 authentication method.

c. Verify the following values:

- On the SP Security Select Authentication Methods smit panel:

  ```
  Authentication method installed/configured:  Kerberos Version 4
  ```

- On the SP Security Select Authorization Files smit panel:

  ```
  Type of authorization to use for root user access:Kerberos Version 4
  ```

- On the SP Security Enabled Authentication Methods smit panel:

  ```
  Authentication method enabled in the partition:  Kerberos Version 4
  ```

Each field may display additional choices. However, in order to use Kerberos Version 4 authentication within the remote commands, Kerberos Version 4 must appear in each of the fields.

4. Check the order of the authentication methods. Issue the **chauthpar** command to change the order of authentication methods for the system partition. The required order is: Kerberos Version 5 (through DCE), Kerberos Version 4, and standard AIX, if enabled.

The following steps apply only to SP-provided remote command support for Kerberos Version 4. Refer to the AIX and DCE documentation for diagnosing problems with the Kerberos Version 5 (through DCE) and standard AIX authentication methods. This documentation includes:

- AIX manpages for the **rlogin**, **rsh**, **rcp**, **krshd**, **telnet**, **ftp**, **rlogind**, **ftpd**, and **telnetd**commands.
- *IBM Distributed Computing Environment 2.2 for AIX: Administration Guide and Reference*

1. Check that the user has a ticket provided by the Kerberos **k4init** command, by issuing the **k4list** command.
2. Check to see if the user has a **.klogin** file in his home directory on the target host. If the user has the file, check it for the user's principal name. It must have this format: **rcmd.***name*@*domain*, where *name* is the short hostname returned by the **hostname** command, and *domain* is the domain name.
3. Run the **/usr/lpp/ssp/rcmd/bin/rsh** *hostname* **date** command on the source host, where *hostname* is the host's name, to see if any Kerberos Version 4 error messages are issued.
4. Verify that an **rcp** file or link is present in the **/usr/kerberos/bin** directory when using the **rcp** command. If the link is present, it should point to the **/usr/lpp/ssp/rcmd/bin/rcp** file. A file should be present only if an MIT version of Kerberos Version 4 is installed.
5. Check to see if the system administrator has run a **ksrvutil change** command. If so, all current tickets are invalid, and the **k4init** command must be run by each user to obtain a current ticket. The **ksrvutil change** command should not be run frequently.
6. Check the **/etc/services** and the **/etc/inetd.conf** files for the **kshell** service. The **kshell** service in the **/etc/inetd.conf** file should point to the **/usr/sbin/krshd** file.
7. If the Kerberos Version 4 authentication method is enabled on the target host, and valid tickets exist, it is possible that the **krshd** subserver is experiencing

problems. Restart the **krshd** subserver. Refer to the AIX manpage for the **krshd start/stop** commands.

Chapter 11. Diagnosing Switch Problems

If your system or system partition shows signs of a switch failure, locate the symptom in the table and perform the recovery action or actions described. All of the actions described require that the user have root access on the specified node.

SP Switch Symptoms and Recovery Actions

Table 11. SP Switch Symptoms

Symptom	Recovery
Estart failure: 1. System cannot find **Estart** command. 2. Primary node is not reachable. 3. **Estart** command times out or fails. 4. Expected number of nodes not initialized. 5. Some links do not initialize.	1. See "Verify Software Installation" 2. See "Verify an SP Switch Node" on page 110 3. See "Diagnose SP Switch Estart Problems" on page 136 4. See "Device and Link Problems" on page 120 5. See "Device and Link Problems" on page 120
Node drops off of the switch (**switch_responds** is off for the node)	See "Verify an SP Switch Node" on page 110
Node fails to communicate over the switch, but its **switch_responds** is on (**ping** or **CSS_test** commands fail)	See "Verify an SP Switch Node" on page 110 and "Isolate Adapter and Switch Error" on page 124
Node crash	See "Node Crash" on page 119
Node fails to **Eunfence**	See "Eunfence Problems" on page 120
Other Ecommand failures	See "Ecommand Problems" on page 123
Oncoming Primary node is **Efenced**	See "Eunfence the Oncoming Primary" on page 119
Note: If the above recovery actions fail to resolve your problem, contact the IBM Support Center.	

Verify Software Installation

Software installation and verification is done using the **CSS_test** script from either the SMIT panel or from the command line. If **CSS_test** is executed following a successful **Estart**, additional verification of the system will be done to determine if each node in the system or system partition can be **ping**ed. If you are using system partitions, **CSS_test** runs in the active partition only. For more information on managing system partitions, see *PSSP: Administration Guide*.

To verify CSS installation from SMIT:

ENTER smit SP_verify

The Installation/Configuration menu appears.

SELECT ssp Communications Subsystem

PRESS Enter

Review the output created to determine the results.

To use **CSS_test** from the command line, you can optionally select the following options:

- **-q** to suppress messages.
- **-l** to designate an alternate log file. The default log file is **/var/adm/SPlogs/css/CSS_test.log**.

Review the log file to determine the results.

Additional items to consider while trying to run **CSS_test** are as follows:

- Each node should have access to the **/usr/lpp/ssp** directory.
- **/etc/inittab** on each node should contain an entry for **rc.switch**.

For complete information on **CSS_test** see *PSSP: Command and Technical Reference.*

Identify the Failing Node

Use this scenario if an application running on several nodes loses connectivity over the switch, or **switch_responds** indicates that several nodes are not on the switch.

1. View the summary log located on the control workstation. See the section entitled: Summary log for SP Switch and SP Switch Adapter Errors in Chapter 4, “Error Logging Overview” on page 69.
2. Locate the first error log entry that indicates a node or connectivity failure.
3. Examine other entries to see if the first failure is the cause of subsequent failures.
4. On the node that experienced the first failure, examine the error log to see the complete version of the error log record described above.
5. Use this as a starting point to debug the problem on this node.

Verify an SP Switch Node

Use this procedure to verify that a single SP Switch node is operating correctly. If the node you are attempting to verify is the primary node, start with Step 1. If it is a secondary node, start with Step 2.

1. Determine which node is the primary by issuing the **Eprimary** command on the control workstation. For complete information on the **Eprimary** command, see *PSSP: Command and Technical Reference*. For our purposes, the following should suffice:

   ```
   Eprimary
   1 - primary
   2 - oncoming primary
   26 - primary backup
   26 - oncoming primary backup
   ```

   If the command returns an oncoming primary value of none, reexecute the **Eprimary** command specifying the node you would like to have as the primary node. Following the execution of the **Eprimary** command (to change the oncoming primary) an **Estart** is required to make the oncoming primary node the primary.

If the command returns a primary value of none, an **Estart** is required to make the oncoming primary node the primary.

The primary node on the SP Switch system can move to another node, if a primary node takeover is initiated by the backup. To determine if this has happened, look at the values of the primary and the oncoming primary backup. If they are the same value, then a takeover has occurred.

2. Ensure that the node is accessible from the control workstation, this can be accomplished by using **rsh** to issue the **date** command on the node as follows:

```
/usr/lpp/ssp/rcmd/bin/rsh <problem hostname> date
TUE Oct 22 10:24:28 EDT 1997
```

   If the current date and time are not returned, refer to Chapter 10, “Diagnosing Remote Command Problems on the SP System” on page 105.

3. Verify that the switch adapter (css0) is configured and is ready for operation on the node. This can be done by interrogating the adapter_config_status attribute in the switch_responds object of the SDR:

```
SDRGetObjects switch_responds node_number==<problem node number>
node_number switch_responds autojoin isolated adapter-config_status
      1           0             0        0          css_ready
```

   If the adapter_config_status object is anything other than css_ready, see “Configure and Diagnose Problems” on page 113.

   Note: To obtain the value to use for problem node number, issue an SDR query of the node_number attribute of the Node object, as follows:

```
SDRGetObjects Node reliable_hostname==<problem hostname> node_number
node_number
    1
```

4. Verify that the **fault_service_Worm_RTG_SP** daemon is running on the node. This can be accomplished by using **rsh** to issue a **ps** command to the problem node as follows:

```
 /usr/lpp/ssp/rcmd/bin/rsh <problem_hostname> ps -e | grep Worm_RTG
18422   -0:00 fault_service_Worm_RTG
```

   If the **fault_service_Worm_RTG_SP** daemon is running, SP Switch node verification is complete.

   If the **fault_service_Worm_RTG_SP** daemon is not running, see “Isolate Adapter and Switch Error” on page 124. The possible reasons why the **fault_service_Worm_RT_SP** deamon is not running are:

   - The daemon exited due to an abnormal error condition.
   - A **SIGTERM**, **SIGBUS** or **SIGDANGER** signal was processed by the daemon.

Recover the Node

You can restart the **fault_service_Worm_RTG_SP** daemon on the node by issuing:

```
/usr/lpp/ssp/css/rc.switch "adapter/mca/tb3"
```

Following the **rc.switch**, do Step 4 again of “Verify an SP Switch Node” on page 110 to determine if the daemon is still running or has died.

At this point you should be able to **Eunfence** the node by issuing:

```
Eunfence <problem_node_number>
```

All nodes successfully unfenced.

If you cannot resolve the problem, contact the IBM Support Center. You should also attempt to gather any log files associated with this failure. See “Collect Information for the IBM Support Center” on page 118.

Verify Switch Topology Configuration

The switch topology file is used to define the hardware configuration to the css support software. It should reflect the number of switches and nodes installed, as well as define how they are connected. The topology file can reside in two places, in the SDR or as **expected.top** in the **/etc/SP** directory of the primary node. The configuration in the SDR is what is commonly used. The configuration in **/etc/SP/expected.top** is generally used for debug purposes only. If **/etc/SP/expected.top** exists on the primary node, it overrides the configuration in the SDR.

To verify that the topology in the SDR is correct, first read it out of the SDR using the **Etopology** command:

```
Etopology -read <file name>
```

The **Etopology** command reads the switch topology from the SDR and places it in to the specified file. For more information on the **Etopology** command, see *PSSP: Command and Technical Reference*. Once the file is extracted, verify that the switch topology is an accurate representation of the installed hardware. If changes to the switch topology file are required, remember to place them back into the SDR using the **Etopology** command.

Verify the System Data Repository (SDR)

To verify that the SDR is installed and operating correctly, you can run **SDR_test** on the control workstation. It can be run either through SMIT or from the command line.

To verify the SDR from SMIT:

ENTER smit SP_verify

The Installation/Configuration menu appears.

SELECT ssp System Data Repository

PRESS Enter

Review the output created to determine the results.

To use **SDR_test** from the command line, issue:

```
/usr/lpp/ssp/bin/SDR_test
```

Review the output created to determine the results.

Next log into the failing node and issue the **SDRGetObjects** command against the switch_responds object:

```
SDRGetObjects switch_responds
```

Examine the output that is returned. If the switch responds bits are returned, this indicates that SDR is operating. You can also determine which nodes are operational on the switch by examining the value returned. A value of 1 indicates the node is operational; a value of 0 indicates that the node is not operational.

Configure and Diagnose Problems

The following table is based on the possible values of the **adapter_config_status** attribute of the **switch_responds** object of the SDR. Use the following command to determine its value:

```
SDRGetObjects switch_responds
```

Use the value of the **adapter_config_status** attribute for the node in question, to index into the table.

Note: The **adapter_config_status** table that follows uses the phrase "adapter configuration command". This refers to the SP Switch adapter configuration method. Use the following syntax to invoke it:

```
/usr/lpp/ssp/css/cfgtb3 -v -l css0 > <output_file_name>
```

Table 12. adapter_config_status Values	
adapter_config_status	**Explanation and Recovery action**
odm_fail genmajor_fail genminor_fail getslot_fail build_dds_fail	An ODM failure has occurred while configuring the CSS adapter. Rerun the adapter configuration command. If the problem persists contact, the IBM Support Center and supply them with the command output.
lname_error	The device logical name specified on the CSS adapter configuration command was invalid. Rerun the adapter configuration. If the problem persists, contact the IBM Support Center and supply them with the command output.
undefine_system_fail define_system_fail xilinx_system_fail	The System Standard C Library Subroutine failed during CSS adapter configuration. Rerun the adapter configuration. If the problem persists, contact the IBM Support Center and supply them with the command output.
undefine_fail define_fail	The current instance of the CSS logical device could not be redefined. Rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply them with the command output.
chkslot_fail	Verify the CSS adapter is properly seated, then rerun the adapter configuration command. If the problem persists, contact the IBM Support Center and supply them with the command output.
busresolve_fail	There are insufficient bus resources to configure the CSS adapter. Contact the IBM Support Center.
xilinx_load_fail dd_load_fail fs_load_fail	See “Verify Software Installation” on page 109. If software installation verification is successful and the problem persists contact the IBM Support Center.
make_special_fail	The CSS device special file could not be created during adapter configuration. Rerun adapter configuration. If the problem persists, contact the IBM Support Center and supply them with the command output.
dd_config_fail fs_init_fail	An internal device driver error occurred during CSS adapter configuration. Use the section “Collect Information for the IBM Support Center” on page 118 to collect pertinent information and contact the IBM Support Center.
diag_fail	See “Adapter Diagnostic Failures” on page 114.

Adapter Diagnostic Failures

The recovery actions to take for adapter diagnostic failure can, in most cases, be determined by examining the Service Request Number (SRN) posted in the error log entry for the diagnostic failure. To get this information, log into the failing node and issue the **errpt** command to view detailed information in the failing entry. The command would be something like:

```
errpt -a | grep "Switch adapter failed POST diagnostics"
```

Use the SRN table that follows to determine the recovery action:

Note: in the following table **x** means any value.

Table 13. Service Request Numbers	
SRN	**Recovery action**
1xx	See “Verify Software Installation” on page 109. If this verification is successful and the problem persists, contact the IBM Support Center.
28x	See “Verify External Clock and Cable” on page 115. If this verification is successful and the problem persists, contact the IBM Support Center.
Axx	See “Verify External Clock and Cable” on page 115. If this verification is successful and the problem persists, contact the IBM Support Center.
All other SRNs	Contact IBM Hardware Service and arrange to have the adapter or cable replaced.

The following is additional information on how to run css adapter diagnostics online:

To run the complete set of css adapter diagnostics on a particular node requires exclusive use of the of the css adapter. You may have to **kill** processes that currently have the css device driver open. This would include the **fault_service_Worm_RTG_SP** daemon or the **fault_service_Worm_RTG** daemon and any other processes that are using it, such as "switch clock reader applications".

To run the same sequence of adapter tests that ran at Power On Self Test (POST) time, enter the following:

```
diag -c -d css0
```

This command runs the css adapter diagnostics in unattended mode, directing the written results to stdout.

When cable or adapter problems are suspected and the POST diagnostics run successfully it may be advisable to run the adapter and adapter cable wrap test found in advanced diagnostics, by entering the following:

```
diag -A -d css0
```

Note: You will need both the card and cable wrap plugs to complete these tests.

Verify External Clock and Cable

The following procedure should not be run on nodes that are operational on the switch. The utilities used for these verifications cannot coexist with normal switch operations on the node.

If clocking problems exist on all nodes in a rack. See “Rack or System Clock Problems” on page 117. Otherwise, start by running the clock verification procedure for the problem node(s).

Verify SP Switch External Clock

Use the following steps to determine if the external clock is operational at the node:

1. Login to the node in question.
2. Execute the following command:

```
/usr/lpp/ssp/css/diags/read_tbic -s
TBIC status register    :78XXXXXX
```

Note: bits are numbered from left to right, starting at 0.

Look at bits 3 and 4:

- if bits 3 and 4 are **ON** the external clock is operational at the node.
- If either bit 3 or 4 is **OFF** the external clock is not operational at the node. To restore the external clock at the node see “Restore SP Switch External Clock.”

Restore SP Switch External Clock

Perform the following steps to restore the external clock at the node:

1. **rc.switch** the node with the following command:

   ```
   /usr/lpp/ssp/css/rc.switch
   "adapter/mca/tb3
   ```

2. Determine if the clock is present using “Verify SP Switch External Clock” on page 115.
3. If the clock is still not present, try to **Eclock** the system. **Eclock** will affect all switch boards in the system and requires exclusive use of the switch (that is, all system partitions). For more information on the **Eclock** command see *PSSP: Command and Technical Reference*.

   To **Eclock** the system use the following command:

   ```
   Eclock -d
   ```

4. Determine if the clock is present by referring to “Verify SP Switch External Clock” on page 115.
5. If the clock is still not present, you have two choices. The first is to follow the instructions presented in “Verify Cable.” The second is to contact IBM Hardware Service and have them perform the cable verification.

Verify Cable

This section covers node to switch and switch to switch cables. For node to switch cables use “Verify Node to Switch Cable,” for switch to switch cables use “Verify Switch to Switch Cable” on page 117 .

Verify Node to Switch Cable

The first step is to visually inspect the cable in question. This is done as follows:

1. Remove the cable from the back of the node and examine the connectors (cable and back of the adapter) for bent pins or other visible damage. If every thing looks OK, reconnect the cable to the adapter. If not, contact IBM Hardware Service and have them replace or repair the damaged components.
2. Remove the cable from the back of the switch and examine the connectors (cable and switch bulkhead jack) for bent pins or other visible damage. If every thing looks OK, reconnect the cable to the switch bulkhead jack. If not, contact IBM Hardware Service and have them replace or repair the damaged components.
3. If every thing visually checks out, run Advanced Diagnostics on the suspect adapter and cable. The procedure for doing this is outlined in “Adapter Diagnostic Failures” on page 114. Follow the online instructions. If the diagnostics detect a failure, contact IBM Hardware Service and have the failing

components replaced. If diagnostics pass and the problem persists, contact the IBM Support Center.

4. As a result of removing the cable, the node may be fenced by the system. After reinstalling the cable, reboot the node or run the **rc.switch** command to reset the switch adapter before you try to **Eunfence** the node.

Verify Switch to Switch Cable

The first step is to visually inspect the cable in question. This should be done as follows:

1. Remove the cable from the back of the switch and examine the connectors (cable and switch bulkhead jack) for bent pins or other visible damage. If every thing looks OK, reconnect the cable to the switch bulkhead jack. If not, contact IBM Hardware Service and have them replace or repair the damaged components.
2. Repeat Step 1 for the other end of the switch to switch cable.
3. If everything visually checks out, contact IBM Hardware Service and have them replace the cable. If the problem persists, contact the IBM Support Center.

Rack or System Clock Problems

The following table list the possible clock loss problems on single racks and systems along with their recovery actions:

Table 14. Clock Problems

Problem	Recovery Action
All nodes in a single rack will not clock.	Cause: Switch is powered off Action: Power the switch on, run **Eclock -d**, then **Estart**. Cause: The switch is not **Eclock**ed Action: Run **Eclock -d**, then **Estart**. Cause: The clock topology file used does not match physical system topology or the switch board is defective. Action: Contact IBM Hardware Service
Some racks in the system will not clock.	Cause: Switches are powered off Action: Power the switches on, run **Eclock -d**, then **Estart**. Cause: The system is not **Eclock**ed Action: Run **Eclock -d**, then **Estart**. Cause: The clock topology file used does not match physical system topology or the master clock switch is bad. Action: Contact IBM Hardware Service

Collect Information for the IBM Support Center

The **css.snap** script collects log files created by switch support code (device driver, worm, fault-service, diags) into a single package.

Collect the **css.snap** information from both the primary node and all nodes that are experiencing SP switch problems. **Do not reboot** the nodes before running **css.snap**, because this causes the loss of valuable diagnostic information.

Under normal circumstances**css.snap** collects the following files:

- cable_miswire
- cable_miswire.old
- core (fault service daemon dump file)
- css.snap.log
- css_dump.out
- daemon.stderr
- daemon.stdout
- dtbx.trace
- dtbx.failed.trace
- errpt.out (most recent **errpt -a** and **errpt** entries)
- flt
- fs_daemon_print.file
- netstat.out (current **netstat -I css0** and **netstat -m**)
- out.top
- rc.switch.log
- regs.out
- router.log
- scan_out.log
- scan_save.log
- tb_dump.out
- vdidl.out
- worm.trace

The files ending in **.out** are produced by running the appropriate command to dump internal (in memory) trace info or dump data to a file. The completed output file is:

`/var/adm/SPlogs/css/`*`hostname`*`.`*`yymmddttttt`*`.css.snap.tar.Z`

where*hostname* is the host name of the node where **css.snap** was issued, and *yymmddttttt* is the date and time that the **css.snap** information was collected.

css.snap avoids filling up **/var** by following these rules:

1. If less than 10% of **/var** is free, **css.snap** exits.

2. If the css portion of **/var** is more than 30% of the total space in **/var**, **css.snap** erases old snap files until the css portion becomes less than 30%. If it is successful, the snap proceeds. If not, it exits.

The **css.snap** command is called automatically from the fault-service daemon when certain serious errors are detected. It can also be issued from the command line when a switch or adapter related problem is indicated.

```
/usr/lpp/ssp/css/css.snap
```

Caution: **css.snap** uses a number of undocumented utilities to collect information. Some of these, like **read_regs** and the **tbXdump** routines, can be destructive when used on a running system. After using **css.snap** to collect diagnostic information, it is advisable to run **/usr/lpp/ssp/css/rc.switch** in order to reset and reload the switch adapter and eliminate the residual effects of these utilities.

Node Crash

A node crash is generally identified by the LED/LCD display on the node flashing 888. The recovery action for this event is to save the dump that was created. The procedure for doing this can be found in Chapter 5, “Producing a System Dump” on page 83. Once the dump is saved, Contact the IBM Support Center and supply them with the tar image in the **/tmp/ibmsupt directory**.

Eunfence the Oncoming Primary

If the oncoming primary node becomes fenced from the switch use the following procedure to **Eunfence** it prior to issuing **Estart**:

- If the switch is up and operational with another primary node in control of the switch, then issue **Eunfence** on the oncoming primary and issue **Estart** to make it the active primary node.

  ```
  Eunfence 1
  All node(s) successfully unfenced.
  Estart
  Switch initialization started on k60n01
  Initialized 14 node(s).
  Switch initialization completed.
  ```

- If the switch is not operational and **Estart** is failing because the oncoming primary's switch port is fenced, you must first change the oncoming primary to another node on the switch and **Estart**. Once the switch is operational you can then **Eunfence** the old oncoming primary node. If you also want to make it the active primary then issue an **Eprimary** command to make it the oncoming primary node and **Estart** the switch once again.

```
Eprimary 3
Eprimary: Defaulting oncoming primary backup node to k60n14
Estart
Switch initialization started on k60n03
Initialized 13 node(s)
Eunfence 1
All node(s) successfully unfenced.
Eprimary 1
Eprimary:  Defaulting oncoming primary backup node to k60n14
Estart
Switch initialization started on k60n01
Initialized 14 node(s)
Switch initialization completed.
```

Eunfence Problems

You can isolate and correct most **Eunfence** problems by referring to “Ecommand Problems” on page 123. The following list provides some additional reasons for a particular node to fail to **Eunfence**:

- The node can no longer be reached through the switch network. The section “Device and Link Problems” can be used to isolate and correcting this problem.
- If an SP Switch node fails to **Eunfence** because the switch topology could not be distributed, see Chapter 9, “Diagnosing Authentication Problems” on page 95.
- The node fails to respond when attempting to **Eunfence** it. The section “Verify an SP Switch Node” on page 110 can be used to isolate and correct the problem.
- If you receive the message "Cannot Unfence node xxxx — timeout", the most likely cause is that the fault_service_Worm_RTG daemon is not running on the node. If the Worm is not running, issue the **/usr/lpp/ssp/css/rc.switch** command to start the Worm.
- You can also receive the above timeout message from **Eunfence** if you have replaced the switch cable. See “Verify Cable” on page 116. Even though the **fault_service_Worm_RTG** daemon may still be running, you must issue the **/usr/lpp/ssp/css/rc.switch** command to reload and reset the switch adapter before you can **Eunfence** the node.
- If any of the above procedures fail to resolve the problem, gather the css logs from both the primary node and the node being unfenced. This can be accomplished by logging into those nodes and using the procedure outlined in “Collect Information for the IBM Support Center” on page 118.

Note: You cannot **Eunfence** a node from a switch that is not **Estart**ed!

Device and Link Problems

When evaluating device and link problems on the system, first examine the **out.top** file in the **/var/adm/SPlogs/css** directory of the primary node. This file looks like a switch topology file, except for the additional comments on lines where either the device or link is not operational.

These additional comments are appended to the file by the fault_service daemon to reflect the current device and link status of the system. If there are no comments

on any of the lines or the only comments are for wrap plug where they actually exist, you should consider all devices and links to be operational. If this is not the case, however, the following information should help to resolve the problem:

The following is an example of a failing entry in the **out.top** file:

```
s 14 2  tb3 9 0 E01-S17-BH-J32 to E01-N10   -4 R: device has been removed
from network-faulty (link has been removed from network or miswired-faulty)
```

This example means the following:

- switch chip 14, port 2 is connected to switch node number 9.
- The switch is located in frame E01 slot 17.
- Its bulkhead connection to the node is jack 32.
- The node is also in frame E01 and its node number is 10.
- The -4R refers to the device status of the right side device (tb0 9), which has the more severe device status of the two devices listed. The device status of the node is "device has been removed from the network - faulty".
- The link status is "link has been removed from the network or miswired - faulty".

The possible device and link status for SP switch systems are listed in the following tables with possible recovery actions:

Table 15. SP Switch Device Status

#	Text	Explanation and Recovery Action
-4	Device has been removed from network - faulty	The device has been removed from the switch network, because of a fault on the device. If the device in question is a node, see “Verify an SP Switch Node” on page 110. Otherwise contact IBM Hardware Service.
-5	Device has been removed from network by system administrator	The device was placed offline by the Systems Administrator (**Efence**). **Eunfence** the device.
-6	Device has been removed from network - no AUTOJOIN	The device was removed and isolated from the switch network. The possible causes are: the node was **Efence** without AUTOJOIN, the node was rebooted or powered off, or the node faulted. First attempt to **Eunfence** the device, if the node fails to rejoin the switch network, see “Isolate Adapter and Switch Error” on page 124. If the problem persists contact the IBM Support Center.
-7	Device has been removed from network for not responding	The device was removed from the switch network. An attempt was made to contact the device, but the device did not respond. If the device in question is a node, see “Verify an SP Switch Node” on page 110. Otherwise contact the IBM Support Center.
-8	Device has been removed from network because of a miswire	The device is not cabled properly. There are 2 possible causes for this condition. The first is that the switch network is miswired, the other is that the frame supervisor s tty is not cabled properly. First view the **/var/adm/SPlogs/css/ cable_miswire** file. Verify and/or correct all links listed in the file. Then issue **Eclock -d** and rerun **Estart.** If the problem persists contact IBM Hardware Service.
-9	Destination not reachable	The device was not reachable through the switch network. This is generally due to other errors in the switch network fabric. Investigate and correct the other problem, then **rc.switch** the primary node and rerun **Estart**.

Table 16. SP Switch Link Status		
#	**Text**	**Explanation and Recovery Action**
-2	Wrap plug is installed	This link is connected to a wrap plug. This is not normally a problem.
-4	Link has been removed from network or miswired - faulty	The link is not operational and was removed from the network. The possible causes are: the link is miswired or the link has faulted. First check the **/var/adm/SPlogs/css** directory for the existence of a **cable_miswire** file. If the file exists verify and correct all links listed in the file. Then execute **Eclock -d** and rerun **Estart.** If the **cable_miswire** file doesn't exist, examine the **/var/adm/SPlogs/css/flt** file for entries relating to this link. If entries are found, verify that the cable is seated at both ends, then **rc.switch** the primary node and rerun **Estart.** If the problem persists contact the IBM Support Center.
-6	Link has been removed from network - no AUTOJOIN	The device was removed and isolated from the switch network. The possible causes are: the node was **Efence** without AUTOJOIN, the node was rebooted or powered off, or the node faulted. First attempt to **Eunfence** the device, if the node fails to rejoin the switch network, see “Isolate Adapter and Switch Error” on page 124. If the problem persists contact the IBM Support Center.
-7	Link has been removed from network - fenced	The device was placed offline by the Systems Administrator (**Efence**). **Eunfence** the associated node.
-8	Link has been removed from network - probable miswire	The link is not cabled properly. View the **/var/adm/SPlogs/ css/cable_miswire** file. Verify and correct all links listed in the file, then **rc.switch** the primary node and rerun **Estart.**
-9	Link has been removed from network - not connected	The link can not be reached by the primary node, so initialization of the link is not possible. This is generally cause by other problems in the switch network , such as a switch chip being disabled. Investigate and correct the root problem, then **rc.switch** the primary node and rerun **Estart.**

Note: If the recovery actions previously listed fail to resolve your problem, contact the IBM Support Center.

Ecommand Problems

The error isolation procedure for any Ecommand (**Eclock**, **Eannotator**, etc.) that fails is as follows:

1. View the error output returned from the command, making note of the error message number and text.
2. Find the message in *PSSP Messages Reference* and execute the recommended recovery action.
3. If the recovery action was taken and the problem persists, contact the IBM Support Center.

Note: Many of the Ecommands are global in nature (communicating with all nodes in the system or partition). These commands can sometimes fail because they are unable to communicate with every node. If you suspect this type of failure, see Chapter 9, “Diagnosing Authentication Problems” on page 95.

Note: Many of the Ecommands update or access the SDR to perform their functions. These commands sometimes fail because they cannot access the SDR or the SDR is set up incorrectly. If you suspect this type of failure, see “Verify the System Data Repository (SDR)” on page 112.

Isolate Adapter and Switch Error

To isolate an adapter or switch error for the SP Switch , first view the AIX error log. For switch related errors, login to the primary node, for adapter problems login to suspect node. Once you are logged in, enter the following:

```
errpt | more
ERROR_ID TIMESTAMP    T CL Res Name  ERROR_Description
34FFBE83 0604140393T  T H  Worm      Switch Fault-detected by switch chip
C3189234 0604135793   T H  Worm      Switch Fault-not isolated
```

The Resource Name (Res Name) in the error log should give you an indication of how the failure was detected.

Table 17. Resource Name Failure Indications

Resource name	**Indication**
Worm	The information was extracted from the switch and/or adapter hardware
css	Incorrect status was detected by the css device driver
css0	The css adapter failed diagnostics.

