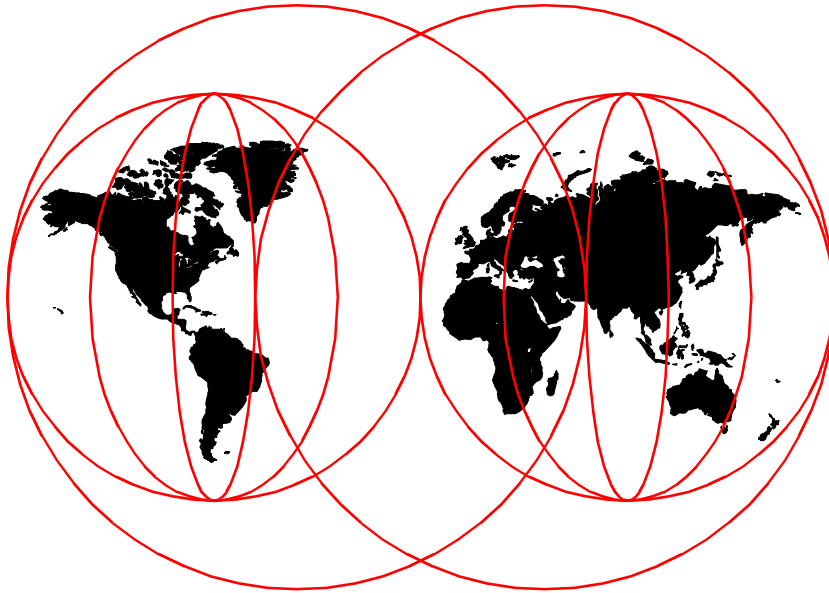


RS/6000 SP System Management: Power Recipes for PSSP 3.1

Yoshimichi Kosuge, Christoph Krafft, Yasuhiro Saitoh, Judy Vesely



International Technical Support Organization

www.redbooks.ibm.com

SG24-5628-00



International Technical Support Organization

**RS/6000 SP System Management:
Power Recipes for PSSP 3.1**

September 1999

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix A, "Special notices" on page 529.

First Edition (September 1999)

This edition applies to IBM Parallel System Support Programs for AIX Version 3, Release 1 (5765-D51) for use with the AIX Version 4, Release 3, Modification 2 (5765-C34).

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. JN9B Mail Station P099
522 South Road
Poughkeepsie, NY 12601-5400

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

© Copyright International Business Machines Corporation 1999. All rights reserved.
Note to U.S Government Users – Documentation related to restricted rights – Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Figures	xiii
Tables	xvii
Preface	xix
The team that wrote this redbook	xix
Comments welcome	xx
<hr/>	
Part 1. Managing hardware and software	1
Chapter 1. Hardware	3
1.1 Numbering schemes	4
1.1.1 The frame numbering scheme	4
1.1.2 The slot numbering scheme	4
1.1.3 The node numbering scheme	5
1.1.4 The switch port numbering scheme	6
1.2 Adding frames/nodes/SP-attached servers	10
1.2.1 Adding SP frames	11
1.2.2 Adding SP nodes	18
1.2.3 Adding SP-attached servers	38
1.3 Deleting frames/nodes/SP-attached servers	57
1.3.1 Before deleting frames/nodes	58
1.3.2 Deleting SP frames	59
1.3.3 Deleting SP nodes	64
1.3.4 Deleting SP-attached servers	69
1.4 Attaching/Detaching SP-attached servers	74
1.4.1 Attaching SP-attached servers	75
1.4.2 Detaching SP-attached servers	81
Chapter 2. Software	87
2.1 System firmware and microcode for devices	87
2.1.1 What is system firmware?	88
2.1.2 Checking the system firmware level	88
2.1.3 Upgrading system firmware	92
2.1.4 What is the microcode for devices?	93
2.1.5 Checking the microcode for devices	93
2.1.6 Upgrading the microcode for devices	95
2.2 Supervisor microcode	97
2.2.1 What is a supervisor microcode?	97
2.2.2 Getting supervisor microcode	97
2.2.3 Checking supervisor microcode levels	98

2.2.4	Upgrading supervisor microcode	100
2.2.5	Downgrading supervisor microcode	101
2.3	Program temporary fixes	102
2.3.1	What is a PTF?	102
2.3.2	Getting a PTF	102
2.3.3	Applying PTFs for AIX to CWS, nodes, and SPOT	108
2.3.4	Applying PTFs for LPPs to CWS and nodes	116
2.3.5	When should you apply PTF?	117
2.4	System backup	117
2.4.1	Backing up rootvg on CWS	118
2.4.2	Restoring rootvg on CWS	118
2.4.3	Backing up /spdata on CWS	119
2.4.4	Restoring /spdata on CWS	120
2.4.5	Backing up/restoring strategy for SP nodes	121
2.4.6	Backing up rootvg on SP node	122
2.4.7	Restoring rootvg on SP node	122
2.4.8	Restoring rootvg mirroring node	125
2.5	Public domain software	126
2.5.1	Included PDS	126
2.5.2	Where is the source code?	126

Part 2. Managing installation, customization, and configuration 131

Chapter 3. Network installation management	133
3.1 Operations	133
3.1.1 NIM in SP system environment	134
3.1.2 What are wrappers?	136
3.1.3 Deleting a NIM client	136
3.1.4 Creating a NIM client	137
3.1.5 What is the lpp_source object?	138
3.1.6 Checking lpp_source object	140
3.1.7 Updating lpp_source object	140
3.1.8 What is the spot object?	141
3.1.9 Checking spot object	141
3.1.10 Checking SPOT log file	142
3.1.11 Updating the spot object	144
3.2 Isolating NIM problems	144
3.2.1 Checking NIM log files	144
3.2.2 Checking NIM configuration files	145
3.2.3 The c_sh_lib file	146
3.2.4 Getting NIM information from ODM	149

Chapter 4. Node installation	153
4.1 Monitoring node installation online	153
4.1.1 Using the s1term command	153
4.1.2 Using LED or LCD messages	155
4.2 Monitoring node installation offline	155
4.2.1 The bootlog file	155
4.2.2 The bosinst.data file	157
4.2.3 The bosinstlog file	158
4.2.4 The devinst.log file	158
4.2.5 The errlog file	159
4.2.6 The image.data file	159
4.3 Isolating problems during node installation	160
4.3.1 Hints and tips	160
Chapter 5. Node customization	163
5.1 Node customization	163
5.1.1 What is customization?	163
5.1.2 When do you need customization?	164
5.1.3 Customizing a node with or without rebooting	164
5.2 After node customization	166
5.2.1 /etc/inittab changes on the CWS	166
5.2.2 /etc/inittab changes on the SP node	167
5.2.3 What does normal node bootup do?	168
5.3 Isolating problems during node customization	169
5.3.1 The customization script files	169
5.3.2 Isolating problems by pssp_script	170
5.3.3 Isolating problems by pspfb_script	172
5.3.4 Meaning of the three digit codes	173
5.3.5 Hints and tips	175
Chapter 6. Disk configuration	177
6.1 Physical configuration	177
6.1.1 Adding a physical disk to a volume group	178
6.1.2 Deleting a physical disk from a volume group	180
6.2 Boot configuration	183
6.2.1 Using the alternate boot system image	183
6.2.2 Defining alternative boot system image	183
6.2.3 Installing alternative boot system image	185
6.2.4 Switching alternative boot system images	186
6.2.5 Booting from external SCSI disks	186
6.2.6 Booting from external SSA disks	188
6.3 Mirroring configuration	190
6.3.1 Configuring root volume group mirroring	190

6.3.2	Initiating root volume group mirroring	192
6.3.3	Discontinuing root volume group mirroring	195
Chapter 7. Network configuration		197
7.1	SP Ethernet	197
7.1.1	Replacing an SP Ethernet adapter on a node	197
7.2	Other networks	199
7.2.1	Adding a network adapter	199
7.2.2	Deleting a network adapter	205
7.2.3	Changing an IP address and host name	207
7.3	Global configuration	207
7.3.1	Changing an initial host name	207
7.3.2	Using Domain Name System	208
<hr/>		
Part 3. Controlling and monitoring the SP system		211
Chapter 8. Hardware information		213
8.1	Information in VPD	213
8.1.1	Accessing SP nodes information	213
8.1.2	Accessing control workstation information	218
8.2	Information in SDR	218
8.2.1	Accessing SP frames information	219
8.2.2	Accessing SP-attached servers information	220
8.2.3	Accessing SP nodes information	225
8.2.4	Accessing SP Switch information	228
Chapter 9. Controlling hardware		231
9.1	Controlling tools	231
9.1.1	Creating a Kerberos principal	232
9.1.2	Authorizing a Kerberos principal	232
9.2	Using Hardware Perspective	233
9.2.1	Starting Hardware Perspective	234
9.2.2	Controlling SP nodes	237
9.3	Using the spmon command	241
9.3.1	Setting the key switch position	242
9.3.2	Querying the key switch position	242
9.3.3	Powering on/off	242
9.3.4	Resetting the nodes	243
9.4	Using the hmcnds command	243
9.4.1	Setting the frame ID into the frame supervisor card	244
9.4.2	Initiating power-on self tests in the frame supervisor card	244
9.4.3	Booting a supervisor card	245
9.4.4	Switching to basecode version	245

9.4.5	Executing the basecode version	246
9.4.6	Downloading supervisor microcode	247
Chapter 10.	Monitoring hardware	249
10.1	Monitoring tools	249
10.1.1	Creating a Kerberos principal	250
10.1.2	Authorizing a Kerberos principal	250
10.2	Using Hardware Perspective	250
10.2.1	Starting Hardware Perspective	251
10.2.2	Monitoring SP frames or SP-attached servers	251
10.2.3	Monitoring SP nodes or SP-attached servers	254
10.2.4	Monitoring SP Switch boards	258
10.2.5	Monitoring your original conditions	262
10.3	Using the spon command	263
10.3.1	Monitoring the SP system	263
10.4	Using the hmmon command	265
10.4.1	Monitoring SP frames	265
10.4.2	Monitoring SP-attached servers	268
10.4.3	Monitoring SP nodes	270
10.4.4	Monitoring SP switch boards	272
<hr/>		
Part 4.	Managing the SP system events	277
Chapter 11.	Managing Events	279
11.1	Event Management subsystem concepts	279
11.1.1	What is a resource variable?	279
11.1.2	What is a resource ID?	280
11.1.3	What is an event expression?	282
11.1.4	What is a rearm expression?	283
11.2	Security Considerations for the Event Perspective	284
11.2.1	To define new conditions	284
11.2.2	To define event or rearm event actions	284
11.2.3	To take action	285
11.3	Using the Event Perspective	285
11.3.1	Starting Event Perspective	286
11.3.2	Viewing an event definition	287
11.3.3	Registering an event definition	292
11.3.4	Checking event notification	292
11.3.5	Checking rearm event notification	293
11.3.6	Unregistering event definition	295
11.4	Using the pmandef command	296
11.4.1	pmandefaults file	296
11.4.2	Subscribing an event	297

11.4.3	Listing events	300
11.4.4	unsubscribing the event	301
11.5	Using the haemqvar command	301
11.5.1	Listing resource variables	302
11.5.2	Getting an explanation of resource variable	302
11.5.3	Getting the value of a resource variable	304
11.5.4	Managing events	304

Part 5. Managing resources 307

Chapter 12. Managing software resources	309
12.1 File systems	309
12.1.1 Considering file systems	309
12.1.2 Getting more available file system space	311
12.1.3 Monitoring file systems	315
12.2 Paging space	317
12.2.1 Sizing paging space	318
12.2.2 Getting paging space information	318
12.2.3 Adjusting paging space size during installation	320
12.2.4 Adjusting paging space size after installation	322
12.2.5 Monitoring page space	327
12.3 Log files	327
12.3.1 Getting authorization	328
12.3.2 Collecting AIX error logs	328
12.3.3 Collecting BSD syslog logs	330
12.3.4 Collecting other logs	332
12.3.5 Eliminating increased log files	337
12.3.6 Monitoring log files	338
Chapter 13. Managing sets of nodes	343
13.1 Managing SP switch	343
13.1.1 What are primary and primary backup nodes?	343
13.1.2 Showing primary and primary backup nodes	344
13.1.3 Changing primary and primary backup nodes	344
13.1.4 Starting the SP switch	345
13.1.5 Stopping the SP switch	346
13.1.6 Joining switch fabric automatically	346
13.1.7 Getting switch buffer pool information	346
13.1.8 Changing switch buffer pool size	347
13.1.9 Using the cssadm daemon	348
13.1.10 Reading the switch topology file	351
13.1.11 Reinitializing clock source	351
13.1.12 Checking a switch log	352