Table 18 (Page 1 of 11). Possible Causes of SP Switch Failures	
Error Description	**Analysis**
Receiver EDC-class error	**Label:** SP_SW_EDC_ERROR_RE **Explanation:** A receiver EDC-class error occurred **Cause:** A transient error in data occurred during transmission over switch links. The EDC error may be one of the following: • A receiver EDC error • A parity error on route • An undefined control character was received • Unsolicited data was received • A receiver lost end-of-packet • A token count miscomparison • A token sequence error • A token count overflow **Cause:** A loose, disconnected, or faulty cable **Action:** • See “Verify Cable” on page 116 • See **/var/adm/SPlogs/css/out.top** for cable information **Cause:** A node was shut down, reset, powered off, or disconnected **Action:** See “Verify an SP Switch Node” on page 110 for additional information **Cause:** A switch adapter hardware failure **Action:** • Run adapter diagnostics • See “Device and Link Problems” on page 120 for additional information
Switch receiver link sync error	**Label:** SP_SW_RCVLNKSYNC_RE **Explanation:** A switch receiver link sync error occurred, or a switch has lost clock synchronization on one of its receive ports **Cause:** A loose, disconnected, or faulty cable **Action:** • Check, reconnect, or replace the cable • See **/var/adm/SPlogs/css/out.top** for cable information • See “Verify Cable” on page 116 for more information **Cause:** A node was shut down, reset, powered off, or disconnected **Action:** • Replace the switch cable from a powered-off node with a wrap plug • See **/var/adm/SPlogs/css/flt** for more information **Cause:** A switch adapter hardware failure **Action:** • Run adapter diagnostics • See **/var/adm/SPlogs/css/flt** for more information **Cause:** A remote switch adapter hardware failure **Action:** • Run adapter diagnostics on remote node • See **/var/adm/SPlogs/css/flt** to identify remote node

Table 18 (Page 2 of 11). Possible Causes of SP Switch Failures

Error Description	Analysis
Switch receiver FIFO overflow error	**Label:** SP_SW_FIFOOVRFLW_RE **Explanation:** A switch receiver FIFO overflow error occurred **Cause:** A loose, disconnected, or faulty cable **Action:** • See "Verify Cable" on page 116 • See **/var/adm/SPlogs/css/flt** for more information **Cause:** A node was shut down, reset, powered off, or disconnected **Action:** • See **/var/adm/SPlogs/css/flt** for more information • See "Verify an SP Switch Node" on page 110 **Cause:** A switch adapter hardware failure occurred **Action:** • Run adapter diagnostics • See **/var/adm/SPlogs/css/flt** for more information • See "Device and Link Problems" on page 120
Switch receiver EDC errors exceed threshold	**Label:** SP_SW_EDCTHRSHLD_RE **Explanation:** Switch receiver EDC errors exceed the threshold level **Cause:** A Loose, disconnected, or faulty cable **Action:** • See "Verify Cable" on page 116 • See **/var/adm/SPlogs/css/flt** for more information • Call IBM Hardware Service if the problem persists
Switch receiver state machine error	**Label:** SP_SW_RECV_STATE_RE **Explanation:** A switch receiver state machine error occurred **Cause:** A switch adapter or switch failure **Action:** Call IBM Hardware Service if the problem persists
Switch sender parity error on data	**Label:** SP_SW_PE_ON_DATA_RE **Explanation:** A switch sender parity error on data occurred **Cause:** A switch board failure **Action:** Call IBM Hardware Service
Switch sender invalid route error	**Label:** SP_SW_INVALD_RTE_RE **Explanation:** A switch sender route that was not valid error **Cause:** A switch adapter microcode or a switch daemon software error **Action:** Call the IBM Support Center
Switch sender token errors exceed threshold	**Label:** SP_SW_SNDTKNTHRS_RE **Explanation:** Switch sender token errors exceed the threshold level **Cause:** A loose, disconnected, or faulty cable **Action:** • See "Verify Cable" on page 116 • Call IBM Hardware Service if the problem persists • See **/var/adm/SPlogs/css/flt** for more information

Table 18 (Page 3 of 11). Possible Causes of SP Switch Failures

Error Description	Analysis
Switch sender link sync error	**Label:** SP_SW_SNDLNKSYNC_RE **Explanation:** A switch chip has lost clock synchronization on one of its send ports **Cause:** A loose, disconnected, or faulty cable **Action:** • Check, reconnect, or replace the cable • See **/var/adm/SPlogs/css/out.top** for cable information • See "Verify Cable" on page 116 **Cause:** A node was shut down, reset, powered off, or disconnected **Action:** • Replace the switch cable from a powered-off node with a wrap plug • See **/var/adm/SPlogs/css/flt** for more information • See "Verify an SP Switch Node" on page 110 **Cause:** A switch adapter hardware failure **Action:** • Run adapter diagnostics • See **/var/adm/SPlogs/css/flt** for more information • See "Device and Link Problems" on page 120 **Cause:** A remote switch adapter hardware failure **Action:** • Run adapter diagnostics on remote node • See **/var/adm/SPlogs/css/flt** to identify remote node • See "Device and Link Problems" on page 120
Switch sender state machine error	**Label:** SP_SW_SND_STATE_RE **Explanation:** A switch sender state machine error occurred **Cause:** A switch adapter or switch failure **Action:** Call IBM Hardware Service if the problem persists
Switch central queue parity error - NMLL	**Label:** SP_SW_PE_ON_NMLL_RE **Explanation:** An NMLL switch central queue parity error occurred **Cause:** A switch board failure **Action:** Call IBM Hardware Service
Switch central queue parity error - NCLL	**Label:** SP_SW_PE_ON_NCLL_RE **Explanation:** An NCLL switch central queue parity error occurred **Cause:** A switch board failure **Action:** Call IBM Hardware Service
Switch central queue NCLL uninitialized	**Label:** SP_SW_NCLL_UNINT_RE **Explanation:** The NCLL switch central queue was not initialized **Cause:** A switch board failure **Action:** Call IBM Hardware Service

Table 18 (Page 4 of 11). Possible Causes of SP Switch Failures

Error Description	Analysis
Switch service logic incorrect CRC	**Label:** SP_SW_CRC_SVCPKT_RE **Explanation:** The switch service logic saw an incorrect CRC on a packet **Cause:** A transient error in data occurred during transmission over switch links **Action:** • See **/var/adm/SPlogs/css/flt** for more information **Cause:** A loose, disconnected, or faulty cable **Action:** • See **/var/adm/SPlogs/css/flt** for more information • See "Verify Cable" on page 116 **Cause:** A node was shut down, reset, powered off, or disconnected **Action:** • See **/var/adm/SPlogs/css/flt** for more information • See "Verify an SP Switch Node" on page 110 for additional information **Cause:** A switch adapter hardware failure **Action:** • Run adapter diagnostics • See **/var/adm/SPlogs/css/flt** for more information • See "Device and Link Problems" on page 120 for additional information
Switch svc logic saw bad packet length	**Label:** SP_SW_SVC_PKTLEN_RE **Explanation:** The switch service logic saw an incorrect packet length **Cause:** A switch adapter microcode error or a switch daemon software error **Action:** • See **/var/adm/SPlogs/css/flt** for more information • Call the IBM Support Center if the problem persists
Switch svc logic bad parity - in FIFO	**Label:** SP_SW_PE_INBFIFO_RE **Explanation:** The switch service logic determined a bad parity in FIFO **Cause:** A switch board failure **Action:** Call IBM Hardware Service
Switch svc logic bad parity - route tbl	**Label:** SP_SW_PE_RTE_TBL_RE **Explanation:** The switch service logic determined that there was a parity error in the route table **Cause:** A switch board failure **Action:** Call IBM Hardware Service
Switch svc logic invalid link enable	**Label:** SP_SW_LNK_ENABLE_RE **Explanation:** The switch service logic saw an invalid link enable value **Cause:** A switch daemon software error occurred **Action:** Call the IBM Support Center

Table 18 (Page 5 of 11). Possible Causes of SP Switch Failures	
Error Description	**Analysis**
Switch svc logic send TOD error	**Label:** SP_SW_SEND_TOD_RE **Explanation:** A switch service logic send TOD error occurred **Cause:** A switch daemon software error **Action:** Call the IBM Support Center
Switch svc logic state machine error	**Label:** SP_SW_SVC_STATE_RE **Explanation:** A switch service logic state machine error occurred **Cause:** A switch adapter or switch failure **Action:** Call IBM Hardware Service if the problem persists
Switch adapter interrupt handler error	**Label:** TB3_SLIH_ER **Explanation:** A switch interrupt handler error occurred **Cause:** A switch adapter or switch failure occurred **Action:** • See neighboring error log entries to determine the cause of the outage • See **/var/adm/SPlogs/css/flt** for more information • Run adapter diagnostics • Call IBM Hardware Service if the problem persists
Switch adapter hardware/microcode error	**Label:** TB3_HARDWARE_ER **Explanation:** A switch adapter hardware or microcode error occurred **Cause:** A switch adapter hardware or microcode error occurred **Action:** • Run adapter diagnostics • Call IBM Hardware Service if the problem persists
Switch adapter hardware/microcode error	**Label:** TB3_MICROCODE_ER **Explanation:** A switch adapter microcode error occurred **Cause:** A switch adapter or switch failure occurred **Action:** • Run adapter diagnostics • Call the IBM Support Center if the problem persists
Switch adapter link outage	**Label:** TB3_LINK_RE **Explanation:** A switch adapter link outage occurred **Cause:** The node is fenced **Action:** Unfence the node **Cause:** A loose, disconnected, or faulty cable **Action:** See “Verify Cable” on page 116

Table 18 (Page 6 of 11). Possible Causes of SP Switch Failures	
Error Description	**Analysis**
Bad packet received	**Label:** TB3_BAD_PACKET_RE **Explanation:** A single bad packet was received **Cause:** A switch cable failure **Action:** See “Verify Cable” on page 116. No action required unless error is frequent **Cause:** A switch adapter or switch failure occurred **Action:** • Run adapter diagnostics • No action required unless error is frequent • Call IBM Hardware Service if the problem persists
Switch adapter transient error	**Label:** TB3_TRANSIENT_RE **Explanation:** A switch adapter transient error occurred **Cause:** A loose, disconnected, or faulty cable **Action:** See “Verify Cable” on page 116
Switch adapter error threshold exceeded	**Label:** TB3_THRESHOLD_ER **Explanation:** Bad packets exceeded the threshold level **Cause:** A loose, disconnected, or faulty cable **Action:** See “Verify Cable” on page 116
Switch adapter svc interface overrun	**Label:** TB3_SVC_QUE_FULL_ER **Explanation:** A switch adapter service interface overrun occurred **Cause:** A switch adapter or switch failure **Action:** Call the IBM Support Center if the problem persists
Failed to update ODM during CSS config	**Label:** TB3_CONFIG1_ER **Explanation:** The ODM was not updated during CSS configuration **Cause:** A software error **Action:** • Run the configuration method with the verbose option for more information • See “Configure and Diagnose Problems” on page 113
I/O error, switch adapter device driver	**Label:** TB3_PIO_ER **Explanation:** An I/O error was received on the switch adapter device driver **Cause:** A switch adapter hardware failure **Action:** • Run adapter diagnostics • Call IBM Hardware Service if the problem persists
Node fence request received	**Label:** SP_SW_OFFLINE_RE **Explanation:** A node fence request was received **Cause:** The operator ran the **Efence** command **Action:** Run the **Eunfence** command to bring the node onto the switch

Table 18 (Page 7 of 11). Possible Causes of SP Switch Failures	
Error Description	**Analysis**
Switch primary node takeover	**Label:** SP_SW_PRI_TAKOVR_RE **Explanation:** A takeover on the switch primary node occurred **Cause:** The switch primary node became inaccessible **Action:** See the error log on the old switch primary node
Switch primary backup node takeover	**Label:** SP_SW_BCKUP_TOVR **Explanation:** A takeover on the switch primary backup node occurred **Cause:** The primary backup node became inaccessible **Action:** • See **/var/adm/SPlogs/css/out.top** for cable information • See the error log on the old switch primary backup node
Primary backup node not responding	**Label:** SP_SW_LST_BUP_CT_RE **Explanation:** The primary backup node is not responding **Cause:** The primary backup node became inaccessible **Action:** See the error log on the current switch primary backup node
Links not initialized during **Estart**	**Label:** SP_SW_UNINIT_LINK_RE **Explanation:** Switch links were not initialized during an **Estart** **Cause:** The switch cable is not wired correctly **Action:** See **/var/adm/SPlogs/css/cable_miswire** to determine if cables were not wired correctly **Cause:** A loose, disconnected, or faulty cable **Action:** See “Verify Cable” on page 116
Process killed due to link outage	**Label:** SP_SW_PROCESS_KILLD_RE **Explanation:** A user process was killed due to a link outage **Cause:** A switch adapter or switch failure **Action:** See neighboring error log entries to determine the cause of the outage **Cause:** The operator fenced this node **Action:** See neighboring error log entries to determine the cause of the outage
Switch cable miswired	**Label:** SP_SW_MISWIRE_ER **Explanation:** The switch cable is not connected to the correct switch jack **Cause:** The switch cable was not wired correctly **Action:** See **/var/adm/SPlogs/css/cable_miswire** to determine if cables were not wired correctly
Eclock command issued by user	**Label:** SP_SW_ECLOCK_RE **Explanation:** The **Eclock** command was executed **Cause:** The **Eclock** command was run by the administrator **Action:** Issue the **Estart** command to initialize the switch network

Table 18 (Page 8 of 11). Possible Causes of SP Switch Failures

Error Description	Analysis
Switch (master oscillator) lost clock	**Label:** SP_MCLCK_MISS_RE **Explanation:** The switch (master oscillator) lost clock **Cause:** A node or a switch board lost power **Action:** • See “Rack or System Clock Problems” on page 117 • Issue the **Estart** command to initialize the switch network **Cause:** A switch board failure **Action:** • See “Rack or System Clock Problems” on page 117 • Issue the **Estart** command to initialize the switch network • Call IBM Hardware Service if the problem persists **Cause:** A user incorrectly clocked the system **Action:** • See “Rack or System Clock Problems” on page 117 • Issue the **Estart** command to initialize the switch network
Switch (non-master) lost clock	**Label:** SP_CLCK_MISS_RE **Explanation:** A switch (non-master) lost clock **Cause:** The switch clock signal is missing **Action:** • Call IBM Hardware Service if the problem persists • See “Rack or System Clock Problems” on page 117
Error writing switch log files	**Label:** SP_SW_LOGFAILURE_RE **Explanation:** An error occurred while writing switch log files **Cause:** The **/var** file system is full **Action:** Obtain free space in the file system or expand the file system **Cause:** There are too many files open in the system **Action:** Reduce the number of open files in the system
Switch daemon initialization failed	**Label:** SP_SW_INIT_FAIL_ER **Explanation:** The switch daemon initialization failed **Cause:** The operating environment could not be established **Action:** • See detail data for specific failure if possible • Correct the problem and restart the daemon • Call the IBM Support Center if the problem persists
Switch daemon received **SIGTERM**	**Label:** SP_SW_SIGTERM_ER **Explanation:** The switch daemon received **SIGTERM** **Cause:** Another process sent a **SIGTERM** **Action:** Run **rc.switch** to restart the switch daemon

Table 18 (Page 9 of 11). Possible Causes of SP Switch Failures

Error Description	Analysis
Switch service send queue full	**Label:** SP_SW_SVC_Q_FULL_RE **Explanation:** The switch service send queue is full **Cause:** There is a traffic backlog on the switch adapter **Action:** Call the IBM Support Center if the problem persists
Switch daemon couldn't get svc request	**Label:** SP_SW_GET_SVCREQ_ER **Explanation:** The switch daemon could not get a service request **Cause:** A switch kernel extension error **Action:** Call the IBM Support Center
Resigning switch primary responsibilities	**Label:** SP_SW_RSGN_PRIM_RE **Explanation:** Resigning switch primary duties **Cause:** Could not communicate over the switch **Action:** See neighboring error log entries to determine the cause of the outage **Cause:** Another node was selected as the primary node **Action:** None
Resigning as switch primary backup	**Label:** SP_SW_RSGN_BKUP_RE **Explanation:** Resigning as the switch primary backup node **Cause:** Could not communicate over the switch **Action:** See neighboring error log entries to determine the cause of the outage **Cause:** Another node was selected as the primary backup node **Action:** None
Switch daemon ACK of svc command failed	**Label:** SP_SW_ACK_FAILED_RE **Explanation:** The switch daemon ACK of svc command failed **Cause:** A switch communications failure **Action:** Call the IBM Support Center if the problem persists **Cause:** A traffic backlog on the switch adapter **Action:** Call the IBM Support Center if the problem persists
Switch daemon SDR communications failed	**Label:** SP_SW_SDR_FAIL_RE **Explanation:** The switch daemon SDR communications failed **Cause:** An Ethernet overload **Action:** Call the IBM Support Center if the problem persists **Cause:** Excessive SDR traffic **Action:** Call the IBM Support Center if the problem persists **Cause:** The SDR daemon or the control workstation is down **Action:** Check to see if the SDR daemon is up **Cause:** A software error **Action:** Call the IBM Support Center if the problem persists

Table 18 (Page 10 of 11). Possible Causes of SP Switch Failures	
Error Description	**Analysis**
Switch adapter i/f system call failed	**Label:** SP_CSS_IF_FAIL_ER **Explanation:** The switch adapter interface system call failed **Cause:** Could not communicate with the switch adapter **Action:** • Check the switch adapter configuration • Run adapter diagnostics • Call the IBM Support Center if the problem persists
Switch scan failed	**Label:** SP_SW_SCAN_FAIL_ER **Explanation:** The switch scan failed **Cause:** Could not communicate over the switch **Action:** Issue the **Estart** command if primary takeover does not occur **Cause:** A switch adapter or switch failure **Action:** Issue the **Estart** command if primary takeover does not occur
Switch node miswired	**Label:** SP_SW_NODEMISW_RE **Explanation:** A switch node was not wired correctly **Cause:** A switch cable was not plugged into the correct node **Action:** See **/var/adm/SPlogs/css/cable_miswire** to determine if cables were not wired correctly
Switch daemon failed to generate routes	**Label:** SP_SW_RTE_GEN_RE **Explanation:** The switch daemon failed to generate routes **Cause:** A software error **Action:** Call the IBM Support Center
Fence of node failed	**Label:** SP_SW_FENCE_FAIL_RE **Explanation:** The fence of a node failed **Cause:** Could not communicate over the switch **Action:** • See **/var/adm/SPlogs/css/flt** for more information • See the error log on the failing node • Issue the **Estart** command to initialize the switch network
Switch daemon reopen windows failed	**Label:** SP_SW_REOP_WIN_ER **Explanation:** The switch daemon could not reopen adapter windows **Cause:** A switch kernel extension error **Action:** Call the IBM Support Center if the problem persists
Estart failed	**Label:** SP_SW_ESTART_FAIL_R **Explanation:** The switch network could not be initialized **Cause:** Could not initialize switch chips or nodes **Action:** • See detail data for specific failure • Issue **Eclock -d** to reset the switch network and reestablish switch clocking • Call the IBM Support Center if the problem persists

Table 18 (Page 11 of 11). Possible Causes of SP Switch Failures	
Error Description	**Analysis**
Switch daemon couldn't reset IP	**Label:** SP_SW_IP_RESET_ER **Explanation:** The switch daemon could not reset IP **Cause:** A switch kernel extension error **Action:** Call the IBM Support Center if the problem persists
Switch daemon DBupdate broadcast failed	**Label:** SP_SW_UBCAST_FAIL_R **Explanation:** The switch daemon DBupdate broadcast failed **Cause:** A switch communications failure **Action:** Call the IBM Support Center if the problem persists **Cause:** A traffic backlog on the switch adapter **Action:** Call the IBM Support Center if the problem persists
Switch daemon command broadcast failed	**Label:** SP_SW_CBCAST_FAIL_R **Explanation:** A switch daemon command broadcast failed **Cause:** Could not communicate over the switch **Action:** Call the IBM Support Center if the problem persists **Cause:** A traffic backlog on the switch adapter **Action:** Call the IBM Support Center if the problem persists
Switch daemon dependent node svc failure	**Label:** SP_SW_DEPNOD_FAIL_R **Explanation:** A switch daemon dependent node svc failure **Cause:** Could not communicate over the switch **Action:** Call the IBM Support Center if the problem persists **Cause:** A traffic backlog on the switch adapter **Action:** Call the IBM Support Center if the problem persists
Switch Daemon Process Terminated	**Label:** HPS_FAULT6_ER **Explanation:** The **fault_service_Worm_RTG_SP** daemon was terminated. **Cause:** Bad switch adapter or missing external clock source **Action:** • See preceding error log entries for failure cause • See “Adapter Diagnostic Failures” on page 114 **Cause:** Bad system planar **Action:** Run complete diagnostics on the node. If diagnostics fail to isolate the problem, contact IBM Hardware Service
Switch Adapter failed POST diagnostics	**Label:** SWT_DIAG_ERROR1_ER **Explanation:** The switch adapter failed Power-On-Self-Test diagnostics **Cause:** Bad switch adapter or missing external clock source **Action:** See “Adapter Diagnostic Failures” on page 114.

Diagnose SP Switch Estart Problems

Refer to the following list of steps to diagnose Estart failures:

1. Login to the primary node.
2. View the bottom of the **/var/adm/SPlogs/css/fs_daemon_print.file**.
3. Use the failure listed to index the following table:

Table 19. SP Switch Estart Problem Possible Causes

Message	Analysis
Error in **buildDeviceDatabase()**	**Explanation:** Unable to build the device database **Cause:** Missing or corrupt Topology file **Action:** See "Verify Switch Topology Configuration" on page 112. **Cause: malloc** failures **Action:** Contact the IBM Support Center
Error in **TBSswitchInit()**	**Explanation:** Unable to initialize the switch network **Cause:** Switch initialization failed **Action:** See "SP Switch Worm Errors" on page 137.
Error in **writeDeviceDatabase()**	**Explanation:** Unable to write **/var/adm/SPlogs/css/out.top** **Cause:** Missing or corrupt Topology file **Action**: See "Verify Switch Topology Configuration" on page 112. **Cause**: **/var** is not large enough to accommodate the new **out.top** file **Action**: Increase the size of **/var**
No valid backup - SDR current Backup being changed to none	**Explanation:** Informational message **Cause:** No node available as a backup **Action:** No action required
Can't access SDR - SDR current Backup not changed	**Explanation:** SDR failure(s) **Cause:** SDR not setup properly **Action:** See "Verify the System Data Repository (SDR)" on page 112.
Error in: **fopen(act.top.PID)** **fprintf(act.top.PID)** **fclose(act.top.PID)** **rename(act.top, act.top.PID)**	**Explanation**: An error occurred accessing **/var/adm/SPlogs/css/act.top.PID** **Cause:** File access problems **Action:** Evaluate the errno returned and take the appropriate action. If the problem persists contact the IBM Support Center

Note: If the message found at the bottom of **/var/adm/SPlogs/css/fs_daemon_print.file** was not found in the previous table or the actions specified did not correct the problem, contact the IBM Support Center.

SP Switch Worm Errors

Refer to the following list of steps to diagnose worm initialization failures:

1. Login to the primary node.
2. View the bottom of the **/var/adm/SPlogs/css/worm.trace** file. You should find a message similar to one of the following:

   ```
   TBSworm_bfs_phase1() failed with rc=xx
   ```

   or

   ```
   TBSworm_bfs_phase2() failed with rc=xx
   ```

3. Use the rc value of either of these messages to index the following table:

Table 20 (Page 1 of 2). SP Switch Worm Return Codes

Return Code	Analysis
-3	**Explanation:** Local adapter receiver port is not enabled **Cause:** The switch is not clocked **Action:** From the control workstation issue **Eclock -d** then **Estart** **Cause:** Oncoming Primary is fenced off of the switch **Action:** See “Eunfence the Oncoming Primary” on page 119.
-4	**Explanation:** Unable to generate routes for the network **Cause:** Corrupt topology file **Action:** See “Verify Switch Topology Configuration” on page 112.
-5	**Explanation:** Send packet from local node failed **Cause:** Bad switch adapter **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem contact the IBM Support Center.
-6	**Explanation:** A switch miswire was detected **Cause:** Switch network cabling does not match the switch topology file **Action:** View the **/var/adm/SPlogs/css/cable_miswire** file to determine which cables are in question. Then check and reconnect the associated cables. If the problem persists contact IBM Hardware Service.
-7	**Explanation:** A node miswire was detected **Cause:** Switch network cabling does not match the switch topology file **Action:** The device is not cabled properly. There are 2 possible causes for this condition. The first is that the switch network is miswired, the other is that the frame supervisor's tty is not cabled properly. First view the **/var/adm/SPlogs/css/cable_miswire** file. Verify and correct all links listed in the file. Then issue **Eclock -d** and rerun **Estart**. If the problem persists contact IBM Hardware Service.
-8	**Explanation:** Receive FIFO is full **Cause:** Bad Switch Adapter **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem contact IBM Hardware Service. **Cause:** The switch is backed up from a node or a switch chip **Action:** Contact the IBM Support Center

Table 20 (Page 2 of 2). SP Switch Worm Return Codes	
Return Code	**Analysis**
-9	**Explanation:** Unable to initialize FIFOs **Cause:** Bad Switch Adapter **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem contact IBM Hardware Service.
-27	**Explanation:** The TBIC was not initialized **Cause:** The switch adapter is uninitialized **Action: rc.switch** the primary node, then reissue **Estart** from the control workstation **Cause:** Bad Switch Adapter **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem contact IBM Hardware Service.
-36	**Explanation:** This node resigned as the primary node **Cause:** The node determined it could no longer control and monitor the switch. The backup primary node is now in control of the switch. **Action:** No action required
-43	**Explanation:** A read or write operation to the switch adapter failed. **Cause:** Bad Switch Adapter **Action:** Run switch adapter diagnostics on the primary node. If diagnostics fails to isolate the problem contact IBM Hardware Service.

Note: If you do not see the return code found at the bottom of: **/var/adm/SPlogs/css/worm.trace** in the previous table or the actions specified did not correct the problem, contact the IBM Support Center.

Chapter 12. SP Switch Advanced Diagnostic Tools

ATTENTION - READ THIS FIRST: Do **not** activate the SP Switch advanced diagnostic facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.

Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

You can run advanced switch diagnostic tests from the control workstation when you suspect that a component in the switch network is not performing properly. The reason for running a diagnostic test is usually an error reported in the system error log, or one of the CSS log files. In many cases, it is difficult to isolate the cause of these errors, since the same error may be caused by a hardware failure of different network components, or even for some other reason (such as the network being initialized by the switch management software).

If you detect such errors, you may decide to invoke an appropriate diagnostic test, as explained in the following sections. These tests help to identify the failing component, and determine whether your system has a "real" hardware problem or not. The "positive" result of a test is a message that indicates a hardware failure. In this case, contact IBM hardware support.

Detailed information about the commands used to invoke these tests can be found in *PSSP: Command and Technical Reference*.

There are cases in which the test reports software errors. This indicates either a temporary error condition (for example, the **Estart** command is invoked while the test is running, in which case the test should be restarted), or a permanent problem in the test itself or in other PSSP software. For these cases, contact the IBM Support Center. The displayed message contains the number of the node reporting the problem. Before calling IBM, save the following information: node number of the primary node and the reporting node; and the log file **/var/adm/SPlogs/css/spd.trace** on the primary node and on the reporting node.

During the test, different messages are displayed on the SPD GUI (graphical user interface for SP switch diagnostics) or to your console. These messages include several fields: node ID of the reporting node, its personality (primary, backup or secondary), timestamp, type of the message, and the message itself. There are several types of messages: informational messages, warnings, and errors. Warnings and error messages are marked by "!" and "E" respectively (in the proper field on the GUI). You can ignore informational messages (which do not have such marks).

For detailed information about warning and error messages, see *PSSP: Messages Reference*.

Adapter Error Log Analyzer (ELA)

Adapter ELA is an extension of adapter diagnostics. It is executed on switch nodes in order to diagnose problems that occurred on the node, but cannot be reproduced by the adapter diagnostics tests. The reason for running the adapter ELA and the commands to invoke it are the same as those for adapter diagnostics and AIX diagnostics in general. Adapter ELA does not have any impact on the system.

When to Run the Adapter ELA

Run adapter ELA in all cases where you suspect that the SP Switch adapter is not functioning properly. These are some examples:

- The system error log contains adapter-related errors.
- The node loses connectivity to the rest of the switch network.
- An SP Switch diagnostic test reports an adapter error.
- An SP Switch diagnostic test recommends that the adapter be checked on a specific node.

How to Run the Adapter ELA

There are several ways to run the adapter ELA using the AIX **diag** command:

1. To run the adapter ELA directly from the command line, issue this command:

   ```
   diag -d css0 -e
   ```

   Wait until the tool gives you progress messages.

2. To run the adapter ELA from a GUI:

   a. Issue the AIX **diag** command without operands.

   b. When the first screen appears, press the enter key.

   c. Highlight the line "Task Selection (Diagnostics, Advanced Diagnostics, Service Aids, etc...)" and press the enter key.

   d. On the next screen, scroll down until you see "Run Error Log Analysis", highlight it and press the enter key.

   e. On the next screen, scroll down until you see **css0**, highlight it and press the enter key.

   f. A plus sign **(+)** appears to the left of **css0**.

   g. Press PF7 to commit. This starts the adapter ELA.

   h. Messages indicating the results of the adapter ELA appear on the top of the screen.

Interpreting the Results of the Adapter ELA

The adapter ELA first notifies you that it is testing the adapter, and asks you to stand by. Then, it displays an **ELA recommendations** screen, which either informs you that no problems were found, or gives a list of recommendations. These messages are self-explanatory. For example, the message may say that an SP Switch adapter problem was detected, and that you should contact IBM Software Support or IBM Hardware Support.

If the adapter ELA results are not displayed, or the adapter ELA reports that it failed to operate correctly (for example, wrong or missing files), the problem is probably due to incorrect PSSP installation. In this case, contact the IBM Support Center.

SP Switch Stress Test

This test verifies the functionality of a specific switch chip or switch port. The **switch_stress** command starts this test. For detailed information about this command, its flags, arguments, and usage examples, see *PSSP: Command and Technical Reference*. The **-g** flag can be used to run the test with the SPD GUI, but you must specify the command operands and flags on the command line.

When to Run the SP Switch Stress Test

Run this test when you suspect that a switch chip is faulty because the system error log on the primary node contains error status reports from the switch chip. This test checks the functionality of this switch chip and the attached links, in order to decide whether the switch chip should be replaced.

How to Run the SP Switch Stress Test

First, decide which switch chip you want to test. Usually this is a switch chip that is reporting errors. Then, decide which nodes to use for the test. **These nodes cannot run parallel applications during the execution of the test.** By default, all nodes are allowed, so be careful to avoid disturbing applications that are running. The nodes that are not allowed will not be affected directly. However, since the test implies stress traffic in the switch network, the performance of applications running on **all** nodes may be affected.

Invoke the test specifying the desired switch chip ID. Also specify the nodes that can execute the test, or alternatively the nodes that are forbidden (because they are running critical applications).

The test does not require user intervention. It executes several iterations, each one having a different combination of switch chip ports. In each iteration, the test sends data through the ports under test. In the beginning of the iteration, the test notifies you as to which nodes are executing it. At the end of the iteration, it displays these statistics: number of packets that were sent and lost, and number of switch errors (reported from all switch chips used for sending data by the test iteration).

Interpreting the Results of the SP Switch Stress Test

Each error reported by a switch chip is displayed. In a stable system, there should be few or no such reports. If the test succeeded to stimulate a critical fault on one of the switch chips, the test decides that its goal has been achieved (the faulty component is isolated). In this case, the test displays an appropriate message and terminates. Contact IBM hardware support to replace the faulty component.

Otherwise, the test just displays the statistics and continues to the next iteration. If the test did not cause critical faults, but did cause some failures (that were recovered), it does not necessarily mean that some hardware component should be replaced, but gives an indication of a possible cause of problems. Contact IBM hardware support in this case also.

Multiple Senders/Single Receiver Test

This test detects nodes that are injecting corrupted packets into the switch network. The **mult_senders_test** command starts this test. For detailed information about this command, its flags, arguments, and usage examples, see *PSSP: Command and Technical Reference*. The **-g** flag can be used to run the test with the SPD GUI, but you must specify the command operands and flags on the command line.

When to Run the Multiple Senders/Single Receiver Test

Run this test if one or more of the nodes are reporting that they received "bad packets." This may indicate a situation where there is a malfunctioning switch adapter in the system that is generating bad packets. You wants to detect such "bad sender" nodes. The Multiple Senders/Single Receiver Test finds the malfunctioning switch adapter among all of the nodes in the system.

How to Run the Multiple Senders/Single Receiver Test

Select a receiver node to be used by the test. The receiver node is usually one of the nodes that are reporting bad packet events. Then decide which nodes can be used as senders. **The receiver node and nodes that will be used as senders cannot run parallel applications during the test execution.** By default, all nodes are allowed, so be careful to avoid disturbing applications that are running. The nodes that are not allowed will not be affected directly. However, since the test implies stress traffic in the switch network, the performance of applications running on **all** nodes may be affected.

Invoke the test specifying the desired receiver node. Also specify the nodes that are allowed to be used as senders, or alternatively the nodes that are forbidden (because they are running critical applications).

The test does not require user intervention. In the beginning, the test notifies the user which sender nodes are executing. The selected senders send data to the receiver and the test progress is monitored until it completes.

Interpreting the Results of the Multiple Senders/Single Receiver Test

The test monitors error reports from all switch network components during the iterations. If a critical fault occurs, the test displays a message about it and terminates. In this case, contact IBM hardware support to replace the faulty component. Otherwise, the test continues until all iterations are done, and then either displays a message that contains the list of bad senders, or notifies you that no errors were found.

SP Switch Wrap Test

This test verifies the functionality of a link. The **wrap_test** command starts this test. For detailed information about this command, its flags, arguments, and usage examples, see *PSSP: Command and Technical Reference*. This command must be used with the SPD GUI.

When to Run the SP Switch Wrap Test

Run the SP Switch Wrap test if:

- The **Estart** command reports that it failed to initialize links or nodes.
- The error log on the primary node contains error reports about a link between two switches.
- The error log on the primary node contains error reports about a link between a node and a switch.

These conditions may indicate a hardware problem in the cable between two ports. Run this test to determine which component should be replaced. The Switch Wrap Test identifies the specific failing component.

How to Run the SP Switch Wrap Test

First, decide which link to test, using the information in the error logs. If the link under test is a link connecting a switch to a node, fence the node before running the test. If the link under test is a link connecting two switches, be aware that during the test the link will be disabled.

The test guides you through several steps. In each step, a message box is displayed asking you to perform some operation, such as installing the wrap plug, according to the displayed information. You can either perform the requested operation and press OK, or press CANCEL. The test then checks the corresponding link component and displays the results (component passed or failed). If the user pressed CANCEL, or in a few other cases, the test cannot check the component and the test terminates. In these cases, it displays a message that explains the situation.

At the end of the test, you are requested to perform tasks to restore the system to the state it was in prior to the test.

Interpreting the Results of the SP Switch Wrap Test

The test displays explicit diagnosis information for the link components.

Chapter 13. Diagnosing System Connectivity Problems

If a node becomes unresponsive or inaccessible, use the following table to diagnose the problem.

Table 21. System Connectivity Symptoms

Symptom	Recovery
Using the Hardware Perspective, bringing up **hostResponds** in a table view for nodes and multiple nodes shows: **Node Not Responding.** Using the Hardware Perspective, bringing up **switchResponds** in a table view for nodes and multiple nodes shows: **Adapter Not Configured.**	"Action 1. Diagnose Multiple Nodes"
Using the Hardware Perspective, either the table view or nodes status page of the notebook shows: **Node Not Responding** for **hostResponds** or **Adapter Not Configured** for **switchResponds.**	"Action 2. Diagnose Individual Nodes"
Cannot access the node using **rsh, telnet, rlogin,** or **ping**.	"Action 3. Diagnose a Network Problem" on page 146
Cannot access the node using **telnet or rlogin,** but can access the node using **ping**.	The is a probable software error. Initiate a dump, record all relevant information and contact the IBM Support Center.
Can access the node using **telnet** or **ping**, but **hostResponds** still shows: **Node Not Responding.**	"Action 4. Diagnose a Topology-Related Problem" on page 146

Actions

Action 1. Diagnose Multiple Nodes

If several node icons in a frame report a failure, (either the nodes are not responding or several adapters are inactive) there may be a network problem.

If the failing nodes or communication adapters are on the same Local Area Network (LAN), verify the LAN hardware. If you determine that the hardware is functioning properly, call the IBM Support Center. Otherwise, follow local procedures for servicing your hardware.

If the nodes are not on the same LAN, diagnose the nodes individually as described in "Action 2. Diagnose Individual Nodes."

Action 2. Diagnose Individual Nodes

If an individual node icon in a frame reports a failure, use the Hardware Perspective to display the Nodes Status page in the Node notebook, for the failing node.

1. Check the node's LCD/LED indicator.
2. If a three-digit code is displayed, check Chapter 37, "SP-Specific LED/LCD Values" on page 257 to see if the code is described there. If the code is not described in this section, refer to *IBM RS/6000 Problem Solving Guide*.
3. Check the **hostResponds** indicator for a failure.

4. Check the node's power indicator.

   If it shows the node power is off, turn the node's power on.

   If it shows the node power is on or if the problem persists, call IBM hardware support.

Action 3. Diagnose a Network Problem

If a node is not responding to a network command, you can access the node using the tty. This can be done by using the Hardware Perspectives, selecting the node and performing an **open tty** action on it. It can also be done by issuing the

```
s1term -w frame number slot number
```

command, where *frame number* is the frame number of the node and *slot number* is the slot number of the node.

Using either method, you can login to the node and check the hostname, network interfaces, network routes, and hostname resolution to determine why the node is not responding. The Appendix entitled 'IP Address and Host Name Changes for SP Systems' in *PSSP: Administration Guide* contains a procedure for changing host names and IP addresses.

Action 4. Diagnose a Topology-Related Problem

If the **ping** and **telnet** commands are successful, but **hostResponds** still shows **Node Not Responding**, there may be something wrong with the Topology Services **(hats)** subsystem. Perform these steps:

1. Examine the en0 (Ethernet adapter) and css0 (switch adapter) addresses on all nodes to see if they match the addresses in **/var/ha/run/hats.***partition_name***machines.lst**.
2. Verify that the netmask and broadcast addresses are consistent across all nodes. Use the **ifconfig en0** and **ifconfig css0** commands.
3. Examine the **hats** log file on the failing node. It is named: **/var/ha/log/hats.***dd.HHMMSS.partition_name*, where *dd.HHMMSS* is the day of the month and time of day when the Topology Services daemon was started, and *partition_name* is the name of the node's system partition.
4. Examine the **hats** log file for the Group Leader nodes. Group Leader nodes are those that host the adapter whose address is listed below the line "Group ID" in the output of the **lssrc -ls hats** command. For more information, see the Topology Services chapter in *PSSP: Administration Guide*.