13.2 System partition	355
13.2.1 Applying a system partition	355
13.2.2 Deleting a system partition	361
13.2.3 Creating your original partition configuration/layout	362
13.3 Node group	368
13.3.1 Starting up nodes by group	369
13.3.2 Shutting down nodes by group	370
13.3.3 Managing nodes using node group	371
13.3.4 Managing nodes using working collective	373

Part 6. Managing administrative tasks 377

Chapter 14. Administration tools	379
14.1 File collection technology	379
14.1.1 Getting information	380
14.1.2 Checking status	382
14.1.3 Checking served file collections	383
14.1.4 Checking resident files	383
14.1.5 Checking file collection server	384
14.1.6 Checking last updated time and date	384
14.1.7 Updating files managed by file collection	385
14.1.8 Changing update cycle	386
14.1.9 Update sequence	387
14.1.10 Checking log files	388
14.2 SP User Management	390
14.2.1 Getting information	390
14.2.2 Adding a user	392
14.2.3 Deleting a user	394
14.2.4 Controlling log in to CWS or nodes	395
14.2.5 Managing users, groups, or password on specific nodes	403
14.2.6 Stop using SP User Management	404
14.2.7 Using Network Information Service	405
14.3 Time synchronization	410
14.3.1 How does it work?	411
14.3.2 Getting information	412
14.3.3 Changing NTP time server	414
14.3.4 Monitoring NTP	415
14.3.5 Changing system time	416
14.4 The Automounter	418
14.4.1 Getting information	419
14.4.2 Changing home directory server and path	420
14.4.3 Using an SP switch for Automounter	422
14.4.4 Stop using Automounter	423

14.4.5	Checking logs	425
14.5	SP Accounting	425
14.5.1	Before using SP accounting	425
14.5.2	Getting information	425
14.5.3	Enabling SP accounting	427
Chapter 15. Online documentations		429
15.1	Man page	429
15.1.1	Installing an AIX man page	429
15.1.2	Installing an SP man page	430
15.1.3	Using man pages	431
15.2	PDF	432
15.2.1	Installing PDF files	433
15.2.2	Installing Adobe Acrobat Reader	434
15.2.3	Reading PDF files	434
15.3	HTML files	436
15.3.1	Installing HTML files	436
15.3.2	Installing Netscape Navigator	436
15.3.3	Reading HTML files	437
15.4	IBM RS/6000 SP Resource Center	438
15.4.1	Installing SP Resource Center	438
15.4.2	Starting SP Resource Center	439
15.4.3	Reading online documentations	440
15.4.4	Using SP man pages	441

Part 7. Managing security 443

Chapter 16. Security		445
16.1	Authentication and authorization methods	446
16.1.1	Enabling authentication methods	447
16.1.2	Selecting authorization methods	448
16.1.3	Listing authentication methods	449
16.2	Daemons	449
16.2.1	kerberos daemon	449
16.2.2	kadmind daemon	450
16.2.3	kproxd daemon	451
16.2.4	Kerberos authenticated-applications	451
16.2.5	Starting daemons	452
16.2.6	Stopping daemons	452
16.3	Authentication database	452
16.3.1	Initializing authentication database	453
16.3.2	Destroying authentication database	453
16.3.3	Backing up authentication database	454

16.3.4	Restoring authentication database	454
16.3.5	Reading an authentication database dump	455
16.3.6	Changing the Kerberos master key	457
16.4	Principal	459
16.4.1	Listing principal	459
16.4.2	Getting principal information	460
16.4.3	Adding a principal	461
16.4.4	Giving access authorization	465
16.4.5	Delete principal.	466
16.4.6	Changing password	468
16.4.7	Changing expiration date	470
16.4.8	Changing maximum ticket lifetime	471
16.4.9	Changing default values	473
16.5	Service key file	475
16.5.1	Creating a service key file.	475
16.5.2	Listing service key information	477
16.5.3	Changing service keys	478
16.6	Ticket cache file	480
16.6.1	Getting a ticket-granting ticket automatically.	480
16.6.2	Destroying tickets	483
16.7	Configuration	483
16.7.1	Configuring the Kerberos system	484
16.7.2	Unconfiguring the Kerberos system	490
16.7.3	Backing up the Kerberos system automatically.	492
16.7.4	Restoring the Kerberos system.	494
16.7.5	Adding an external RS/6000 system to Kerberos realm	495

Part 8. Managing shared disks 507

Chapter 17. IBM Recoverable Virtual Shared Disk	509
17.1 Configuration	509
17.1.1 Installing IBM Virtual Shared Disk.	510
17.1.2 Installing IBM Recoverable Virtual Shared Disk	510
17.1.3 Getting authorization	510
17.1.4 Starting IBM Virtual Shared Disk Perspective	511
17.1.5 Designating a node as an IBM Virtual Shared Disk node	513
17.1.6 Creating IBM Virtual Shared Disk	515
17.1.7 Configuring and activating IBM VSD and IBM RVSD	519
17.1.8 Verify IBM Virtual Shared Disk	520
17.2 Monitoring IBM Recoverable Virtual Shared Disk	522
17.2.1 The recovery subsystem.	523
17.2.2 Using resource variables	523
17.2.3 Using log files	526

Appendix A. Special notices	529
Appendix B. Related publications	531
B.1 International Technical Support Organization Publications	531
B.2 Redbooks on CD-ROMs	531
B.3 Other publications	532
How to get ITSO redbooks	535
IBM Redbook fax order form	536
List of abbreviations	537
Index	539
ITSO redbook evaluation	553

Figures

1. Frame, slot, and node numbers	6
2. Base configuration without expansion frame	7
3. Configuration 1 with one switchless expansion frame	7
4. Configuration 2 with two switchless expansion frames	8
5. Configuration 3 with three switchless expansion frames	8
6. SP Switch-8 configuration	10
7. Adding a new SP frame	11
8. Verify frame information by Hardware Perspective	16
9. Adding two new SP nodes	18
10. Verify node information by Hardware Perspective	22
11. Adding a new SP-attached server	38
12. Verify SP-attached server information by Hardware Perspective	44
13. Deleting SP frame	59
14. Deleting SP node	65
15. Deleting the SP-attached server	70
16. Attach SP-attached server	75
17. The /etc/bootptab.info File	76
18. Detaching SP-Attached Servers	82
19. RS/6000 microcode updates	91
20. Download RS/6000 microcode updates	92
21. SSA customer support	96
22. Supervisor microcodes in /spdata/sys1/ucode directory	98
23. RS/6000 SP Supervisor Manager	99
24. List status of supervisors (report form)	100
25. PTF Set 05 available on October 26, 1998	105
26. Get PTF from anonymous FTP server (1 of 3)	106
27. Get PTF from anonymous FTP server (2 of 3)	107
28. Get PTF from anonymous FTP server (3 of 3)	108
29. /usr/lpp/ssp/README/ssp.public.README (1 of 3)	128
30. /usr/lpp/ssp/README/ssp.public.README (2 of 3)	129
31. /usr/lpp/ssp/README/ssp.public.README (3 of 3)	130
32. The log file for creating the SPOT (1 of 2)	143
33. The log file for creating the SPOT (2 of 2)	143
34. Output during netboot phase	154
35. Output of lsvg command	179
36. Create volume group information	184
37. Boot/Install server information	185
38. Configuring root volume group mirroring	191
39. Check the current root volume group status	192
40. Hardware Perspective frame properties	220

41. Hardware Perspective frame properties of an SP-attached server	222
42. Node SDR class contents	223
43. NodeControl SDR class contents	224
44. Hardware Perspective node properties of an SP-attached server.	225
45. Node SDR class contents	226
46. Hardware Perspective node properties	228
47. Hardware Perspective SP Switch board properties.	230
48. Hardware Perspective	235
49. Add Pane dialog box	236
50. Hardware Perspective with Frames and Switches pane	237
51. Hardware Perspective SP node notebook.	238
52. Power Off, Reset, Shutdown or Fence Nodes dialog box	239
53. Fence or Unfence Nodes dialog box	240
54. Console window.	240
55. Network Boot Nodes dialog box.	241
56. Frame page of Set Monitoring notebook	252
57. Monitored conditions page of frame notebook.	253
58. Dynamic Resource Variables page of frame notebook	254
59. Node page of set monitoring notebook	255
60. Monitored conditions page of node notebook	257
61. All Dynamic Resource Variables page of node notebook	258
62. SwitchBoard page of set monitoring notebook	259
63. Monitored conditions page of SwitchBoard notebook	261
64. Dynamic Resource Variables page of SwitchBoard notebook.	262
65. The hmmon command output for an SP frame	266
66. The hmmon command output for an SP-attached server	269
67. The hmmon command output for an SP node.	270
68. The hmmon command output for an SP switch board.	273
69. Event Perspective window.	287
70. Event Definition Notebook	288
71. Condition notebook	290
72. Resource Variable Details window	291
73. Event Notification Log Window	292
74. Event Notification window	293
75. Event Notification log window	294
76. Event Notification (rearm) window.	295
77. Event Notification Window	299
78. Rearm Notification window	300
79. Event Notification window	306
80. /usr/lpp/ssp/samples/script.cust file (1 of 2).	321
81. /usr/lpp/ssp/samples/script.cust file (2 of 2).	322
82. The sysman.tab log table file	333
83. The errpt command output.	354

84. System Partitioning Aid Perspective window	363
85. Define System Partition dialog box	365
86. New system partition	366
87. Display existing configurations.	368
88. The user.admin file collection	385
89. Site environment database information for SP user management	391
90. The /etc/security/user file.	397
91. Change NTP installation mode (smitty site_env_dialog)	414
92. PSSP online documentations in PDF format	435
93. PSSP online documentations in HTML format.	438
94. SP Perspectives Launch Pad	439
95. RS/6000 SP Resource Center	440
96. Reading PSSP online document	441
97. Reading SP man page.	442
98. The setup_authent script (1 of 3)	485
99. The setup_authent script (2 of 3)	486
100. The setup_authent script (3 of 3)	488
101. Sample configuration.	497
102. IBM Virtual Shared Disk Perspective	512
103. Designate as an IBM VSD node dialog box	514
104. Designated nodes as IBM VSD nodes	515
105. Add Pane dialog box	516
106. IBM Virtual Shared Disk Perspective with IBM VSDs pane	517
107. Create IBM VSDs dialog box.	518
108. IBM Virtual Shared Disk vsd1n1 created.	519
109. Control IBM RVSD subsystem	520
110. Run Command on Nodes dialog box	521

Tables

1. FRU number and processor card	90
2. Anonymous FTP servers	105
3. NIM objects in the SP system resource	135
4. The naming convention for lpp_source objects	139
5. The naming convention for spot objects	141
6. Customization phase codes	173
7. Relationship between physical partitions and volumes	179
8. Supported adapters for nodes with SCSI disk boot	187
9. Supported adapters for nodes with SSA disk boot	188
10. Related Resource ID for Resource Variable Name	281
11. Structured byte string definitions for IBM.PSSP.pm.Errlog	340
12. System Partitioning Aid Perspective tool bar icons	363
13. Lifetime operand and approximate duration	472
14. Network adapter attributes	496
15. File sets for IBM Virtual Shared Disk (CWS and nodes)	510
16. File sets for IBM Virtual Shared Disk (CWS only)	510
17. File sets for IBM Recoverable Virtual Shared Disk (CWS and nodes)	510
18. IBM Virtual Shared Disk Perspective tool bar icons	512

Preface

Managing an RS/6000 SP system running AIX plus IBM Parallel System Support Programs for AIX (PSSP) Version 3, Release 1 can be a challenge because of the complex hardware and software interrelationships. This redbook supplies you with step-by-step instructions, sample operations, and screen captures to help make RS/6000 SP system management easier and more effective.

This redbook is intended as a task-oriented supplement to PSSP publications.

The team that wrote this redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization Poughkeepsie Center.

Yoshimichi Kosuge is an IBM RS/6000 SP project leader at the International Technical Support Organization, Poughkeepsie Center. Since he joined IBM, he has worked in the following areas: LSI design, S/390 CP microcode, VM, MVS, OS/2, and AIX. After joining the ITSO in 1998, he has been involved in writing redbooks and teaching IBM classes worldwide on all areas of RS/6000 SP system.

Christoph Krafft is an IT specialist in the Systems Management Services division of IBM Global Services in Germany. He joined IBM in 1992 after studying Technical Computer Science at the Berufsakademie in Stuttgart. While finishing his studies, Christoph worked in the IBM Lab in Boeblingen as a software developer.

Yasuhiro Saitoh is an IT specialist in IBM Japan Systems Engineering Co., Ltd. in Makuhari, Japan. After joining IBM, he has been working for a division that provides second-level support for AIX and RS/6000.

Judy Vesely is an IT specialist working in the AIX Support Center at IBM Canada. She has worked for IBM since 1996 and is a professional engineer. She provides support for AIX and RS/6000, specializing in SPs. Her main expertise is in troubleshooting and problem determination for an SP environment. Judy holds a degree in Electrical Engineering from the University of Toronto.

Thanks to the following people for their invaluable contributions to this project:

Dave Barton
Michael Chase-Salerno
Endy Chiakpo
Janet Ellsworth
Debra Kessler
Ronald Lember
Norman Nott
Keshav Ranganathan
Richard Russell
IBM Poughkeepsie

Scott Trent
IBM Austin

Marcelo Barrios
International Technical Support Organization, Poughkeepsie

Steve Gardner
International Technical Support Organization, Austin

Alison Chandler
Carol Dixon
John Owczarzak
Al Schwab
International Technical Support Organization, Editing Team

Comments welcome

Your comments are important to us!

We want our redbooks to be as helpful as possible. Please send us your comments about this or other redbooks in one of the following ways:

- Fax the evaluation form found in “ITSO redbook evaluation” on page 553 to the fax number shown on the form.
- Use the online evaluation form found at <http://www.redbooks.ibm.com/>
- Send your comments in an internet note to redbook@us.ibm.com

Part 1. Managing hardware and software

Chapter 1. Hardware

The IBM RS/6000 SP system provides you with great scalability. You can start with a configuration of only one SP node and then expand your system with numerous SP nodes. An IBM RS/6000 SP also has flexibility. If you have a couple of small SP systems, you can integrate them into one big SP system. The same is true in the other direction: You can divide one big SP system into a couple of small SP systems. Using IBM Parallel System Support Programs for AIX (PSSP) Version 3.1, you can integrate the RS/6000 Enterprise Server to your SP system. You can enable an SP Switch connection and a single point of control by a control workstation (CWS) to the RS/6000 Enterprise Server, thus, giving your SP system more power. All these features mean your IBM RS/6000 SP system is scalable and reconfigurable as your business requirements change.

This chapter describes how to address your SP system's hardware components and how to reconfigure the system. It covers the following topics:

- Adding frames/nodes/SP-attached servers
- Deleting frames/nodes/SP-attached servers
- Attaching/Detaching SP-attached servers

Adding SP hardware components is a necessary operation when you expand your SP system. Deleting SP hardware components is used when some of your SP hardware components are no longer used or are relocated to other SP systems.

Attention

The procedures introduced in this chapter assume simplified conditions (for example, that the SP system has only one partition, uses only SP Switch as an additional adapter, and so on).

This chapter provides you with *basic* reconfiguration procedures and shows you sample operations.

When you reconfigure your SP system, refer to the following IBM publication and follow the steps described:

- *RS/6000 SP: Planning, Volume 1, Hardware and Physical Environment, GA22-7280*

1.1 Numbering schemes

First of all, you need to know about numbering schemes. When you identify an individual frame, slot, node, or switch port, you need a number with which to refer to it. This section covers the following numbering schemes:

- The frame numbering scheme
- The slot numbering scheme
- The SP-attached server numbering scheme
- The node numbering scheme
- The switch port numbering scheme

For more information, refer to Chapter 3, “Defining the Configuration that Fits Your Needs” in *RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281.

1.1.1 The frame numbering scheme

The administrator establishes the frame numbers when the system is installed. Each frame is referenced by the TTY port to which the frame supervisor is attached and is assigned a numeric identifier. The order in which the frames are numbered defines the sequence in which they are examined during the configuration process. This order is used to assign global identifiers to the switch ports and nodes. This is also the order used to determine which frames share a switch.

An SP-attached server appears similar to a regular processor node to the PSSP. However, it must be assigned a frame number. There are a couple of rules that you must follow when selecting frame numbers:

- Do not make the SP-attached server the first frame in the SP system.
- The assigned frame number cannot be a number that would come between the number assigned to a switched frame and the frame numbers assigned to any switchless expansion frames attached to the switched frame.

1.1.2 The slot numbering scheme

A 1.93 m frame has a total of 16 slots; a 1.25 m frame has a total of eight slots. When viewing an SP frame from the front, the slots are numbered sequentially from bottom left to top right. The slot number starts with one, and the maximum number is 16 for the case of an 1.93 m frame.

The position of a node in an SP frame is sensed by the hardware, and this position is the slot to which it is wired. This slot is the slot number of the node as follows:

- A thin node's slot number is the corresponding slot.
- A wide node's slot number is the odd-numbered slot.
- A high node's slot number is the first (lowest-number) slot.
- An SP-attached server's slot number is always one (1).

The slot numbers used for other than nodes or SP-attached servers are:

- Slot number zero (0) is used for the frame supervisor card.
- Slot number 17 is used for the SP Switch board supervisor card.

1.1.3 The node numbering scheme

A node number is a global ID assigned to a node. It is the method by which an administrator can reference a specific node in the SP system. Node numbers are assigned for all nodes, including SP-attached servers, regardless of node or frame type by the following formula:

```
node_number = ((frame_number - 1) x 16) + slot_number
```

Node numbers are assigned independently of whether the frame is fully populated.

The node number used on items other than nodes or SP-attached servers is:

- The CWS uses node number 0.

Figure 1 on page 6 shows an example of frame, slot, and node numbers. Frame 3 is an SP-attached server. The number to the left of the comma is the slot number, to the right of the comma is the node number.

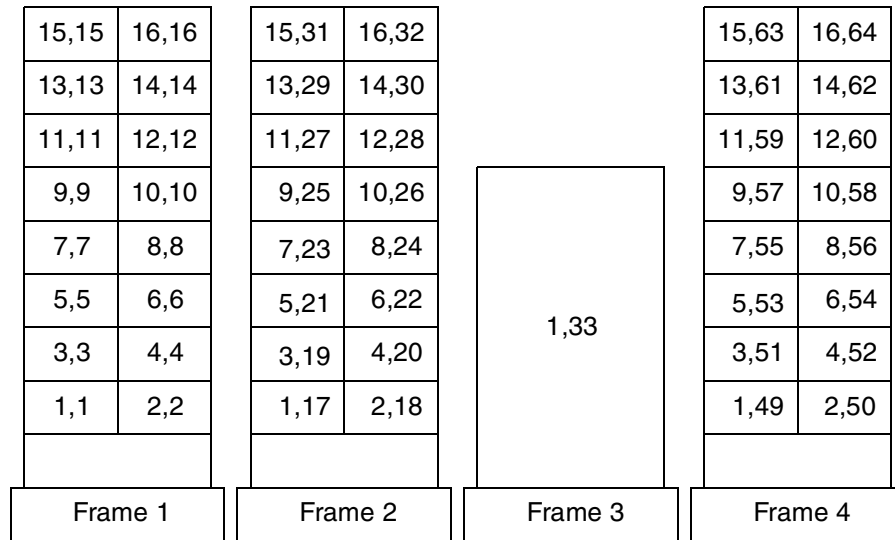


Figure 1. Frame, slot, and node numbers

1.1.4 The switch port numbering scheme

This scheme is also known as the node placement scheme because you can place nodes only in the slots that have a switch port number.

For the SP Switch-8, different algorithms are used for assigning nodes their switch port numbers.

SP Switch

For the 16-port SP switches, you can use the following formula to determine the switch port number to which a node is attached:

$$\text{switch_port_number} = (\text{switch_number} - 1) \times 16 + \text{port_number}$$

Here, switch_number is the number of the switch board to which the node is connected, and port_number is the port position on the switch board to which the node is connected. In Figure 2 on page 7, the model frame has an SP Switch that uses all 16 of its switch ports. Since all switch ports are used, the frame does not support switchless expansion frames.

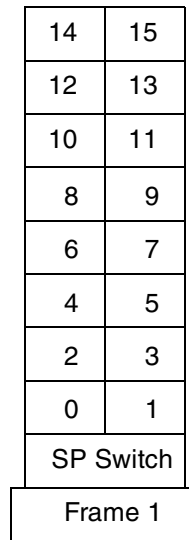


Figure 2. Base configuration without expansion frame

When you add switchless 1.93 m expansion frames to 1.93 m switched model frames, the SP system currently supports three frame configurations. Figure 3 through Figure 5 illustrate the supported configurations for switch port numbers.

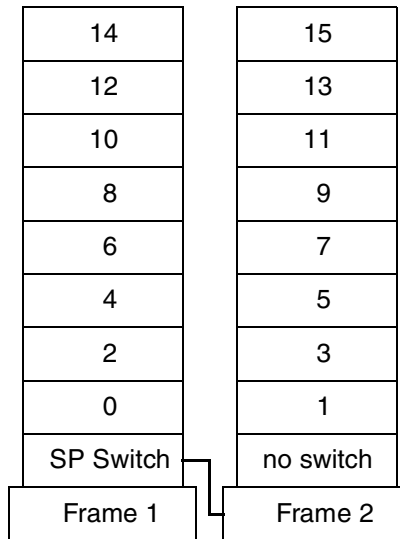


Figure 3. Configuration 1 with one switchless expansion frame

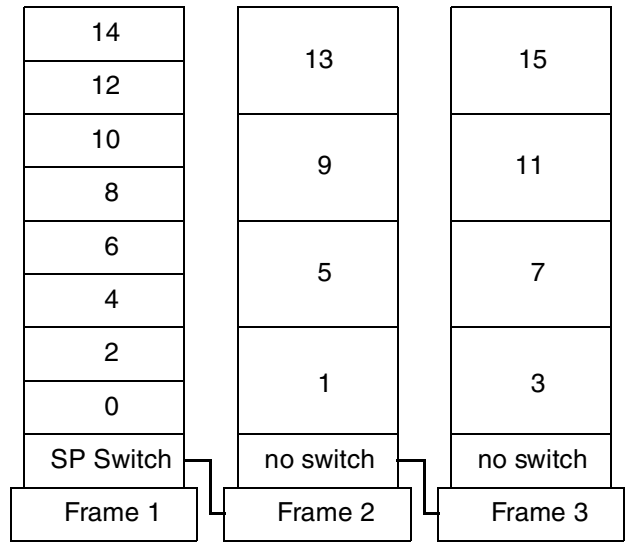


Figure 4. Configuration 2 with two switchless expansion frames

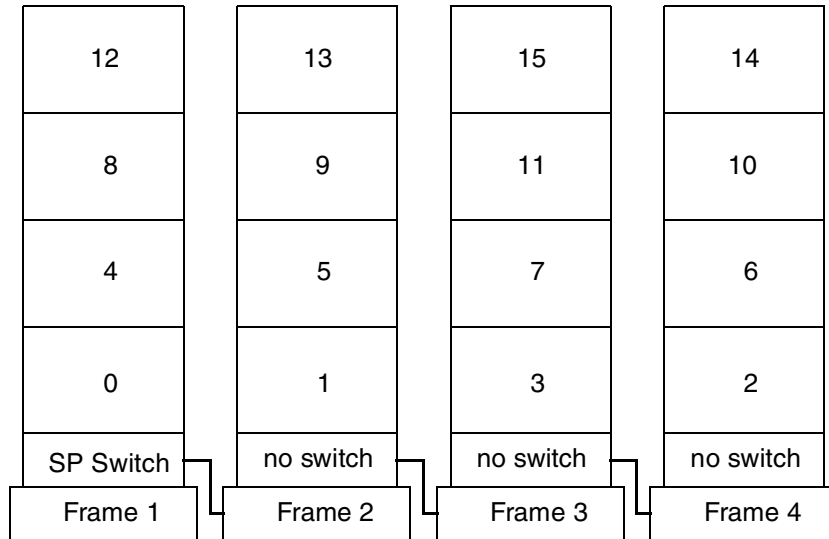


Figure 5. Configuration 3 with three switchless expansion frames

Configuration 1, shown in Figure 3 on page 7, has a single switchless expansion frame. If thin nodes are used, they must be placed in the wide node address points.