Chapter 14. Diagnosing System Monitor Problems

- If you cannot monitor or operate hardware controls (such as powering on/off frames, nodes, or switches), check **/var/adm/splogs/SPdaemon.log** for messages that specifically indicate a hardware problem. Search for messages with resource name **sphwlog**.

  If you cannot find any messages related to hardware problems, check the table below for other symptoms.

- If you have a problem with the **spmon**, **hmmon**, **hmcmds**, **hmadm**, and **s1term** System Monitor commands, check for the following symptoms to diagnose the problem:

Table 22. System Monitor Symptoms

Symptom	Recovery
System Monitor commands not found	"Action 1. Verify Installation"
System Monitor commands fail	"Action 2. Verify Authorization" on page 148
The Hardware Monitor daemon dies	"Action 4. Check for a Core Dump" on page 148 "Action 7. Check the Hardware Monitor Daemon (hardmon)" on page 149
The Hardware Monitor **s70d** daemon dies	"Action 11. Check for Core Dump of SP- Attached Server" on page 150 "Action 12. Check the s70d daemon" on page 150
The logging daemon dies	"Action 4. Check for a Core Dump" on page 148
Performance problems on the control workstation	"Action 8. Check Performance" on page 149
Log file **/var/adm/SPlogs/SPdaemon.log** is no longer being updated or SP hardware error messages are no longer being written to the AIX error log	"Action 6. Check Logging Daemon" on page 148 "Action 7. Check the Hardware Monitor Daemon (hardmon)" on page 149
A logging daemon user exit is not being called when it should	"Action 10. Start State Change Logging" on page 150
Note: If the actions in this table do not resolve your problem, if you can re-create the problem, document the procedure and call IBM support for further assistance.	

Refer to the chapter on using the System Monitor in *PSSP: Administration Guide* for more detail on performing various actions.

Actions

Action 1. Verify Installation

1. Make sure the **ssp.basic** option of the **pssp.installp** image was installed. The **ssp.basic** option provides system monitoring function.

   - Run verification tests using SMIT or the command line to ensure installation is complete.

     Using SMIT:

TYPE **smit SP_verify**

- The Installation/Configuration Menu appears.

SELECT SSP System Monitor Installation

PRESS Enter.

Using the command line, enter:

```
/usr/lpp/ssp/bin/spmon_itest
```

2. Add **/usr/lpp/ssp/bin** to your **PATH** environment variable or use the full path name with the command.

Action 2. Verify Authorization

1. If your Kerberos ticket has expired, reissue **k4init**. Make sure the Kerberos principal name and optional instance for your ID is in the Hardware Monitor Action Control List (ACL) file, **/etc/hmacls**.
2. If you are unable to resolve the Kerberos problem refer to Chapter 8, “Diagnosing SDR Problems” on page 93.

Action 3. Export Windows

If you want to export your Xwindows to another X-server, be sure that the **DISPLAY** environment variable is set properly and that the X-server has authorized you to use it.

Action 4. Check for a Core Dump

Check **/var/adm/SPlogs/spmon/hardmon** or **/var/adm/SPlogs/spmon/splogd** for a core dump. If these files exist, save them.

Action 5. Frame Supervisor Communication Diagnosis

If this is a problem communicating with a frame, perform the diagnosis procedure described in Chapter 7, “Diagnosing Frame Supervisor Communication Problems” on page 91.

Action 6. Check Logging Daemon

If logging stops working, check the following:

1. If only **syslog** stops working:
   a. **/etc/syslog.conf** must have an entry for **daemon.notice** (or **daemon.debug** or **daemon.info**) in order to write to **/var/adm/SPlogs/SPdaemon.log**
   b. **/var/adm/SPlogs/SPdaemon.log** file must exist.
   c. **/var/adm/SPlogs/SPdaemon.log** file must have permissions set to **777**.

   Once this is set up, sending the HUP signal to **syslogd** will cause **syslogd** to reread its configuration file and start logging.
2. If only error logging stops working:
   a. Make sure the error record templates for SPMON have been defined:

      ```
      errpt -t | grep SPMON
      ```

   b. Make sure the error daemon is running:

```
ps -ef | grep errdemon
```

c. Make sure hardware events are happening by shutting down a node and bringing it back on.

3. If no logging is being done:

   a. Make sure the logging daemon is running:

   ```
   ps -ef | grep splogd
   ```

   b. Make sure the **/usr/lpp/ssp/config/hwevents** file is set up to do logging (with a valid SP_ERROR_LOG entry).

   c. Turn debugging on by editing **/usr/lpp/ssp/bin/splog.start** and adding a **-d** option to the line that starts with **splogd**. Then, **kill** the logging daemon so that it will restart.

Action 7. Check the Hardware Monitor Daemon (hardmon)

Try a hardmon monitor command from an ID that has monitor authority and a valid ticket-granting ticket from **k4init**. For example, to check frame 1, node 1, enter:

```
hmmon -Q 1:1
```

If this command does not work, check the hardmon log **/var/adm/SPlogs/spmon/hmlogfile**.*nnn*, where *nnn* is the Julian date of when the file was created.

Action 8. Check Performance

1. Check that the paging space is adequate and adjust if necessary.
2. Check the overall CPU utilization. You can use the **vmstat** command to do this.
3. Check the CPU utilization of the hardmon and logging daemon. One method is issuing:

   ```
   ps gvc | grep hardmon
   ps gvc | grep splogd
   ```

If the CPU utilization rate is very high and cannot be attributed to the hardmon or logging daemon, look for other processes which are consuming the CPU resources. If you are using your control workstation as a boot file server, then check if the NFS daemons are using all of the processor time.

Action 9. Check Logs

Check the **/var/adm/SPlogs/SPdaemon.log** for hardware error messages. If the file is not found there, check the **/etc/syslog.conf** file to see where messages written to the daemon facility are being directed. You can also find the same messages in the AIX error log. Issue the following command to get a detailed report of the SP hardware errors (found under resource name **sphwlog**):

```
errpt -aN sphwlog. | pg
```

Action 10. Start State Change Logging

To get more information about what hardware variables are changing, state change logging can be enabled. Edit the **/usr/lpp/ssp/config/hwevents** file and uncomment the line that defines the function SP_STATE_LOG. Then, kill the logging daemon (**splogd**) so that it rereads the **/usr/lpp/ssp/config/hwevents** file.

Action 11. Check for Core Dump of SP- Attached Server

Check **/var/adm/SPlogs/spmon/s70d** for a core dump. If these files exist, save them.

Action 12. Check the s70d daemon

Try a hardmon monitor command from your SP-attached server using an ID that has monitor authority and a valid ticket-granting ticket from **k4init**. For example, to check frame 2, node 1, enter:

```
hmmon -Q 2:1
```

If this command does not work, check the hardmon log **/var/adm/SPlogs/spmon/hmlogfile**.*nnn*, where *nnn* is the Julian date when the file was created. Also check **/var/adm/SPlogs/spmon/s70d/s70d**.*frame*.**log.***nnn*, where *frame* is the frame number and *nnn* is the Julian date when the file was created.

Chapter 15. Diagnosing SP Perspectives Problems

Perspectives problems may arise in one of the following categories. See the accompanying table to diagnose these problems.

Table 23. Perspectives SymptomTypes

Type of problem	**Table to Reference**
General problems running the Launch Pad or any of the Perspectives	Table 24
General problems running the Hardware Perspective	Table 25
General problems running the Event Perspective	Table 26 on page 152
General problems running the IBM Virtual Shared Disk Perspective	Table 27 on page 152

Table 24. Launch Pad and General Perspectives Symptoms

Symptom	**Recovery**
Perspectives command not found (**perspectives**, **sphardware**, **spevent**, **spvsd**, **spsyspar** or **spperfmon**).	See “Action 1. Verify SP Perspectives Installation” on page 153
The Launch Pad or Perspectives fails to come up.	See “Action 2. Export the DISPLAY Variable” on page 153
The Launch Pad or Perspectives terminates prematurely.	See “Action 7. Check For a Core Dump” on page 155
The Launch Pad or Perspectives hangs.	See “Action 8. Check Performance of the System” on page 155
During startup, you receive a message stating that you cannot run this application directly.	See “Action 9. Run Perspectives From /usr/lpp/ssp/bin” on page 155
During startup, you receive a message stating that you have only read access to the SDR.	See “Action 3. Obtain Root Access to the Control Workstation” on page 153
The SP TaskGuides icon does not appear.	Ensure that the **ssp.tguides** option of the **pssp.installp** image was installed.
When trying to open the Perspectives online help, you receive the following message: 'The requested online help is either not installed or not in the proper search path. The Help Volume is: **Help4Help**, Location ID: **QUICK-HELP**.' For information on installing online help, consult *PSSP: Installation and Migration Guide*.	See “Action 15. Install the File Set Needed for Perspectives Online Help” on page 157

Table 25 (Page 1 of 2). Hardware Perspectives Symptoms

Symptom	**Recovery**
The Hardware Perspective will not permit you to power on or off: nodes, frames or switches.	See “Action 4. Check SP Hardware Monitor Authorization to Control Hardware” on page 153

Table 25 (Page 2 of 2). Hardware Perspectives Symptoms	
Symptom	**Recovery**
You receive a message that the connection to the Event Manager was lost. If you were monitoring hardware, all icons now have question marks, indicating an unknown state.	See “Action 5. Check the Event Manager Daemon” on page 154
The Node Environment page of the Node Notebook is blank. This means that the hardmon resource monitor is down or locked.	See “Action 6. Check the Resource Monitors” on page 154
One or more of the following: • The option to Cluster Power On is missing from the Power On dialog. • The options for Shutdown and Fence are missing from the Power Off dialog. • The action to Fence or Unfence is missing from the Actions menu.	See “Action 14. Obtain Sysctl Authority” on page 157

Table 26. Event Perspectives Symptoms	
Symptom	**Recovery**
You receive a message in a dialog box that the connection to the Event Manager was lost. The Event Perspective closes when you press the OK button.	See “Action 5. Check the Event Manager Daemon” on page 154
During startup, you receive a message that an error occurred while trying to retrieve your Kerberos principal.	See “Action 10. Check Your Kerberos Authorization” on page 155
During startup, you receive a message that you do not have update authority for the problem manager.	See “Action 11. Obtain Authority to the Problem Management Subsystem” on page 156
Buttons and fields of the Actions page of the Event Definition Notebook are not selectable.	See “Action 11. Obtain Authority to the Problem Management Subsystem” on page 156
The actions you defined on the Actions page of the Event Definition Notebook fail.	See “Action 10. Check Your Kerberos Authorization” on page 155
You are not allowed to create conditions.	See “Action 3. Obtain Root Access to the Control Workstation” on page 153

Table 27 (Page 1 of 2). IBM Virtual Shared Disk Perspectives Symptoms	
Symptom	**Recovery**
You receive a message dialog stating that the connection to the Event Manager was lost. If you were monitoring hardware, all icons now have question marks, indicating an unknown state.	See “Action 5. Check the Event Manager Daemon” on page 154
During startup, you are not authorized to run the Perspective.	See “Action 12. Set Up Correct Authorization to Run the IBM Virtual Shared Disk Perspective” on page 156
You are not able to create IBM Virtual Shared Disks or IBM Hashed Shared Disks (HSDs).	See “Action 13. Prepare Disks for the createvsd or createhsd Commands” on page 156

Table 27 (Page 2 of 2). IBM Virtual Shared Disk Perspectives Symptoms	
Symptom	**Recovery**
IBM Virtual Shared Disk information in notebooks is not being updating. For example, information in notebooks and the table view is not being update.	See "Action 6. Check the Resource Monitors" on page 154
SDR information is not being updated automatically. For example, information in notebooks and the table view is not being updated.	"Action 6. Check the Resource Monitors" on page 154

For more detail on performing various actions, see the chapter on using the SP Perspectives in *PSSP: Administration Guide*.

Note: If the actions in these tables do not resolve your problem, determine if you can re-create the problem. If you can, document the procedure and call the IBM Support Center for further assistance.

Actions

Action 1. Verify SP Perspectives Installation

To verify that SP Perspectives is installed properly:

1. Ensure that the **ssp.gui**, **ssp.csd.gui** and **ssp.top.gui** options of the **pssp.installp** image were installed. If the **spperfmon** command was not found, ensure that the **ssp.ptpegui** option of the **pssp.installp** image was installed. For further information, see *PSSP: Installation and Migration Guide*.
2. Add **/usr/lpp/ssp/bin** to your **PATH** environment variable or use the full path name with the command. The full path name of a perspectives command is: **/usr/lpp/ssp/bin/***command-name*.

Action 2. Export the DISPLAY Variable

In order to login to the control workstation and display to another machine, do the following:

1. Export your **DISPLAY** variable to the X-server of the machine that you want to use.
2. Use the **xhost +** command to ensure that you have authorization to use that X-server.

Action 3. Obtain Root Access to the Control Workstation

Some Perspectives require write access to the SDR. In order to write to the SDR, you must have **root** authority on the control workstation.

Action 4. Check SP Hardware Monitor Authorization to Control Hardware

In order to perform certain actions, such as power on and off hardware, you must have your user ID registered in the SP Hardware Monitor Access Control List (ACL).

Use the following procedure to authorize your user ID to use the SP Hardware Monitor. If you are not authorized to make these changes, ask the person who administers security on your system to perform these steps:

1. Add your user ID to the **/spdata/sys1/spmon/hmacls** file on the control workstation.
2. Refresh the **hardmon** daemon by executing the **hmadm setacls** command on the control workstation.

Action 5. Check the Event Manager Daemon

If you receive messages that Perspectives has lost its connection to the Event Manager, the problem could be that the Event Manager daemon, **haemd** terminated or that the network connection to the Event Manager daemon was lost. Perform the following steps:

1. Exit Perspectives.
2. Check that the Event Manager daemon is up and running by issuing the **lssrc -a | grep haem** command.

   This lists the Event Manager daemon. If the system is partitioned, the daemon will be listed for each system partition.
3. If any **haem** daemon is listed as inoperative, restart the daemon by issuing the **startsrc -g haem** command.
4. Issue the **lssrc -a | grep haem** command again to verify that the Event Manager daemon is now up and running.
5. Restart Perspectives.

Action 6. Check the Resource Monitors

Resource monitors are software components that provide resource variables to the Event Manager daemon. Here are some examples of resource variables and the resource monitors that supply them to the Event Manager daemon:

Table 28. Perspectives Resource Variables

Resource Variable Name	Resource Monitor
IBM.PSSP.Response.Host.State	Response
IBM.PSSP.Response.Switch.State	Response
IBM.PSSP.SP_HW.Node.envLED	IBM.PSSP.hmrmd
IBM.PSSP.SP_HW.Node.lcd1	IBM.PSSP.hmrmd
IBM.PSSP.aixos.FS.%totused	aixos

To check if any of the resource monitors are down or locked, issue the **lssrc -ls haem.***syspar_name* command for each system partition.

For example, if you have two system partitions named k4s and k4sp1, you would type: **lssrc -ls haem.k4s** to check resource monitors in the first system partition, and: **lssrc -ls haem.k4sp1** to check resource monitors in the second system partition.

You will see a listing of **haem** information similar to the following:

```
Resource Monitor Information
Resource Monitor Name    Inst  Type   FD     SHMID      PID     Locked
IBM.PSSP.CSSLogMon        0      C    -1      -1         -2    No  00/00
IBM.PSSP.SDR              0      C    -1      -1         -2    No  00/00
IBM.PSSP.harmld           0      S    19    8202      38118    No  01/01
IBM.PSSP.harmpd           0      S    17      -1      41572    No  01/01
IBM.PSSP.hmrmd            0      S    18      -1      26882    No  01/01
IBM.PSSP.pmanrmd          0      C    15      -1         -2    No  00/00
Membership                0      I    -1      -1         -2    No  00/00
Response                  0      I    -1      -1         -2    No  00/00
aixos                     0      S    12   16393         -2    No  00/01
```

All nine resource monitors should be listed. They should all be unlocked and up.

If any resource monitors are locked, the entry in the Locked column will be **yes**. Issue the **haemunlkrm** command to unlock the resource monitor. For example, if the hardmon resource monitor (IBM.PSSP.hmrmd) is locked, issue the **haemunlkrm -s haem -a IBM.PSSP.hmrmd** command.

Action 7. Check For a Core Dump

Check the directory you are running from for a file named **core** with a current time/date stamp. If the file exists, save it. If you are able to re-create the problem, save all relevant logs, messages and other information and call the IBM Support Center.

Action 8. Check Performance of the System

Check the overall CPU utilization of the control workstation to see if any processes are consuming a large amount of time. See if any of the following are consuming a large amount of CPU time:

- The Perspectives processes (**perspectives**, **sphardware**, **spevent**, **spvsd**, **spsyspar** or **spperfmon**)
- The underlying subsystems: **hardmon**, **s70d**, **hatsd**, **sdrd**, **haemd**, or **hagsd**

If any of these processes are continually consuming a large amount of CPU time, they may need to be brought down and restarted.

Action 9. Run Perspectives From /usr/lpp/ssp/bin

You tried to run one of the Perspectives commands: **perspectives**, **sphardware**, **spevent**, **spvsd**, **spsyspar** or **spperfmon** from **/usr/lpp/ssp/perspectives/bin**. These are the executables and cannot be run directly. Run the command again from the **/usr/lpp/ssp/bin** directory, or add **/usr/lpp/ssp/bin** to your **PATH** variable.

Action 10. Check Your Kerberos Authorization

Perform the following steps to detect and correct SP Perspectives problems with Kerberos.

1. Check to see if you have a Kerberos principal. If not, establish one by issuing the **k4init** command.
2. Check to see if your Kerberos ticket has expired. If so, reissue the **k4init** command.

3. Check to see if the **sysctld** daemon, used to retrieve Kerberos principals, is running. If not, start the daemon by issuing the **sysctld** command.

Action 11. Obtain Authority to the Problem Management Subsystem

Add your Kerberos principal name to the **/etc/sysctl/pman.acl** file. If you are not authorized to make these changes, ask the person who administers security on your system to do so.

Action 12. Set Up Correct Authorization to Run the IBM Virtual Shared Disk Perspective

Set up the authorization needed to run the IBM Virtual Shared Disk Perspective by following these steps:

1. Login as **root** on the control workstation.
2. Edit the **/etc/sysctl.acl** file to include the **root.admin** principal. The line should look something like:

   ```
   _PRINCIPAL root.admin@PPD.POK.IBM.COM
   ```

   There should be no spaces at the beginning of the line.
3. Edit the **/etc/sysctl.vsd.acl** file to include the **root.admin** principal. The line should look something like:

   ```
   _PRINCIPAL root.admin@PPD.POK.IBM.COM
   ```

   There should be no spaces at the beginning of the line.
4. Copy the first file to all nodes with this command: **pcp -a /etc/sysctl.acl**.
5. Copy the second file to all nodes with this command: **pcp -a /etc/sysctl.vsd.acl**.
6. Issue the **k4init root.admin** command on each node that does not have Kerberos tickets.
7. Run the **k4list** command on each node to verify that the tickets were created successfully by issuing the **dsh -a k4list** command.
8. Verify that the setup is correct by issuing the **vsdsklst -a** command.

   If data is returned, the permissions have been set up correctly. Otherwise, you will receive **Kerberos** or **sysctl** error messages.
9. Repeat 1 through 8 for each system partition.

Action 13. Prepare Disks for the createvsd or createhsd Commands

1. On the control workstation, check which disks are free for creating new IBM Virtual Shared Disks. Issue the **vsdsklst -a** command.
2. On nodes that have the message: **No allocated physical disks**, issue the **lspv** command. If the Physical Volume ID (pvid) is **None** and the volume group is **None**, issue the **chdev -l** *disk_name* **-a pv=yes** command . Then issue the **lspv** command again to verify that the Physical Volume ID was successfully assigned.

   Note: Twin-tailed disks have the same Physical Volume ID on each node.

Action 14. Obtain Sysctl Authority

Add your user ID to the **/etc/sysctl.acl** file. If you are not authorized to make these changes, ask the person who administers security on your system to do so.

Action 15. Install the File Set Needed for Perspectives Online Help

In order for the Perspectives online help information help window to be available, the **X11.Dt.helpinfo** file set must be installed. For information on installing Perspectives, see *PSSP: Installation and Migration Guide*.

Chapter 16. Diagnosing SP TaskGuides Problems

Use this table to analyze problems encountered when executing SP TaskGuides.

Table 29. SP TaskGuides Symptoms

Symptom	Recovery
The **sptg** command is not found.	1. Ensure that the **ssp.tguides** option of the **pssp.installp** image was installed. 2. Add **/usr/lpp/ssp/bin** to your **PATH** environment variable or use the full path name with the command, **/usr/lpp/ssp/bin/sptg**.
The SP TaskGuide selection GUI or one of the specific TaskGuides fails to come up.	Export your **DISPLAY** variable to the X-server of the machine that you want to use. Use the **xhost +** command to ensure that you have authorization to use that X-server.
An SP TaskGuide hangs.	Ensure that the SP TaskGuide is really hung, as opposed to executing a time-consuming command. Start the SP TaskGuide again and examine the log of the previous invocation for errors. If you cannot resolve the problem, record all relevant information and contact the IBM Support Center.
An SP TaskGuide displays an error dialog and will not proceed.	This is a problem with an AIX or PSSP command issued by the SP TaskGuide. Determine why the command failed from the error panel displayed.
An SP TaskGuide displays an error dialog with a 'panel not found' message.	This is an internal error. Record all relevant information and contact the IBM Support Center.

Chapter 17. Diagnosing Job Switch Resource Table Services Problems

Use the following table to diagnose problems with the Job Switch Resource Table (JSRT) Services component of PSSP. Locate the symptom and perform the action described in the following table.

Table 30. Job Switch Resource Table (JSRT) Services Symptoms

Symptom	Recovery
Cannot load or unload a Job Switch Resource Table (JSRT) on a node.	"Action 1. Verify JSRT Services Installation" "Action 2. Check the JSRT Services Log File" on page 162 "Action 3: Request More Detailed Log Information" on page 163 "Action 5: Check the switch_node_number File" on page 164 "Action 6: Check the Current Status of JSRT Services for a Node" on page 164
Cannot obtain the status of a JSRT window.	"Action 1. Verify JSRT Services Installation" "Action 2. Check the JSRT Services Log File" on page 162 "Action 4: Check the JSRT Services Data Files" on page 164
Cannot run the **switchtbld** daemon.	"Action 1. Verify JSRT Services Installation" "Action 2. Check the JSRT Services Log File" on page 162

Actions

Action 1. Verify JSRT Services Installation

Run installation verification tests using SMIT or the command line to ensure installation is complete.

Using SMIT:

```
TYPE    smit SP_verify
        (the Installation/Configuration Menu appears)
SELECT  Job Switch Resource Table Services Installation
PRESS   Enter
```

Using the command line, enter: **/usr/lpp/ssp/bin/st_verify**

When executed on the control workstation, the **st_verify** script checks the installation of the JSRT on every node that is defined in the current system partition. When executed on a single node, it verifies the installation of the JSRT only on that node.

The **st_verify** script checks that the correct files and directories were installed and that the necessary entries exist in the files. The files and directories are **/etc/services**, **/etc/inittab**, and **/etc/inetd.conf**.

Installation Verification Test Output

The **st_verify** script produces an output log, located in **/var/adm/SPlogs/st/st_verify.log** (by default) or in a location that you specify. After execution, a message is written to stdout stating whether the verification passed or failed. If a failure occurred, inspect the log for a list of the errors that were found.

If JSRT Services has been installed correctly on the control workstation, a message similar to the following is written to stdout:

```
Verifying installation of the Job Switch Resource Table Services on node 0.
Job Switch Resource Table Services installation verification SUCCESSFUL on node 0.
Check /var/adm/SPlogs/st/st_verify.log file for further details.
```

Action 2. Check the JSRT Services Log File

The JSRT Services maintains a single log file, **st_log**, which is located in: **/var/adm/SPlogs/st**. This log is located on every node where the services are used. For example, if the **swtbl_load_job** API is used, entries are found on the local node where the API was invoked and entries are also be found on the nodes that were being loaded by the **swtbl_load_job** API.

Examine the logs and correct any obvious problems that have been identified.

The following table indicates return codes that may appear in the log. They are defined in **/usr/lpp/ssp/include/st_client.h**

Table 31. JSRT Services Return Codes

Return Code	Name	Explanation
0	ST_SUCCESS	The service request was successful.
1	ST_INVALID_TASK_ID	An invalid task ID is specified as input.
2	ST_NOT_AUTHOR	The caller is not authorized to perform the service.
3	ST_NOT_AUTHEN	The caller is not authenticated to perform the service.
4	ST_SWITCH_IN_USE	The JSRT is already loaded or in use.
5	ST_SYSTEM_ERROR	A system error occurred.
6	ST_SDR_ERROR	An SDR error occurred.
7	ST_CANT_CONNECT	The **connect** system call failed.
8	ST_NO_SWITCH	No css device is installed.
9	ST_INVALID_PARAM	An invalid parameter was specified as input.
10	ST_INVALID_ADDR	The **inet_ntoa** command failed on the **st_addr** input value.
11	ST_SWITCH_NOT_LOADED	No JSRT is currently loaded.
12	ST_UNLOADED	A previously successful load was unloaded because of an error.
13	ST_NOT_UNLOADED	No unload request was issued.
14	ST_NO_STATUS	No status request was issued.
15	ST_DOWNON_SWITCH	The node is down on the switch.
16	ST_ALREADY_CONNECTED	The node has already been loaded.
17	ST_LOADED_BYOTHER	The JSRT was loaded outside of the API.
18	ST_SWNODENUM_ERROR	An error occurred when processing the switch node number.

Action 3: Request More Detailed Log Information

To facilitate debugging, you can set an environment variable before invoking a JSRT service. The variable provides more detailed information in the **st_log** file. The environment variable is **SWTBLAPIERRORMSGS** and it must be set to **"yes"**.

For example, as a **ksh** user, enter:

```
export SWTBLAPIERRORMSGS=yes
```

Here is an example of the more detailed log information for a call to **swtbl_load_table**:

```
Thu Jun 18 10:12:22 1998: swtbl_load_table: INPUT PARAMETERS: uid – 0
pid – 19118 job_key – 1 requestor_node – k10n11.ppd.pok.ibm.com
num_tasks – 1 job_desb – load client test
Thu Jun 18 10:12:22 1998: swtbl_load_table: INPUT PARAMETERS virtual
 task id=0 switch_node_number=10 window_id=1
```

Action 4: Check the JSRT Services Data Files

The JSRT Services maintains a set of datafiles that are located in the **/spdata/sys1/st** directory on every node. Verify that the directory exists and that the files have **root** access. Note that the files are not created until the load or unload services have been invoked.

Action 5: Check the switch_node_number File

The **/spdata/sys1/st/switch_node_number** file contains a single integer that represents the switch node number of the node. This number should match the **switch_node_number** attribute in the SDR Node class for that node. Issue the **/usr/lpp/ssp/bin/st_set_switch_number** command on every node to create the switch_node_number file and set the correct value.

Action 6: Check the Current Status of JSRT Services for a Node

The **st_status** command shows you the current status of the JSRT windows on all nodes or on the node specified. This tells you whether the JSRT windows are loaded, unloaded, loaded by another subsystem, or in error.

To show the status of JSRT windows on all nodes within the current system partition, issue **st_status**.

To show the status of all JSRT windows on node **k10n15**, issue: **st_status k10n15.** Output similar to the following appears:

```
*****************************
Status from node: k10n15  User: root
Load request from: k10n15  Pid:12494  Uid:0
Job Description: No_job_description_given
Time of request: Wed_Jan_24_13:38:21_1998 Window_id:0
*****************************
Node k10n15 Window 1 ST_SWITCH_NOT_LOADED
*****************************
Node k10n15 Window 2 ST_SWITCH_NOT_LOADED
*****************************
Node k10n15 Window 3 ST_SWITCH_NOT_LOADED
```

AIX Error Logs and Templates for JSRT Services

The JSRT services component uses the AIX error log facility to record specific events. To view the JSRT error events, issue the following command on the node where an error is suspected:

```
errpt -aN Switch_Table | more
```

When you retrieve an error log entry, look for the “Diagnostic Explanation” section near the bottom of the entry.

Table 32 on page 165 shows the error log templates used by JSRT services.

Table 32. AIX Error Log Templates for JSRT Services

ID	Error Type	Diagnostic Explanation and Action
ST_TRUNCATE_ST	UNKN	**Explanation:** check_size: Copied *logname* to *logname*.**previous** and truncated *logname.* **Cause:** The st_log exceeded 100 KB. It was copied and truncated. **Action:** None required.
ST_SWITCH_ERR	PERM	**Explanation:** An error occurred during the processing of a JSRT service. **Cause:** The request for a load, unload, or clean failed. **Actions:** Perform these tasks: • Check the file **/var/adm/SPlogs/st/st_log** for further information. • Perform Switch diagnostics. • Issue the **/usr/lpp/ssp/css/fs_dump** command and look for messages about the **UNLOAD_ST**, **LOAD_ST** or **QUERY_ST ioctl** commands.

Chapter 18. Diagnosing User Access Problems

If your users are having problems logging into the SP System or accessing their home directories, locate the symptom and perform the action described in the following table.

Table 33. User Access Symptoms

Symptom	Recovery
No one can login to an SP node	Reboot the node in maintenance mode. If no one is able to login, including **root**, follow the steps in Chapter 24, "Diagnosing Devices Including the Hard Disk" on page 195 for rebooting a node in maintenance mode.
User unable to login to SP node	"Action 1. Check the /etc/security/passwd File" "Action 2. Check Login Control"
User unable to access directories managed by the automounter	"Action 3: Verify that the Automount Daemon is Running"

Actions

Action 1. Check the /etc/security/passwd File

If a user is having problems logging in to nodes in the SP System, check the **login** and **rlogin** attributes for the user in the **/etc/security/passwd** file on the SP node.

Action 2. Check Login Control

Check the Login Control facility to see whether the user's access to the node has been blocked.

The System Administrator should verify that the user is allowed access. The System Administrator may have blocked interactive access so that parallel jobs could run on a node.

Action 3: Verify that the Automount Daemon is Running

On AIX 4.3.1 and later systems, the AutoFS function replaces the automount function of AIX 4.3.0 and earlier systems. All automount functions are compatible with AutoFS. With AutoFS, file systems are mounted directly to the target directory instead of using an intermediate mount point and symbolic links.

Review the commands in the following table and issue the ones that are appropriate for diagnosing the problem.

Table 34. Automounter Related Commands	
Command	**Comments**
ps -ef | grep automount	Verifies that the automount daemon is running on the system on which you are having problems accessing directories.
lssrc -g autofs	For AIX 4.3.1 and later systems, the automounter is controlled by the System Resource Controller (SRC). This command indicates whether the **automountd** daemon is active or not.
mount	For AIX 4.3.0 and earlier systems, provides the process id of the automount daemon if it is running, the names of the file systems controlled by the automount daemon, and the active mounts under the **/tmp_mnt** directory. For AIX 4.3.1 and later systems, provides the names of the file systems controlled by the automounter daemon and lists any currently active mounts under the **/tmp_mnt** directory.
view /var/adm/SPlogs/auto/auto.log	Contains error messages generated by PSSP.
view /var/adm/SPlogs/SPdaemon.log	Contains error messages generated by the automount daemon.
splstdata -e | grep amd_config	Informs whether SP automounter has been configured.
splstdata -e | grep usermgmt_config	Informs whether SP user management support has been configured.
splstdata -e | grep filecoll_config	Informs whether SP file collections have been configured.
view /etc/auto.master	Lists the file systems to be controlled by automount and their associated map files.
ls -l /etc/auto/maps	Lists of map files and whether they are readable. Specifically, the existence of **auto.u** map file.
view /etc/auto/maps/auto.u	Lists the user map entries for the **/u** file system.
ls -l /etc/auto/cust	Lists customization files and whether they are executable.
view /var/sysman/sup/lists/user.admin	Lists automounter files that are distributed through file collections.

It may be that the automounter daemon is not running. It is also possible that automount is running but that there is another problem. For AIX 4.3.0 or earlier systems, issue:

```
ps -ef | grep automount
```

For AIX 4.3.1 or later systems, issue:

```
lssrc -g autofs
```

1. Automount Is Not Running

If issuing the previous command did not show that the automount process was running, issue:

```
mount
```

to see if any automount points are still in use. If you see an entry similar to the following one, there is still an active automount mount point. This is for AIX 4.3.0 and earlier systems:

```
luna.pok.ibm.com (pid234500@/u) /u afs Nov 07 15:41 ro,noacl,ignore
```

For AIX 4.3.1 and later systems, the output is:

```
/etc/auto/maps/auto.u /u autofs Aug 07 11:16 ignore
```

Attempt to unmount the file system by issuing:

```
unmount /u
```

If the file system is busy, issue the following command to determine the processes accessing the file system. Stop all of these processes and attempt to unmount the file system again.

```
fuser /u
```

If the **mount** command does not show any active mounts for automount, issue the following command to start the autmounter:

```
/etc/auto/startauto
```

Proceed as follows:

- **startauto** succeeds

  If this command succeeds, issue the previous **ps** or **lssrc** command again to verify that the automount daemon is actually running. If so, verify that the user directories can be accessed or continue with "2. Automounter Is Running, But the User Cannot Access User Files" on page 170.

  Note that the automount daemon should be started automatically during boot. Check to see if your SP system is configured for automounter support by issuing:

  ```
  splstdata -e | grep amd_config
  ```

  If the result is *true*, you have automounter support configured for the SP in your Site Environment options.

  If the **startauto** command was successful but the automount daemon is still not running, check to see if the SP automounter function has been replaced by issuing:

  ```
  ls -l /etc/auto/cust
  ```

  If the result of this command contains an entry similar to:

  ```
  -rwx ----- 1 root system 0 Nov 12 13:20 startauto.cust
  ```

  the SP function to start the automounter has been replaced. View this file to determine which automounter was started and follow local procedures for diagnosing problems for that automounter.

  If the result of the **ls** command does not show any executable user customization script, check both the automounter log file **/var/adm/SPlogs/auto/auto.log** and the daemon log file **/var/adm/SPlogs/SPdaemon.log** for error messages. Find the recorded error messages in *PSSP: Messages Reference* or in the AIX error message documentation and follow the recommended actions.

- **startauto** fails

  If the **startauto** command fails, find the reported error messages in *PSSP: Messages Reference* and follow the recommended actions. Check the automounter log file **/var/adm/SPlogs/auto/auto.log** for additional messages. Also, check the daemon log file **/var/adm/SPlogs/SPdaemon.log** for messages that may have been written by the automounter daemon itself.

If no error messages were recorded, the failure may be due to problems with the automount map files, master map file, or the **/u** directory. Check the following:

- Verify that all entries in the automount master map file **/etc/auto.master** are correct and follow the format specified in the AIX publication *System Management Guide: Communications and Networks*. If there are no entries in this file, the automount daemon invocation will fail.
- Verify that each map file the master map references is correct and follows the format specified in the same AIX publication.
- Verify that each file system listed in the master map file is a local directory and is not a symbolic link to another directory. *PSSP: Administration Guide* contains a chapter on managing the Automount that contains useful information for understanding your SP automounter installation.

2. Automounter Is Running, But the User Cannot Access User Files

For an AIX 4.3.0 or earlier system, if the result of issuing the **ps -ef | grep automount** command is similar to:

```
root 21430    1   0  10:37:41    0:00 /usr/sbin/automount
/etc/auto.master  -m -D HOST=k22n11
```

then the automount daemon is running.

For an AIX 4.3.1 or later system, if the result of issuing the **lssrc -g autofs** command is similar to:

```
Subsystem       Group       PID     Status
automountd      autofs      12126   active
```

then the automount daemon is running.

The problem may be that automount is waiting for a response from an NFS server that is not responding, or that there is a problem with a map file.

Check the **/var/adm/SPlogs/SPdaemon.log** for information relating to NFS servers not responding. If a user's files are mounted with NFS and the server is not responding, then automount may hang on the NFS mount request. This NFS failure should be resolved prior to continuing. After you resolve the NFS failure, you can restart the automount daemon.

If the problem does not appear to be related to an NFS failure, you will need to check your automount maps. Look at the **/ect/auto/maps/auto.u** map file to see if an entry for the user exists in this file. If the user's name is test, and the command **cd /u/test** results in:

```
ksh:  /u/test:   not found
```

you can look at the **auto.u** map to see if there is an entry defined for the user by issuing:

```
grep test /etc/auto/maps/auto.u
```

The result may indicate that there is no entry for this user in the automounter map. This can happen if the user was recently added and the maps have not been distributed in the file collections. To check if the file gets updated with the new map

copied from the control workstation, issue the following command on the node experiencing the problems:

```
supper update -v user.admin
```

Note that the automount maps are automatically distributed to the nodes each hour by command in the cron. You can look at these commands with **crontab -l**.

After you updated the **auto.u** map with the version that contains the user information, reread the **auto.u** map by issuing:

```
grep test /etc/auto/maps/auto.u
```

If the result appears as follows:

```
test   luna:/home/luna:&
```

issue the following:

```
cd /u/test
```

If the **cd** command does not work, there may not be a route to the hostname specified in the host field of the user's automount map entry. This can happen on file servers where there are multiple interfaces and the routing has not been defined for all of the interfaces from the systems attempting to access the server. You can verify this by attempting to **ping** the server specified.

Another possible problem is that the server is exporting the file system to an interface that is not the interface from which the client is requesting the mount. This problem can be found by attempting to mount the file system manually on the system where the failure is occurring. For the map in the previous example, you could issue:

```
mount luna:/home/luna /mnt
```

to mount the file system on **/mnt**. Listing the contents would show the user's files. If the map file information is incorrect use the **spchuser** command from the control workstation to update the map file entry for the user. For example, if the test user's home directory moves from *luna* to *starship*, you would issue:

```
spchuser home=starship:/home/starship/test test
```

This will update the automount map file. You must then wait up to five minutes with no access attempts to the **/u/test** directory. This will allow the automount daemon to time out any old access attempts to the previously mounted **luna:/home/luna** exported file system. Do not attempt to force the unmount by manually issuing the **unmount** command on the previously mounted file system. This will put the automounter daemon in an inconsistent state and you will need to stop and restart the daemon to recover access to the **/u/test** directory.

Stopping and Restarting automount

If you have determined that you need to stop and restart the automount daemon, the cleanest and safest way is to reboot the system. However, this may not be desired due to other processes currently running on the system. If you cannot reboot the system, follow the steps in 171 for an AIX 4.3.0 or earlier system, or the steps in 172 for an AIX 4.3.1 or later system.

1. Determine whether any users are already working in directories mounted by the automount daemon. Issue:

```
mount
```

If the automount daemon is running, you will see an entry similar to:

```
luna.pok.ibm.com (pid23450@/u) nfs Nov 07 15:41
rc,noacl,ignore
```

Also, if a user is working in a directory mounted by the automount daemon, you will see an entry similar to:

```
luna.pok.ibm.com /home/luna /tmp_mnt/u/test nfs Nov 18 11:30
rw,hard,noacl,rsize=4096,wsize=4096,timeo=40,retry=3
```

You can request the user to either logoff or **cd** out of their home directory so that the directory will not longer be in use. If any directory managed by the automount daemon is in use at the time the daemon is stopped, the mount will continue to exist until it is manually removed or the system is rebooted.

2. Stop the automount daemon

   kill -term *process_id*

   where *process_id* is the process number listed by the previous **mount** command. You can also determine the automount process number by issuing:

   ```
   ps -ef | grep automount

   root 23450     1   0 11:35:49   -  0:00 /usr/sbin/automount -f
    /etc/auto.master -m -D HOST=luna
   ```

   To stop the process shown in this example, you would issue:

   ```
   kill -15 23450
   ```

   The commands **kill -term pid** and **kill -15 pid** are identical.

   Note that it is important that you DO NOT stop the daemon with the **kill -kill** or **kill -9**. This will prevent the automount daemon from cleaning up its mounts and releasing its hold on the file systems. It may cause file system hangs and force you to reboot your system to recover those file systems.

3. Start the automount daemon

   ```
   /etc/auto/startauto
   ```

   You can verify that the daemon is running by issuing the previous **mount** or **ps** commands.

1. Determine whether any users are already working on the directories mounted by the **autmountd** daemon. Issue: **mount**

   If automounter is controlling any file systems, you will see an entry similar to:

   ```
   /etc/auto/maps/auto.u /u  autofs Aug 07 11:16 ignore
   ```

   Also, if a user is working in a directory mounted by the atumounter, you will see an entry similar to:

   ```
   luna.pok.ibm.com luna.pok.ibm.com:/home/luna/test /u/test nfs Aug 10 10:37
   ```

   You can request the user to either logoff or **cd** out of their home directory so that the directory will no longer be in use. If any directory managed by the automounter is accessed while the daemon is stopped, the file system may hang.

2. Stop the **automountd** daemon with this command:

```
stopsrc -g autofs
```

Note that it is important that you DO NOT stop the deamon with the **kill -kill** or **kill -9** command. This may cause file system hangs and force you to reboot your system to recover those file systems.

3. Restart the autmounter

```
/etc/auto/startauto
```

You can verify that the daemon is running by issuing the previous **lssrc** command.

Chapter 19. Diagnosing NIM Problems

This chapter discusses problems installing and configuring SP nodes with the Network Installation Management (NIM) service. See *PSSP: Installation and Migration Guide* for information about the installation process, *PSSP: Administration Guide* for information about system management, and *IBM AIX Version 4.3 Network Installation Management Guide and Reference* for more NIM information.

Useful NIM Commands

NIM diagnosis and recovery tasks commonly require you to reset or recover NIM objects. Before and after you do this, you should list the NIM database information to find out the current NIM status and contents. Several NIM commands can help you determine how NIM is configured and can reset or remove some NIM objects.

Listing NIM Database Information

Use the **lsnim** command to list NIM database information. Use **lsnim** to list only the objects, and **lsnim -l** to list all the objects and their attributes. An example of **lsnim** usage is:

```
> lsnim
master              machines        master
boot                resources       boot
nim_script          resources       nim_script
spnet               networks        ent
lppsource           resources       lpp_source
psspspot            resources       spot
noprompt            resources       bosinst_data
prompt              resources       bosinst_data
psspscript          resources       script
mksysb_1            resources       mksysb
spnet_en1           networks        ent
k47n01              machines        standalone
k47n02              machines        standalone
⋮
k47n12              machines        standalone
k47n13              machines        standalone
```

Managing NIM Objects in the NIM Database

Issue the **nim** command to manage NIM objects in the NIM database. See *IBM AIX Version 4.3 Network Installation Management Guide and Reference* for the **nim** command and all its options.

Using the nim Command to Unconfigure the NIM Master

Unconfigure the NIM master only if you need to start completely over. Do this only after other recovery attempts have failed.

1. Issue:

   ```
   nim -o unconfig master
   ```

2. To uninstall the NIM master lpp issue:

   ```
   installp -u bos.sysmgt.nim.master
   ```

3. To reinstall the lpp and redefine and configure the NIM master issue:

   ```
   setup_server
   ```

setup_server reinstalls the **bos.sysmgmt.nim.master** lpp and redefines and configures the NIM master. This command may take some amount of time to complete depending on the number of lpps installed.

Reviewing NIM Client Definition

A closer look at the NIM client definition on the NIM master can often lead you to the source of a problem. Follow these steps to review the NIM client definition.

1. Determine the client's NIM master by issuing:

   splstdata -b -l *client_node_number* and looking at the **srvr** field which lists the NIM master's *node_number*.

2. Determine the NIM master host name by issuing:

   splstdata -b -l *server_node_number*.

   (You can skip this step if this is the control workstation.)

3. Login to the NIM master using **telnet** or **rsh** and the reliable host name.
4. Use **lsnim** to list the objects in the NIM master's database.
5. Use **lsnim -l** *client_name* to list the NIM client definition for the node that is having the problem.
6. Review the information in the following table and ensure that the necessary resources are allocated for bootp_response. If not, run the **mknimres** command or the **setup_server** command on either the control workstation or the server.

Table 35. NIM Client Definition Information

bootp_response	**Cstate**	**Allocations**	
install	BOS installation has been enabled	spot	psspspot
		lpp_source	lppsource
		bosinst_data	noprompt
		script	psspscript
		mksysb	mksysb_1
diag	diagnostic boot has been enabled	spot	psspspot
		bosinst_data	prompt
maintenance	BOS installation has been enabled	spot	psspspot
		bosinst_data	prompt
disk or customize	ready for a NIM operation		
migrate	BOS installation has been enabled	spot	psspspot
		lpp_source	lppsource
		bosinst_data	migrate
		script	psspscript
		mksysb	mksysb_1

Export Problems

NIM exports the directories that are locations of the resources the NIM client needs to perform the installation.

Occasionally NIM may be unable to configure a NIM client when a NIM client definition is not entirely removed from the exported directories it manages. Here is an example of the successful export, by the **exportfs** command, of a NIM client, k47tpn09, which is ready to be installed:

exportfs

```
/spdata/sys1/install/pssplpp   -ro
/spdata/sys1/install/pssp/noprompt
/spdata/sys1/install/pssp/pssp_script
/spdata/sys1/install/images/bos.obj.min.41
                               -ro,root=k47tpn09.hpssl.kgn.ibm.com
/export/nim/scripts/k47tpn09.script
                               -ro,root=k47tpn09.hpssl.kgn.ibm.com
/spdata/sys1/install/lppsource -ro
```

Note: The **/spdata/sys1/install/images** directory entry for your node reflects your **mksysb** image name and the **noprompt** name depends on the installation disks you have selected.

The **/spdata/sys1/install/pssplpp** and **/spdata/sys1/install/lppsource** directories are exported as *read-only* by everyone.

A problem occurs if the client you are attempting to define is listed in some of these directories but the resource has not been allocated to the NIM client as described above. This may happen if NIM has not successfully removed the NIM client in a previous NIM command.

To determine if this is a problem with the definition of your nodes as NIM clients, use the **exportfs** command to list the exported directories on the NIM master where the problem is occurring.

1. List the exports

   ```
   > exportfs

   /spdata/sys1/install/pssplpp              -ro
   /spdata/sys1/install/lppsource            -ro
   /spdata/sys1/install/images/bos.obj.min.41 -ro,root=k47tpn09
   ```

2. To correct this situation you must issue the **rmnfsexp** command to remove the client from the exported list.

   rmnfsexp -d /spdata/sys1/install/images/bos.obj.min.41

   ```
   unexported /spdata/sys1/install/images/bos.obj.min.41
   ```

3. List the exports again to verify that the exported directory has been removed from the export list.

   exportfs

   ```
   /spdata/sys1/install/pssplpp   -ro
   /spdata/sys1/install/lppsource -ro
   ```

Once the NFS export has been corrected, you can issue **setup_server** on the NIM master to redefine the NIM client.

Conflicting NIM Cstate and SDR information

A NIM client may be in a state that conflicts with your intentions for the node. You may intend to install a node, but **setup_server** returns a message that the **nim -o bos_inst** command failed for this client. When **setup_server** runs on the NIM master to configure this node it detects that the node is busy installing and does not reconfigure it. This can happen for several reasons.

- During a node NIM **mksysb** installation the client node being installed was interrupted before the successful completion of the node installation.
- A node was booted in diagnostics or maintenance mode and now you would like to reinstall it.
- The node was switched from one **bootp_response** to another.

Each of these occurrences causes the client to be in a state which appears that the node is still installing.

To correct this problem, issue the following command for the NIM client.

nim -Fo reset *client_name*

Allocation of a Resource to a Client Fails

In some situations an allocation for a resource may fail when **setup_server** is attempting to allocate those resources to the client.

Allocating the SPOT Resource Fails

If allocating the SPOT resources fails follow these steps to determine and correct the problem:

1. Perform a check on the SPOT by issuing **nim -o check psspspot**. This check should inform you if there is a problem.
2. If you are unable to determine the problem with the SPOT, you can update the SPOT by issuing **nim -o cust psspspot**
3. Finally, you can remove the SPOT with **nim -Fo remove psspspot** and then run **setup_server** to recreate the SPOT.

   Note: **setup_server** can take some amount of time to create the SPOT.

Creation of the mksysb Resource Fails

If **setup_server** cannot create the **mksysb** resource, verify that the specified mksysb image is in the **/spdata/sys1/install/images** directory.

Creation of the lppsource Resource Fails

If **setup_server** is unable to create the lppsource resource, verify that the minimal required file sets reside in the **lppsource** directory

```
/spdata/sys1/install/lppsource
```

The file **/usr/lpp/bos.sysmgt/nim/methods/c_sh_lib** contains a variable SIMAGES_OPTIONS which lists required file sets. NIM uses this variable when creating an lpp_source.

This is the listing of the SIMAGES_OPTIONS variable:

```
SIMAGES_OPTIONS="\
    bos \
    bos.info.any \
    bos.net \
    bos.rte.up \
    bos.rte.mp \
    bos.diag \
    bos.powermgt \
    bos.sysmgt \
    bos.terminfo.all \
    devices.all \
    X11.apps \
    X11.base \
    X11.compat \
    X11.Dt \
    X11.fnt \
    X11.loc \
    X11.motif \
    X11.msg.all \
    X11.vsm"
```

To successfully create the lppsource resource on a boot/install server, **setup_server** must acquire a lock in the lppsource directory on the control workstation. Failure to acquire this lock may mean that the lock was not released properly. This lock file contains the hostname of the system that currently has the lock. The lock file is: **/spdata/sys1/install/lppsource/lppsource.lock**.

Login to the system specified in the lock file and determine if **setup_server** is currently running. If it is not running, remove the lock file and run **setup_server** again on the system that failed to create the lppsource resource.

Allocation Failures

The following error sequence :

- 0042— 001 nim: processing error encountered on "master" :
- rshd: 0826–813 Permission is denied. rc=6.
- 0042–006 m_allocate: (From_Masster) rcmd Error 0
- allnimres: 0016–254: Failure to allocate lpp_source resource lppsource_defualt from server (node_number) (node_name) to client (node_number) (node_name)
- (nim —o allocate ; rc=1)

This failure is caused by incorrect or missing **rcmd** support on the control workstation, in the **./rhosts** file, for the Boot Install Server (BIS) nodes. The **./rhosts** file needs to have an entry for the BIS hostname when trying to execute the **allnimres** command.

The **setup_server** command on the BIS node should correct this problem.

Missing installp images

If you determine that you are missing installp images do the following actions:

1. Download them from the AIX 4.3 installation tape to the appropriate lppsource resource directory
2. Either
   a. remove the lppsource resource and run **setup_server**

   or
   b. unconfigure the NIM master, uninstall the bos.sysmgt.nim.master lpp, and run **setup_server**.

Note: This command may take some amount of time to complete depending on the number of lpps installed.

Creation of the SPOT Resource Fails

If **setup_server** fails to create the SPOT (Shared Product Object Tree) resource, verify that the following resources are available:

1. First use the **df** command to verify that none of the file systems are full.
   a. Verify that the **root** (**/**) file system has enough space. The NIM network install images will reside in **/tftpboot**. NIM recommends that you have enough room for eight 2.5MB images or a total of 20MB.
   b. There should be at least 1MB of free space in **/tmp**.
2. List the lppsource resource to see if it is available. Here is a sample of a valid lppsource resource:

```
> lsnim -l lppsource

lppsource:
   class       = resources
   type        = lpp_source
   server      = master
   location    = /spdata/sys1/install/lppsource
   alloc_count = 0
   Rstate      = ready for use
   prev_state  = unavailable for use
   simages     = yes
```

   The Rstate is **ready for use** and the simages is **yes**.

If the simages attribute on the lppsource resource is **no** then the required images for the support images needed to create the SPOT were not available in the lppsource resource. The required images needed for the SPOT creation are listed in the REQUIRED_SIMAGES variable in the **/usr/lpp/bos.sysmgt/nim/methods/c_sh_lib** file. Here is a listing of the REQUIRED_SIMAGES variable:

```
REQUIRED_SIMAGES="\
    bos \
    bos.net \
    bos.rte.up \
    bos.rte.mp \
    bos.diag \
    bos.sysmgt \
    bos.terminfo \
    bos.terminfo.all.data \
    devices.base.all \
    devices.buc.all \
    devices.graphics.all \
    devices.mca.all \
    devices.scsi.all \
    devices.sio.all \
    devices.sys.all \
    devices.tty.all"
```

If you determine that you are missing installp images from the lppsource directory, follow the steps in “Missing installp images” on page 180 to correct the problem.

Using a NIM Debug SPOT to Diagnose Install Problems

To diagnose a hang during installation (LED 611), it may be necessary to create and use a debug version of the Shared Product Object Tree (SPOT). This section describes how to do this.

Note: To run diagnostics on a node, the node supervisor card must be at microcode version **1294** or later. To determine the microcode level of the card, issue this command on the control workstation, substituting # for the frame number and node number respectively.

```
/usr/lpp/ssp/bin/spmon –G –q – l frame#/node#/codeVersion/value
```

If your card is not at **1294** or later, then debug installation may loop issuing this message: **032-001 You entered a command** *command_name* **that is not valid.**

1. On the control workstation, issue the **spbootins** command to set the boot response to **disk**. For example, for frame 1 node 15, issue:

   ```
   spbootins –r disk 1 15 1
   ```

   This will issue the necessary NIM commands to prepare for reallocation of the debug SPOT for frame 1 node 15.

2. From the boot/install server issue:

   ```
   nim –Fo check –a debug=yes spot_name
   ```

   where *spot_name* is the name of the SPOT you are using.

   To find the SPOT name, issue:

   ```
   lsnim –l | grep spot_
   ```

   The SPOT name should match the AIX level on that node. For example: **spot_aix42**.

   This creates the SPOT in debug mode. When the above command completes, issue:

   ```
   lsnim -l spot_name
   ```

Look for lines that start with: **enter_dbg =**. Choose a line as follows:

- If there is only one line, use that line.
- If there is more than one line, determine which one to use as follows:
  a. If the node is a 332 MHz SMP wide node or 332 MHz SMP thin node, use the line that has **chrp.mp**.
  b. For an SMP node, use the line that has **rs6k.mp**.
  c. Otherwise, use the line with **rs6k.up**.

The line chosen contains an address, such as **0x0013afa0**. Omit the **0x** part and record the remainder of the address.

3. On the control workstation issue:

   ```
   splstdata –b
   ```

   Verify that the node's *next_install_image* is set correctly. If not, use:

   ```
   spchvgobj -i
   ```

   to correct it.

   Now issue the **spbootins** command to set the node's boot response to **install**. For example, for frame 1 node 15 issue:

   ```
   spbootins –r install 1 15 1
   ```

   This command is similar to Step 1 on page 181, but it will issue the **nim** commands that will allocate the debug SPOT created in Step 2 on page 181.

4. Condition the node. For example, for frame 1 node 15 issue:

   ```
   nodecond 1 15 &
   ```

   Open a read-only tty. For frame 1 node 15 issue:

   ```
   s1term 1 15
   ```

   This may take a few minutes to complete. Do not enter anything until it finishes with:

   ```
   Trap instruction interrupt
   ```

   and a **0>** prompt is displayed. Type <**Ctrl-c>** to stop your s1term.

5. Now start a console log to capture debug output:

   ```
   script filename
   ```

   where *filename* is a name you choose for the file.

   Open a read-write tty to the node:

   ```
   spmon –o nodexx
   ```

   where *xx* is the node number. This may take a few minutes to complete.

6. You should see the **0>** prompt again. Issue:

   ```
   st hex_number 2
   ```

   (where *hex_number* is the address recorded in Step 2 on page 181, omitting the 0x). If this gives an error, enter it again. Then, enter: **g**.

   The netboot will now be displayed as **live**. Chapter 38, “Network Installation Progress” on page 261 and Chapter 37, “SP-Specific LED/LCD Values” on page 257 give the meanings of the LED codes and help determine approximately where in the boot process your node is.

As the node boots, it may hang with LED C46. This does not indicate a problem, but the debug netboot needs to be restarted by issuing <**Ctrl-q>**. If there is a hang at any other LED value, stop logging by going to Step 7 on page 183.

7. When the node hangs, exit the tty by typing <**Ctrl-x>** and stop the logging by sending a **kill** signal to the script process from Step 4 on page 182, by issuing:

   ```
   kill pid
   ```

   To get the *pid*, issue:

   ```
   ps –ef | grep script
   ```

   If there are two scripts, killing the child process will stop both of them, or the **kill** command may be used on both.

   Now view the log file to determine what went wrong with the installation. If you contact the IBM Support Center, make sure that you have the log file available.

8. Finally, you will need to re-create a regular version of the SPOT. From the control workstation issue the

   ```
   nim -Fo check spot_name
   ```

   command, where *spot_name* is the name of the SPOT used in Step 2 on page 181.

NIM Errors in a Multiple Boot/Install Server (BIS) Environment

In a multiple BIS environment, due to an AIX constraint regarding the **inutoc** command and the **.toc** file, the following actions can only be performed on one BIS at a time:

1. Installation of NIM master file sets
2. Creation of the SPOT

This error message is reported from the **setup_server** command. This message is also displayed on the console and recorded in the **/tmp/spot.out** file.

```
0503-005 installp: The format of the .toc file is invalid.
```

This details of this message are:

- **Explanation:** Only 1 instance of the **setup_server** command on one of the Boot/Install Servers can be run at a time. All others started after the initial one will fail.
- **User Response:** Rerun the **setup_server** command on one Boot/Install Server at a time.

Chapter 20. Diagnosing Node Installation Problems

Table 36. Node Installation Symptoms	
Symptom	**Recovery**
Node fails to install from a **mksysb** image from its boot/install server.	"Action 1. Verify That the Boot/Install Server is Available" "Action 2. Open a Write Console to Check for Console Messages" "Action 3. Check Image Availability" on page 186 "Action 4. Review the NIM Configuration and Perform NIM Diagnostics for this Node" on page 186

Actions

Action 1. Verify That the Boot/Install Server is Available

Perform the following steps to verify that the boot/install server is available:

1. Determine the clients' boot/install server by issuing:

   splstdata -b -l *client_node_number*

   The **srvr** field in the returned information lists the boot/install server node number.

2. Determine the boot/install server's host name by issuing:

   splstdata -b -l *server_node_number*

   Look at the hostname field in the returned information.

You can skip the following step if this is the control workstation.

3. Login to the boot/install server using **telnet** or **rsh** and the hostname.
4. Look at the **/etc/bootptab** to make sure the node you are installing is listed in this file. If the node is not listed in this file, see Chapter 19, "Diagnosing NIM Problems" on page 175.
5. If the node is listed in this file, continue to "Action 2. Open a Write Console to Check for Console Messages."

Action 2. Open a Write Console to Check for Console Messages

1. At the control workstation open a write console for the node with the install problem by issuing:

   s1term -w *frame_number node_number*

2. Wait several minutes to see if any error messages are printed to the console which might help determine the cause of the problem. Also, look for NIM messages that might suggest that the installation is proceeding. An example of a NIM progress message is:

   / *step_number* of *total_steps* complete

which tells how many installation steps have completed. This message is accompanied by an LED code of u54. Another NIM message of interest is:

```
%% NIM Customization in Process
```

which is displayed after the **mksysb** installation is complete and further customization is taking place.

Action 3. Check Image Availability

Check to see if the image is available and the permissions are appropriate. Issue:

```
/usr/lpp/ssp/bin/splstdata -b
```

The **next_install_image** field lists the name of the image to be installed. If the field for this node is set to **default**, the default image specified by the **install_image** attribute of the SP object, will be installed. The images are found in the **/spdata/sys1/install/images** directory. You can check the images and their permissions by issuing:

```
ls -l /spdata/sys1/install/images
```

This should return:

```
total 857840
-rw-r--r--   1 root     sys      130083840 Jan 14 11:15 bos.obj.ssp.4.3
```

The important things to check are that the **images** directory has execute (x) permissions by all and that the image is readable (r) by all.

The **setup_server** script tries to clean up obsolete images on install servers. If it finds an image in the **/spdata/sys1/install/images** directory that is not needed by an install client, it deletes the image. However, **setup_server** deletes images on the control workstation only if the site environment variable REMOVE_IMAGES is **true**.

Action 4. Review the NIM Configuration and Perform NIM Diagnostics for this Node

Chapter 19, “Diagnosing NIM Problems” on page 175 describes how to review the NIM configuration and perform NIM diagnostics.

Chapter 21. Diagnosing Root Volume Group Problems

Table 37. Root Volume Group Symptoms

Symptom	Recovery Action
Creating the root volume group failed during installation because an invalid disk was specified. The specified disk does not exist.	"Action 1. Check Disks" on page 188
Creating the root volume group failed during installation because the disk is in use by another volume group.	"Action 2. Check Disk Allocation" on page 188
Mirroring failed because a disk specified in the physical volume list is already in use by another volume group. **Mirroring** is defined in "Root Volume Group Terminology" on page 190.	"Action 2. Check Disk Allocation" on page 188 "Action 3. Force the Root Volume Group Extension" on page 188
Mirroring failed because the root volume group is locked.	"Action 4. Unlock the Root Volume Group" on page 188
Mirroring failed because there is insufficient space.	"Action 5. Add Space to Physical Volumes" on page 188
Mirroring failed because there is insufficient space for strictness. **Strictness** is defined in "Root Volume Group Terminology" on page 190.	"Action 6. Add Physical Volumes to the Root Volume Group" on page 188
Mirroring failed because an incorrect number of copies was specified.	"Action 7. Verify the Number of Copies of AIX on the Node for Mirroring" on page 188
Unmirroring failed because the volume group is locked. **Unmirroring** is defined in "Root Volume Group Terminology" on page 190.	"Action 4. Unlock the Root Volume Group" on page 188
Unmirroring failed because an incorrect number of copies was specified.	"Action 8. Verify the Number of Copies of AIX on the Node for Unmirroring" on page 189
Unmirroring failed because the **reducevg** command could not remove a physical volume from the root volume group.	"Action 9. Check for User Logical Volumes on the Physical Volume" on page 189
Verification of mirroring or unmirroring is required.	"Action 10. Verify Mirroring or Unmirroring" on page 189

Some root volume group terms are defined in "Root Volume Group Terminology" on page 190. To understand more about root volume groups and mirroring concepts, see the appendix on Mirroring a Root Volume Group in *PSSP: Administration Guide*.

Actions

Action 1. Check Disks

Check to see which disks were used for installation using the **splstdata** command. If the disks are not valid, change the set of disks using the **spchvgobj** command, and reinstall the node.

Action 2. Check Disk Allocation

A physical volume can belong to only one volume group at a time. If you specified a list of disks for installation and one of the disks is in use by another volume group, you must remove the disk from the other volume group using the **reducevg** command, and then reinstall the node.

Action 3. Force the Root Volume Group Extension

A physical volume can belong to only one volume group at a time. If you attempt to extend the root volume group with a disk that is in use by another volume group, the **extendvg** command fails. If the disk is part of an inactive volume group, you can force the extension of the root volume group by specifying the **-f** (force) option. Use the **spmirrorvg -f** command.

Action 4. Unlock the Root Volume Group

If a process terminated and left the volume group in a locked state, do the following:

1. Unlock the root volume group using the **chvg -r** command.
2. Rerun the desired mirroring function, using the **spmirrorvg** command if mirroring is desired or the **spunmirrorvg** command if unmirroring is desired.

Action 5. Add Space to Physical Volumes

When mirroring, there must be enough total space on the additional physical volumes to contain all of AIX's logical volumes and still maintain strictness. Strictness is defined in “Root Volume Group Terminology” on page 190. If there is not enough space in the additional physical volumes, add additional physical volumes to the root volume group using the **spchvgobj** command. Then rerun mirroring using the **spmirrorvg** command.

Action 6. Add Physical Volumes to the Root Volume Group

For each copy of AIX, you must have at least one physical volume in the root volume group. For example, if you have specified three copies of the root volume group (the original and two copies), you must have at least three physical volumes in the root volume group. If you have fewer physical volumes than copies, add additional disks to the physical volume list using the **spchvgobj** command, and rerun mirroring using the **spmirrorvg** command.

Action 7. Verify the Number of Copies of AIX on the Node for Mirroring

To mirror successfully, you must specify more copies of AIX than are currently in effect on the node. For example, if there is one copy of the root volume group in effect on the node, you must specify two or three as the desired number of copies for mirroring. If there are two copies of the root volume group in effect, you must specify three copies for mirroring. If you specify the same number or fewer copies than are currently in effect on the node, mirroring has no effect.

Determine how many copies of the root volume group are currently in effect for the node and correct this number with these commands:

1. Use the **splstdata** command to find out how many copies of the root volume group are in effect for the node.
2. Use the **spchvgobj** command to change the number of desired copies.
3. Rerun the **spmirrorvg** command.

Action 8. Verify the Number of Copies of AIX on the Node for Unmirroring

To unmirror successfully, you must specify fewer copies of AIX than are currently in effect on the node. For example, if there are three copies of the root volume group in effect on the node, you must specify one or two as the desired number of copies for unmirroring. If there are two copies of the root volume group in effect, you must specify one for unmirroring. If you specify the same number or more copies than are currently in effect on the node, unmirroring has no effect.

Determine how many copies of the root volume group are currently in effect for the node and correct this number with these commands:

1. Use the **splstdata** command to find out how many copies of the root volume group are in effect for the node.
2. Use the **spchvgobj** command to change the number of desired copies.
3. Rerun the **spunmirrorvg** command.

Action 9. Check for User Logical Volumes on the Physical Volume

The root volume group may not be reduced by a physical volume unless all the logical volumes have been removed from the physical volume. During the unmirroring operation, all the AIX logical volumes are removed from the mirror's physical volumes. However, if additional (user) logical volumes were created on the physical volume, you cannot reduce the root volume group by the physical volume until all user logical volumes are moved or deleted.

Action 10. Verify Mirroring or Unmirroring

Use this table to perform verification tasks for the root volume group.

Table 38. Verification of Mirroring or Unmirroring for Root Volume Groups	
Verification Task	**Use These Commands**
Display the number of copies of the root volume group on the node.	Use the **lslv** command to show the number of copies of a logical volume in the root volume group. In AIX, mirroring is done on a logical volume basis. For example, **lslv hd2** will show the number of copies of the hd2 logical volume, which is used for the **/usr** file system.
Display the set of physical volumes in the root volume group.	Use the **lspv** command.
Display the boot list.	Use the AIX **diag** command, select the "Task Selection" option, and display the normal mode boot list.
Display the state of the quorum. **Quorum** is defined in “Root Volume Group Terminology” on page 190.	Use the **lsvg rootvg** command to list information about the root volume group. If the Quorum attribute is equal to 1, quorum is off. If the Quorum attribute is equal to anything else, quorum is on.

Root Volume Group Terminology

Some root volume group terms are defined below:

mirroring. Mirroring is used to provide redundant copies of AIX to prevent single points of failure. AIX provides for the original copy of AIX, and one or two additional copies. AIX places the data on the physical volumes in such a way that no two copies of the same data are ever on the same physical volume. This provides the redundancy necessary so that a single physical volume failure does not cause its volume group to fail.

quorum. A vote of the number of Volume Group Descriptor Areas (VGDAs) and Volume Group Status Areas (VGSAs) that are active. For a volume group of one disk, there are two VGDA/VGSAs. For a volume group of two disks, there are two VGDA/VGSAs on the first disk and one on the second disk. For a volume group of three or more disks, there is one VGDA/VGSA per disk.

A quorum ensures data integrity in the event of a disk failure. When a majority (51 %) of the VGDA/VGSAs in the volume group cannot be accessed, the group varies itself offline to prevent data loss or incorrect I/O operations. An error log entry is produced when this situation occurs.

strictness. A rule that AIX applies to the disks in a mirrored volume group. The strictness rule requires that enough physical volumes with sufficient space are part of a volume group, so that AIX can allocate the data among the physical volumes in such a way that the loss of a single physical volume does not constitute the loss of the volume group. Refer to the definition of mirroring and unmirroring.

unmirroring. Unmirroring is used to reduce the number of copies of a root volume group. For example, if there are currently three copies of the root volume group (the original and two copies), unmirroring would be used to reduce the number of copies from three to two or one. Optionally, physical volumes used for mirroring may be removed from the root volume group after unmirroring. When physical volumes are removed from the root volume group, they are made available for use by other volume groups.

Chapter 22. Diagnosing Boot Problems

This chapter helps if you have a node that you are unable to boot. Usually this situation is caused by one of the following problems:

- A configuration change has been made to the node that affects the boot process.
- There is a hardware problem.

Table 39. Boot Symptoms

Symptom	**Recovery.**
Unable to boot node	"Action 1. Boot a Node in Maintenance Mode"

Action

Action 1. Boot a Node in Maintenance Mode

To boot a node in maintenance mode you must change the bootp response in the SDR and then issue **setup_server** on the boot/install server to configure the node for a network boot. Follow these procedures to boot the node over the network to enter a maintenance shell:

1. To change the SDR information (bootp_response) for this node issue

   spbootins -r maintenance -l *node_number*

   then issue **setup_server** on the node's boot/install server. Watch for any error messages at this point. If you receive any, be sure to fix the error before continuing.