Configuration 2, shown in Figure 4 on page 8, has two switchless expansion frames. If thin nodes are used in the model frame, they must be placed in the wide node address points. If thin or wide nodes are used in the expansion frames, they must be placed in the high node address points.

Configuration 3, shown in Figure 5 on page 8, has three switchless expansion frames. If thin or wide nodes are used, they must be placed in the high node address points.

The following configuration scheme applies to these configurations:

- Nodes in switchless expansion frames will be attached to the switch in the first preceding frame equipped with an SP Switch.

Therefore, you have to skip numbers between frames equipped with an SP Switch if you plan to expand your SP system at a later time with the addition of switchless expansion frames. Switchless expansion frames must be numbered consecutively from the model frame number that contains an SP Switch.

SP-attached server

Regardless of whether your SP system is a switched SP system or switchless SP system, you must assign a switch port number to the SP-attached server.

The following is the assignment scheme for both cases:

Switched SP system You can assign any valid and unused switch port in the SP system.

Switchless SP system An used node slot is required in the frame associated with the SP-attached server, and you must calculate the switch port number as if you used a switched SP system.

SP Switch-8

A system with SP Switch-8 contains only switch port number 0 through 7.

The following algorithm is used to assign nodes their switch port numbers for systems with SP Switch-8:

1. Assign the node in slot 1 to `switch_port_number = 0`.
2. Increment `switch_port_number` by 1.
3. Check the next slot. If there is a node in the slot, assign it the current `switch_port_number`.
4. Then repeat from Step 2.

Repeat these steps until you reach the last slot in the frames or the switch_port_number is 7, whichever comes first.

Figure 6 shows a sample configuration.

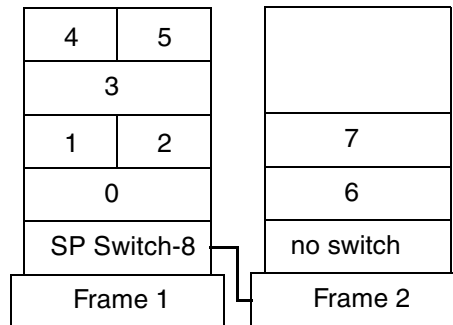


Figure 6. SP Switch-8 configuration

1.2 Adding frames/nodes/SP-attached servers

This section describes how you can expand your SP system. This discussion assumes that you have new SP hardware components and need to install them to your existing SP system. The section covers the following three topics:

- Adding SP frames
- Adding SP nodes
- Adding SP-attached servers

In 1.2.1, “Adding SP frames” on page 11, we assume that you have one SP frame with an SP Switch. You are going to add one new SP frame with an SP Switch to your SP system. At this point, the new SP frame has no SP nodes.

In 1.2.2, “Adding SP nodes” on page 18, you are going to add two new SP nodes to the empty SP frame installed in the previous section.

In 1.2.3, “Adding SP-attached servers” on page 38, you have one SP frame with an SP Switch. You are going to add one new RS/6000 Enterprise Server as an SP-attached server.

For more information, refer to Chapter 5, “Reconfiguring the RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347.

1.2.1 Adding SP frames

This section uses the assumption that you have one SP frame with an SP Switch. You are going to add one new SP frame with an SP Switch. At this point, the new SP frame has no SP nodes.

Figure 7 illustrates how you are going to add a new SP frame to your existing SP system.

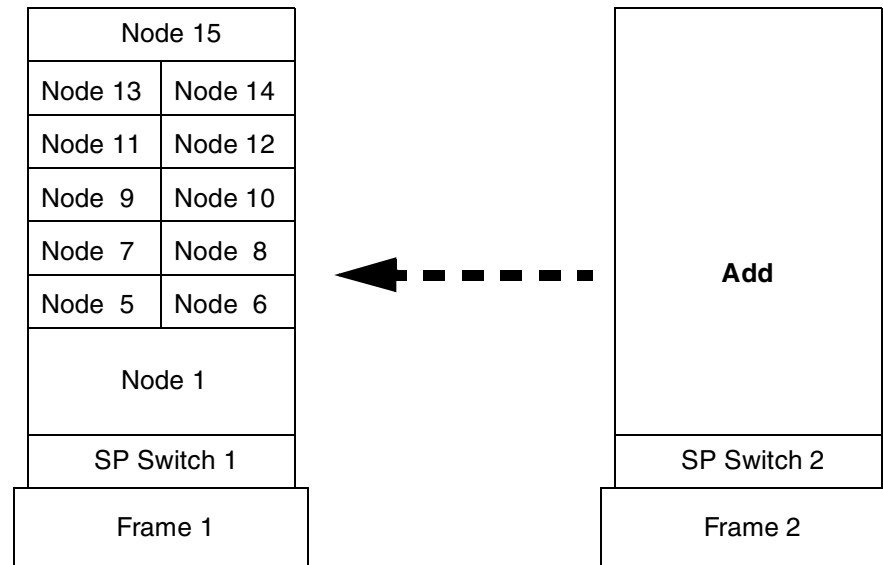


Figure 7. Adding a new SP frame

To add an SP frame, use the following steps:

Step 1: Archive the SDR

Before adding a frame to your SP system, you should back up the System Data Repository (SDR) by issuing the `SDRArchive` command:

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2: Connect frame to CWS

Your IBM Customer Engineer performs this step.

Step 3: Configure RS-232 control line

Each frame in your SP system requires a serial port on the CWS configured to accommodate the RS-232 control line.

To configure RS-232 control line, issue the `smitty maktty fast path`. After selecting **tty rs232 Asynchronous Terminal**, then selecting the parent adapter for the serial port you configure, you will see the following SMIT menu:

```

                                Add a TTY

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
TTY type                             tty
TTY interface                         rs232
Description                           Asynchronous Terminal
Parent adapter                         sa2
* PORT number                          [0] +
Enable LOGIN                           disable +
BAUD rate                              [9600] +
PARITY                                  [none] +
BITS per character                      [8] +
Number of STOP BITS                     [1] +
TIME before advancing to next port setting [0] +#
TERMINAL type                           [dumb]
FLOW CONTROL to be used                 [xon] +
[MORE...29]

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit           Enter=Do
```

Enter the required information, or issue the `mkdev` command to perform the same operation:

```
# mkdev -c tty -t tty -s rs232 -p sa2 -w 0
```

Step 4: Enter frame information and reinitialize the SDR

Each frame in your SP system requires an RS-232 control line assignment. In this example, you assign `/dev/tty1` tty port to the second frame.

To enter the SP frame information to the SDR, issue the `smitty sp_frame_dialog fast path`:

```

                                Frame Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* Start Frame                    [2]                #
* Frame Count                    [1]                #
* Starting Frame tty port        [/dev/tty1]
Re-initialize the System Data Repository  yes                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Alternatively, issue the `spframe` command to perform the same operation:

```

# spframe -r yes 2 1 /dev/tty1
0513-044 The stop of the splogd Subsystem was completed successfully.
0513-044 The stop of the hardmon Subsystem was completed successfully.
0513-059 The hardmon Subsystem has been started. Subsystem PID is 13220.
0513-059 The splogd Subsystem has been started. Subsystem PID is 31772.
#

```

Step 5: Test the RS-232 control line

Issue the `spmon_ctest` command. This command tests the RS-232 control line. The output should look like the following:

```

# spmon_ctest
spmon_ctest: Start spmon configuration verification test
spmon_ctest: Verification Succeeded
#

```

If problems are reported, check the RS-232 control line and look at the `/var/adm/SPlogs/spmon_ctest.log` log file. If the information in the log file is not sufficient, try tracing the `spmon_ctest` command. This command is a shell

script; so, the `set -x` command will not harm anything, and may give you a clue.

Step 6: Verify frame information

To verify that the SP System Monitor can detect the frame correctly, issue the `smon` command:

```
# smon -G -d
1. Checking server process
   Process 25670 has accumulated 0 minutes and 1 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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
2      yes      yes   on    0      on on on on

5. Checking nodes
----- Frame 1 -----
Frame Slot Node Node Host/Switch Key Env Front Panel LCD/LED is
      Number Type Power Responds Switch Fail LCD/LED  Flashing
-----
1      1    high on yes yes normal no LCDs 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    thin on yes yes normal no LEDs are blank no
8      8    thin on yes yes normal no LEDs are blank no
9      9    thin on yes yes normal no LEDs are blank no
10     10   thin on yes yes normal no LEDs are blank no
11     11   thin on yes yes normal no LEDs are blank no
12     12   thin on yes yes normal no LEDs are blank no
13     13   thin on yes yes normal no LEDs are blank no
14     14   thin on yes yes normal no LEDs are blank no
15     15   wide on yes yes normal no LEDs are blank no

Node information not found for frame 2.
#
```

Alternatively, issue the `splstdata` command:

```
# splstdata -f
      List Frame Database Information

frame#      tty      s1_tty      frame_type  hardware_protocol
-----
      1      /dev/tty0      ""          switch      SP
      2      /dev/tty1      ""          switch      SP
#
```

You can also use the Hardware Perspective shown in Figure 8 on page 16. To learn how to use the Hardware Perspective, refer to 9.2, “Using Hardware Perspective” on page 233.

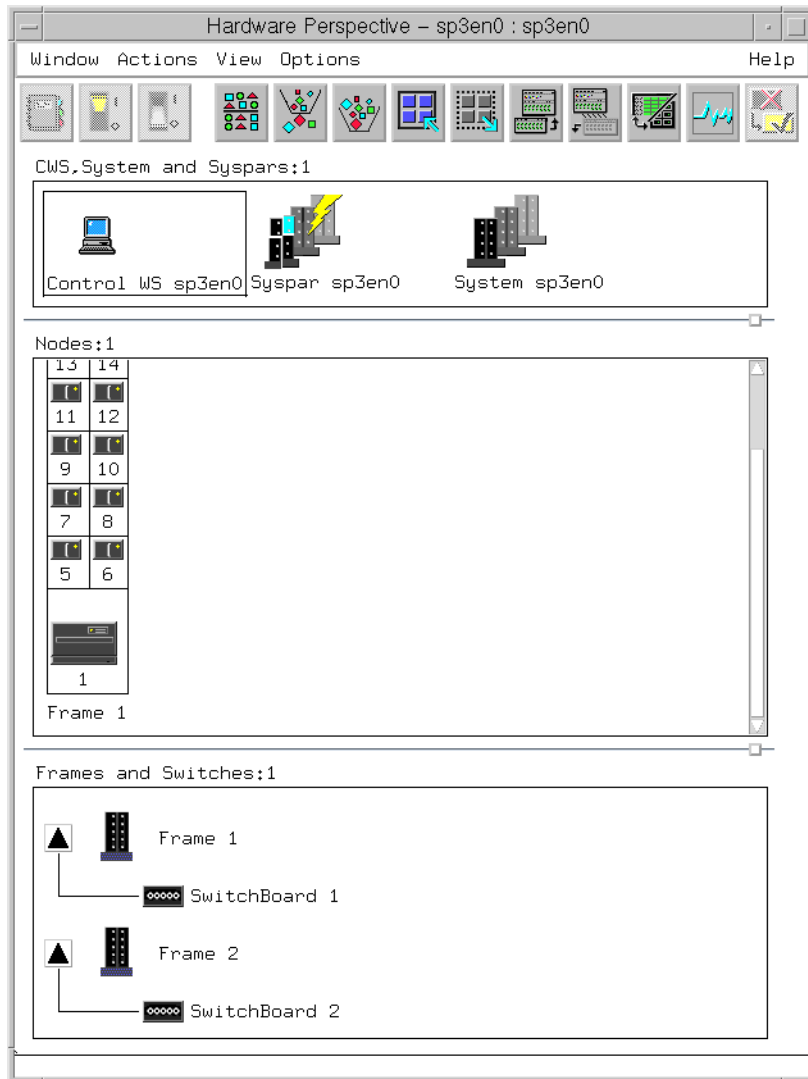


Figure 8. Verify frame information by Hardware Perspective

If your SP system does not recognize the frame, check the following items:

- Check that the frame supervisor card is working correctly. You can reset the supervisor card in the frame:

```
# hmcnds -G setid 2:0
# hmcnds -G boot_supervisor 2:0
```

Refer to 9.4, “Using the hmcnds command” on page 243, for details about the hmcnds command.

- Using the SDR_config command may give you some useful hints. To call it in verbose and debug mode, issue the -v and -d flags:

```
# SDR_config -v -d
```

Attention

The SDR_config command is to be used by the PSSP components. Use of the SDR_config command without the -d flag can cause corruption of system configuration data.

- Check that the hardmon daemon is working correctly. You can try to stop and restart it:

```
# stopsrc -s hardmon  
# startsrc -s hardmon
```

Step 7: Update the state of the supervisor microcode

After the frame has been detected by the hardmon daemon, check the frame supervisor microcode level by issuing the spsvmgr command:

```
# spsvmgr -G -r status all  
  
spsvmgr: Frame Slot Supervisor State Media Versions Installed Version Required Action  
-----  
1 0 Active u_10.1c.0709 u_10.1c.070c u_10.1c.070c None  
u_10.1c.070c  
1 1 Active u_10.3a.0614 u_10.3a.0615 u_10.3a.0615 None  
u_10.3a.0615  
17 17 Active u_80.19.060b u_80.19.060b u_80.19.060b None  
-----  
2 0 Active u_10.3c.0709 u_10.3c.070c u_10.3c.070c None  
u_10.3c.070c  
17 17 Active u_80.19.060b u_80.19.060b u_80.19.060b None  
-----  
#
```

If you need to upgrade the frame supervisor microcode level, issue the spsvmgr command:

```
# spsvmgr -G -u all
```

For more information about the supervisor microcode, refer to 2.2, “Supervisor microcode” on page 97.

1.2.2 Adding SP nodes

This section uses the assumption that you are going to add two new SP nodes to the empty SP frame installed in the preceding example. Figure 9 illustrates how you are going to add two new SP nodes to your existing SP system.

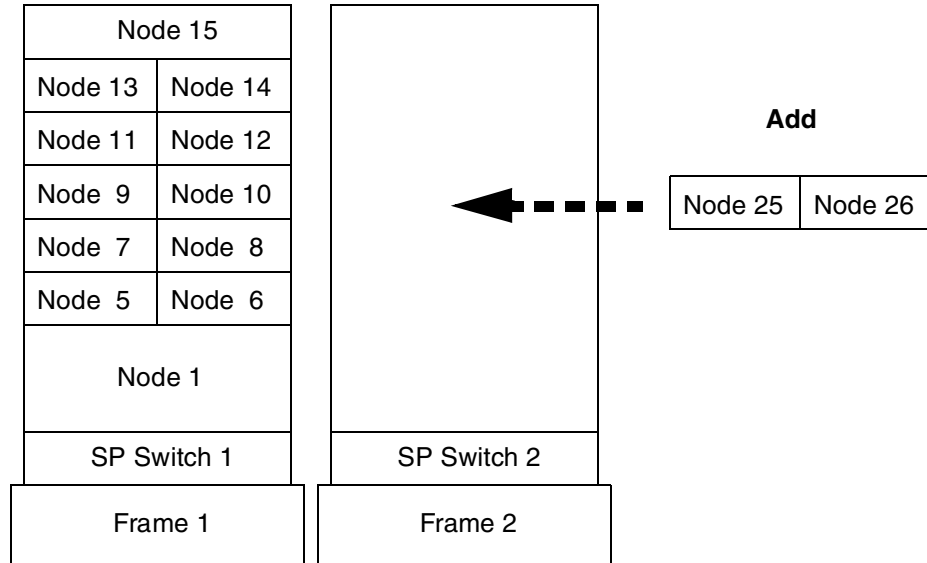


Figure 9. Adding two new SP nodes

To add the SP nodes, use the following steps:

Step 1: Archive the SDR

Before adding the nodes to your SP system, you should back up the SDR by issuing the `SDRArchive` command:

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```


Note the location and the name of the file created after you issue this command. In this example, the backup.99064.1459.mybackup file is created in the /spdata/sys1/sdr/archives directory.

Step 2: Connect nodes to the frame

Your IBM Customer Engineer performs this step.

Step 3: Verify the node information

To verify that the SP System Monitor can detect the nodes correctly, issue the `smon` command:

```

# spmon -G -d
1. Checking server process
   Process 25670 has accumulated 0 minutes and 4 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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
   2     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    high    on  yes  yes  normal  no  LCDs 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    thin    on  yes  yes  normal  no  LEDs are blank  no
   8     8    thin    on  yes  yes  normal  no  LEDs are blank  no
   9     9    thin    on  yes  yes  normal  no  LEDs are blank  no
  10    10    thin    on  yes  yes  normal  no  LEDs are blank  no
  11    11    thin    on  yes  yes  normal  no  LEDs are blank  no
  12    12    thin    on  yes  yes  normal  no  LEDs are blank  no
  13    13    thin    on  yes  yes  normal  no  LEDs are blank  no
  14    14    thin    on  yes  yes  normal  no  LEDs are blank  no
  15    15    wide    on  yes  yes  normal  no  LEDs are blank  no

----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   9     25    thin    off  no  notcfg  N/A    no  LCDs are blank  no
  10     26    thin    off  no  notcfg  N/A    no  LCDs are blank  no
#

```

Alternatively, issue the `splstdata` command:

```

# splstdata -n 2 8 2
                List Node Configuration Information

node# frame# slot# slots initial_hostname reliable_hostname dcehostname
      default_route processor_type processors_installed description
-----
  25     2     9     1 ""          ""          1 ""          ""
      ""          MP
  26     2    10     1 ""          ""          1 ""          ""
      ""          MP
#

```

You can also use the Hardware Perspective shown in Figure 10 on page 22. To learn how to use the Hardware Perspective, refer to 9.2, “Using Hardware Perspective” on page 233.

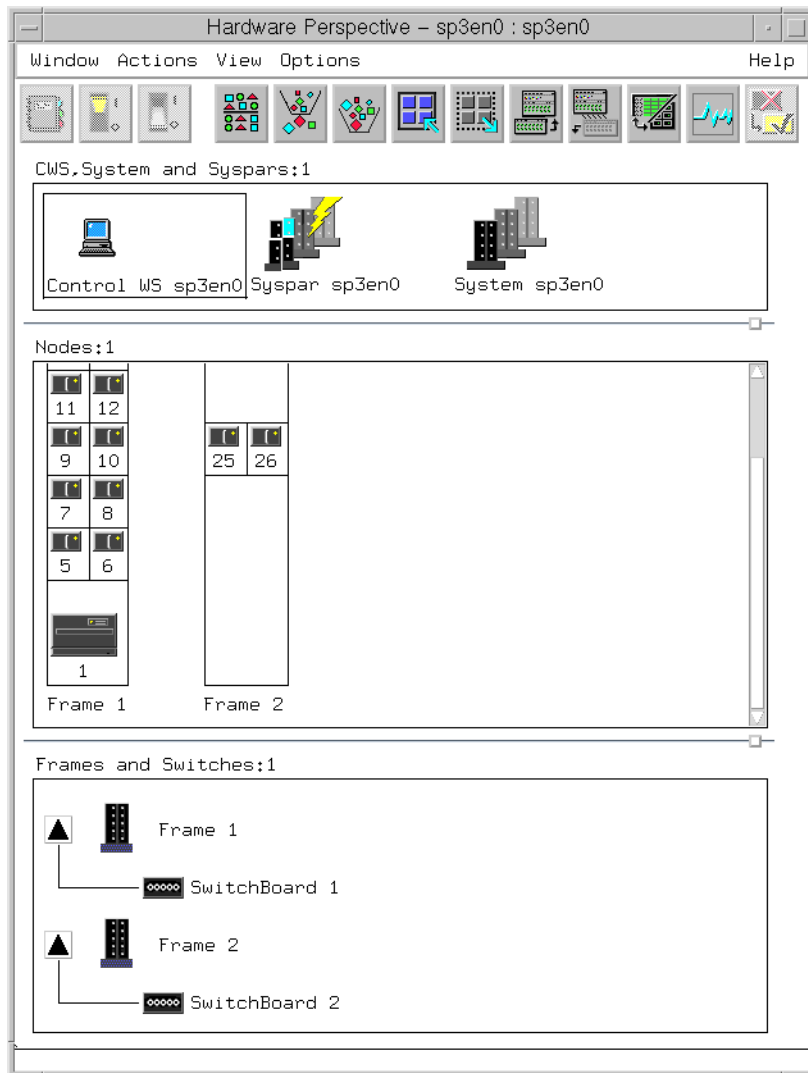


Figure 10. Verify node information by Hardware Perspective

Whichever method you used, you will find that two new nodes, node 25 and node 26, are created on your SP system.

If the system does not recognize the nodes, perform the following:

- Check the LEDs on the node supervisor cards. If there is trouble, try resetting the node supervisor cards by issuing the `hmreinit` command:

```
# hmreinit
```

- If the node supervisor cards have no problem, but the problem persists, using the `SDR_config` command may give you some useful hints. To call it in verbose and debug mode, issue the `-v` and `-d` flags:

```
# SDR_config -v -d
```

Attention

The `SDR_config` command is to be used by the PSSP components. Use of the `SDR_config` command without the `-d` flag can cause corruption of system configuration data.

- Check that the `hardmon` daemon is working correctly. Stop and start it.

```
# stopsrc -s hardmon
# startsrc -s hardmon
```

Step 4: Update the state of the supervisor microcode

To ensure that you have the latest level of microcode required by the nodes, issue the `spsvmgr` command:

```
# spsvmgr -G -r status all

spsvmgr: Frame Slot Supervisor Media Installed Required
           |  | State Versions Version Action
-----|---|-----|-----|-----|-----
          |  | Active | u_10.1c.0709 | u_10.1c.070c | None
          |  |       | u_10.1c.070c |
          |  |-----|-----|-----|-----
          |  | Active | u_10.3a.0614 | u_10.3a.0615 | None
          |  |       | u_10.3a.0615 |
          |  |-----|-----|-----|-----
          |  | Active | u_80.19.060b | u_80.19.060b | None
          |  |-----|-----|-----|-----
          |  | Active | u_10.3c.0709 | u_10.3c.070c | None
          |  |       | u_10.3c.070c |
          |  |-----|-----|-----|-----
          |  | Active | u_10.3e.0704 | u_10.3e.0708 | None
          |  |       | u_10.3e.0706 |
          |  |       | u_10.3e.0708 |
          |  |-----|-----|-----|-----
          |  | Active | u_10.3e.0704 | u_10.3e.0708 | None
          |  |       | u_10.3e.0706 |
          |  |       | u_10.3e.0708 |
          |  |-----|-----|-----|-----
          |  | Active | u_80.19.060b | u_80.19.060b | None
          |  |-----|-----|-----|-----

#
```

The output shows the status in report form of all of your SP frames, SP nodes, and SP switches.

If you need to update the microcode of the node supervisor of node 25, issue the `spsvrmgr` command:

```
# spsvrmgr -G -u 2:9
```

For more details about supervisor microcode, refer to 2.2, “Supervisor microcode” on page 97.

Step 5: Host name resolution

Add the host name and IP address for the nodes’ SP Ethernet interface and SP Switch interface to the `/etc/hosts` file. The following is an excerpt from the `/etc/hosts` file:

```
# CWS
192.168.3.130 sp3en0.msc.itso.ibm.com sp3en0
# SP Ethernet for New Nodes
192.168.3.25 sp3n25.msc.itso.ibm.com sp3n25
192.168.3.26 sp3n26.msc.itso.ibm.com sp3n26
# SP Switch for New Nodes
192.168.13.25 sp3sw25.msc.itso.ibm.com sp3sw25
192.168.13.26 sp3sw26.msc.itso.ibm.com sp3sw26
```

If your SP system uses Domain Name System (DNS) for host name resolution, add the information about the nodes correctly.

Attention

- The `/etc/hosts` file is not managed by file collection. After modifying the file on the CWS, distribute it to all the nodes manually.
- The same operation is required for the `/etc/netsvc.conf` file.

Step 6: Enter the required node information

To add IP address-related information to the Node SDR class, use the `smitty sp_eth_dialog` fast path:

```

                                SP Ethernet Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                      #
Start Slot                       [9]                      #
Node Count                       [2]                      #

OR

Node Group                       []                       +

OR

Node List                        []

* Starting Node's en0 Hostname or IP Address [192.168.3.25]
* Netmask                         [255.255.255.0]
* Default Route Hostname or IP Address [192.168.3.130]
Ethernet Adapter Type             bnc                      +
Duplex                            half                      +
Ethernet Speed                    10                      +
Skip IP Addresses for Unused Slots? no                      +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit         Enter=Do

```

Alternatively, issue the `spethernt` command:

```

# spethernt 2 9 2 192.168.3.25 255.255.255.0 192.168.3.130
#

```

Step 7: Verify the SP Ethernet address is added

To verify that the SP Ethernet address is added, issue the `splstdata` command:

```

# splstdata -a 2 9 2
                List IAN Database Information

node#  adapt          netaddr          netmask          hostname type t/r r
ate
enet_rate duplex          other_addr
-----
---
 25   en0            192.168.3.25     255.255.255.0   sp3n25.msc.itso.  bnc
NA
      10   half          ""
 26   en0            192.168.3.26     255.255.255.0   sp3n26.msc.itso.  bnc
NA
      10   half          ""
#

```

Step 8: Acquire the hardware Ethernet address

To acquire the hardware Ethernet address, issue the `smitty hrdwrad_dialog` fast path:

```

                Get Hardware Ethernet Addresses

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                      #
Start Slot                       [9]                      #
Node Count                       [2]
OR
Node Group                       []                      +
OR
Node List                       []

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit            Enter=Do

```

Alternatively, issue the `sphrdwrad` command to perform the same operation:


```
# sphrdwrad 2 9 2
Acquiring hardware Ethernet address for node 25
Acquiring hardware Ethernet address for node 26
Hardware ethernet address for node 25 is 0004AC4947E9
Hardware ethernet address for node 26 is 10005AFA07DF
#
```

Step 9: Verify the hardware Ethernet address is acquired

To verify that a hardware Ethernet address is in place on the SDR, issue the `splstdata` command:

```
# splstdata -b 2 9 2
List Node Boot/Install Information

node#      hostname  hdw_enet_addr  svr      response  install_disk
  last_install_image  last_install_time  next_install_image  lppsource_name
  pssp_ver          selected_vg
-----
  25 sp3n25.msc.itso.  0004AC4947E9    0    install  hdisk0
      initial          initial          default  default
      PSSP-3.1          rootvg
  26 sp3n26.msc.itso.  10005AFA07DF    0    install  hdisk0
      initial          initial          default  default
      PSSP-3.1          rootvg

#
```

Step 10: Configure the SP Switch adapter

To configure SP Switch adapters to the nodes, issue the `smitty add_adapt_dialog` fast path:

Additional Adapter Database Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

```

                                     [Entry Fields]
Start Frame                          [2]                #
Start Slot                            [9]                #
Node Count                            [2]                #

OR

Node Group                            []                 +

OR

Node List                             []

* Adapter Name                        [css0]
* Starting Node's IP Address or Hostname [192.168.13.25]
* Netmask                              [255.255.255.0]
Additional IP Addresses                 []
Ethernet Adapter Type                  +
Duplex                                 +
Ethernet Speed                         +
Token Ring Data Rate                   +
Skip IP Addresses for Unused Slots?    no                 +
Enable ARP for the css0 Adapter?       yes                +
Use Switch Node Numbers for css0 IP Addresses? no            +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

Alternatively, issue the `spadaptrs` command:

```
# spadaptrs -n no 2 9 2 css0 192.168.13.25 255.255.255.0
#
```

Step 11: Verify the SP Switch address is added

To verify that the SP Switch address is added, issue the `splstdata` command:

```

# splstdata -a 2 9 2
List LAN Database Information

node#  adapt          netaddr          netmask          hostname  type  t/r r
ate
enet_rate duplex          other_addr
-----
---
25  css0      192.168.13.25   255.255.255.0   sp3sw25.msc.itso  NA
NA
   NA      NA      ""
26  css0      192.168.13.26   255.255.255.0   sp3sw26.msc.itso  NA
NA
   NA      NA      ""
25  en0       192.168.3.25    255.255.255.0   sp3n25.msc.itso.  bnc
NA
   10 half      ""
26  en0       192.168.3.26    255.255.255.0   sp3n26.msc.itso.  bnc
NA
   10 half      ""
#

```

Step 12: Annotate and store an SP Switch topology file

Annotate an SP Switch topology file and store it in the SDR, then update the SP Switch topology file's connection labels with their correct physical locations. Using an annotated file makes debugging the SP Switch easier because the SP Switch diagnostics information is based on physical locations.

To do this, issue the `smitty annotator fast path`:

```

Topology File Annotator

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* Input Topology File Name           [/etc/SP/expected.top.2> /
* Output Topology File Name          [/etc/SP/expected.top.a> /
* Save Output File to SDR            [yes] +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do

```

Alternatively, issue the `Eannotator` command with the `-O yes` flag to store the SP Switch topology file in the SDR:

```

# Eannotator -F /etc/SP/expected.top.2nsb.0isb.0 \
> -f /etc/SP/expected.top.annotated -O yes
#

```

Then, to initialize the SP Switch clock inputs for all SP Switches in the SP system, issue the `Eclock` command:

```

# Eclock -d
#

```

Step 13: Verify that the SP Switch information is added

To verify that the SP Switch information is stored in the SDR correctly, issue the `splstdata` command:

```

# splstdata -s
List Node Switch Information

switch switch switch switch switch
node# initial_hostname node# protocol number chip chip_port
-----
 1 sp3n01                0      IP      1      5      3
 5 sp3n05                4      IP      1      5      1
 6 sp3n06                5      IP      1      5      0
 7 sp3n07                6      IP      1      6      2
 8 sp3n08                7      IP      1      6      3
 9 sp3n09                8      IP      1      4      3
10 sp3n10                9      IP      1      4      2
11 sp3n11               10      IP      1      7      0
12 sp3n12               11      IP      1      7      1
13 sp3n13               12      IP      1      4      1
14 sp3n14               13      IP      1      4      0
15 sp3n15               14      IP      1      7      2
25 sp3n25.msc.itso.    24      IP      2      4      3
26 sp3n26.msc.itso.    25      IP      2      4      2

switch frame slot switch_partition switch clock switch
number number number number type input level
-----
 1      1      17                1    129    0
 2      2      17                1    129    3

switch_part topology primary arp switch_node
number filename name enabled nos._used
-----
 1 expected.top.an sp3n01.msc.itso. yes no
#

```

Step 14: Specify the `lppsource_name`

If your SP system does not use the default `lppsource_name`, you need to specify that name, for example, `aix432`.

To change the `lppsource_name`, issue the `smitty changevg_dialog fast path`:

Change Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]	
Start Frame	[2]	#
Start Slot	[9]	#
Node Count	[2]	#
OR		
Node List	[]	
Volume Group Name	[rootvg]	
Physical Volume List	[]	
Number of Copies of Volume Group	1	+
Set Quorum on the Node		+
Boot/Install Server Node	[]	#
Network Install Image Name	[]	
LPP Source Name	[aix432]	
PSSP Code Version	PSSP-3.1	+

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

Alternatively, issue the `spchvgobj` command:

```
# spchvgobj -r rootvg -c 1 -v aix432 -p PSSP-3.1 2 9 2
spchvgobj: Successfully changed the Node and Volume_Group objects for node number 25,
volume group rootvg.
spchvgobj: Successfully changed the Node and Volume_Group objects for node number 26,
volume group rootvg.
spchvgobj: The total number of changes successfully completed is 2.
spchvgobj: The total number of changes which were not successfully completed is 0.
#
```

Step 15: Verify `lppsource_name` changed

To check the `lppsource_name`, issue the `splstdata` command:

```

# splstdata -v 2 9 2
                List Volume Group Information

node# name          boot_server quorum copies  code_version lppsource_name
      last_install_image  last_install_time  last_bootdisk
      pv_list
-----
 25 rootvg          0          true    1          PSSP-3.1 aix432
      initial
      hdisk0
 26 rootvg          0          true    1          PSSP-3.1 aix432
      initial
      hdisk0
#

```

Step 16: Configure initial host name for the nodes

If you want to use the short host name as the initial host name for the nodes, you need this step. The default is long host name. This step changes the initial host name information in the SDR. It is used during customization to set up the host name on each node.

To change the initial host name for the nodes, issue the `smitty hostname_dialog fast` path:

```

                Hostname Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                #
Start Slot                       [9]                #
Node Count                       [2]                #

OR

Node Group                       []                  +

OR

Node List                         []

Adapter Name used for Initial Hostname  en0                +
Use Short or Long Hostnames           short              +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit            Enter=Do

```

Alternatively, issue the `sphostnam` command:

```
# sphostnam -a en0 -f short 2 9 2
#
```

Step 17: Verify initial host name for the nodes are changed

To verify that the initial host names for the nodes are changed, issue the `splstdata` command:

```
# splstdata -n 2 9 2
List Node Configuration Information

node# frame# slot# slots initial_hostname reliable_hostname dcehostname
      default_route processor_type processors_installed description
-----
  25     2     9     1  sp3n25             sp3n25.msc.itso.  ""
      192.168.3.130      MP                4 332_MHz_SMP_Thin
  26     2    10     1  sp3n26             sp3n26.msc.itso.  ""
      192.168.3.130      MP                4 332_MHz_SMP_Thin
#
```

Step 18: Issue the `spbootins` command

To install PSSP to the nodes, set their mode to install. Do this by issuing the `smitty server_dialog fast path`:


```

                                Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                      #
Start Slot                       [9]                      #
Node Count                       [2]                      #

OR

Node List                        []

Response from Server to bootp Request  install          +
Volume Group Name                 []
Run setup_server?                  yes                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell    F10=Exit         Enter=Do

```

Alternatively, issue the `spbootins` command:

```
# spbootins -r install 2 9 2
```

Step 19: Network boot the nodes

Issue the `nodecond` command to network boot the nodes:

```
# nodecond 2 9
```

And:

```
# nodecond 2 10
```

For discussion on how to monitor the installation process, refer to Chapter 4, “Node installation” on page 153.

Step 20: Verify node installation

To check the `hostResponds` of nodes 25 and 26, issue the `spmon` command:

```

# sponon -d -G
1. Checking server process
   Process 25670 has accumulated 2 minutes and 44 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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
   2     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   high    on  yes  yes  normal  no  LCDs 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   thin    on  yes  yes  normal  no  LEDs are blank  no
   8     8   thin    on  yes  yes  normal  no  LEDs are blank  no
   9     9   thin    on  yes  yes  normal  no  LEDs are blank  no
  10    10  thin    on  yes  yes  normal  no  LEDs are blank  no
  11    11  thin    on  yes  yes  normal  no  LEDs are blank  no
  12    12  thin    on  yes  yes  normal  no  LEDs are blank  no
  13    13  thin    on  yes  yes  normal  no  LEDs are blank  no
  14    14  thin    on  yes  yes  normal  no  LEDs are blank  no
  15    15  wide    on  yes  yes  normal  no  LEDs are blank  no

----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   9     25  thin    on  yes no  N/A  no  LCDs are blank  no
  10    26  thin    on  yes no  N/A  no  LCDs are blank  no
#

```

Step 21: Start the SP Switch

The nodes are fenced; so, to start the SP Switch, issue the `Estart` command:

```
# Estart
```

Alternatively, use the `Eunfence` command:

```
# Eunfence 25 26
```

Step 22: Verify that the SP Switch is running

To check the switchResponds on nodes 25 and 26, issue the `spmon` command:

```
# spmon -d -G
1. Checking server process
   Process 25670 has accumulated 2 minutes and 56 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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
  2     yes      yes   on      3   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   high  on yes yes normal no LCDs 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   thin  on yes yes normal no LEDs are blank no
  8     8   thin  on yes yes normal no LEDs are blank no
  9     9   thin  on yes yes normal no LEDs are blank no
 10    10   thin  on yes yes normal no LEDs are blank no
 11    11   thin  on yes yes normal no LEDs are blank no
 12    12   thin  on yes yes normal no LEDs are blank no
 13    13   thin  on yes yes normal no LEDs are blank no
 14    14   thin  on yes yes normal no LEDs are blank no
 15    15   wide  on yes yes normal no LEDs are blank no
----- Frame 2 -----
Frame Node Node Host/Switch Key Env Front Panel LCD/LED is
Slot Number Type Power Responds Switch Fail LCD/LED Flashing
-----
  9     25  thin  on yes yes N/A no LCDs are blank no
 10    26  thin  on yes yes N/A no LCDs are blank no
#
```

1.2.3 Adding SP-attached servers

This section uses the assumption that you have one SP frame with an SP Switch. You are going to add one new RS/6000 Enterprise Server as an SP-attached server.