2. Use the Hardware Perspective to netboot this node. The node should boot over the network and startup the network boot image from its boot/install server.

3. Open a read console for the node so you can monitor the boot process. Issue:

   s1term *frame_number node_number*

   When the node reaches the main installation panel, close the read console by issuing:

   <Ctrl-c>

   If node LED stays at the same number for longer than two minutes, there may be a hardware problem. In this case, contact IBM support.

4. Open a write tty for the node by issuing:

   s1term -w *frame_number node_number*

5. Answer the prompts to request a mount of the root volume group. A maintenance shell will start.

6. Inspect system files, starting with any that may have been changed recently. A good starting place is **/etc/inittab**.

Chapter 23. Diagnosing IP Routing Problems

Ensure that the routing tables on the nodes accurately reflect the network. If they do, commands specified in **script.cust** should run correctly with all permissions and routes available.

Issue the following AIX commands to check the routing:

- To display network statistics and routing, enter:

  ```
  netstat -nr
  ```

- To see the routing to a specific point, enter:

  ```
  traceroute IP_address | host_name
  ```

IP Source Routing

Note that the **hats** script, which controls the operation of the Topology Services subsystem, issues the **no -o nonlocsrcroute=1** command. This command enables IP source routing. Do **not** change this setting, because the Topology Services subsystem requires this setting to work properly. If you change the setting, the Topology Services subsystems and a number of other subsystems that depend on it will no longer operate properly.

Chapter 24. Diagnosing Devices Including the Hard Disk

Table 40. Device Symptoms	
Symptom	**Recovery**
Problems with devices, including the hard disk	“Action 1. Boot a Node in Diagnostic Mode”

Actions

Action 1. Boot a Node in Diagnostic Mode

You can boot a node from Ethernet network in diagnostic mode when you want to run diagnostics on any device, including the hard disk, attached to that node. When a node is booted in diagnostics mode, it brings up the diagnostics menu, just as if the **diag** command were issued from AIX. But, because the hard disk of the node is not used as the boot device, you can format the hard disk, certify, diagnose, and download microcode.

Caution: Formatting destroys all data on that disk.

To boot a node in diagnostics mode, use the **spbootins** command. For example, to boot node 12 in diagnostics mode, issue:

```
spbootins -r diag -l 12
```

After the **spbootins** command has been executed, the next time the node is network booted, it will boot using the diagnostic image served over the network. The tty console will open on the display and you will be able to select actions as in the AIX **diag** command. See *AIX Version 4.3 Commands Reference* for a full description of the **diag** command.

A node in Diagnostic mode will NFS mount the Shared Product Object Tree (SPOT) from the boot/install server for use by the diagnostic image on the node. If device support is not present in the SPOT of the boot/install server, the device will not be supported by diagnostics on the node.

Chapter 25. Verifying System Management Installation

PSSP includes verification tests you can run to check installation of this software. You must be the **root** user to execute the verification test script. You can invoke the system management verification test on the control workstation from SMIT by selecting SSP System Management on the RS/6000 SP Installation/Configuration Verification menu.

Alternatively, you can invoke it from the command line by entering

```
/usr/lpp/ssp/bin/SYSMAN_test
```

When you invoke the verification test from the control workstation, it uses the **dsh** command to execute on all responding nodes. **SYSMAN_test** issues a message if one or more nodes are not tested. You can test a single node by executing the script on that node directly by local login or using **dsh**.

Verification Test Output

SYSMAN_test has three output modes: normal, verbose, and quiet. In all cases the full log of test activity is stored in **/var/adm/SPlogs/SYSMAN_test.log** (or a user-specified alternative log file). Each node's log file is stored locally on the node. In normal mode the test displays all error messages and a summary message reporting success or failure, grouped by node. In verbose mode, the output includes all information recorded in the log files. In quiet mode, only the success or failure message appears. You must examine the log files to see the errors that occurred.

The system management verification SMIT log that follows contains a sample log showing one failing node.

```
HOST: tserv11.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: 0037-031 The xntpd daemon is not running, but the ntp option was configured
SYSMAN_test: Verification failed. The number of errors found was 1

SYSMAN_test: Executing test on all active nodes

r05n11.hpssl.kgn.ibm.com: r05n11.hpssl.kgn.ibm.com: Connection timed out
dsh:  5025-509 r05n11.hpssl.kgn.ibm.com rsh had exit code 1
HOST: r05n01.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n03.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n05.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: 0037-005 Required ntp entry missing from /etc/services
SYSMAN_test: 0037-002 File /etc/rc.ntp does not exist
SYSMAN_test: 0037-003 Directory /etc/auto does not exist
SYSMAN_test: 0037-003 Directory /etc/amd does not exist
SYSMAN_test: 0037-028 /u should not be linked when using automounter
SYSMAN_test: 0037-031 The automount daemon is not running, but the Automount option was configured
SYSMAN_test: 0037-005 Required supfilesrv entry missing from /etc/services
SYSMAN_test: 0037-002 File /var/sysman/supper does not exist
SYSMAN_test: 0037-002 File /var/sysman/file.collections does not exist
SYSMAN_test: 0037-002 File /var/sysman/etc/sup does not exist
SYSMAN_test: 0037-002 File /var/sysman/etc/supfilesrv does not exist
SYSMAN_test: 0037-002 File /var/sysman/etc/supscan does not exist
```

```
SYSMAN_test: 0037-003 Directory /var/sysman/etc does not exist
SYSMAN_test: 0037-003 Directory /var/sysman/logs does not exist
SYSMAN_test: 0037-003 Directory /var/sysman/sup does not exist
SYSMAN_test: 0037-003 Directory /var/adm/acct/nite does not exist
SYSMAN_test: 0037-003 Directory /var/adm/acct/sum does not exist
SYSMAN_test: 0037-003 Directory /var/adm/acct/fiscal does not exist
SYSMAN_test: 0037-002 File /var/adm/acct/nite/jobcharge does not exist
SYSMAN_test: 0037-005 Required acct/startup entry missing from /etc/rc
SYSMAN_test: 0037-004 Directory /var/adm/acct was not exported
SYSMAN_test: Verification failed. The number of errors found was 21

HOST: r05n07.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n09.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n10.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n12.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n13.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n14.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n15.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded

HOST: r05n16.hpssl.kgn.ibm.com
------------------------------
SYSMAN_test: Verification succeeded
SYSMAN_test: The number of nodes that were not tested is 1
SYSMAN_test: The total number of errors found was 22
```

What System Management Verification Checks

This test is intended to verify that the PSSP software on the control workstation and nodes and the configuration of the system management facilities completed normally. Since those activities, for the most part, do not result directly in the establishment of testable user interfaces, the verification test concentrates on verifying that various objects relating to node installation and configuration are in the expected state.

The following tables show the key objects that are tested by **SYSMAN_test**. Some objects need to be tested only on the control workstation, some on boot/install servers, some on **/usr** servers, some on **/usr** clients, and others on all SP nodes.

Objects Tested by SYSMAN_test on the Control Workstation Only

Table 41. Objects Tested by SYSMAN_test on the Control Workstation Only

Object	Verification
CMI SMIT stanzas	Defined in ODM
/etc/SP (directory)	Created
install images	Exported
/.klogin	Entry for each node
nfs (daemon)	Active

Objects Tested by SYSMAN_test on the Control Workstation and Boot/Install Servers

Table 42. Objects Tested by SYSMAN_test on the Control Workstation and Boot/Install Servers

Object	Verification
/etc/services	tftp, bootpc, bootps
/etc/inetd.conf	tftp, bootps, instsrv
netinst user	Defined, home directory and files
install image	Exported
bootp_response	Set to disk for clients
tftpd	Daemon running
install_info	File exists for each client
config_info	File exists for each client
net.image file	Link exists for each client
/.klogin	Entry exists for client
/etc/bootptab	Entry exists for client
/etc/exports	Client access to images
install image	Each client's image available

Objects Tested by SYSMAN_test on all SP Nodes (not the Control Workstation)

Table 43. Objects Tested by SYSMAN_test on the SP Nodes

Object	Verification
Node number	Exists in ODM
server_name file	Exists
/etc/hosts	Contains server_name
/.klogin	Contains CW and its boot-install server entries

Objects Relating to Optional System Management Services

These are objects relating to Optional System Management Services that are tested on all SP nodes and the control workstation.

Table 44. Optional System Management Objects Tested by SYSMAN_test

Object	Verification
/etc/ntp.conf	Exists, content matches config option
xntpd daemon	Running or not
supfilesrv daemon	Running or not
crontab	Set for filec updates or not
/etc/inittab	Set up for filec or not
/etc/services	Set up for filec or not
supman	User defined for filec or not
Predefined file collections	Exist or not
Admin SMIT stanzas	Installed in ODM or not
Svcs.daemons	Contains entries for selected options
Automounter maps	In user.admin if automounter used
Automount	Daemon running or not
sp_passwd	Linked from /bin/passwd or not
pmswitch	Links from print commands or not
Print id	Defined and in password file or not
pmbec	Permissions set appropriately if exists
acct	Directories and files set up or not
crontab	Accounting entries added or not
Jobcharge	File matches SDR attribute if reqd
/etc/exports	Mount and root access for acct master if reqd

Additional Tests

The following table shows the additional tests that are performed to verify the functionality of System Management services.

Table 45. Additional System Management Tests

Service	Verification
ntp	Execute a **xntpdc -l** command
Automount daemon	Execute a **ps -e | grep automount** command
filec	Execute a **supper where** command

Interpreting the Test Results

If you find that all information was entered correctly, the tests that **SYSMAN_test** performs should never fail. If errors are reported in the log files, first review the log files created by the various installation steps and use the Hardware Perspective as recommended to review the hardware configuration information. Use the SMIT menus that display SP configuration and customization data.

If you find no erroneously entered information or incorrect procedures, you should report the error to the IBM Support Center as an error in either the installation/configuration process or in the test itself (the error indication could be spurious). In some cases, for example a missing directory or file, you may want to try to manually correct the problem as a circumvention. Nevertheless, it should still be reported to the IBM Support Center as a possible defect.

Chapter 26. Diagnosing Group Services Problems

This chapter contains information about diagnosing problems with the Group Services (GS) subsystem. Additional information about the GS subsystem is contained in the following publications:

- *PSSP: Administration Guide*
- *RSCT: Group Services Programming Guide and Reference*

Using the lssrc Command

The **lssrc** command is a standard AIX command and works with any subsystem defined to the AIX SRC (System Resource Controller). You do **not** have to be logged in as **root** to use the **lssrc** command. The following examples illustrate the **lssrc** command.

On a node other than the control workstation, issue:

```
lssrc -l -s hags
```

The output is similar to:

```
Subsystem         Group            PID     Status
 hags             hags             11938   active
 4 locally-connected clients.  Their PIDs:
 21344 17000 18852 23486
 HA Group Services domain information:
 Domain established by node 9.
 Number of groups known locally: 3
                    Number of   Number of local
 Group name         providers   providers/subscribers
 cssMembership            5           1           0
 WomSchg_1                5           1           1
 ha_em_peers              7           1           0
```

For the same domain on the control workstation, issue:

```
lssrc -l -s hags.k21sp2
```

The output is similar to:

```
Subsystem         Group            PID     Status
 hags.k21sp2      hags             24804   active
 1 locally-connected clients.  Their PIDs:
 44146
 HA Group Services domain information:
 Domain established by node 9.
 Number of groups known locally: 1
                    Number of   Number of local
 Group name         providers   providers/subscribers
 ha_em_peers              7           1           0
```

The differences between the output from the **lssrc** command on the node and the control workstation are:

1. The subsystem name is **hags** on the node, and **hags.***syspar_name* (hags.k21sp2) on the control workstation. This is because GS is identified differently on the control workstation and on a node.
2. The list of client groups with providers and subscribers is different.

This output is interpreted as follows:

```
Subsystem         Group            PID     Status
 hags             hags             11938   active
```

This is the same output as the **lssrc** short (**lssrc -s hags**) command. It verifies that Group Services is running, and that the process ID (PID) is the GS daemon's PID.

```
4 locally-connected clients.  Their PIDs:
21344 17000 18852 23486
```

These are the PIDs of the Group Services client processes.

```
HA Group Services domain information:
Domain established by node 9
```

This indicates that the node with node_number 9 in the system partition is acting as the GS name server. To get the hostname of node number 9, search the SDR using the **SDRGetObjects Node node_number==9** command.

To determine the node_number of a node, execute the program **/usr/lpp/ssp/install/bin/node_number**. The section “GS lssrc Output” discusses some of the other messages that may appear in this output.

```
Number of groups known locally: 3
                    Number of   Number of local
Group name          providers   providers/subscribers
cssMembership             5           1           0
WomSchg_1                 5           1           1
ha_em_peers               7           1           0
```

This lists the client groups that have providers or subscribers on this node. The "Number of providers" is the total number of providers joined to the group within the domain. The "Number of local providers/subscribers" is the number of providers on this node joined to each group and the number of subscribers on this node subscribed to each group.

GS lssrc Output

In the examples above, the GS domain is fully active and functional. However, there are other possibilities. It may take a few minutes for the GS name server to become established. During this period the command:

```
lssrc -l -s hags
```

produces output similar to:

```
Subsystem         Group            PID     Status
 hags.k21sp2      hags             24653   active
 1 locally-connected clients.  Their PIDs:
 54234
 HA Group Services domain information:
 Domain not established.
 Number of groups known locally: 0
```

The message "Domain not established." specifies that no GS name server has yet been established in this system partition. Normally, this is a temporary condition, and within approximately 30 seconds the GS name server should be established.

Another situation is when the GS name server fails (either the GS daemon process acting as the GS name server fails, or the node on which it is executing fails). In this case, GS must perform recovery actions within the domain to establish a new GS name server. The command:

```
lssrc -l -s hags
```

produces output similar to:

```
Subsystem         Group            PID     Status
 hags.k21sp2      hags             24653   active
 4 locally-connected clients.  Their PIDs:
 21344 17000 18852 23486
 HA Group Services domain information:  Domain is recovering.
 Number of groups known locally: 3
                    Number of   Number of local
 Group name         providers   providers/subscribers
 cssMembership            5           1           0
 WomSchg_1                5           1           1
 ha_em_peers              7           1           0
```

As for the case where the domain is not established, this should be a temporary condition, and a new GS name server should be established within approximately 30 seconds. AIX error log records are recorded repeatedly in certain periods while the domain is not established or recovered.

GS Failure Conditions

When the "Domain not established" or "Domain is recovering" message remains in the **lssrc** command output for an extended period of time, there may be a problem with the GS subsystem in the system partition.

Possible causes are:

1. The Topology Services subsystem (**hats**) is active on one or more of the nodes in the system partition, but the GS subsystem (**hags**) is not active on all of these nodes.

   a. To determine the status of these subsystems on the control workstation, issue:

      ```
      lssrc -s hags.syspar_name
      ```

      and

```
lssrc -s hats.syspar_name
```

b. To determine the status of these subsystems on each of the nodes in the system partition, issue:

```
lssrc -s hags
```

and

```
lssrc -s hats
```

You can use the **dsh** command to issue these commands from the control workstation.

```
dsh -a "lssrc -s hats"
```

The output is similar to:

```
c47n01.ppd.pok.ibm.com: Subsystem         Group        PID      Status
c47n01.ppd.pok.ibm.com:  hats             hats         22324    active
c47n03.ppd.pok.ibm.com: Subsystem         Group        PID      Status
c47n03.ppd.pok.ibm.com:  hats             hats         20386    active
c47n05.ppd.pok.ibm.com: Subsystem         Group        PID      Status
c47n05.ppd.pok.ibm.com:  hats             hats         6870     active
```

```
dsh -a "lssrc -s hags"
```

The output is similar to:

```
c47n01.ppd.pok.ibm.com: Subsystem         Group        PID      Status
c47n01.ppd.pok.ibm.com:  hags             hags             inoperative
c47n03.ppd.pok.ibm.com: Subsystem         Group        PID      Status
c47n03.ppd.pok.ibm.com:  hags             hags         19358    active
c47n05.ppd.pok.ibm.com: Subsystem         Group        PID      Status
c47n05.ppd.pok.ibm.com:  hags             hags             inoperative
```

c. For any nodes where **hats** is active but **hags** is inactive, start **hags** using the

```
startsrc -s hags
```

command on the nodes, and the

```
startsrc -s hags.syspar_name
```

command on the control workstation. You can use the **dsh** command to issue the commands on each node.

In the above sample output, since **hats** is active but **hags** is inactive on nodes 1 and 5, **hags** can be started on these nodes by issuing this command from the control workstation:

```
dsh -w c47n01.ppd.pok.ibm.com, c47n05.ppd.pok.ibm.com "startsrc -s hags"
```

2. Group Services and Topology Services are active on all nodes in the system partition, but GS still does not establish its domain. Record all relevant information and contact the IBM Support Center.

The relevant information includes the output of the **lssrc -l -s hags** command, and the **hags** log files **/var/ha/logs/hags*** on all nodes, as well as any information that is requested by the IBM Support Center.

The Group Services Trace

> **ATTENTION - READ THIS FIRST:** Do **not** activate the Group Services trace facility until you have read this section completely, and understand this material. If you are not certain how to properly use this facility, or if you are not under the guidance of IBM Service, do **not** activate this facility.
>
> Activating this facility may result in degraded performance of your system. Activating this facility may also result in longer response times, higher processor loads, and the consumption of system disk resources. Activating this facility may also obscure or modify the symptoms of timing-related problems.

The GS trace provides a method to record more detailed information about the internal workings of the GS daemon. This information is used to debug GS problems. The trace is controlled via the AIX **traceson** and **tracesoff** commands. The output of the trace is recorded in a log file.

Setting the Trace Level

Under normal circumstances, the GS daemon records a defined set of informational messages in its log file. It is possible to modify the level of tracing to record more detailed information, which is used to debug GS problems. The AIX **traceson** and **tracesoff** commands may be used to increase and decrease the level of internal tracing in the GS daemon.

Note that the internal tracing is level-based. It is **not** specific to any one group or event. When you increase the level of tracing, the activity for all groups is recorded, in detail, in the log. You must search the log for information specific to the group that you are interested in.

The default level of tracing (informational messages only) is equivalent to the level set by the **tracesoff** command. The **traceson** command has two settings, **short** and **long**. To change the trace level, execute one of the following commands:

- **traceson -s** *subsystem_name*, which is the **short** option
- **traceson -l -s** *subsystem_name*, which is the **long** option
- **tracesoff -s** *subsystem_name*, which is the default. This provides informational messages only.

where *subsystem_name* is **hags.***syspar_name* on the control workstation and **hags** on a node.

The long option causes tracing of: all internal GS functions executed, debug output, and all message activity. The log file may fill up and wrap quickly.

Trace Log File

The GS daemon log file is named:
/var/ha/log/*subsystem_name_nodenum_instnum***.***syspar_name*, where:

- *subsystem_name* is the name of the GS subsystem. On the control workstation, this is **hags**.*syspar_name*. On nodes, it is **hags**.
- *nodenum* is the node number on which the GS daemon is running.
- *instnum* is the instance number of the GS daemon.
- *syspar_name* is the name of the system partition in which the GS daemon is running.

Data is written to the log file in a wrapping fashion. The log file has a default size of 5000 lines. If the trace logging level is increased, the log file will wrap too quickly to be useful. To increase the size of the log file, start the daemon with the environment variable PGSD_LOGSIZE set to a larger number, such as 10000. This command can be used:

```
startsrc -s hags -e 'PGSD_LOGSIZE=10000'
```

Note that each 5000 lines of log file space occupies 750KB of disk space.

Chapter 27. Diagnosing IBM Virtual Shared Disk Problems

Table 46. IBM Virtual Shared Disk Symptoms	
Symptom	**Recovery**
Buddy Buffer Mismatch	"Action 1. Recover From a Buddy Buffer Mismatch"
EMSGSIZE Error	"Action 2. Fix EMSGSIZE Error on a Running IBM Virtual Shared Disk System"

Actions

Action 1. Recover From a Buddy Buffer Mismatch

When you define IBM Virtual Shared Disks you should specify the same **max_buddy_buffer_size** for both the client nodes and the server nodes. A buddy buffer mismatch could also come about if you redefine IBM Virtual Shared Disks, or migrate nodes to another software level or partition.

If the request size for data on the client is greater than the **max_buddy_buffer_size** on the server, the I/O request can fail and return the system error code, EMSGSIZE.

If you receive this error code, compare the **max_buddy_buffer_size** on all nodes with IBM Virtual Shared Disks. If there is a mismatch do one of the following sets of steps.

- You can postpone stopping your IBM Virtual Shared Disks by following "Action 2. Fix EMSGSIZE Error on a Running IBM Virtual Shared Disk System."
- To make a permanent correction, follow the steps in "A Permanent Fix for the EMSGSIZE Error" on page 210.

Note: Fixing this error permanently is a disruptive change.

Action 2. Fix EMSGSIZE Error on a Running IBM Virtual Shared Disk System

To recover from the EMSGSIZE error without disrupting your IBM Virtual Shared Disks, do the following:

1. Issue the **vsdnode** command from the control workstation or use the **vsdnode_dialog** SMIT panels to change the **max_buddy_buffer_size**.
2. Issue the **ctlvsd** –**M** command from the node with the smallest **max_buddy_buffer_size** size. This change may degrade your performance until you reconfigure the IBM Virtual Shared Disk.

This change is in effect until you reboot the changed node or until you do the steps in "A Permanent Fix for the EMSGSIZE Error" on page 210.

A Permanent Fix for the EMSGSIZE Error

This step requires that you stop the IBM Virtual Shared Disks. Note that this is a permanent fix.

1. Suspend all IBM Virtual Shared Disks.
2. Stop all IBM Virtual Shared Disks.
3. Unconfigure all IBM Virtual Shared Disks.
4. Update the **max_buddy_buffer_size** values on the nodes with the smaller **max_buddy_buffer_size** specifications. Issue the **vsdnode** command or use the SMIT panels to make the change. (Making this change has no effect on a running system.)
5. Configure the IBM Virtual Shared Disks.
6. Move the IBM Virtual Shared Disks to the active state.

See *PSSP: Administration Guide* for more information. See *PSSP: Command and Technical Reference* for complete IBM Virtual Shared Disk command descriptions.

Using errpt for IBM Virtual Shared Disk Diagnosis

The following table lists the IBM Virtual Shared Disk Error Log templates. To see all the templates for a specific error type, starting with the most recent, use the command:

errpt -aj *template_ID_number*

Table 47. IBM Virtual Shared Disk Error Log Template IDs

name	ID_number	Reason
ERRID_VSD_CALL_ER	ea97c090	IBM Virtual Shared Disk error occurred
ERRID_VSD_BAD_REQUEST	63b3bcf9	IBM Virtual Shared Disk request has been dropped
ERRID_VSD_ERROR_ER	9f40f6c5	Error satisfying an IBM Virtual Shared Disk request
ERRID_VSD_GENERAL_UN	891f14c9	Unexpected IBM Virtual Shared Disk error occurred

Some of these errors are intermittent even though they are recorded in the Error Log as PERM. When an IBM Virtual Shared Disk attempts to send or read data but cannot because of a resource constraint, that failed attempt is recorded as a permanent error. You can safely ignore these errors if they are few, infrequent, and you have suffered no performance problems. If, however, you see that there are many errors, the errors are clustered on only one node, or the errors are associated with performance problems, then you should investigate further.

- The **Resource Name** is **vsdd**.
- The **DETECTING MODULE** tells where the error occurred.
- The **RETURN CODE** gives more information about the error.

Follow the **Recommended Actions** or use the **User Causes** for more information.

The example following is a general error. It gives no indication of the severity of the error.

```
VSD_GENERAL_UN
IDENTIFIER: 891F14C9
Date/Time:       Thu Jun 23 10:10:15
Sequence Number: 1365
Machine Id:      000204315700
Node Id:         k43n09
Class:           S
Type:            UNKN
Resource Name:   vsdd
Description
SOFTWARE ERROR
Probable Causes
SUBSYSTEM
User Causes
SERVER IS SATURATED
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Failure Causes
SEE DETAILED DATA
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Detail Data
DETECTING MODULE
LPP=PSSP,Fn=vsdd.c,SID=SID,L#=235
RETURN CODE
```

The following template is created if an error causes the VSDD to drop a request.

```
LABEL:  VSD_ERROR_ER
IDENTIFIER: 9F40F6C5
Date/Time:       Thu Jun 23 10:10:15
Sequence Number: 1364
Machine Id:      000204315700
Node Id:         k43n09
Class:           S
Type:            PERM
Resource Name:   vsdd
Description
REQUEST PROCESSING ERROR
Probable Causes
SUBSYSTEM
User Causes
VSD DRIVERS ARE NOT CONSISTENT
INVALID DISK REQUEST
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Failure Causes
UNEXPECTED RETURN CODE...SERVICE CALL
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Detail Data
DETECTING MODULE
LPP=PSSP,Fn=vsdd.c,SID=SID,L#=234
RETURN CODE
function name: reason
```

The IBM Virtual Shared Disk places the following entry in the Error Log if the IBM Virtual Shared Disk device drivers receive a bad request as a server or as a server's reply.

```
LABEL:  VSD_BAD_REQUEST_UN
IDENTIFIER: 63B3BCF9
Date/Time:       Thu Jun 23 10:10:15
Sequence Number: 1363
Machine Id:      000204315700
Node Id:         k43n09
Class:           S
Type:            UNKN
Resource Name:   vsdd
Description
DETECTED INVALID REQUEST/CONTENT
Probable Causes
SOFTWARE DEVICE DRIVER
BUDDY BUFFER SHORTAGE
User Causes
SERVER IS SATURATED
CPU / DMA PAGE CONFLICT
Recommended Actions
DETERMINE IF BUDDY BUFFERS POOL IS LOW
IF HSD IS USED, USE 4096 BYTE ALIGNED REQUESTS
Failure Causes
DEVICE DRIVER ERROR.
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Detail Data
DETECTING MODULE
LPP=PSSP,Fn=vsdd.c,SID=SID,L#=233
RETURN CODE
function name: reason
```

The following is inserted in the Error Log if the IBM Virtual Shared Disk driver fails to process a request because a service call had failed, for example **malloc()** failed because there was no memory.

```
LABEL:  VSD_CALL_ER
IDENTIFIER: EA97C090
Date/Time:       Thu Jun 23 10:10:15
Sequence Number: 1362
Machine Id:      000204315700
Node Id:         k43n09
Class:           S
Type:            PERM
Resource Name:   vsdd
Description
SERVICE CALL FAILED
Probable Causes
SUBSYSTEM
User Causes
SERVICE CALL CANNOT COMPLETE FOR
Recommended Actions
VERIFY SYSTEM CONFIGURATION IS VALID
Failure Causes
UNEXPECTED RETURN CODE...SERVICE CALL
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Detail Data
DETECTING MODULE
LPP=PSSP,Fn=vsdd.c,SID=SID,L#=232
RETURN CODE
(rc=nnn) function name: service
```

The following Error Log entry is written if IBM Virtual Shared Disk could not get an mbuf cluster from the IP send pool (communications pool shortage).

```
LABEL: VSD_CALL_ER
IDENTIFIER: EA97C090
Date/Time:       Fri Mar  1 08:11:51
Sequence Number: 255
Machine Id:      000245375700
Node Id:         k7n11
Class:           S
Type:            PERM
Resource Name:   vsdd
Description
SERVICE CALL FAILED
Probable Causes
SUBSYSTEM
User Causes
SERVICE CALL CANNOT COMPLETE FOR
Recommended Actions
VERIFY SYSTEM CONFIGURATION IS VALID
Failure Causes
UNEXPECTED RETURN CODE...SERVICE CALL
Recommended Actions
PERFORM PROBLEM DETERMINATION PROCEDURES
Detail Data
DETECTING MODULE
LPP=PSSP,Fn=todo.c,SID=1.29,L#=601
RETURN CODE
GetCommAdapterBuffer: IFIOCTL_MGET(): send pool shortage
```

Using the VSD_ROLLBACK File

Both the **createvsd** and **createhsd** commands create **vsd_rollback** files in the **/usr/lpp/csd/vsdfiles** directory on the control workstation. At the same time, the nodes specified in the **createvsd** or **createhsd** command create local **vsd_rollback** files in the **/usr/lpp/csd/vsdfiles** directory.

The **vsd_rollback** files contain sub-commands executed on the control workstation and nodes. Most lines begin with either **OK** or **Fail**. The second field separated by a colon is the node number. Node 0 represents the control workstation. The third field is the command and operands.

If the **createvsd** and **createhsd** commands successfully create new IBM Virtual Shared Disks and hsds, these **vsd_rollback** files will be marked **COMMIT** at the end. If the commands fail in the middle of creation, you should review these rollback files to find which step and which command failed. If there is no error message saved in the rollback files, you can try the failed command on the failing node to obtain more information regarding the failure.

For example, assume **importvg** failed on the secondary node 4. You could logon node 4 to read the **vsd_rollback** file and run **lspv** to see if the **/usr/lpp/csd/sysctl/doimportvg.perl** command correctly invoked the right command.

Or, assume you forgot to configure the twin-tailed disks on the secondary node and there was no PV_ID with those disks. After configuring the disks on the secondary node, you can rerun the failed **createvsd** command on the control workstation again.

The **createvsd** and **createhsd** commands use **vsd_rollback** files to undo those operations on both the control workstation and IBM Virtual Shared Disk server nodes if these **vsd_rollback** files are not committed. In other words, you do not need to manually undo what the failed command has done. However, if you have undone what the command had done previously, you should edit the **vsd_rollback** file to remove those correlative lines starting with **OK**. Otherwise, **createvsd** and **createhsd** will fail to do the rollback.

Sometimes, you cannot get rollback to work because the environment has been changed, or the **vsd_rollback** files are left from a long time ago. In this case, you should remove these rollback files before issuing the **createvsd** or **createhsd** command.

Note that **removevsd** and **removehsd** do not support rollback.

If **createvsd**, **createhsd**, **removevsd**, or **removehsd** fails, you may have to use LVM commands or IBM Virtual Shared Disk commands to manually fix the problem. In general, you should follow these steps:

1. Use **mkvg** to create a local VG on the IBM Virtual Shared Disk server node.
2. Use **mklv** to create a logical volume on the server node.
3. If this is a twin-tailed IBM Virtual Shared Disk:
    a. Issue **chvg** to specify the vg will not be automatically varyon, and run varyoffvg on the primary node.
    b. Issue **exportvg**, **chvg**, and **varyoffvg** on the secondary node.
    c. Issue **varyonvg** on the primary node.
4. On the control workstation (or the IBM Virtual Shared Disk server node if ssp.csd.cmi is installed), issue **vsdvg** to create IBM Virtual Shared Disk global volume group.
5. Issue **defvsd** to define IBM Virtual Shared Disk.
6. After defining the underlying IBM Virtual Shared Disks, issue **defhsd** to define HSDs.

To remove an IBM Virtual Shared Disk, follow these steps:

1. Issue **suspendvsd**, **stopvsd**, and **ucfgvsd** to unconfigure the IBM Virtual Shared Disk on all IBM Virtual Shared Disk nodes.
2. Issue **undefvsd** to remove its definition in the SDR and devices on the IBM Virtual Shared Disk nodes.
3. If an IBM Virtual Shared Disk is part of an HSD, issue **undefvsd** to remove its definition in the SDR and devices. Use **ucfghsd** and **undefhsd** to remove HSD definition before running **ucfgvsd** and **undefvsd**.

Using createvsd and Recoverable Virtual Shared Disks

It is better to keep IBM Virtual Shared Disks dormant when running **createvsd**. In order to update ODM and LVM database in the secondary node, the logical volume group must be imported on the secondary node. This requires you to varyoff the logical volume group on the primary node. Since the associated logical volume is opened when an IBM Virtual Shared Disk is moved to ACTIVE state, the IBM Virtual Shared Disk must be moved to suspendvsd to allow the logical volume

group which contains the associated logical volume to varyoff. Therefore, it is better to keep the Recoverable Virtual Shared Disk inactive, in order to avoid IBM Virtual Shared Disk states being automatically changed when any membership changes in the SP system. If Recoverable Virtual Shared Disk is inactive, you need to manually run varyonvg on the primary node if the volume group exists.

Another way for **createvsd** to avoid disabling Recoverable Virtual Shared Disks is to use the **-x** option of **createvsd** and **createhsd**. In this case, **createvsd** will not import twin-tailed volume group to the secondary node nor varyoffvg on the primary node. Remember, however, to manually apply these steps described previously to make sure the secondary node will be able to update LVM information to the ODM. Otherwise, it might slow down the recovery process since the Recoverable Virtual Shared Disk will run **exportvg** and **importvg** at takeover time when it discovers disk time stamps that are not the same on the primary node and the secondary node.

Both **vsdchgserver** and RVSD might move the tail from one server node to another. If I/O with AIX EIO error code, it might be due to the bad connection of the cable. **vsdchgserver** will flip the primary and secondary nodes. The system administrator should issue **vsdatalst -g** to make sure that the current primary node is the same as the desired primary node. If they do not match, **createvsd** will fail to **mklv** on the primary node.

Chapter 28. Diagnosing SP-Attached Server Problems

This chapter provides information on:

- SP-attached server characteristics, including:
  - How PSSP views SP-attached servers
  - How SP-attached servers are attached
  - How SP-attached servers are controlled
- Error symptoms and appropriate actions to correct the problem.

Examples of an SP-attached server are the IBM RS/6000 Enterprise Server Model S70 and Model S70 Advanced.

SP-Attached Server Characteristics

Unless otherwise explicitly stated, the SP-attached server is treated as a standard SP node by PSSP. The operation of the SP-attached server is different from other nodes in several areas:

- The SP-attached servers do not physically reside in an SP frame. They are connected to the control workstation through two serial lines, one for hardware controller support and one for serial terminal support.
- The SP-attached servers do not have SP frame or node supervisors. Each server is controlled by the Hardware Monitor using the serial line to the operator panel on the server.
- There is one frame object defined in the SDR for each SP-attached server. The hardware protocol for this frame must be set to **SAMI.**
- There is one Hardware Monitor child process that runs on the control workstation for each SP-attached server. This process is called the **s70d** daemon. It emulates the SP frame and node supervisors by receiving SP supervisor commands from the Hardware Monitor and converting them to commands that are sent across the serial line to the SP-attached server operator panel. Output from the operator panel is received and converted into an equivalent supervisor response format, which is sent back to the Hardware Monitor.
- The SP-attached servers are attached to the SP Switch through the RS/6000 SP System Attachment Adapter , which is cabled to a switch port in an existing SP frame.

SP-Attached Server Error Symptoms and Recovery Actions

Table 48. SP-Attached Server Symptoms	
Symptom	**Recovery Action**
The SP-Attached server is not defined in the SDR.	"Action 1. Verify the SDR Frame Object Definition" on page 218 "Action 2. Verify the SDR Node Object Definition" on page 219
The SP-Attached server is experiencing frame supervisor communication problems.	See Chapter 7, "Diagnosing Frame Supervisor Communication Problems" on page 91. Although the SP-Attached server does not have an SP frame supervisor, the **controllerResponds** indicator from the System Monitor reflects whether communication across the serial line to the server for hardware controls has been established. The same symptoms and recovery actions apply in this case.
The **s70d** daemon is not responding.	"Action 3. Verify That the s70d Daemon Is Not Responding" on page 219 "Action 4. Check the Logs for Messages" on page 220 "Action 5. Check for a Core Dump" on page 221 "Action 6. Stop and Restart the s70d Daemon" on page 221 "Action 7. Stop and Restart the Hardware Monitor" on page 221
The **s70d** daemon has died.	"Action 4. Check the Logs for Messages" on page 220 "Action 5. Check for a Core Dump" on page 221 "Action 7. Stop and Restart the Hardware Monitor" on page 221

Actions

Action 1. Verify the SDR Frame Object Definition

Verify that the frame object has been created correctly. Issue the following command: **splstdata -f.** You should see output similar to the following, with one frame entry for every SP-attached server installed on your SP system :

```
frame#              tty         s1_tty        frame_type hardware_protocol
------------------------------------------------------------------------
     1         /dev/tty0         /dev/tty1     switch                SP
     5         /dev/tty2         /dev/tty1     switch              SAMI
```

If no entry exists for your SP-attached server, create one using the **spframe** command.

Verify that the tty (the serial port for the connection to the operator panel) and s1_tty (the serial port for the connection to the serial terminal) values are correct. If either of these values is incorrect, issue the **spframe** command to update the information. Each of these tty ports must be defined on the control workstation and

must not be locked (having an entry in **/etc/locks**). Otherwise, the **s70d** daemon will not start.

Verify that the hardware_protocol value is correct. If it is not set to **SAMI**, you must first delete the definition with the **spdelfram** command and then create a new definition with the **spframe** command.

Note: The hardware_protocol field **must** be set to the value of **SAMI** for each SP-attached server entry. If it is set to **SP**, the Hardware Monitor will attempt to send SP frame supervisor commands to the SP-attached server operator panel. The operator panel does not understand this protocol and can become hung, making it impossible to control the server either from the SP system through software, or physically from the operator panel.

Action 2. Verify the SDR Node Object Definition

Verify that the node object has been created correctly. Issue the following command: **splstdata -n -l** *node_number*, where *node_number* is the node number of your SP-attached server.

You should see information about node. If no data is displayed, the node has not been defined in the SDR. If you have verified that the frame information is correct in “Action 1. Verify the SDR Frame Object Definition” on page 218, there are several things that may be wrong:

1. The **SDR_config** command was not able to define the node. Check the **SDR_config** log **/var/adm/SPlogs/sdr/SDR_config.log** for error messages and take the appropriate action.
2. The switch port number specified on the **spframe** command was not valid. Since **SDR_config** is invoked asynchronously from the **spframe** command, it is possible that an incorrect switch port on the **spframe** command for the SP-attached server could cause **SDR_config** to incorrectly configure the server.

   Check the **SDR_config** log for messages. If the switch port number is incorrect, delete the incorrect frame definition with **spdelfram** and invoke **spframe** to create a new, correct definition.
3. There are Frame Supervisor communication problems, or the **s70d** daemon is not responding. Refer to the appropriate action for those problems.

Action 3. Verify That the s70d Daemon Is Not Responding

Verify that the **s70d** daemon is responding correctly. Issue the following command: **hmmon –G –Q** *frame* **:0,1**, where *frame* is the frame number associated with the SP-attached server. You should see a response similar to the following:

```
frame 005, slot 00:
  node 01 I2C not responding   FALSE  node 01 serial link open     FALSE
  diagnosis return code       0x0000  supervisor timer ticks      0x5fd1
  supervisor type             0x0002  supervisor code version     0x0301
  hardware monitor poll rate  0x0005

frame 005, slot 00:
  frame responding to polls     TRUE

frame 005, slot 01:
  DC-DC power on                TRUE  serial link is open          FALSE
  SRC contains a message       FALSE  SPCN contains a message      FALSE
LED/LCD contains a message     FALSE  System Reference Code        BLANK
  System Power Cntl Network    BLANK  hardware status byte        0x0000
  diagnosis return code       0x0000  supervisor timer ticks      0x0171
  supervisor type             0x000a  supervisor code version     0x0301
  LCD line 1                   BLANK  LCD Line 2                   BLANK
```

Wait at least 5 seconds and issue the command again. You should see the values for the **supervisor timer ticks** for both slots 00 and 01 increase. If these values do not change, the **s70d** daemon is not responding to the Hardware Monitor. Attempt to stop and restart the **s70d** daemon. See “Action 6. Stop and Restart the s70d Daemon” on page 221.

Another indication that the **s70d** daemon is not responding is that for slot 00, the **node 01 I2C not responding** indicator consistently stays **TRUE**. During normal operation, it may occasionally switch to **TRUE**. This may simply mean that the daemon happens to be busy and cannot respond to an individual Hardware Monitor request in a timely manner. However, if the value stays **TRUE**, it may indicate that the **s70d** daemon is not responding to the Hardware Monitor.

Action 4. Check the Logs for Messages

Several components of PSSP involved with the operation of the SP-attached server write data to log files:

1. The **s70d** daemon writes all messages to the following SP log file: **/var/adm/SPlogs/spmon/s70d/s70d**.*frame*.**log**.*julian_date*, where *frame* is the frame number of the SP-attached server, and *julian_date* is the day number in the form: nnn.

   Check this log on the control workstation for error messages and take appropriate action.

2. The Hardware Monitor logs errors that it encounters with the **s70d** daemon in the following SP log file: **/var/adm/SPlogs/spmon/hmlogfile**.*julian_date*. Check this log on the control workstation for error messages and take appropriate action.

3. SP hardware errors are written to the following log file: **/var/adm/SPlogs/SPdaemon.log**. Check this log file on the control workstation for error messages and take appropriate action. SP hardware problems have a resource name of **sphwlog**.

   The same messages are found in **errpt**. To get full details of all SP hardware messages in **errpt**, enter: **errpt –aN sphwlog**. Redirect the output of this command into a file because it could be very large.

4. The **SDR_config** command may have had problems defining the frame or node objects. **SDR_config** appends messages to the following log file: **/var/adm/SPlogs/sdr/SDR_config.log**. Check this log on the control workstation for error messages and take appropriate action.

Action 5. Check for a Core Dump

Check the **/var/adm/SPlogs/spmon/s70d** directory for a core dump. If the core file exists, save it and all other relevant log files, messages and other information. Contact the IBM Support Center.

Action 6. Stop and Restart the s70d Daemon

To stop and restart the **s70d** daemon, issue the following command: **hmcmds runpost** *frame***:0**, where *frame* is the frame number of the SP-attached server. This stops the **s70d** daemon and notifies the Hardware Monitor that it has stopped. The Hardware Monitor then restarts the daemon.

Note: You may issue this command no more than 3 times for a given **s70d** daemon during a single Hardware Monitor session. After that, you must stop and restart the Hardware Monitor before you issue the **hmcmds** command for the **s70d** daemon again.

DO NOT stop the **s70d** daemon with the **kill –kill** (**kill —9**) command. The daemon will **NOT** be restarted by the Hardware Monitor. In this case, you must stop and restart the Hardware Monitor in order to restart the **s70d** daemon.

After you restart the **s70d** daemon, verify that it is responding. See “Action 3. Verify That the s70d Daemon Is Not Responding” on page 219.

Action 7. Stop and Restart the Hardware Monitor

If you have attempted to stop and restart the **s70d** daemon, and it still does not responding , issue the following command to stop and restart the Hardware Monitor: **hmreinit**. This command stops the **s70d** daemon if it is running, and stops the Hardware Monitor daemon. When the Hardware Monitor daemon is restarted by the System Resource Controller, it also causes the **SDR_config** command to run, updating the SDR as necessary.

Chapter 29. Diagnosing 604 High Node Problems

This section provides information on:

- 604 high node characteristics, including:
  - Addressing power and fan failures in the 604 high node
  - Rebooting the 604 high node after a system failure
- Error conditions and performance considerations
- Using SystemGuard and BUMP programs

604 High Node Characteristics

The 604 high node operation is different from other nodes in several areas:

- A power feature is available which adds a redundant internal power supply to the node. In this configuration, the node will continue to run in the event of a power supply failure. Error notification for a power supply failure is done through the AIX Error Log on the node.
- The cooling system on the node also has redundancy. In the event that one of the cooling fans fails, the node will continue to run. Error notification for a power supply failure is done through the AIX Error Log on the node.
- If a hardware related crash occurs on the node, SystemGuard will re-IPL the node using the **long** IPL option. During **long** IPL, some CPU or memory resources may be deconfigured by SystemGuard to allow the re-IPL to continue.

Error Conditions and Performance Considerations

You need to be aware of the following conditions that pertain to the unique operation of this node:

- An error notification object should be set up on the node for the label **EPOW_SUS**. The **EPOW_SUS** label is used on AIX Error Log entries that may pertain to the loss of redundant power supplies or fans.
- If the node is experiencing performance degradation, you should use the **lscfg** command to verify that none of the CPU resources have been deconfigured by SystemGuard if it may have re-IPLed the node using the **long** IPL option.

Using SystemGuard and BUMP Programs

SystemGuard is a collection of firmware programs that run on the bringup microprocessor (BUMP). SystemGuard and BUMP provide service processor capability. They enable the operator to manage power supplies, check system hardware status, update various configuration parameters, investigate problems, and perform tests.

The BUMP controls the system when the power is off or the AIX operating system is stopped. The BUMP releases control of the system to AIX after it is loaded. If AIX stops or is shut down, the BUMP again controls the system.

To activate SystemGuard, the key mode switch must be in the service position during the standby or init phases. The standby phase is any time the system power is off. The init phase is the time when the system is being initialized. The PSSP software utilizes SystemGuard IPL flags such as the **FAST IPL** default. The following example shows how to modify the **FAST IPL** flag.

1. Open two windows: one for invoking **hmcmds** and one for the **s1term** tty console for the 604 high node.
2. Invoke **hmcmds** to power off the 604 high node and set the key switch to **Service**:

   ```
   hmcmds -G off frame:node
   hmcmds -G service frame:node
   ```

3. Open the tty console in write mode for the target 604 high node:

   ```
   s1term -G -w frame:node
   ```

   OR

   ```
   spmon -o node_number
   ```

4. Working with the tty console, press <**Enter>** and make sure the BUMP prompt (>) appears in the window. Invoke **sbb** in the tty console to bring up the Stand-By Menu.
5. Select **1** to check the current flag settings on the tty console.
6. If the flag setting for **FAST IPL** is **disabled**, enter **5** to change the value from **disabled** to **enabled**
7. Press **x** a few times to exit out of Stand-By menu on tty console and close the console.
8. Invoke **hmcmds** to power on the 604 high node and set the key switch to **Normal**:

   ```
   hmcmds -G on frame:node
   hmcmds -G normal frame:node
   ```

9. The 604 high node will now initialize through the **BUMP** interface and will IPL the 604 high node using the **FAST IPL**, bypassing the low-level BUMP diagnostics.

Chapter 30. Diagnosing 332 MHz SMP Thin and Wide Node Problems

This section provides information on:

- 332 MHz Symmetric MultiProcessing (SMP) node characteristics
- The boot sequence for the 332 MHz SMP Node
- Error conditions and performance considerations
- Service processor surveillance

332 MHz SMP Node Characteristics

The 332 MHz node is the first SP system SMP node offered in either a thin or wide footprint. Previous SMP nodes were available only as high nodes.

In order to provide an SMP in these more dense packages, the power and cooling design is more like that of earlier thin and wide uniprocessor nodes. N+1 power and cooling is not offered for these nodes.

The 332 MHz SMP node is different from other nodes due to these characteristics:

- RPA (RS/6000 Platform Architecture), formerly known as CHRP. Node firmware performs low-level tasks, allowing operating systems like AIX to be less hardware-dependent.

  To see the installed hardware of an RPA node, as discovered and recognized by the firmware, issue this AIX command:

  ```
  /usr/lib/boot/bin/dmpdt_chrp
  ```

- The RPA architecture dictates that there is no key switch on a 332 MHz SMP node. In order for a node to be ready to produce a dump when the node is reset, the AIX command:

  ```
  sysdumpdev -K
  ```

  must be issued prior to a dump-on-demand condition arising.

  The **-K** (upper case K) flag is most likely the default for the node. The AIX command:

  ```
  sysdumpdev -k
  ```

  (lower case k) resets the setting.

The Boot Sequence for the 332 MHz SMP Node

The progress of a node's boot may be observed by watching its LCD change either using the SP Hardware Perspective or via the command line using the command:

```
hmmon -Gq frame:slot
```

where *frame:slot* is the frame and slot number for the node.

If a power on or reset has been performed, the first phase of the node boot is the initialization of the node's Service Processor. This phase has completed once an LCD of the form **E3xy** is displayed, where **x** and **y** may be any values.

If the node boot stops during this phase, display the Service Processor logs by performing these steps:

1. Power off the node from the Control Workstation.
2. Open a TTY session to the node.
3. Hit the return key to go to the Service Processor **Main Menu**.
4. Select the **System Information Menu**.
5. Select **Service Processor Error Logs**.

The second phase of node boot for a power on or reset is the initialization of node firmware. This begins with a memory test, denoted by an LCD of **E3xy**. This phase has completed once the LCD value **E175** is displayed. If progress stops during this phase, check the meaning of the final LCD code in "Appendix B, 332 MHz and POWER3 SMP Thin and Wide Nodes Messages and Codes" of *RS/6000 SP: Maintenance Information, Volume 3: Locations and Service Procedures*.

The final phase of node boot is the booting of the AIX operating system. This begins with a BootP request, which is denoted by the **E175** LCD code. Control is later handed over to AIX under the **E05** LCD code, after which AIX codes are seen, which all begin with **0xxx**. If the AIX boot has difficulties, record the LCD codes and consult "Appendix B, 332 MHz and POWER3 SMP Thin and Wide Nodes Messages and Codes" of *RS/6000 SP: Maintenance Information, Volume 3: Locations and Service Procedures* for their meaning. When the AIX boot completes, review the AIX error log on the node.

Error Conditions and Performance Considerations

When a checkstop or machine check occurs, it is indicated by an LCD of **4B246101**, **4B246102**, **4B246110**, **4B24A101**, **4B24A102**, **4B24A110**. Firmware routines are automatically executed, which result in one of two new AIX error log entries:

- MACHINE_CHECK_CHRP - an immediate report of a hardware failure
- SCAN_ERROR_CHRP - a failure reported via periodic scan of NVRAM (non-volatile RAM)

Error data is encoded within these logs in accordance with the RPA architecture. Run the Error Log Analysis (ELA) to get AIX to process the error log data.

During boot, a processor may fail diagnostic tests and be deleted from the system. This could cause a decrease in the node's performance. When this occurs, an AIX error log entry is written.

Service Processor Surveillance

If surveillance is enabled, the Service Processor provides a 'processor-well' checking function for the node. If a processor fails to perform for a length of time, it is assumed by the Service Processor to be hung, and the Service Processor posts an LCD code of **40B0010p**, where *p* is the processor number of the hung processor. If this condition occurs, call IBM hardware service, or consult the hardware manual for a procedure to collect pertinent data from the Service Processor.

Chapter 31. Diagnosing POWER3 SMP High Node Problems

This chapter provides information on:

- POWER3 SMP (Symmetric MultiProcessor) High Node characteristics
- The boot sequence for this node
- Error conditions and performance considerations
- Service processor surveillance
- The SP Expansion I/O Unit

POWER3 SMP High Node Characteristics

The POWER3 SMP High Node is different from other nodes in several areas:

- The architecture of these nodes conforms to the RS/6000 Platform Architecture (RPA). Node firmware performs low-level tasks, allowing operating systems like AIX to be less hardware-dependent.

  To see the installed hardware of an RPA node, as discovered and recognized by the firmware, issue this AIX command:

  ```
  /usr/lib/boot/bin/dmpdt_chrp
  ```

- The RPA dictates that there is no key switch on these nodes. In order for a node to be ready to produce a dump when the node is reset, the AIX command:

  ```
  /usr/bin/sysdumpdev -K
  ```

  must be issued prior to a dump-on-demand condition arising.

  The -K (upper case K) flag is the default for the node. The AIX command:

  ```
  /usr/bin/sysdumpdev -k
  ```

  (lower case k) resets the setting.

The Boot Sequence for the POWER3 SMP High Node

The progress of a node's boot may be observed by watching its LCD change either using the SP Hardware Perspective or via the command line using the command:

```
/usr/lpp/ssp/bin/nmmon -Gq frame:slot
```

If a power on or reset has been performed, the first phase of the node boot is the initialization of the node's Service Processor. This phase has completed once an LCD of the form **E1xy** is displayed, where **x** and **y** may be any values.

If the node boot stops during this phase, display the Service Processor logs by performing these steps:

1. Power off the node from the Control Workstation.
2. Open a TTY session to the node.
3. Hit the return key to go to the Service Processor **Main Menu**.
4. Select the **System Information Menu**.

5. Select **Service Processor Error Logs**.

The second phase of node boot for a power on or reset is the initialization of node firmware. This begins with a memory test, denoted by an LCD of **E1xy**, where **x** and **y** may be any values. This phase has completed once the LCD value **E175** is displayed. If progress stops during this phase, check the meaning of the final LCD code in the appropriate *RS/6000 SP: Maintenance Information Manual*.

The final phase of node boot is the booting of the AIX operating system. This begins with a bootp request, which is denoted by the **E175** LCD code. If the AIX boot has difficulties, record the LCD codes and consult the appropriate *RS/6000 SP: Maintenance Information Manual* for their meaning. When the AIX boot completes, review the AIX error log on the node.

Error Conditions and Performance Considerations

When a checkstop or machine check occurs, it is indicated by an LCD of **45800001**. Firmware routines are automatically executed, which result in one of two AIX error log entries:

- MACHINE_CHECK_CHRP
- SCAN_ERROR_CHRP

Error data is encoded with these logs in accordance with RPA. Run the Error Log Analysis (ELA) to get AIX to process the error log data.

During boot, a processor may fail diagnostic tests and be deleted from the system. This could cause a decrease in the node's performance. When this occurs, an AIX error log entry is written.

Service Processor Surveillance

If surveillance is enabled, the Service Processor provides a 'processor-well' checking function for the node. If a processor fails to perform for a length of time, it is assumed by the Service Processor to be hung, and the Service Processor posts an LCD code of **40B00000**. If this condition occurs, call IBM hardware service, or consult the hardware manual for a procedure to collect pertinent data from the Service Processor.

SP Expansion I/O Unit

The POWER3 SMP High Node can optionally have one or more SP Expansion I/O Units attached to the node. An SP Expansion I/O Unit is normally cabled to the node in a loop (with up to two expansion units per loop) which returns to the node. This creates a redundant data path from the expansion unit to the node.

An LCD value of **203w0xyz** indicates a cabling configuration between the POWER3 SMP High node and the SP Expansion I/O Unit which does not result in a complete loop. In this LCD, the **w** is the expansion unit loop number, the **x** is always 0, the **y** indicates the expansion I/O port number on the node's system rack, and the **z** can have the following values:

z=0 Incorrectly cabled SP Expansion I/O Unit configuration

z=B	A missing return link from the SP Expansion I/O Unit to the node's system rack
z=C	A missing cable between two SP Expansion I/O Units
z=D	A BIST (built-in self test) failure in the SP Expansion I/O Unit
z=E	AN SP Expansion I/O Unit was found connected to port 1, 3, or 5 with no return to the node's system rack, and no SP Expansion I/O unit was found connected to port 2, 4, or 6 (respectively). In this case, the expansion unit connected to port 1, 3, or 5 is removed from the configuration since the cause of the error and the proper location of the expansion unit cannot be determined.

Chapter 32. Diagnosing POWER3 SMP Thin and Wide Node Problems

This section provides information on:

- POWER3 Symmetric MultiProcessing (SMP) node characteristics
- The boot sequence for the POWER3 SMP Thin and Wide Node
- Error conditions and performance considerations
- Service processor surveillance

POWER3 SMP Node Characteristics

The POWER3 SMP node is offered in either a thin or wide footprint. The power and packaging is similar to the 332 MHz SMP Thin and Wide nodes. The POWER3 nodes also have the same hardware monitor supervisor card types as the 332 MHz SMP Thin and Wide nodes.

In order to provide an SMP in these more dense packages, the power and cooling design is more like that of earlier thin and wide uniprocessor nodes. N+1 power and cooling is not offered for these nodes.

The POWER3 SMP node has the following characteristics like the 332 MHz SMP node:

- RPA (RS/6000 Platform Architecture) architecture, formerly known as CHRP. Node firmware performs low-level tasks, allowing operating systems like AIX to be less hardware-dependent.

  To see the installed hardware of an RPA node, as discovered and recognized by the firmware, issue this AIX command:

  ```
  /usr/lib/boot/bin/dmpdt_chrp
  ```

- The RPA architecture dictates that there is no key switch on a POWER3 SMP node. In order for a node to be ready to produce a dump when the node is reset, the AIX command:

  ```
  sysdumpdev -K
  ```

  must be issued prior to a dump-on-demand condition arising.

  The **-K** (upper case K) flag is most likely the default for the node. The AIX command:

  ```
  sysdumpdev -k
  ```

  (lower case k) resets the setting.

The Boot Sequence for the POWER3 SMP Thin and Wide Node

The progress of a node's boot may be observed by watching its LCD change either using the SP Hardware Perspective or via the command line using the command:

```
hmmon -Gq frame:slot
```

where *frame:slot* is the frame and slot number for the node.

If a power on or reset has been performed, the first phase of the node boot is the initialization of the node's Service Processor. This phase has completed once an LCD of the form **E3xy** is displayed, where **x** and **y** may be any values.

If the node boot stops during this phase, display the Service Processor logs by performing these steps:

1. Power off the node from the Control Workstation.
2. Open a TTY session to the node.
3. Hit the return key to go to the Service Processor **Main Menu**.
4. Select the **System Information Menu**.
5. Select **Service Processor Error Logs**.

The second phase of node boot for a power on or reset is the initialization of node firmware. This begins with a memory test, denoted by an LCD of **E3xy**. This phase has completed once the LCD value **E175** is displayed. If progress stops during this phase, check the meaning of the final LCD code in "Appendix B, 332 MHz and POWER3 SMP Thin and Wide Nodes Messages and Codes" of *RS/6000 SP: Maintenance Information, Volume 3: Locations and Service Procedures*.

The final phase of node boot is the booting of the AIX operating system. This begins with a BootP request, which is denoted by the **E175** LCD code. Control is later handed over to AIX under the **E105** LCD code, after which AIX codes are seen, which all begin with **0xxx**. If the AIX boot has difficulties, record the LCD codes and consult "Appendix B, 332 MHz and POWER3 SMP Thin and Wide Nodes Messages and Codes" of *RS/6000 SP: Maintenance Information, Volume 3: Locations and Service Procedures* for their meaning. When the AIX boot completes, review the AIX error log on the node.

Error Conditions and Performance Considerations

When a checkstop or machine check occurs, it is indicated by an LCD of **4B265401**, **4B265402** or **4B265410**. Firmware routines are automatically executed, which result in one of two new AIX error log entries:

- MACHINE_CHECK_CHRP - an immediate report of a hardware failure
- SCAN_ERROR_CHRP - a failure reported via periodic scan of NVRAM (non-volatile RAM)

Error data is encoded within these logs in accordance with the RPA architecture. Run the Error Log Analysis (ELA) to get AIX to process the error log data.

During boot, a processor may fail diagnostic tests and be deleted from the system. This could cause a decrease in the node's performance. When this occurs, an AIX error log entry is written.

Service Processor Surveillance

If surveillance is enabled, the Service Processor provides a 'processor-well' checking function for the node. If a processor fails to perform for a length of time, it is assumed by the Service Processor to be hung, and the Service Processor posts an LCD code of **40B00000**. If this occurs, reboot the node and check **errpt** for hardware errors. If there are hardware errors, call IBM hardware service. Otherwise, collect information using a SNAP tool and call the IBM Support Center to open a software PMR and have the collected data analyzed.

Chapter 33. Diagnosing Dependent Node Configuration Problems

The implementation of *dependent* nodes deals directly with RS/6000 SP Switch Routers. Therefore, configuration problems with a dependent node may be caused by either problems on the router node or on the SP. This chapter helps you diagnose these problems and determine how to correct configuration problems on the SP and on the SP Switch Router.

SP Configuration Diagnosis

The first step in diagnosing a dependent node problem is verifying that both the node and associated adapter have been properly configured and are operating correctly. To do this, perform steps 1, 2, and 3 below. You can optionally perform step 4.

1. Issue the following command to verify the definition of the extension node:

   ```
   splstnodes -t dependent node_number reliable_host_name
   management_agent_hostname extension_node_identifier snmp_community_name
   ```

2. Issue the following command to verify the definition of the extension node adapter:

   ```
   splstadapters -t node_number netaddr netmask
   ```

3. Issue the following command to verify that the extension node is connected to the switch:

   ```
   SDRGetObjects switch_responds
   ```

   You should receive output similar to the following:

   ```
   node_number switch_responds autojoin  isolated     adapter_config_status
   1           0               0         0            css_ready
   3           0               0         0            css_ready
   4           0               0         0            css_ready
   5           0               0         0            css_ready
   6           0               0         0            css_ready
   7           0               0         0            css_ready
   ```

4. Use the Hardware Perspectives to verify the definitions of the extension node and extension node adapter, and to verify that the extension node is connected to the switch.

   a. Issue **perspectives &** to bring up Perspectives as a background process.

   b.

      Double Click on: manage/control hardware

      Double Click on: IP Node

      Double Click on: Monitoring

      See the online help for an explanation of the notebook associated with the IP Node. Review the information associated with the node to verify that it is operating correctly.

Verify that these configuration parameters are correct. If you find errors here with any values other than the node number, use the **endefnode** and **endefadapter**

commands to change those values. If the node_number is incorrect, you have to remove the node_number and replace it with another node_number. To remove it, use the **enrmnode** command or the **enrmadapter** command.

The next step is to determine what state your dependent node is in on the SP Switch. Use the **SDRGetObjects switch_responds** command to determine this. Here is an example of the output from this command showing several dependent nodes in various states:

```
node_number  switch_responds autojoin  isolated  adapter_config_status
          1                0        0         0  not_configured
          2                0        0         0  css_ready
          3                0        0         0  micro_code_load_failed
          4                0        0         1  css_ready
          5                0        1         1  css_ready
          6                1        0         0  css_ready
```

An examination of each of these six samples and the state they are in will help to diagnose problems.

Dependent Node 1, in the previous example, is in the state of a newly-defined dependent node. This state means that the configuration is not complete or there is no communication via SNMP to the node. See “SNMP Configuration Diagnosis” on page 238 for more information on diagnosing and correcting SNMP communication problems.

Dependent Node 2 is in the state of a dependent node after its dependent adapter has been defined and the node has been reconfigured.

Dependent Node 2 now shows the **adapter_config_status** is **css_ready**. This means that the configuration information has been delivered to and accepted by the node via SNMP. This node is ready to become active on the SP Switch network. If this reconfiguration was not successful, the **adapter_config_status** would either remain **not_configured** or become **micro_code_load_failed**, as seen with Dependent Node 3.

Dependent Node 3's **micro_code_load_failed adapter_config_status** could also indicate that the node is still resetting the SP Switch Router Adapter in its chassis and has not finished. This reconfiguration can take some time and should be checked again. If you suspect that your configuration information is not being transmitted via SNMP to the node properly, check that the SP Switch Router Adapter is in the same slot as defined in the adapter definition.

Dependent Nodes 4 and 5 are fenced and ready to be unfenced and brought onto the SP Switch network. The only difference is that Dependent Node 5 has **autojoin** turned on, and whenever a reconfigure of the SP Switch Router Adapter occurs in this node , it will automatically unfence. Dependent Node 4 will have to be unfenced manually by issuing the **Eunfence 4** command.

Dependent Node 6 indicates the active state of a properly configured dependent node. If you are unable to attain this state with a dependent node, and you cannot discover any SP–related configuration or SNMP network -related problems, then you will have to open an administrative telnet session to the node to diagnose problems there.

SP Switch Router Configuration Diagnosis

More information pertaining to the diagnosis of problems with the SP Switch Router and the SP Switch Router Adapter can be found in the documentation that ships along with the hardware. This documentation is listed in the Bibliography.

Some basic information about diagnosing problems with the SP Switch Router Adapter is included in the following section for ease of use, but you should also reference the documentation for the SP Switch Router.

Once you have exhausted configuration problems on the SP system, **telnet** to the SP Switch Router. To do this, you need the system userid (with **root** privileges) and password. You can open a **telnet** session from any network terminal screen or from Perspectives **only if remote logins have been enabled for the router node.** For more information about enabling remote logins, see *GRF Configuration Guide - Enable telnet access.* If you cannot remotely login, you can obtain local access to the router node from its RS-232 terminal if the user has left it enabled (this terminal is only necessary for initial configuration).

After logging in with **root** privileges to the router node, there are several commands that are helpful in diagnosing problems. The first of these is the **grcard** command. Here is sample output from this command as run on a router node with 4 SP Switch Router Adapter cards installed:

```
# grcard -v
Slot    HWtype  State
----    ------  -----
0       DEV1_V1 held-reset
1       DEVI_V1 loading
2       DEV1_V1 dumping
3       DEV1_V1 running
```

The SP Switch Router Adapter is referred to as a DEV1_V1 media card to the router node, as you can see it listed under the HWtype column. Some of the common states of this media card include the four from the example, described as:

loading

The media card is loading its configuration information. The SP Switch Router Adapter will remain in this state until configuration information from the SDR in the SP System is communicated to the router node via SNMP. If your card never leaves this state, you should read the section later in this chapter about “SNMP Configuration Diagnosis” on page 238.

dumping

A media card that has been reset, or is failing, will dump its memory to a file before resetting if the MIB configuration field for this feature is turned on. By default, all SP Switch Router Adapter cards have this feature enabled. For more information on this feature see the GRF publications.

held-reset

This is the state that the SP Switch Router Adapter will be in if the **grreset -h** (discussed later) command has been issued on it, or if the card has been put in a reset state from the SP System with the **enadmin** command. For more information about the **enadmin** command, refer to *PSSP: Command and Technical Reference*.

running
This is the normal active state of the SP Switch Router Adapter.

Another useful command on the router node is the **grreset** command. Use this command to reset the SP Switch Router Adapter. When this is done, configuration information for the adapter is loaded from the control board on the router node. It is like rebooting the adapter. For more complete information about the **grreset** command, see *GRF Reference Guide.*

The **grconslog** command is useful in diagnosing problems on the router node. This command opens and displays the console log for the router node. Here you can see all configuration and network activity on the node. The most useful method for running this command is to run **grconslog -pdf** to open the console log, display information with port and date stamps, and keep the log open in flow mode. Occasionally, the static file used for this log fills up, and it then wraps or opens a new file depending on how the system is configured. When this happens, it will appear that the flow mode of your console log hangs. To recover, issue <**Ctrl-c>** out of the console log and start it up again. For more information on this command, see the GRF publications.

The primary use of the **grconslog -pdf** command is to open the console log and then issue commands from the SP System to see if they are getting to the router node and being executed. You can find more information on this in the "SNMP Configuration Diagnosis" section in this chapter.

SNMP Configuration Diagnosis

The following section will aid you in diagnosing communication problems which may occur between the SNMP Agent administering the dependent node (residing on the router node) and the SP Manager residing on the control workstation. You should run with tracing enabled for the SPMGR subsystem during a dependent node configuration.

When you configure a dependent node within an extension node class, you create attribute values in the SDR DependentNode class which are used by the SP SNMP Manager to communicate with the SNMP Agent on the router node. These attributes are:

node_number
The node number for the dependent node

extension_node_identifier
The identifier assigned to the dependent node (this is the 2-digit slot number of the dependent node adapter on the router node)

management_agent_hostname
The fully qualified hostname of the node on which the SNMP Agent administering the dependent node resides. This is used to communicate with the router node. It must resolve to an IP address.

snmp_community_name
The SNMP community name placed within SNMP messages passed between the SNMP Agent and the SNMP SP Manager for authentication. This value must match the community name value configured on the SNMP Agent host for communicating with the SP Manager on the control workstation.

If the **node_number** is specified in error, the configuration data may get sent to the SNMP Agent administering the dependent node successfully. However. problems will occur when attaching the switch adapter to the switch network.

When you have completed the definition of the dependent node on the SP control workstation and have installed the SP Switch Router Adapter on the router node, check to see if the SDR *adapter_config_status* attribute value for the dependent node in the *switch_responds* class remains *configured.* If so, then trap messages from the router node are not being processed successfully by the SNMP Manager on the control workstation. This can be caused by one of several problems:

1. If the **spmgr** subsystem trace file in the directory **/var/adm/SPlogs/spmgr** contains an entry indicating *init_io failed: udp port in use*, then the UDP port specified for service name **spmgrd-trap** in the **/etc/services** file on the control workstation is already in use. This error will also appear in an AIX error log entry written by the **spmgrd** daemon.

   Solution: Change the UDP port number for the **spmgrd-trap** service to an unused port number. The router node **snmpd** daemon configuration file, **/etc/snmpd.conf** on the router node, must also be updated to specify this same port number when sending trap messages to the control workstation. Both the **snmpd** daemon on the router node and the **spmgr** subsystem on the control workstation must be restarted after this change is made.

2. If the **lssrc -ls spmgr** command response contains zeros for both the number of switchInfoNeeded traps processed successfully and the number processed unsuccessfully, then trap messages sent by the SNMP Agent on the router node are not being received by the SNMP Manager on the control workstation.

   Either the control workstation IP address or the UDP port number may have been specified in error in the **/etc/snmpd.conf** file on the router node. The UDP port number associated with the control workstation in file **/etc/snmpd.conf** on the router node must match the UDP port number specified for the **spmgrd-trap** service in the **/etc/services** file on the control workstation.

   Solution: Correct the erroneous value and restart the **spmgrd** daemon on the router node and the **spmgr** subsystem on the control workstation.