Figure 11 illustrates how you are going to add the new SP-attached server to your existing SP system.

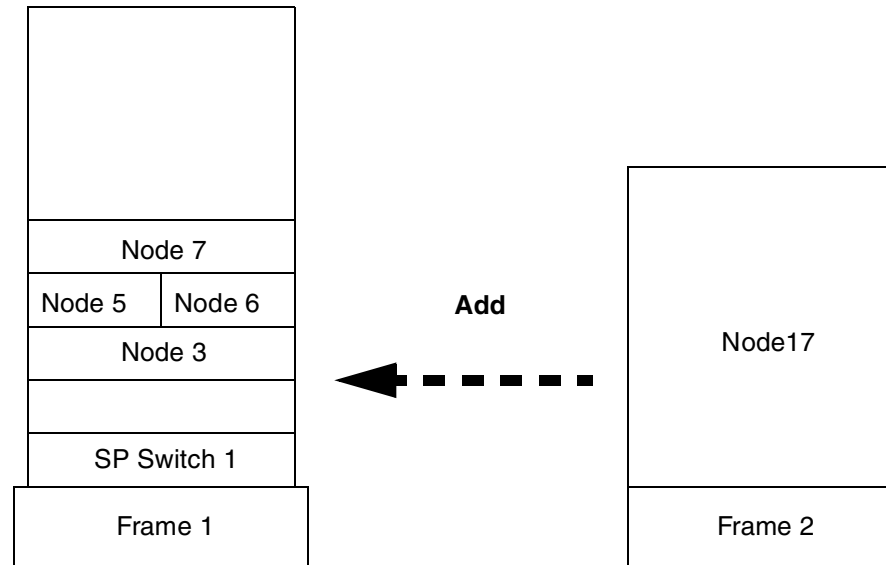


Figure 11. Adding a new SP-attached server

To add the SP-attached server, use the following steps:

Step 1: Archive the SDR

Before adding the SP-attached server to your SP system, you should back up the SDR by issuing the `SDRArchive` command:

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2: Connect the SP-attached server to CWS

Your IBM Customer Engineer performs this step.

Step 3: Configure RS-232 control line

Each frame in your SP system requires a serial port on the CWS configured to accommodate the RS-232 control line. In the case of an SP-attached server, it requires two serial ports: One for Service and Manufacturing Interface (SAMI) connection and the other for s1term connection.

To configure an RS-232 control line, issue the `smitty maktty` fast path. After selecting **tty rs232 Asynchronous Terminal**, then selecting the parent adapter for the serial port you configure, you will see the following SMIT menu:

```

                                Add a TTY

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
TTY type                             tty
TTY interface                         rs232
Description                           Asynchronous Terminal
Parent adapter                         sa2
* PORT number                          [0] +
Enable LOGIN                           disable +
BAUD rate                               [9600] +
PARITY                                  [none] +
BITS per character                       [8] +
Number of STOP BITS                      [1] +
TIME before advancing to next port setting [0] +#
TERMINAL type                            [dumb]
FLOW CONTROL to be used                  [xon] +
[MORE...29]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

Enter the required information, or issue the `mkdev` command to perform the same operation:

```
# mkdev -c tty -t tty -s rs232 -p sa2 -w 0
```

You need to create one more serial port. Issue the `smitty maktty` fast path or the `mkdev` command:

```
# mkdev -c tty -t tty -s rs232 -p sa2 -w 2
```

Step 4: Enter the SP-attached server information

To enter the SP-attached server information to the SDR, issue the `smitty nonsp_frame_dialog fast path:`

```
Non-SP Frame Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* Start Frame                        [2]                #
* Frame Count                        [1]                #
* Starting Frame tty port            [/dev/tty1]
* Starting Switch Port Number        [1]                #
s1 tty port                          [/dev/tty2]
* Frame Hardware Protocol            [SAMI]
Re-initialize the System Data Repository  yes                +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell        F10=Exit            Enter=Do
```

Make sure you input two tty ports. The `/dev/tty1` tty port is used for the SAMI connection and the `/dev/tty2` tty port is used for the s1term connection.

In this example you are going to associate switch port number 1 with the SP-attached server because this slot is not used by any other nodes in your SP system.

Alternatively, issue the `spframe` command to perform the same operation:

```
# spframe -n 1 -s /dev/tty2 -p SAMI -r yes 2 1 /dev/tty1
0513-044 The stop of the splogd Subsystem was completed successfully.
0513-059 The splogd Subsystem has been started. Subsystem PID is 11410.
#
```

Attention

You can install an SP-attached server to the switchless SP system. However, you still need to provide Starting Switch Port Number as if it uses an SP switch. For more information, refer to 1.1.4, “The switch port numbering scheme” on page 6.

Step 5: Test the RS-232 control line

Issue the `spmon_ctest` command to test the RS-232 control line. The output should look like the following:

```
# spmon_ctest
spmon_ctest: Start spmon configuration verification test
spmon_ctest: Verification Succeeded
#
```

If there are problems reported, check the RS-232 control line and look at the `/var/adm/SPlogs/spmon_ctest.log` log file. If the information given in the log file is not sufficient, try tracing the `spmon_ctest` command. This command is a shell script; so, the `set -x` command will not harm anything and may give you a clue.

Step 6: Verify frame/node information

To verify that the SP System Monitor can detect the SP-attached server correctly, issue the `spmon` command:

```

# spmon -G -d
1. Checking server process
   Process 24518 has accumulated 0 minutes and 1 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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 N/A
   2    yes      no     N/A    N/A       N/A N/A N/A N/A

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
-----
   3    3    wide    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  no   normal  no  LEDs are blank  no
----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   1    17   extrn   on   no  notcfg N/A    N/A  F00D0000  N/A
                                     LCD2 is blank
#

```

Alternatively, issue the `splstdata` command for both frame and node information:


```

# splstdata -f
List Frame Database Information

frame#          tty          s1_tty          frame_type hardware_protocol
-----
   1      /dev/tty0          ""          switch          SP
   2      /dev/tty1      /dev/tty2          ""          SAMI
# splstdata -n
List Node Configuration Information

node# frame# slot# slots initial_hostname reliable_hostname dcehostname
      default_route processor_type processors_installed description
-----
   3     1     3     2 f01n03.itsc.aust f01n03.itsc.aust ""
      9.3.187.202          UP          1 77_MHz_PWR2_Wide-2
   5     1     5     1 f01n05.itsc.aust f01n05.itsc.aust ""
      9.3.187.202          UP          1 66_MHz_PWR2_Thin-2
   6     1     6     1 f01n06.itsc.aust f01n06.itsc.aust ""
      9.3.187.202          UP          1 66_MHz_PWR2_Thin-2
   7     1     7     2 f01n07.itsc.aust f01n07.itsc.aust ""
      9.3.187.202          UP          1 77_MHz_PWR2_Wide-2
  17     2     1     1 ""          ""          ""
      ""          MP          1 ""
#

```

You can also use the Hardware Perspective shown in Figure 12 on page 44. To learn how to use the Hardware Perspective, refer to 9.2, “Using Hardware Perspective” on page 233.

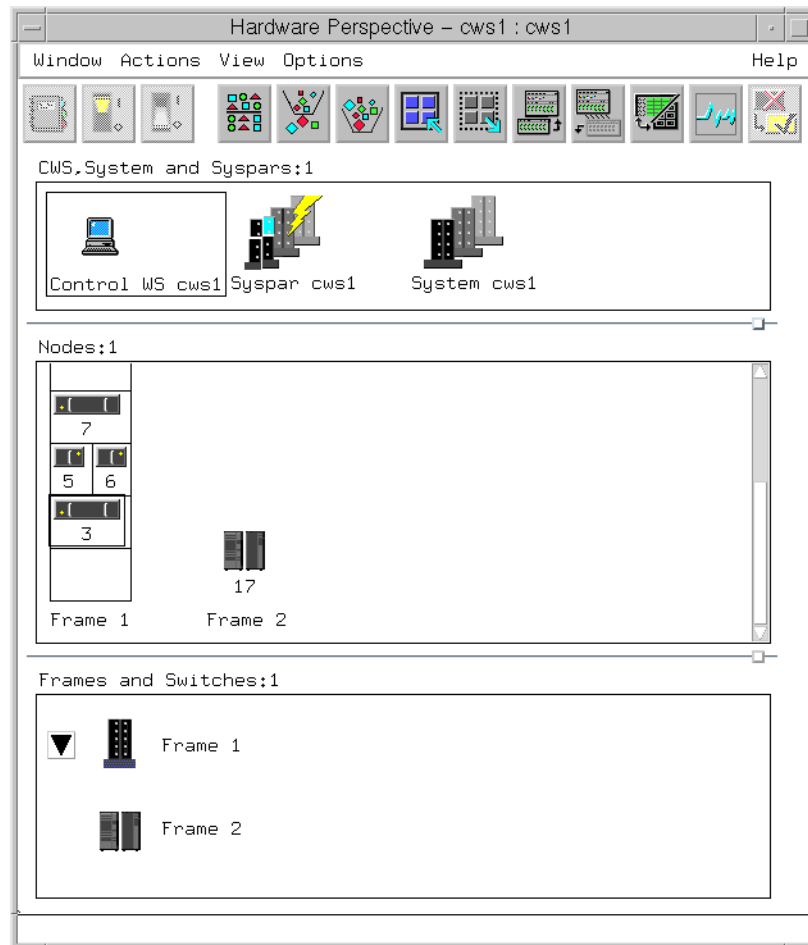


Figure 12. Verify SP-attached server information by Hardware Perspective

If your SP system does not recognize the SP-attached server, check the following items:

- Check that the s70d daemon is working correctly. You can try to stop and start it:

```
# hmcmds runpost 2:0
```

This stops the s70d daemon and notifies the SP System Monitor that it has stopped. The SP System Monitor then restarts the daemon.

Do not stop the s70d daemon with the `kill` command. The daemon will not be restarted by the SP System Monitor. In this case, you must stop and restart the SP System Monitor in order to restart the s70d daemon.

- Using the `SDR_config` command may give you some useful hints. To call it in verbose and debug mode, issue the `-v` and `-d` flags:

```
# SDR_config -v -d
```

Attention

The `SDR_config` command is to be used by the PSSP components. Use of the `SDR_config` command without the `-d` flag can cause corruption of system configuration data.

- Check that the hardmon demon is working correctly. You can try to stop and start it:

```
# stopsrc -s hardmon
# startsrc -s hardmon
```

Step 7: Host name resolution

Add the host name and IP address for the SP-attached server's SP Ethernet interface and SP Switch interface to the `/etc/hosts` file. The following is an excerpt from the `/etc/hosts` file:

```
# CWS
9.3.187.202 cws1.itsc.austin.ibm.com cws1
# SP Ethernet for SP-attached server
9.3.187.219 f01n17.itsc.austin.ibm.com f01n17
# SP Switch for SP-attached server
9.3.187.247 f01n17s.itsc.austin.ibm.com f01n17s
```

If your SP system uses Domain Name System (DNS) for host name resolution, add the information about the nodes correctly.

Attention

- The `/etc/hosts` file is not managed by file collection. After modifying the file on the CWS, you need to distribute it to all the nodes manually.
- The same operation is required for `/etc/netsvc.conf` file.

Step 8: Enter the required node information

To add IP address-related information to the SDR, issue the `smitty sp_eth_dialog` fast path:

```
SP Ethernet Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                     [Entry Fields]
Start Frame                               [2]                #
Start Slot                                [1]                #
Node Count                                 [1]                #

OR

Node Group                                 []                 +

OR

Node List                                  []