3. If the **lssrc -ls spmgr** command response contains zeros for the number of switchInfoNeeded traps processed successfully and the number processed unsuccessfully is greater than zero, then trap messages sent by the SNMP Agent on the router node are being received by the SNMP Manger on the control workstation but they are not being successfully processed. This may be the result of one of the following errors:

   a. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** contains an entry indicating: *'Dependent node <ext_id>* managed by the SNMP Agent on *<router_node_hostname>* is not configured in the SDR - switchInfoNeeded trap ignored', then either the *extension_node_identifier* or the *management_agent_hostname* attribute value for the corresponding extension node in the SDR DependentNode class is incorrect.

      Solution: Correct the attribute value.

   b. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** contains an entry indicating: *'SDR attribute <attrname>* for *dependent node <ext_id>* in *class <classname>* has a null value for SNMP Agent on *host <router_node_hostname>'*, or an entry indicating: *'SDRGetAllObjects()*

DependentAdapter failed with return code 4', then required configuration values are missing from the indicated SDR class.

Solution: Supply the missing attribute values.

c. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** contains an entry indicating: '*Dependent node <ext_id>* managed by the SNMP Agent on *host <router_node_hostname>* is configured with a bad community name - switchInfoNeeded trap ignored', then the *snmp_community_name* attribute value specified for the corresponding node in the SDR DependentNode class does not match the community name specified for the control workstation in the **/etc/snmpd.conf** file on the router node.

Note that if the *snmp_community_name* attribute value is null, the community name to be specified in the router node is documented in the Ascend documentation.

Solution: Correct the community names in the **/etc/snmpd.conf** file on the router node and/or the *snmp_community_name* attribute for the corresponding SDR DependentNode class so that they match.

Some SNMP-related configuration problems occur when data is changed in the SDR after an initial configuration. Most of these problems are detected by the configuration-related commands and messages are issued to the operator.

If you attempt to reconfigure a dependent node **after** doing one of the following:

- Issuing either the **endefnode** or **endefadapter** command with the **-r** operand.
- Selecting the reconfigure option from a SMIT extension node configuration panel.
- Issuing an **enadmin** command.

These problems could occur:

1. A time_out occurs on the **enadmin** command (invoked internally from the SMIT panels, **endefnode**, and **endefadpter** commands). This could be caused by one of the following configuration problems:

   a. If the **spmgr** subsystem trace file in directory **/var/adm/SPlogs/spmgr** or the AIX error log contains an entry indicating '2536-007 An authentication failure notification was received from an SNMP Agent running on host <router_node_hostname> which supports Dependent Nodes', then the SDR *snmp_community_name* attribute value in the DependentNode class for the extension node contains a name that does not match the community name specified for the control workstation in the **/etc/snmpd.conf** file on the router node.

   Solution: Correct the community names in the **/etc/snmpd.conf** file on the router node and/or the *snmp_community_name* attribute for the corresponding SDR DependentNode class so that they match.

   b. If no authentication error exists in either the trace file or the AIX error log, then the value specified for the SDR *management_agent_hostname* attribute in the DependentNode class for the extension node must not be the correct fully–qualified name for the router node.

   Solution: Correct the *management_agent_hostname* attribute value in the DependentNode class for the extension node.

Note: if the *extension_node_identifier* attribute value for an extension node is erroneously set to the ID of another existing extension node on the router node managed by another SP system, then the results are unpredictable since two SNMP managers are trying to configure the same SP Switch Router Adapter.

Chapter 34. Diagnosing File Collections Problems

File Collections depends on several areas of the system to work properly, including the network, the appropriate ID being created and used, and entries in specific files. Depending on the type of problem you have, you may need to perform one or more of the following actions:

1. Verify that the **supman** ID is in the **/etc/passwd** file with a password of ***** (second field). The asterisk must be located in the **/etc/passwd** file and not in the **/etc/security/password** file.

   The ID must have the **uid** of 102 (third field). If you are using File Collections to distribute **/etc/passwd** and other user management files, the **supman** ID must also belong to the security group. The ID may have been removed if **pwdck** was run and was set to remove passwords that are not valid.

   Run **services_config** to add the supman ID again.

2. Verify that the **/etc/services** file contains the following line:

   ```
   supfilesrv     8431/tcp
   ```

3. Verify that the network is up and running properly.

4. Check your problem against the table below.

Table 49. File Collection Problems

Problem	Recovery Action
File collections not updated, permission denied messages.	Verify that the **host** file for the collection that received the error message has not been corrupted. This file should contain the hostnames of all the adapters in the SP system as stored in the SDR. These entries are followed by a special comment and should not be deleted. If the file has been corrupted, use the **dsh** command to access the File Collection server host (the Control Workstation or the boot/install server) and run the **/usr/lpp/ssp/bin/filec_host** script. If you are distributing the per-collection **host** file through file collections, be sure to run the **filec_host** script on the Control Workstation as well.
host file not being distributed.	This file is not normally distributed. If you wish the **host** file to be distributed, update the **/var/sysman/sup/sup.admin/list** file on the Control Workstation to contain the following: `always ./var/sysman/sup/*/host`
Non-SP host not receiving a file collection.	The hostname of the non-SP host must be present in the **host** file of the collection to be obtained by the non-SP host. Verify that the hostname used to invoke the **supper** commands is the same as the hostname in the **host** file. If the non-SP host has more than one adapter with an associated hostname, you may need to place each hostname of the non-SP host in the **host** file.
Customer-defined collection not restricted to SP nodes.	The **host** file was not created in the collection. Verify that the **/var/sysman/collection/host.list** file has been updated with the name of the customer-defined collection.

Chapter 35. Diagnosing SP-Controlled Netfinity Server Software

The information provided in this chapter addresses problems specific to the SP-controlled Netfinity server software. Information on how to contact IBM for service is located at the end of this chapter.

Diagnosing Netfinity Server Monitoring Problems

Symptom: Netfinity server cannot be monitored or controlled by the control workstation.

Action

Step 1. Verify that the Netfinity server itself is functioning properly. Make a visual inspection of the server. If it is not functioning properly, follow the Netfinity diagnostic procedures that accompanied your server.

Step 2. Identify the frame number of the Netfinity you need to diagnose. Refer to steps 4 and 5 of the section "Configuring SP-Controlled Netfinity Servers in the SP Environment" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*.

Step 3. Verify that the **nfd** daemon is running for the logical frame, by issuing this command:

```
hmmon -GQ frame#:0
```

Where *frame#* is the logical frame associated with the Netfinity server.

- If the **hmmon** command returns a prompt with no response, there may be a problem with **hardmon**. For diagnosis information see Chapter 14, “Diagnosing System Monitor Problems” on page 147. If the steps provided there do not resolve your problem, document your procedures and call the IBM Support Center for further assistance.
- If `supervisor type` is not equal to 3, the protocol was specified incorrectly.
  - Delete the Frame object using the **spdelfram** command.
  - Re-enter the frame information by following step 4 of the section "Configuring SP-Controlled Netfinity Servers in the SP Environment" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*.
  - If the problem still exists, return to the beginning of step 3.
- If `frame responding to polls` is TRUE, go to step 4.
- If `frame responding to polls` is FALSE, the **nfd** daemon is either not running or not responding to **hardmon**.
  - An error may have occurred while **nfd** was starting.

- Check the AIX error log (**errpt** command) for **hardmon** and **nfd** entries:

  ```
  errpt -a -N hardmon -N nfd | more
  ```

- Correct any errors. If this does not completely solve your problem, return to step 3.
- If `frame responding to polls` is still FALSE, contact the IBM Support Center.

Step 4. Check the communication path with the Netfinity server by entering the following command:

```
hmmon -GQ frame#:0
```

Where *frame#* is the logical frame associated with the Netfinity server.

- If `node 01 I2C not responding` is TRUE, a communication path has not been established by the **nfd** daemon with the Netfinity server. Follow the procedures in step 5.
- If `node 01 I2C not responding` is FALSE, the **nfd** daemon is reporting that it is monitoring and controlling a Netfinity server. Follow the procedures in step 6 to determine which server is being controlled.
- If the problem cannot be resolved, call the IBM Support Center for further assistance.

Step 5. Check the AIX error log for **nfd** entries by issuing this command:

```
errpt -a -N nfd | more
```

You may receive the following messages:

- `0026-533 Nothing is connected to device "/dev/tty#": The tty cable may be unplugged, or the Netfinity service processor may be powered off.`

  Perform the diagnostic procedures listed in step 6.

- `0026-535 Device "/dev/tty#" appears to be connected to something other than a Netfinity service processor.`

  Perform the diagnostic procedures listed in step 6.

- `0026-532 Cannot obtain node power status from the Netfinity service processor: Check for configuration error or hardware error.`

  Perform the diagnostic procedures listed in step 6.

- `0026-536 Netfinity service processor has stopped responding: Check for defective serial cable or problem on Netfinity server.`

  This could indicate a problem with the userid and password. Perform the diagnostic procedures listed in step 6.

- `0026-537 Cannot logon to the Netfinity service processor: Check for invalid user name, configuration error or hardware error.`

  Perform the diagnostic procedures listed in step 7. If that does not find the problem, then perform the diagnostic procedures listed in step 6.

- `0026-543 Cannot logon to the Netfinity service processor: Invalid user name or password.`

  Perform the diagnostic procedures listed in step 7.

- `0026-534 Access to device "/dev/tty#" has been revoked: Make sure that getty is not running on this device.`

  Perform the diagnostic procedures listed in step 8.

It may take several minutes for the **nfd** daemon to diagnose certain errors. If there are no **nfd** entries in the AIX error log after **nfd** has been running for 15 minutes, then call the IBM Support Center.

If the cause of the communication problem between the control workstation and the Netfinity cannot be determined or, there are no **nfd** entries in the Log, call the IBM Support Center.

Step 6. Verify the RS–232 cable connection by performing the following:

- Retrieve the tty number by using the **splstdata –t** command.
- Retrieve the adapter and port number using the command:

  ```
  smit lstty
  ```

- Verify that the control workstation end of the RS-232 cable is connected to the correct adapter port.
- Verify that the Netfinity server end of the RS-232 cable is connected to the correct Netfinity server.
- Verify that all connectors are clean and tight.
- Verify that the RS-232 cable is part #31L7197 and that it is not damaged. This cable is a db25 to db9 null modem cable.
- Check that the RS-232 cable is plugged in to the correct db9 connector on the back of the Netfinity server.
  - Model 5500 the connector is labeled *management* and is on the system board.
  - Model 7000–M10 the connector is located on the cable plugged into the service processor card labeled *com aux.* (Be sure you do not have it plugged into the connector labeled *modem*).

Step 7. Verify that the userid and password are correctly set in the Netfinity server, and that they match those set in the password file on the control workstation.

- Verify that the file **/spdata/sys1/spmon/netfinity_passwd** contains the userid and password that the **nfd** daemon is expecting.
- Use the **cu** command (explained in step 9) to logon to the Netfinity server.
  - If the Netfinity service processor logon was not successful, follow the instructions in step 3 of the section "Configuring SP-Controlled Netfinity Servers in the SP Environment" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*.
  - If the logon was successful, call the IBM Support Center for further assistance.

Step 8. Check for AIX processes interfering with **nfd**.

- Using the **smit chtty** command, determine if the tty connected to the Netfinity server is enabled for logins. If it is enabled, getty will attempt to use this tty. To correct this problem, disable logins on the tty using the **smit chtty** command.

Step 9. Check that all hardware connections have been made correctly and the service processor is responding to the userid and password set in the installation steps. Follow the following steps to enable **cu** for this purpose.

Note: The **cu** command invokes a standard AIX program that emulates an ascii terminal. It provides a way to control and monitor a Netfinity server from the control workstation without using **hardmon** and **nfd**. It is part of the `bos.net.uucp` file set available on your operating system installation media.

- Use the tty number found in step 6.
- Issue the command,

  ```
  ps -ef | grep nfd
  ```

  This will tell you if the **nfd** is running for the Netfinity server. The logical frame number is the third parameter after the program name. Example:

  ```
  root 16928 23622  0 11:39:53  - 0:00 /usr/lpp/ssp/install/bin/nfd
  -d 0 4 1 8 /dev/tty4/dev/tty4
  root 27370 24240  2 13:46:29  pts/8  0:00 grep nfd
  root 29998 23622  0 11:39:52  - 0:01 /usr/lpp/ssp/install/bin/nfd
  -d 0 2 1 6 /dev/tty2/dev/tty2
  root 40024 23622  0 11:39:53  - 0:01 /usr/lpp/ssp/install/bin/nfd
  -d 0 3 1 7 /dev/tty3/dev/tty3
  ```

  The pid for frame two is 29998.

- If **nfd** is running, stop it with `kill -30 pid` using the pid obtained in the previous step.
- Change the permissions of the tty device using the command:

  ```
  chmod 666 /dev/tty#
  ```

  where # is the tty number obtained in step 6. This step is needed because **nfd** protects itself by changing the permissions to 600.

- Edit the file **/etc/uucp/Devices** and add a line to the end, enabling a direct connection to the tty desired. You may copy and paste a previous line in the file.
- Issue the command,

  ```
  cu -l/dev/tty#
  ```

  where # is the tty number. It will indicate that you are connected.

- If everything is working correctly, a repeating message `press escape for the menu` will appear. After pressing escape, a prompt will come up enabling you to login to the service processor. If you are able to login successfully, the userid and password are correct.
- End the **cu** session by entering a '~' (tilde) followed by a '.' (period) then hit enter.

- Restart the daemon by issuing the following command:

  ```
  hmcmds -G boot_supervisor frame#:0
  ```

Diagnosing Problems from Within Perspectives

Symptom: Netfinity Nodes and/or Frames are not displayed in Netfinity Nodes or Frames pane.

Action

Step 1. Verify that you are running in the default partition by performing the following:

- Make the system partition with the same name as the control workstation current by selecting it, then choose the **Set Current System Partition** from the **Action** menu. This is done from within the hardware perspective.

OR

- Verify that the **SP_NAME** variable is not set to a non-default system partition before starting the hardware perspective.

Step 2. Verify that the Netfinity Node and Netfinity Frame data are entered correctly in the ProcessorExtensionNode and Frame classes of the SDR.

Symptom: Unknown state is presented when monitoring Netfinity Nodes, Frames, System Partitions and System.

Action

Step 1. Verify that **hardmon** and **nfd** are working correctly. See step 3 of *Diagnosing Netfinity Server Monitoring Problems, Symptom: Netfinity server cannot be monitored or controlled by the control workstation, Action*.

Step 2. Check frame status. See *Diagnosing Netfinity Server Monitoring Problems, Symptom: Netfinity server cannot be monitored or controlled by the control workstation, Action*.

Note: For more detailed information about steps 3-6 see *RS/6000 Cluster Technology, Event Management Programming Guide and Reference*.

Step 3. Verify that the event management subsystem **haem** is running successfully in the default partition.

Step 4. Verify that the group services subsystem **hags** is running successfully in the default partition.

Step 5. Verify that the topology services subsystem **hats** is running successfully in the default partition.

Step 6. Verify that the event monitor subsystem **emon** is running successfully in the default partition.

Step 7. Verify that **hmrmd** is not locked by doing the following:

- Use lssrc -ls haem.[default_partition_name]
- Find IBM.PSSP.hmrmd under **Resource Monitor Information**.
- If **locked column** displays **yes**, you must unlock it by issuing the following command:

  ```
  /usr/sbin/rsct/bin/haemunlkrm -s haem.[default_partition_name]
  -a IBM.PSSP.hmrmd
  ```

Step 8. Ensure that /spdata/sys1/spmon/hmacls has an entry for each Netfinity frame.

Step 9. Ensure that /spdata/sys1/spmon/hmthresholds.spnf exists and has an entry for Netfinity nodes.

Symptom: Launching of Netfinity Services Manager, Web Administration for Windows NT, or Windows NT Desktop applications fail.

Action

See the following section, “Diagnosing Web-Based Interface Error Messages.”

Diagnosing Web-Based Interface Error Messages

Messages From the Windows NT Desktop Icon

- You must configure the Windows NT Desktop icon according to the instructions in Chapter 3 of the SP-Controlled Netfinity Servers online documentation before attempting to run this program.

  An attempt was made to use the Windows NT Desktop icon before it was set up. Setup directions are in the section "Configuring SP-Controlled Netfinity Servers in the SP Environment" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*..

- Unable to run the Windows NT Desktop because the program specified either does not exist on your system, or you do not have the necessary permissions to run it.

  The Windows NT Desktop icon was set up incorrectly. Verify that the remote display program name provided when the icon was set up is correct. See the section "Configuring SP-Controlled Netfinity Servers in the SP Environment" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*.

Messages from Netfinity Services Manager and Web Administration for Windows NT Icons.

- Unable to display the web interface you have requested. You must insure that Netscape Navigator 4.0.4 or later is correctly installed and configured before launching this web interface.

  The system is unable to invoke Netscape. Possible causes:

  – Netscape is not installed

- **MOZILLA_HOME** environment variable was not set up as indicated by the Netscape launcher script
- Path to Netscape launcher script is not included in the **PATH** environment variable

- `Unable to display the web interface you have requested because IP communication could not be established with the Netfinity server using the hostname specified.`

  The system is unable to communicate with the Netfinity server.

  Perform the following:

  - Verify that the Netfinity server has power and is operational.
  - Verify that the Netfinity server is reachable over a TCP/IP network connection from the control workstation by running:

    ```
    ping hostname
    ```

    Where *hostname* is the IP address of your Netfinity node.

    If you do not receive any packet statistics, the Netfinity server is not reachable.

- `Unable to display the web interface you have requested because the Netfinity server hostname specified is unknown.  You must insure that the Netfinity server's hostname is correctly defined in the SDR and that your network is correctly configured.`

  Verify the following:

  - The Netfinity server hostname provided in the SDR ProcessorExtensionNode class matches the actual hostname of the Netfinity server. To view the SDR data use the command:

    ```
    splstdata –t
    ```

  - The Netfinity server hostname is entered in your network's nameserver.
  - The control workstation has been configured to use your network's nameserver when it attempts to resolve hostnames. See your network administrator for verification. Refer to the section "Configuring SP-Controlled Netfinity Servers in the SP Environment" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*.

- `Unable to display the web interface you have requested due to a problem updating the SP Netfinity web pages.  Insure that you are logged in as root or a userid with membership in the "bin" group and that the html files in /usr/lpp/ssp/perspectives/web have read-write permissions for owner and group.`

  The web page containing the Netfinity node list could not be opened for reading or writing.

  Note: This message will only appear for the Netfinity Services Manager icon.

  - Verify that you are logged in as root or a userid that has membership in the **bin** group.
  - Verify that the permissions on the html files in **/usr/lpp/ssp/perspectives/web** have not been changed.

– Verify that the files have not been deleted or moved to another directory.

- `Unable to display the web interface you have requested due to a problem updating the SP Netfinity web pages. Insure that you are logged in as root or a userid with membership in the "bin" group and that the html files in /usr/lpp/ssp/perspectives/web/os_mgt have read-write permissions for owner and group.`

  The web page containing the Netfinity node list could not be opened for reading or writing.

  Note: This message will only appear for the Windows NT Administration icon.

  – Verify that you are logged in as root or a userid that has membership in the **bin** group.

  – Verify that the permissions on the html files in **/usr/lpp/ssp/perspectives/web/os_mgt** have not been changed.

  – Verify that the files have not been deleted or moved to another directory.

Messages From Netscape

If Netscape gives you an error similar to the following:

- `Netscape is unable to locate the server. The server does not have a DNS entry. Check the server name in the Location (URL) and try again.`

  This indicates that either the hostname or IP address of the Netfinity server is incorrect or not set up properly in the user's nameserver.

  – Verify that the Netfinity server hostname provided in the SDR ProcessorExtensionNode class matches the actual hostname of the Netfinity server. To view the SDR data use the following command:

  ```
  splstdata –t
  ```

  – Verify that the Netfinity server hostname for the server in question is entered in your network's nameserver.

- If Netscape gets a 403 error or a `Network error: connection refused` message when trying to load either the Netfinity Manager web interface or Web Administration for Windows NT for a particular Netfinity server, you must verify that the web interface in question has been correctly configured to grant access to the control workstation's hostname or IP address.

- If Netscape gets a 404 error when trying to load the Netfinity Manager web interface for a particular Netfinity server, you must verify that the Netfinity Manager web interface has been installed and enabled on the Netfinity server.

- If Netscape gets a 404 error when trying to load the Web Administration for Windows NT page for a particular Netfinity server, do the following:

  1. Verify that the Web Administration for Windows NT software has been installed and configured on the Netfinity server.
  2. If the Web Administration for Windows NT software on the Netfinity server was installed from the BackOffice server installation media instead of the Windows NT Resource Kit, verify that the Windows NT Administration icon on the SP Perspectives launchpad was properly customized. Customization directions are in the section "Configuring Optional NT Administration

Products" of the "Installing and Configuring an SP-Controlled Netfinity Server" chapter in *PSSP: Installation and Migration*.

IBM Support Contact Numbers

IBM Service

For information to obtain prior to contacting IBM, see “Making Effective Use Of the IBM Support Center” on page 9.

See “How To Contact the IBM Support Center” on page 14 for the phone number of the IBM Support Center.

PC Help Center

1–800–772–2227

Please have your machine type and serial number ready when you place a call.

Chapter 36. Diagnosing PSSP T/EC Event Adapter Problems

If the Tivoli Enterprise Console (T/EC) event adapter fails to send events to the SP system, do the following:

1. Check your event subscription and test the event generation by forcing the event.
2. Verify that the **tecad_pssp** command is being executed by issuing this command on the control workstation: **lssrc -ls pman.***your_partition_name*, where *your_partition_name* is the name of the system partition of the node for which you are subscribed.

   The output from this command shows whether the event is being properly triggered at this point. If not, check your subscription again.
3. Use the **wtdumprl** command in the T/EC side to see if you are getting any event notifications from the PSSP side.
   - If you are, the problem is not in the SP system. Check the event source, event group, and event filter definitions, as well as the event group assignments in the T/EC.
   - If you are not getting any event notifications from the PSSP side, then the problem may be on the SP system side.
4. If you suspect that the **tecad_pssp** command is being run, but nothing is being generated at the T/EC side, check to see if you have the proper configuration file **/usr/lpp/ssp/tecad/tecad_pssp.cfg** installed, and that it points to the T/EC server.
5. Use the **/usr/lpp/ssp/tecad/test_agent** shell script to force the execution of the **tecad_pssp** command. Check the results.
6. Check the network connectivity. See Chapter 13, “Diagnosing System Connectivity Problems” on page 145.

Chapter 37. SP-Specific LED/LCD Values

LED and LCD values generated by the SP system and Parallel System Support Programs can be of two types: those that convey information and signal status, or those that indicate a problem. These are the PSSP-specific LED/LCD Values, generated during the NIM installation of the node. This list is displayed in chronological order, or the order in which the they occur during processing.

Table 50 (Page 1 of 2). SP-Specific LED/LCD Values (Chronological Order)

u20 Create log directory (enter function create_directories).
u21 Establish working environment (enter function setup_environment).
u03 Get the *node*.install_info file from the master.
u04 Expand *node*.install_info file.
u22 Configure node (enter function configure_node).
u57 Get the *node*.config_info file from the master.
u59 Get the cuat.sp template from the master.
u23 Create/update /etc/ssp files (enter function create_files).
u60 Create/update /etc/ssp files.
u24 Update /etc/hosts file (enter function update_etchosts).
u25 Get configuration files (enter function get_files).
u61 Get /etc/SDR_dest_info from boot/install server.
u79 Get script.cust from boot/install server.
u50 Get tuning.cust from boot/install server.
u54 Get spfbcheck from boot/install server.
u56 Get psspfb_script from boot/install server.
u58 Get psspfb_script from control workstation.
u26 Get authentication files (enter function authent_stuff).
u67 Get /etc/krb.conf from boot/install server.
u68 Get /etc/krb.realms from boot/install server.
u69 Get krb-srvtab from boot/install server.
u27 Update /etc/inittab file (enter function update_etcinittab).
u28 Perform MP-specific functions (enter function upmp_work).
u52 Processor is "MP".
u51 Processor is "UP".
u55 Fatal error in bosboot.
u29 Install prerequisite file sets (enter function install_prereqs).
u30 Install ssp.clients (enter function install_ssp_clients).
u80 Mount lppsource and install ssp.clients.
u31 Install ssp.basic (enter function install_ssp_basic).
u81 Install ssp.basic.
u32 Install ssp.ha (enter function install_ssp_ha).
u53 Install ssp.ha.

Table 50 (Page 2 of 2). SP-Specific LED/LCD Values (Chronological Order)

u33 Install ssp.sysctl (enter function install_ssp_sysctl).
u82 Install ssp.sysctl.
u34 Install ssp.pman (enter function install_ssp_pman).
u41 Configure switch (enter function config_switch).
u35 Install ssp.css (enter function install_ssp_css).
u84 Install ssp.css.
u36 Install ssp.jm (enter function install_ssp_jm).
u85 Install ssp.jm.
u37 Delete master .rhosts entry (enter function delete_master_rhosts).
u38 Create new dump logical volume (enter function create_dump_lv).
u86 Create new dump logical volume.
u39 Run customer's tuning.cust (enter function run_tuning_cust).
u40 Run customer's script.cust (enter function run_script_cust).
u87 Run customer's script.cust script file.
u42 Run psspfb_script (enter function run_psspfb_script).

The following list contains the same PSSP LEDs/LCDs as in the previous list, but sorted numerically for reference use:

Table 51 (Page 1 of 2). SP-Specific LED/LCD Values (Numerical Order)

u03 Get the *node*.install_info file from the master.
u04 Expand *node*.install_info file.
u20 Create log directory (enter function create_directories).
u21 Establish working environment (enter function setup_environment).
u22 Configure node (enter function configure_node).
u23 Create/update /etc/ssp files (enter function create_files).
u24 Update /etc/hosts file (enter function update_etchosts).
u25 Get configuration files (enter function get_files).
u26 Get authentication files (enter function authent_stuff).
u27 Update /etc/inittab file (enter function update_etcinittab).
u28 Perform MP-specific functions (enter function upmp_work).
u29 Install prerequisite file sets (enter function install_prereqs).
u30 Install ssp.clients (enter function install_ssp_clients).
u31 Install ssp.basic (enter function install_ssp_basic).
u32 Install ssp.ha (enter function install_ssp_ha).
u33 Install ssp.sysctl (enter function install_ssp_sysctl).
u34 Install ssp.pman (enter function install_ssp_pman).
u35 Install ssp.css (enter function install_ssp_css).
u36 Install ssp.jm (enter function install_ssp_jm).
u37 Delete master .rhosts entry (enter function delete_master_rhosts).
u38 Create new dump logical volume (enter function create_dump_lv).

Table 51 (Page 2 of 2). SP-Specific LED/LCD Values (Numerical Order)

u39 Run customer's tuning.cust (enter function run_tuning_cust).
u40 Run customer's script.cust (enter function run_script_cust).
u41 Configure switch (enter function config_switch).
u42 Run psspfb_script (enter function run_psspfb_script).
u50 Get tuning.cust from boot/install server.
u51 Processor is "UP".
u52 Processor is "MP".
u53 Install ssp.ha.
u54 Get spfbcheck from boot/install server.
u55 Fatal error in bosboot.
u56 Get psspfb_script from boot/install server.
u58 Get psspfb_script from control workstation.
u57 Get the *node*.config_info file from the master.
u59 Get the cuat.sp template from the master.
u60 Create/update /etc/ssp files.
u61 Get /etc/SDR_dest_info from boot/install server.
u67 Get /etc/krb.conf from boot/install server.
u68 Get /etc/krb.realms from boot/install server.
u69 Get krb-srvtab from boot/install server.
u79 Get script.cust from boot/install server.
u80 Mount lppsource and install ssp.clients.
u81 Install ssp.basic.
u82 Install ssp.sysctl.
u84 Install ssp.css.
u85 Install ssp.jm.
u86 Create new dump logical volume.
u87 Run customer's script.cust script file.

The following LEDs/LCDs are produced after NIM installation has occurred, and during the initial post-installation reboot of the node. This list is sorted chronologically, or in the order of which the LEDs/LCDs occur during processing.

u90 Setup working environment (enter function setup_environment).
u92 Configure adapters (enter function config_adapters).
u93 Configure inet0 (enter function config_inet0).
u94 Run cfgmgr (enter function run_cfgmgr).
u95 Run complete_node on boot/install server (enter function complete_node).
u78 Set the KRBTKFILE variable and get an rcmd ticket.
u96 Run customer's firstboot.cust (enter function run_firstboot_cust).

The following list contains the same LEDs/LCDs as in the previous list, but sorted numerically for reference use:

u78 Set the KRBTKFILE variable and get an rcmd ticket.

u90 Setup working environment (enter function setup_environment).

u92 Configure adapters (enter function config_adapters).

u93 Configure inet0 (enter function config_inet0).

u94 Run cfgmgr (enter function run_cfgmgr).

u95 Run complete_node on boot/install server (enter function complete_node).

u96 Run customer's firstboot.cust (enter function run_firstboot_cust).

The following LEDs/LCDs occur during a node IPL.

762 SP Switch Adapter configuring on node

763 SP Switch MX Adapter configuring on node

764 RS/6000 SP System Attachment Adapter configuring on node

Chapter 38. Network Installation Progress

When a network installation is in progress, the LED/LCD for the nodes involved show various values. These values indicate the installation stage.

Table 52 lists the sequence of LED/LCD values a node goes through. This table is only a sample list of LED/LCD values, and a node may not show all values listed. Depending on the adapters installed on the node, there may be additional values displayed. The table also lists the approximate time, after you start to install a node, when the LED/LCD values may be shown. Since the elapsed time to reach a specific LED/LCD value can vary, you should use these times as a gauge to determine the installation progress of the node. The factors that can affect the actual time it takes a node to reach a specific LED/LCD value include the following:

- The number of nodes being installed
- The size of the network installation image you are using
- The amount of traffic on the SP Ethernet
- The amount of work being conducted on the network installation node

Places in the node installation where a single LED/LCD value is displayed for an extended period of time are as follows:

1. c40
2. c54

These values are periods of high volumes of network traffic and are only a problem on a node when the value does not change for a very extended period of time. You can use Table 52 as a debugging tool when a node is stuck on a specific LED/LCD value.

Table 52 (Page 1 of 4). Sample NIM Installation Trace

Time (min:sec)	LED/LCD Value	Description
0:00		
	124	BIST started a CRC check on the OCS area of NVRAM.
	151	BIST started AIPGM test code.
	214	Power status register failed.
	219	Generating RAM POST bit map.
	291	Running standard I/O POST.
	200	Attempting IPL with key in secure position.
	102	BIST started following power-on reset.
	153	BIST started ACLST test code.
0:30		
	154	BIST started AST test code.
	100	BIST completed successfully; control was passed to IPL ROS.
	219	Generating RAM POST bit map.
	292	Running SCSI POST.

Table 52 (Page 2 of 4). Sample NIM Installation Trace

Time (min:sec)	LED/LCD Value	Description
1:00		
	291	Running standard I/O POST.
	262	No keyboard connected to the system.
	260	Displaying information on the display console.
1:30		
	231	Attempting a normal mode IPL from Ethernet specified in IPL ROM.
2:30		
	Blank	
	606	Running the **if_config** command to bring up network interface.
	610	Attempting to NFS mount a remote file system.
	299	IPL ROM passed control to the loaded program code.
	608	Attempting to tftp the **.info** file from client's SPOT server.
	612	Accessing remote configuration files.
	622	Returning control to the **/sbin/rc.boot** program.
	520	Running bus configuration.
	890	SCSI-2 differential fast/wide adapter.
	620	Updating special device files.
3:00		
	890	SCSI-2 differential fast/wide adapter.
	570	Configuring virtual SCSI devices.
3:30		
	622	Returning control to the **/sbin/rc.boot** program
4:00		
	811	Identifying or configuring processor complex.
	Blank	
	570	Configuring virtual SCSI devices.
	727	Identifying or configuring unknown asynchronous device.
	538	The configuration manager is invoking a configuration method.
	622	Returning control to the **/sbin/rc.boot** program.
	c40	Restoring configuration files.
8:00		
	c42	Extracting data from diskette.

Table 52 (Page 3 of 4). Sample NIM Installation Trace

Time (min:sec)	LED/LCD Value	Description
	c33	Selecting a tty terminal attached to serial ports S1 or S2.
	c44	Initializing installation database with target disk info.
8:30		
	c46	Normal installation processing.
	Blank	
9:00		
	c50	Creating root volume group on target disks.
	c46	Normal installation processing.
10:30		
	c54	Installing either BOS or additional packages.
19:00		
	c52	Changing from RAM environment to disk environment.
21:30		
	c46	Normal installation processing.
	570	Configuring virtual SCSI devices.
22:00		
	c46	Normal installation processing.
	731	Identifying or configuring PTY.
	539	The configuration method has terminated, returning to config.
	811	Identifying or configuring processor complex.
	538	The configuration manager is invoking a configuration method.
	570	Configuring virtual SCSI devices.
22:30		
	Blank	
23:00		
	u78	Running rsh to complete node processing on boot/install server.
	u60	Creating **/etc/ssp/server_name** and updating **/etc/hosts**.
	u68	Copying kerberos realms file from boot/install server.
	u59	Running **config_node** to define adapters.
27:00		

Table 52 (Page 4 of 4). Sample NIM Installation Trace

Time (min:sec)	LED/LCD Value	Description
	u65	SSP completing the **install/customize/maint**. Issuing shutdown.
27:30		
	c46	Normal installation processing.
28:00		
	292	Running SCSI POST.
	298	Attempting a software IPL.
	291	Running standard I/O POST.
	299	IPL ROM passed control to the loaded program code.
	Blank	
	890	SCSI-2 differential fast/wide adapter.
	820	
	538	The configuration manager is invoking a configuration method.
	570	Configuring Virtual SCSI devices.
28:30		
29:00		
	551	Running IPL varyon.
	517	Mounting client remote file system during network IPL.
29:30		
	553	IPL phase 1 is complete.
	570	Configuring virtual SCSI devices.
	538	The configuration manager is invoking a configuration method.
	c33	Selecting a tty terminal attached to serial ports S1 or S2.
30:00		
	Blank	
33:00		
	762	Running SSP configuration method for SP Switch Adapter
34:00		
	570	Configuring virtual SCSI devices.
	538	The configuration manager is invoking a configuration method.
34:30		
	Blank	

Glossary of Terms and Abbreviations

This glossary includes terms and definitions from:

- The *IBM Dictionary of Computing*, New York: McGraw-Hill, 1994.
- The *American National Standard Dictionary for Information Systems*, ANSI X3.172-1990, copyright 1990 by the American National Standards Institute (ANSI). Copies can be purchased from the American National Standards Institute, 1430 Broadway, New York, New York 10018. Definitions are identified by the symbol (A) after the definition.
- The *ANSI/EIA Standard - 440A: Fiber Optic Terminology* copyright 1989 by the Electronics Industries Association (EIA). Copies can be purchased from the Electronic Industries Association, 2001 Pennsylvania Avenue N.W., Washington, D.C. 20006. Definitions are identified by the symbol (E) after the definition.
- The *Information Technology Vocabulary* developed by Subcommittee 1, Joint Technical Committee 1, of the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC JTC1/SC1). Definitions of published parts of this vocabulary are identified by the symbol (I) after the definition; definitions taken from draft international standards, committee drafts, and working papers being developed by ISO/IEC JTC1/SC1 are identified by the symbol (T) after the definition, indicating that final agreement has not yet been reached among the participating National Bodies of SC1.

The following cross-references are used in this glossary:

> **Contrast with.** This refers to a term that has an opposed or substantively different meaning.
> **See.** This refers the reader to multiple-word terms in which this term appears.
> **See also.** This refers the reader to terms that have a related, but not synonymous, meaning.
> **Synonym for.** This indicates that the term has the same meaning as a preferred term, which is defined in the glossary.

This section contains some of the terms that are commonly used in the SP publications.

IBM is grateful to the American National Standards Institute (ANSI) for permission to reprint its definitions from the American National Standard *Vocabulary for Information Processing* (Copyright 1970 by American National Standards Institute, Incorporated), which was prepared by Subcommittee X3K5 on Terminology and Glossary of the American National Standards Committee X3. ANSI definitions are preceded by an asterisk (*).

Other definitions in this glossary are taken from *IBM Vocabulary for Data Processing, Telecommunications, and Office Systems* (SC20-1699) and *IBM DATABASE 2 Application Programming Guide for TSO Users* (SC26-4081).

A

adapter. An adapter is a mechanism for attaching parts. For example, an adapter could be a part that electrically or physically connects a device to a computer or to another device. In the SP system, network connectivity is supplied by various adapters, some optional, that can provide connection to I/O devices, networks of workstations, and mainframe networks. Ethernet, FDDI, token-ring, HiPPI, SCSI, FCS, and ATM are examples of adapters that can be used as part of an SP system.

address. A character or group of characters that identifies a register, a device, a particular part of storage, or some other data source or destination.

AFS. A distributed file system that provides authentication services as part of its file system creation.

AIX. Abbreviation for Advanced Interactive Executive, IBM's licensed version of the UNIX operating system. AIX is particularly suited to support technical computing applications, including high function graphics and floating point computations.

Amd. Berkeley Software Distribution automount daemon.

API. Application Programming Interface. A set of programming functions and routines that provide access between the Application layer of the OSI seven-layer model and applications that want to use the network. It is a software interface.

application. The use to which a data processing system is put; for example, a payroll application, an airline reservation application.

application data. The data that is produced using an application program.

ARP. Address Resolution Protocol.

ATM. Asynchronous Transfer Mode. (See *TURBOWAYS 100 ATM Adapter*.)

Authentication. The process of validating the identity of a user or server.

Authorization. The process of obtaining permission to perform specific actions.

B

batch processing. * (1) The processing of data or the accomplishment of jobs accumulated in advance in such a manner that each accumulation thus formed is processed or accomplished in the same run. * (2) The processing of data accumulating over a period of time. * (3) Loosely, the execution of computer programs serially. (4) Computer programs executed in the background.

BMCA. Block Multiplexer Channel Adapter. The block multiplexer channel connection allows the RS/6000 to communicate directly with a host System/370 or System/390; the host operating system views the system unit as a control unit.

BOS. The AIX Base Operating System.

C

call home function. The ability of a system to call the IBM support center and open a PMR to have a repair scheduled.

CDE. Common Desktop Environment. A graphical user interface for UNIX.

charge feature. An optional feature for either software or hardware for which there is a charge.

CLI. Command Line Interface.

client. * (1) A function that requests services from a server and makes them available to the user. * (2) A term used in an environment to identify a machine that uses the resources of the network.

Client Input/Output Sockets (CLIO/S). A software package that enables high-speed data and tape access between SP systems, AIX systems, and ES/9000 mainframes.

CLIO/S. Client Input/Output Sockets.

CMI. Centralized Management Interface provides a series of SMIT menus and dialogues used for defining and querying the SP system configuration.

connectionless. A communication process that takes place without first establishing a connection.

connectionless network. A network in which the sending logical node must have the address of the receiving logical node before information interchange can begin. The packet is routed through nodes in the network based on the destination address in the packet. The originating source does not receive an acknowledgment that the packet was received at the destination.

control workstation. A single point of control allowing the administrator or operator to monitor and manage the SP system using the IBM AIX Parallel System Support Programs.

css. Communication subsystem.

D

daemon. A process, not associated with a particular user, that performs system-wide functions such as administration and control of networks, execution of time-dependent activities, line printer spooling and so forth.

DASD. Direct Access Storage Device. Storage for input/output data.

DCE. Distributed Computing Environment.

DFS. distributed file system. A subset of the IBM Distributed Computing Environment.

DNS. Domain Name Service. A hierarchical name service which maps high level machine names to IP addresses.

E

Error Notification Object. An object in the SDR that is matched with an error log entry. When an error log entry occurs that matches the Notification Object, a user-specified action is taken.

ESCON. Enterprise Systems Connection. The ESCON channel connection allows the RS/6000 to communicate directly with a host System/390; the host operating system views the system unit as a control unit.

Ethernet. (1) Ethernet is the standard hardware for TCP/IP local area networks in the UNIX marketplace. It is a 10-megabit per second baseband type LAN that allows multiple stations to access the transmission medium at will without prior coordination, avoids contention by using carrier sense and deference, and resolves contention by collision detection (CSMA/CD). (2) A passive coaxial cable whose interconnections contain devices or components, or both, that are all active. It uses CSMA/CD technology to provide a best-effort delivery system.

Ethernet network. A baseband LAN with a bus topology in which messages are broadcast on a coaxial cabling using the carrier sense multiple access/collision detection (CSMA/CD) transmission method.

event. In Event Management, the notification that an expression evaluated to true. This evaluation occurs each time an instance of a resource variable is observed.

expect. Programmed dialogue with interactive programs.

expression. In Event Management, the relational expression between a resource variable and other elements (such as constants or the previous value of an instance of the variable) that, when true, generates an event. An example of an expression is `X < 10` where X represents the resource variable `IBM.PSSP.aixos.PagSp.%totalfree` (the percentage of total free paging space). When the expression is true, that is, when the total free paging space is observed to be less than 10%, the Event Management subsystem generates an event to notify the appropriate application.

F

failover. Also called fallover, the sequence of events when a primary or server machine fails and a secondary or backup machine assumes the primary workload. This is a disruptive failure with a short recovery time.

fall back. Also called fallback, the sequence of events when a primary or server machine takes back control of its workload from a secondary or backup machine.

FDDI. Fiber Distributed Data Interface.

Fiber Distributed Data Interface (FDDI). An American National Standards Institute (ANSI) standard for 100-megabit-per-second LAN using optical fiber cables. An FDDI local area network (LAN) can be up to 100 km (62 miles) and can include up to 500 system units. There can be up to 2 km (1.24 miles) between system units and concentrators.

file. * A set of related records treated as a unit, for example, in stock control, a file could consist of a set of invoices.

file name. A CMS file identifier in the form of 'filename filetype filemode' (like: TEXT DATA A).

file server. A centrally located computer that acts as a storehouse of data and applications for numerous users of a local area network.

File Transfer Protocol (FTP). The Internet protocol (and program) used to transfer files between hosts. It is an application layer protocol in TCP/IP that uses TELNET and TCP protocols to transfer bulk-data files between machines or hosts.

foreign host. Any host on the network other than the local host.

FTP. File transfer protocol.

G

gateway. An intelligent electronic device interconnecting dissimilar networks and providing protocol conversion for network compatibility. A gateway provides transparent access to dissimilar networks for nodes on either network. It operates at the session presentation and application layers.

H

HACMP. High Availability Cluster Multi-Processing for AIX.

HACWS. High Availability Control Workstation function, based on HACMP, provides for a backup control workstation for the SP system.

HAL. Hardware Abstraction Layer, a communication device interface that provides communication channels for processes.

Hashed Shared Disk (HSD). The data striping device for the IBM Virtual Shared Disk. The device driver lets application programs stripe data across physical disks in multiple IBM Virtual Shared Disks, thus reducing I/O bottlenecks.

help key. In the SP graphical interface, the key that gives you access to the SP graphical interface help facility.

High Availability Cluster Multi-Processing. An IBM facility to cluster nodes or components to provide high availability by eliminating single points of failure.

HiPPI. High Performance Parallel Interface. RS/6000 units can attach to a HiPPI network as defined by the ANSI specifications. The HiPPI channel supports burst rates of 100 Mbps over dual simplex cables; connections can be up to 25 km in length as defined by the standard and can be extended using third-party HiPPI switches and fiber optic extenders.

home directory. The directory associated with an individual user.

host. A computer connected to a network, and providing an access method to that network. A host provides end-user services.

I

instance vector. Obsolete term for resource identifier.

Intermediate Switch Board. Switches mounted in the Sp Switch expansion frame.

Internet. A specific inter-network consisting of large national backbone networks such as APARANET, MILNET, and NSFnet, and a myriad of regional and campus networks all over the world. The network uses the TCP/IP protocol suite.

Internet Protocol (IP). (1) A protocol that routes data through a network or interconnected networks. IP acts as an interface between the higher logical layers and the physical network. This protocol, however, does not provide error recovery, flow control, or guarantee the reliability of the physical network. IP is a connectionless protocol. (2) A protocol used to route data from its source to it destination in an Internet environment.

IP address. A 32-bit address assigned to devices or hosts in an IP internet that maps to a physical address. The IP address is composed of a network and host portion.

ISB. Intermediate Switch Board.

K

Kerberos. A service for authenticating users in a network environment.

kernel. The core portion of the UNIX operating system which controls the resources of the CPU and allocates them to the users. The kernel is memory-resident, is said to run in “kernel mode” and is protected from user tampering by the hardware.

L

LAN. (1) Acronym for Local Area Network, a data network located on the user's premises in which serial transmission is used for direct data communication among data stations. (2) Physical network technology that transfers data a high speed over short distances. (3) A network in which a set of devices is connected to another for communication and that can be connected to a larger network.

local host. The computer to which a user's terminal is directly connected.

log database. A persistent storage location for the logged information.

log event. The recording of an event.

log event type. A particular kind of log event that has a hierarchy associated with it.

logging. The writing of information to persistent storage for subsequent analysis by humans or programs.

M

mask. To use a pattern of characters to control retention or elimination of portions of another pattern of characters.

menu. A display of a list of available functions for selection by the user.

Motif. The graphical user interface for OSF, incorporating the X Window System. Also called OSF/Motif.

MTBF. Mean time between failure. This is a measure of reliability.

MTTR. Mean time to repair. This is a measure of serviceability.

N

naive application. An application with no knowledge of a server that fails over to another server. Client to server retry methods are used to reconnect.

network. An interconnected group of nodes, lines, and terminals. A network provides the ability to transmit data to and receive data from other systems and users.

NFS. Network File System. NFS allows different systems (UNIX or non-UNIX), different architectures, or vendors connected to the same network, to access remote files in a LAN environment as though they were local files.

NIM. Network Installation Management is provided with AIX to install AIX on the nodes.

NIM client. An AIX system installed and managed by a NIM master. NIM supports three types of clients:

- Standalone
- Diskless
- Dataless

NIM master. An AIX system that can install one or more NIM clients. An AIX system must be defined as a NIM master before defining any NIM clients on that system. A NIM master managers the configuration database containing the information for the NIM clients.

NIM object. A representation of information about the NIM environment. NIM stores this information as objects in the NIM database. The types of objects are:

- Network
- Machine
- Resource

NIS. Network Information System.

node. In a network, the point where one or more functional units interconnect transmission lines. A computer location defined in a network. The SP system can house several different types of nodes for both serial and parallel processing. These node types can include thin nodes, wide nodes, 604 high nodes, as well as other types of nodes both internal and external to the SP frame.

Node Switch Board. Switches mounted on frames that contain nodes.

NSB. Node Switch Board.

NTP. Network Time Protocol.

O

ODM. Object Data Manager. In AIX, a hierarchical object-oriented database for configuration data.

P

parallel environment. A system environment where message passing or SP resource manager services are used by the application.

Parallel Environment. A licensed IBM program used for message passing applications on the SP or RS/6000 platforms.

parallel processing. A multiprocessor architecture which allows processes to be allocated to tightly coupled multiple processors in a cooperative processing environment, allowing concurrent execution of tasks.

parameter. * (1) A variable that is given a constant value for a specified application and that may denote the application. * (2) An item in a menu for which the operator specifies a value or for which the system provides a value when the menu is interpreted. * (3) A name in a procedure that is used to refer to an argument that is passed to the procedure. * (4) A particular piece of information that a system or application program needs to process a request.

partition. See system partition.

Perl. Practical Extraction and Report Language.

perspective. The primary window for each SP Perspectives application, so called because it provides a unique view of an SP system.

pipe. A UNIX utility allowing the output of one command to be the input of another. Represented by the | symbol. It is also referred to as filtering output.

PMR. Problem Management Report.

POE. Formerly Parallel Operating Environment, now Parallel Environment for AIX.

port. (1) An end point for communication between devices, generally referring to physical connection. (2) A 16-bit number identifying a particular TCP or UDP resource within a given TCP/IP node.

predicate. Obsolete term for expression.

Primary node or machine. (1) A device that runs a workload and has a standby device ready to assume the primary workload if that primary node fails or is taken out of service. (2) A node on the SP Switch that initializes, provides diagnosis and recovery services, and performs other operations to the switch network. (3) In IBM Virtual Shared Disk function, when physical disks are connected to two nodes (twin-tailed), one node is designated as the primary node for each disk and the other is designated the secondary, or backup, node. The primary node is the server node for IBM Virtual Shared Disks defined on the physical disks under normal conditions. The secondary node can become the server node for the disks if the primary node is unavailable (off-line or down).

Problem Management Report. The number in the IBM support mechanism that represents a service incident with a customer.

process. * (1) A unique, finite course of events defined by its purpose or by its effect, achieved under defined conditions. * (2) Any operation or combination of operations on data. * (3) A function being performed or waiting to be performed. * (4) A program in operation. For example, a daemon is a system process that is always running on the system.

protocol. A set of semantic and syntactic rules that defines the behavior of functional units in achieving communication.

R

RAID. Redundant array of independent disks.

rearm expression. In Event Management, an expression used to generate an event that alternates with an original event expression in the following way: the event expression is used until it is true, then the

rearm expression is used until it is true, then the event expression is used, and so on. The rearm expression is commonly the inverse of the event expression (for example, a resource variable is on or off). It can also be used with the event expression to define an upper and lower boundary for a condition of interest.

rearm predicate. Obsolete term for rearm expression.

remote host. *See foreign host.*

resource. In Event Management, an entity in the system that provides a set of services. Examples of resources include hardware entities such as processors, disk drives, memory, and adapters, and software entities such as database applications, processes, and file systems. Each resource in the system has one or more attributes that define the state of the resource.

resource identifier. In Event Management, a set of elements, where each element is a name/value pair of the form `name=value`, whose values uniquely identify the copy of the resource (and by extension, the copy of the resource variable) in the system.

resource monitor. A program that supplies information about resources in the system. It can be a command, a daemon, or part of an application or subsystem that manages any type of system resource.

resource variable. In Event Management, the representation of an attribute of a resource. An example of a resource variable is `IBM.AIX.PagSp.%totalfree`, which represents the percentage of total free paging space. `IBM.AIX.PagSp` specifies the resource name and `%totalfree` specifies the resource attribute.

RISC. Reduced Instruction Set Computing (RISC), the technology for today's high performance personal computers and workstations, was invented in 1975. Uses a small simplified set of frequently used instructions for rapid execution.

rlogin (remote LOGIN). A service offered by Berkeley UNIX systems that allows authorized users of one machine to connect to other UNIX systems across a network and interact as if their terminals were connected directly. The rlogin software passes information about the user's environment (for example, terminal type) to the remote machine.

RPC. Acronym for Remote Procedure Call, a facility that a client uses to have a server execute a procedure call. This facility is composed of a library of procedures plus an XDR.

RSH. A variant of RLOGIN command that invokes a command interpreter on a remote UNIX machine and passes the command line arguments to the command interpreter, skipping the LOGIN step completely. See also *rlogin*.

S

SCSI. Small Computer System Interface.

Secondary node. In IBM Virtual Shared Disk function, when physical disks are connected to two nodes (twin-tailed), one node is designated as the primary node for each disk and the other is designated as the secondary, or backup, node. The secondary node acts as the server node for the IBM Virtual Shared disks defined on the physical disks if the primary node is unavailable (off-line or down).

server. (1) A function that provides services for users. A machine may run client and server processes at the same time. (2) A machine that provides resources to the network. It provides a network service, such as disk storage and file transfer, or a program that uses such a service. (3) A device, program, or code module on a network dedicated to providing a specific service to a network. (4) On a LAN, a data station that provides facilities to other data stations. Examples are file server, print server, and mail server.

shell. The shell is the primary user interface for the UNIX operating system. It serves as command language interpreter, programming language, and allows foreground and background processing. There are three different implementations of the shell concept: Bourne, C and Korn.

Small Computer System Interface (SCSI). An input and output bus that provides a standard interface for the attachment of various direct access storage devices (DASD) and tape drives to the RS/6000.

Small Computer Systems Interface Adapter (SCSI Adapter). An adapter that supports the attachment of various direct-access storage devices (DASD) and tape drives to the RS/6000.

SMIT. The System Management Interface Toolkit is a set of menu driven utilities for AIX that provides functions such as transaction login, shell script creation, automatic updates of object database, and so forth.

SNMP. Simple Network Management Protocol. (1) An IP network management protocol that is used to monitor attached networks and routers. (2) A TCP/IP-based protocol for exchanging network management information and outlining the structure for communications among network devices.

socket. (1) An abstraction used by Berkeley UNIX that allows an application to access TCP/IP protocol functions. (2) An IP address and port number pairing. (3) In TCP/IP, the Internet address of the host computer on which the application runs, and the port number it uses. A TCP/IP application is identified by its socket.

standby node or machine. A device that waits for a failure of a primary node in order to assume the identity of the primary node. The standby machine then runs the primary's workload until the primary is back in service.

subnet. Shortened form of subnetwork.

subnet mask. A bit template that identifies to the TCP/IP protocol code the bits of the host address that are to be used for routing for specific subnetworks.

subnetwork. Any group of nodes that have a set of common characteristics, such as the same network ID.

subsystem. A software component that is not usually associated with a user command. It is usually a daemon process. A subsystem will perform work or provide services on behalf of a user request or operating system request.

SUP. Software Update Protocol.

switch capsule. A group of SP frames consisting of a switched frame and its companion non-switched frames.

Sysctl. Secure System Command Execution Tool. An authenticated client/server system for running commands remotely and in parallel.

syslog. A BSD logging system used to collect and manage other subsystem's logging data.

System Administrator. The user who is responsible for setting up, modifying, and maintaining the SP system.

system partition. A group of nonoverlapping nodes on a switch chip boundary that act as a logical SP system.

T

tar. Tape ARchive, is a standard UNIX data archive utility for storing data on tape media.

TaskGuides. SP TaskGuides are a form of advanced online assistance designed to walk you through complex or infrequently performed tasks. Each TaskGuide does not simply list the required steps. It actually performs the steps for you, automating the steps to the highest degree possible and prompting you for input only when absolutely necessary. You might recognize them as *wizards*.

Tcl. Tool Command Language.

TclX. Tool Command Language Extended.

TCP. Acronym for Transmission Control Protocol, a stream communication protocol that includes error recovery and flow control.

TCP/IP. Acronym for Transmission Control Protocol/Internet Protocol, a suite of protocols designed to allow communication between networks regardless of the technologies implemented in each network. TCP provides a reliable host-to-host protocol between hosts in packet-switched communications networks and in interconnected systems of such networks. It assumes that the underlying protocol is the Internet Protocol.

Telnet. Terminal Emulation Protocol, a TCP/IP application protocol that allows interactive access to foreign hosts.

Tk. Tcl-based Tool Kit for X Windows.

TMPCP. Tape Management Program Control Point.

token-ring. (1) Network technology that controls media access by passing a token (special packet or frame) between media-attached machines. (2) A network with a ring topology that passes tokens from one attaching device (node) to another. (3) The IBM Token-Ring LAN connection allows the RS/6000 system unit to participate in a LAN adhering to the IEEE 802.5 Token-Passing Ring standard or the ECMA standard 89 for Token-Ring, baseband LANs.

transaction. An exchange between the user and the system. Each activity the system performs for the user is considered a transaction.

transceiver (transmitter-receiver). A physical device that connects a host interface to a local area network, such as Ethernet. Ethernet transceivers contain electronics that apply signals to the cable and sense collisions.

transfer. To send data from one place and to receive the data at another place. Synonymous with move.

transmission. * The sending of data from one place for reception elsewhere.

TURBOWAYS 100 ATM Adapter. An IBM high-performance, high-function intelligent adapter that provides dedicated 100 Mbps ATM (asynchronous transfer mode) connection for high-performance servers and workstations.

U

UDP. User Datagram Protocol.

UNIX operating system. An operating system developed by Bell Laboratories that features multiprogramming in a multiuser environment. The UNIX operating system was originally developed for use on minicomputers, but has been adapted for mainframes and microcomputers. **Note:** The AIX operating system is IBM's implementation of the UNIX operating system.

user. Anyone who requires the services of a computing system.

User Datagram Protocol (UDP). (1) In TCP/IP, a packet-level protocol built directly on the Internet Protocol layer. UDP is used for application-to-application programs between TCP/IP host systems. (2) A transport protocol in the Internet suite of protocols that provides unreliable, connectionless datagram service. (3) The Internet Protocol that enables an application programmer on one machine or process to send a datagram to an application program on another machine or process.

user ID. A nonnegative integer, contained in an object of type *uid_t*, that is used to uniquely identify a system user.

V

Virtual Shared Disk, IBM. The function that allows application programs executing at different nodes of a system partition to access a raw logical volume as if it were local at each of the nodes. In actuality, the logical volume is local at only one of the nodes (the server node).

W

workstation. * (1) A configuration of input/output equipment at which an operator works. * (2) A terminal or microcomputer, usually one that is connected to a mainframe or to a network, at which a user can perform applications.

X

X Window System. A graphical user interface product.

Bibliography

This bibliography helps you find product documentation related to the RS/6000 SP hardware and software products.

You can find most of the IBM product information for RS/6000 SP products on the World Wide Web. Formats for both viewing and downloading are available.

PSSP documentation is shipped with the PSSP product in a variety of formats and can be installed on your system. The man pages for public code that PSSP includes are also available online.

You can order hard copies of the product documentation from IBM. This bibliography lists the titles that are available and their order numbers.

Finally, this bibliography contains a list of non-IBM publications that discuss parallel computing and other topics related to the RS/6000 SP.

Finding Documentation on the World Wide Web

Most of the RS/6000 SP hardware and software books are available from the IBM RS/6000 Web site at:

http://www.rs6000.ibm.com

You can view a book or download a Portable Document Format (PDF) version of it. At the time this manual was published, the Web address of the "RS/6000 SP Product Documentation Library" page was:

http://www.rs6000.ibm.com/resource/aix_resource/sp_books

However, the structure of the RS/6000 Web site can change over time.

Accessing PSSP Documentation Online

On the same medium as the PSSP product code, IBM ships PSSP man pages, HTML files, and PDF files. In order to use these publications, you must first install the **ssp.docs** file set.

To view the PSSP HTML publications, you need access to an HTML document browser such as Netscape. The HTML files and an index that links to them are installed in the **/usr/lpp/ssp/html** directory. Once installed, you can also view the HTML files from the RS/6000 SP Resource Center.

If you have installed the SP Resource Center on your SP system, you can access it by entering the **/usr/lpp/ssp/bin/resource_center** command. If you have the SP Resource Center on CD-ROM, see the **readme.txt** file for information about how to run it.

To view the PSSP PDF publications, you need access to the Adobe Acrobat Reader 3.0.1. The Acrobat Reader is shipped with the AIX Version 4.3 Bonus Pack.

To successfully print a large PDF file (approximately 300 or more pages) from the Adobe Acrobat reader, you may need to select the “Download Fonts Once” button on the Print window.

Manual Pages for Public Code

The following manual pages for public code are available in this product:

SUP	/usr/lpp/ssp/man/man1/sup.1
NTP	/usr/lpp/ssp/man/man8/xntpd.8
	/usr/lpp/ssp/man/man8/xntpdc.8
Perl (Version 4.036)	/usr/lpp/ssp/perl/man/perl.man
	/usr/lpp/ssp/perl/man/h2ph.man
	/usr/lpp/ssp/perl/man/s2p.man
	/usr/lpp/ssp/perl/man/a2p.man
Perl (Version 5.003)	Man pages are in the /usr/lpp/ssp/perl5/man/man1 directory

Manual pages and other documentation for **Tcl**, **TclX**, **Tk**, and **expect** can be found in the compressed **tar** files located in the **/usr/lpp/ssp/public** directory.

RS/6000 SP Planning Publications

This section lists the IBM product documentation for planning for the IBM RS/6000 SP hardware and software.

IBM RS/6000 SP:

- *Planning, Volume 1, Hardware and Physical Environment*, GA22-7280
- *Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281

RS/6000 SP Hardware Publications

This section lists the IBM product documentation for the IBM RS/6000 SP hardware.

IBM RS/6000 SP:

- *Planning, Volume 1, Hardware and Physical Environment*, GA22-7280
- *Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281
- *Maintenance Information, Volume 1, Installation and Relocation*, GA22-7375
- *Maintenance Information, Volume 2, Maintenance Analysis Procedures*, GA22-7376
- *Maintenance Information, Volume 3, Locations and Service Procedures*, GA22-7377
- *Maintenance Information, Volume 4, Parts Catalog*, GA22-7378

RS/6000 SP Switch Router Publications

The RS/6000 SP Switch Router is based on the Ascend GRF switched IP router product from Ascend Communications, Inc.. You can order the SP Switch Router as the IBM 9077.

The following publications are shipped with the SP Switch Router. You can also order these publications from IBM using the order numbers shown.

- *Ascend GRF Getting Started*, GA22-7368
- *Ascend GRF Configuration Guide*, GA22-7366
- *Ascend GRF Reference Guide*, GA22-7367
- *IBM SP Switch Router Adapter Guide*, GA22-7310.

RS/6000 SP Software Publications

This section lists the IBM product documentation for software products related to the IBM RS/6000 SP. These products include:

- IBM Parallel System Support Programs for AIX (PSSP)
- IBM LoadLeveler for AIX (LoadLeveler)
- IBM Parallel Environment for AIX (Parallel Environment)
- IBM General Parallel File System for AIX (GPFS)
- IBM Engineering and Scientific Subroutine Library (ESSL) for AIX
- IBM Parallel ESSL for AIX
- IBM High Availability Cluster Multi-Processing for AIX (HACMP)
- IBM Client Input Output/Sockets (CLIO/S)
- IBM Network Tape Access and Control System for AIX (NetTAPE)

PSSP Publications

IBM RS/6000 SP:

- *Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281

PSSP:

- *Installation and Migration Guide*, GA22-7347
- *Administration Guide*, SA22-7348
- *Managing Shared Disks*, SA22-7349
- *Performance Monitoring Guide and Reference*, SA22-7353
- *Diagnosis Guide*, GA22-7350
- *Command and Technical Reference*, SA22-7351
- *Messages Reference*, GA22-7352

RS/6000 Cluster Technology (RSCT):

- *Event Management Programming Guide and Reference*, SA22-7354
- *Group Services Programming Guide and Reference*, SA22-7355

As an alternative to ordering the individual books, you can use SBOF-8587 to order the PSSP software library.

LoadLeveler Publications

LoadLeveler:

- *Using and Administering*, SA22-7311
- *Diagnosis and Messages Guide*, GA22-7277

GPFS Publications

GPFS:

- *Installation and Administration Guide*, SA22-7278

Parallel Environment Publications

Parallel Environment:

- *Installation Guide*, GC28-1981

- *Hitchhiker's Guide*, GC23-3895
- *Operation and Use, Volume 1*, SC28-1979
- *Operation and Use, Volume 2*, SC28-1980
- *MPI Programming and Subroutine Reference*, GC23-3894
- *MPL Programming and Subroutine Reference*, GC23-3893
- *Messages*, GC28-1982

As an alternative to ordering the individual books, you can use SBOF-8588 to order the PE library.

Parallel ESSL and ESSL Publications

- *ESSL Products: General Information*, GC23-0529
- *Parallel ESSL: Guide and Reference*, SA22-7273
- *ESSL: Guide and Reference*, SA22-7272

HACMP Publications

HACMP:

- *Concepts and Facilities*, SC23-4276
- *Planning Guide*, SC23-4277
- *Installation Guide*, SC23-4278
- *Administration Guide*, SC23-4279
- *Troubleshooting Guide*, SC23-4280
- *Programming Locking Applications*, SC23-4281
- *Programming Client Applications*, SC23-4282
- *Master Index and Glossary*, SC23-4285
- *HANFS for AIX Installation and Administration Guide*, SC23-4283
- *Enhanced Scalability Installation and Administration Guide*, SC23-4284

CLIO/S Publications

CLIO/S:

- *General Information*, GC23-3879
- *User's Guide and Reference*, GC28-1676

NetTAPE Publications

NetTAPE:

- *General Information*, GC23-3990
- *User's Guide and Reference*, available from your IBM representative

AIX and Related Product Publications

For the latest information on AIX and related products, including RS/6000 hardware products, see *AIX and Related Products Documentation Overview*, SC23-2456. You can order a hard copy of the book from IBM. You can also view it online from the “AIX Online Publications and Books” page of the RS/6000 Web site at:

http://www.rs6000.ibm.com/resource/aix_resource/Pubs

Red Books

IBM's International Technical Support Organization (ITSO) has published a number of redbooks related to the RS/6000 SP. For a current list, see the ITSO Web site at:

http://www.redbooks.ibm.com

Non-IBM Publications

Here are some non-IBM publications that you may find helpful.

- Almasi, G., Gottlieb, A., *Highly Parallel Computing*, Benjamin-Cummings Publishing Company, Inc., 1989.
- Foster, I., *Designing and Building Parallel Programs*, Addison-Wesley, 1995.
- Gropp, W., Lusk, E., Skjellum, A., *Using MPI*, The MIT Press, 1994.
- Message Passing Interface Forum, *MPI: A Message-Passing Interface Standard, Version 1.1*, University of Tennessee, Knoxville, Tennessee, June 6, 1995.
- Message Passing Interface Forum, *MPI-2: Extensions to the Message-Passing Interface, Version 2.0*, University of Tennessee, Knoxville, Tennessee, July 18, 1997.
- Ousterhout, John K., *Tcl and the Tk Toolkit*, Addison-Wesley, Reading, MA, 1994, ISBN 0-201-63337-X.
- Pfister, Gregory, F., *In Search of Clusters*, Prentice Hall, 1998.

Index

Special Characters

Numerics

A

B

C

D

E

F

G

H

I

J

L

M

N

O

P

R

S

T

U

V

W

Communicating Your Comments to IBM

Parallel System Support Programs for AIX
Diagnosis Guide
Version 3 Release 1.1

Publication No. GA22-7350-01

If you especially like or dislike anything about this book, please use one of the methods listed below to send your comments to IBM. Whichever method you choose, make sure you send your name, address, and telephone number if you would like a reply.

Feel free to comment on specific errors or omissions, accuracy, organization, subject matter, or completeness of this book. However, the comments you send should pertain to only the information in this manual and the way in which the information is presented. To request additional publications, or to ask questions or make comments about the functions of IBM products or systems, you should talk to your IBM representative or to your IBM authorized remarketer.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you.

If you are mailing a reader's comment form (RCF) from a country other than the United States, you can give the RCF to the local IBM branch office or IBM representative for postage-paid mailing.

- If you prefer to send comments by mail, use the RCF at the back of this book.
- If you prefer to send comments by FAX, use this number:
  - FAX: (International Access Code)+1+914+432-9405
- If you prefer to send comments electronically, use one of these network IDs:
  - IBM Mail Exchange: USIB6TC9 at IBMMAIL
  - Internet e-mail: mhvrcfs@us.ibm.com

Make sure to include the following in your note:

- Title and publication number of this book
- Page number or topic to which your comment applies

Optionally, if you include your telephone number, we will be able to respond to your comments by phone.

Reader's Comments — We'd Like to Hear from You

Parallel System Support Programs for AIX
Diagnosis Guide
Version 3 Release 1.1

Publication No. GA22-7350-01

You may use this form to communicate your comments about this publication, its organization, or subject matter, with the understanding that IBM may use or distribute whatever information you supply in any way it believes appropriate without incurring any obligation to you. Your comments will be sent to the author's department for whatever review and action, if any, are deemed appropriate.

Note: Copies of IBM publications are not stocked at the location to which this form is addressed. Please direct any requests for copies of publications, or for assistance in using your IBM system, to your IBM representative or to the IBM branch office serving your locality.

Today's date: _____

What is your occupation?

Newsletter number of latest Technical Newsletter (if any) concerning this publication:

How did you use this publication?

[ ]	As an introduction	[ ]	As a text (student)
[ ]	As a reference manual	[ ]	As a text (instructor)
[ ]	For another purpose (explain)		

_____

_____

Is there anything you especially like or dislike about the organization, presentation, or writing in this manual? Helpful comments include general usefulness of the book; possible additions, deletions, and clarifications; specific errors and omissions.

Page Number: Comment:

_____ _____
Name Address

_____ _____
Company or Organization

_____ _____
Phone No.

Reader's Comments — We'd Like to Hear from You
GA22-7350-01

Cut or Fold
Along Line

Fold and Tape | **Please do not staple** | Fold and Tape

NO POSTAGE
NECESSARY
IF MAILED IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 40 ARMONK, NEW YORK

POSTAGE WILL BE PAID BY ADDRESSEE

IBM Corporation
Department 55JA, Mail Station P384
522 South Road
Poughkeepsie NY 12601-5400

Fold and Tape | **Please do not staple** | Fold and Tape

Cut or Fold
Along Line

Program Number: 5765-D51

Printed in the United States of America
on recycled paper containing 10%
recovered post-consumer fiber.

GA22-7350-01