* Starting Node's en0 Hostname or IP Address [9.3.187.219]
* Netmask                                   [255.255.255.224]
* Default Route Hostname or IP Address      [9.3.187.202]
Ethernet Adapter Type                       bnc                +
Duplex                                       half               +
Ethernet Speed                              10                +
Skip IP Addresses for Unused Slots?         no                 +
[BOTTOM]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

Alternatively, issue the `spethernt` command:

```
# spethernt 2 1 1 9.3.187.219 255.255.255.224 9.3.187.202
#
```

Step 9: Verify the SP Ethernet address is added

To verify that the SP Ethernet address is added to the SDR, issue the `splstdata` command:

```

# splstdata -a 2 1 1
                List LAN Database Information

node#  adapt          netaddr          netmask          hostname type t/r r
ate
enet_rate duplex          other_addr
-----
---
  17   en0           9.3.187.219      255.255.255.224 f01n17.itsc.aust  bnc
NA
          10  half          ""
#

```

Step 10: Acquire the hardware Ethernet address

To acquire the hardware Ethernet address, issue the `smit` `hrdwrad_dialog` fast path:

```

                Get Hardware Ethernet Addresses

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                #
Start Slot                       [1]                #
Node Count                       [1]
OR
Node Group                       []                +
OR
Node List                        []

F1=Help          F2=Refresh      F3=Cancel      F4=List
F5=Reset         F6=Command     F7=Edit       F8=Image
F9=Shell        F10=Exit       Enter=Do

```

Alternatively, issue the `sphrdwrad` command to perform the same operation:

```
# sphrdwrad 2 1 1
Acquiring hardware Ethernet address for node 17
Hardware ethernet address for node 17 is 02070123F8E1
#
```

Step 11: Verify that the hardware Ethernet address is acquired

To verify that the hardware Ethernet address is added to the SDR, issue the `splstdata` command:

```
# splstdata -b 2 1 1
List Node Boot/Install Information

node#      hostname  hdw_enet_addr  svr  response  install_disk
last_install_image  last_install_time  next_install_image  lppsource_name
pssp_ver          selected_vg
-----
17 f01n17.itsc.aust  02070123F8E1  0    install  hdisk0
      initial          initial          default  default
      PSSP-3.1          rootvg
```

Step 12: Configure the SP Switch adapter

To configure the SP Switch adapter to the SP-attached server, issue the `smitty add_adapt_dialog fast` path:

Additional Adapter Database Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

```
[TOP]                                     [Entry Fields]
Start Frame                             [2]                #
Start Slot                               [1]                #
Node Count                               [1]                #

OR

Node Group                               []                 +

OR

Node List                                []

* Adapter Name                           [css0]
* Starting Node's IP Address or Hostname  [9.3.187.247]
* Netmask                                 [255.255.255.224]
Additional IP Addresses                   []
Ethernet Adapter Type                    +
Duplex                                    +
Ethernet Speed                            +
Token Ring Data Rate                     +
Skip IP Addresses for Unused Slots?      no                 +
Enable ARP for the css0 Adapter?         yes                +
Use Switch Node Numbers for css0 IP Addresses? no             +
[BOTTOM]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Alternatively, issue the `spadaptrs` command:

```
# spadaptrs -n no 2 1 1 css0 9.3.187.247 255.255.255.224
#
```

Step 13: Verify the SP Switch address is added

To verify that the SP Switch address is added to the SDR, issue the `splstdata` command:

```

# splstdata -a 2 1 1
                List IAN Database Information

node#  adapt          netaddr          netmask          hostname type t/r r
ate
enet_rate duplex          other_addr
-----
---
 17  css0          9.3.187.247    255.255.255.224  f01n17s.itsc.aus  NA
NA
      NA    NA    ""
 17  en0          9.3.187.219    255.255.255.224  f01n17.itsc.aust  bnc
NA
      10  full    ""
#

```

Step 14: Annotate and store an SP Switch topology file

Annotate an SP Switch topology file and store it in the SDR. You need to update the SP Switch topology file's connection labels with their correct physical locations. Using annotated file makes the debug of SP Switch easier because the SP Switch diagnostics information is based on physical locations.

To do this, issue the `smitty annotator fast path`:

```

                                Topology File Annotator

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Input Topology File Name          [ /etc/SP/expected.top.1 > /
Output Topology File Name         [ /etc/SP/expected.top.a > /
Save Output File to SDR           [ yes ] +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit            Enter=Do

```


Alternatively, issue the `Eannotator` command with the `-O yes` flag to store the SP Switch topology file in the SDR:

```
# Eannotator -F /etc/SP/expected.top.lnsb.0isb.0 \  
> -f /etc/SP/expected.top.annotated -O yes  
#
```

Step 15: Verify that the SP Switch information is added

To verify that the SP Switch information is stored in the SDR correctly, issue the `splstdata` command:

```
# splstdata -s  
List Node Switch Information  
  
switch switch switch switch switch  
node# initial_hostname node# protocol number chip chip_port  
-----  
3 f01n03.itsc.aust 2 IP 1 6 0  
5 f01n05.itsc.aust 4 IP 1 5 1  
6 f01n06.itsc.aust 5 IP 1 5 0  
7 f01n07.itsc.aust 6 IP 1 6 2  
17 f01n17.itsc.aust 1 IP 1 5 2  
  
switch frame slot switch_partition switch clock switch  
number number number number type input level  
-----  
1 1 17 1 129 0  
  
switch_part topology primary arp switch_node  
number filename name enabled nos._used  
-----  
1 expected.top.an f01n03.itsc.aust yes no  
#
```

Step 16: Specify the `lppsource_name`

If your SP system does not use the default `lppsource_name`, you need to specify that name, for example, `aix432`.

To change the `lppsource_name`, issue the `smitty changevg_dialog` fast path:

Change Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]	
Start Frame	[2]	#
Start Slot	[1]	#
Node Count	[1]	#
OR		
Node List	[]	
Volume Group Name	[rootvg]	
Physical Volume List	[]	
Number of Copies of Volume Group	1	+
Set Quorum on the Node		+
Boot/Install Server Node	[]	#
Network Install Image Name	[]	
LPP Source Name	[aix432]	
PSSP Code Version	PSSP-3.1	+

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

Alternatively, issue the `spchvgobj` command:

```
# spchvgobj -r rootvg -c 1 -v aix432 -p PSSP-3.1 2 1 1
spchvgobj: Successfully changed the Node and Volume_Group objects for node number 17,
volume group rootvg.
spchvgobj: The total number of changes successfully completed is 1.
spchvgobj: The total number of changes which were not successfully completed is 0.
#
```

Step 17: Verify `lppsource_name` changed

To check that the `lppsource_name` is changed, issue the `splstdata` command:

```

# splstdata -v 2 1 1
                List Volume Group Information

node# name          boot_server quorum copies  code_version lppsource_name
      last_install_image  last_install_time  last_bootdisk
      pv_list
-----
  17 rootvg          0          true    1          PSSP-3.1 aix432
      initial
      hdisk0
#

```

Step 18: Configure initial host names

If you want to use the short host names as the initial host name for the SP-attached server, you need this step. The default is long host names. This step changes the host name information in the SDR. It is used during customization to set up the host name for the SP-attached server.

To change the initial host names, issue the `smitty hostname_dialog` fast path:

```

                                Hostname Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                      #
Start Slot                       [1]                      #
Node Count                       [1]                      #

OR

Node Group                       []                          +

OR

Node List                        []

Adapter Name used for Initial Hostname  en0                      +
Use Short or Long Hostnames           short                    +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit           Enter=Do

```

Alternatively, issue the `sphostnam` command:

```
# sphostnam -a en0 -f short 2 1 1
#
```

Step 19: Verify initial host name is changed

To verify that the initial host name for the SP-attached server is changed, issue the `splstdata` command:

```
# splstdata -n 2 1 1
          List Node Configuration Information

node# frame# slot# slots  initial_hostname  reliable_hostname  dcehostname
      default_route  processor_type  processors_installed  description
-----
   17     2     1     1  f01n17          f01n17.itsc.aust  ""
                        9.3.187.219      MP                4 7017-S70
#
```

Step 20: Issue the `spbootins` command

To install PSSP to the SP-attached server, you need to set its mode to install. Do this by issuing the `smitty server_dialog fast path`:

```

                                Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
Start Frame                          [2]                #
Start Slot                            [1]                #
Node Count                            [1]                #

OR

Node List                             []

Response from Server to bootp Request  install            +
Volume Group Name                      []
Run setup_server?                      yes                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Alternatively, issue the `spbootins` command:

```
# spbootins -r install 2 1 1
```

Step 21: Network boot the SP-attached server

Issue the `nodecond` command to network boot the SP-attached server:

```
# nodecond 2 1
```

For discussion on how to monitor the installation process, refer to Chapter 4, “Node installation” on page 153.

Step 22: Verify the hostResponds

To check the hostResponds for the SP-attached server, issue the `spmon` command:

```

# spmon -d -G
1. Checking server process
   Process 20388 has accumulated 3 minutes and 56 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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 N/A
   2    yes      no     N/A    N/A      N/A N/A N/A N/A

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
-----
   3    3    wide    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
----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   1    17   extrn   on  yes no    N/A    N/A  080C          N/A
                                     LCD2 is blank
#

```

Step 23: Start the SP Switch

The SP-attached server is fenced; so, to start the SP Switch, issue the `Estart` command:

```
# Estart
```

Alternatively, use the `Eunfence` command:

```
# Eunfence -G 17
```

Step 24: Verify the SP Switch is working

To check the `switchResponds` for the SP-attached server, issue the `spmon` command:

```

# sponon -d -G
1. Checking server process
   Process 20388 has accumulated 3 minutes and 56 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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 N/A
   2    yes      no     N/A    N/A       N/A N/A N/A N/A

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
-----
   3    3    wide    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
----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   1    17   extrn   on  yes  yes  N/A    N/A  080C          N/A
                                     LCD2 is blank
#

```

1.3 Deleting frames/nodes/SP-attached servers

There are many cases in which you might need to delete some of the SP hardware components from your SP system. After giving you some general information about deleting frames or nodes, this section covers the following three topics:

- Deleting SP frames
- Deleting SP nodes
- Deleting SP-attached servers

In 1.3.2, “Deleting SP frames” on page 59, we use an example of two SP frames in your SP system. This example was configured in 1.2.2, “Adding SP nodes” on page 18. In this section, you will delete the second SP frame.

In 1.3.3, “Deleting SP nodes” on page 64, we use an example of two SP frames, and the second SP frame has two SP nodes in your SP system. This example was configured in 1.2.2, “Adding SP nodes” on page 18. In this section, you will delete one of the two SP nodes in the second SP frame.

In 1.3.4, “Deleting SP-attached servers” on page 69, we use an example of one SP frame and one SP-attached server in your SP system. This example was configured in 1.2.3, “Adding SP-attached servers” on page 38. In this section, you will delete the SP-attached server.

For more information, refer to Chapter 5, “Reconfiguring the RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347.

1.3.1 Before deleting frames/nodes

When you delete frames or nodes from your SP system, you should first evaluate how the change will affect the remainder of your SP system. Consider the work load and applications currently running on the hardware to be deleted and plan how to transfer the work load and applications to equivalent SP resources.

If the node you are deleting is a server node, you have to reassign the server functions to another node that is not being deleted. The following list gives some examples of servers you may need to pay attention to:

- Boot/install server (BIS)
- Kerberos authentication server
- SP Switch primary/primary backup node

To reassign SP Switch primary/primary backup node, refer to 13.1.4, “Starting the SP switch” on page 345.

- Network Time Protocol (NTP) server

To reassign NTP server, refer to 14.3.3, “Changing NTP time server” on page 414.

- Home directory server

To reassign home directory server, refer to 14.4.2, “Changing home directory server and path” on page 420.

1.3.2 Deleting SP frames

This section uses the assumption that you have two SP frames in your SP system, which is the current configuration of your SP system right after you completed the operation described in 1.2.2, “Adding SP nodes” on page 18. You are going to delete the second SP frame.

When you delete an SP frame, all of the SP nodes contained in the SP frame will be deleted automatically.

If you are deleting an SP frame that contains an SP Switch with SP-attached server connected to it, you must first delete the SP-attached server. Refer to 1.3.4, “Deleting SP-attached servers” on page 69, to delete an SP-attached server.

Figure 13 illustrates how you are going to delete the second SP frame from your SP system:

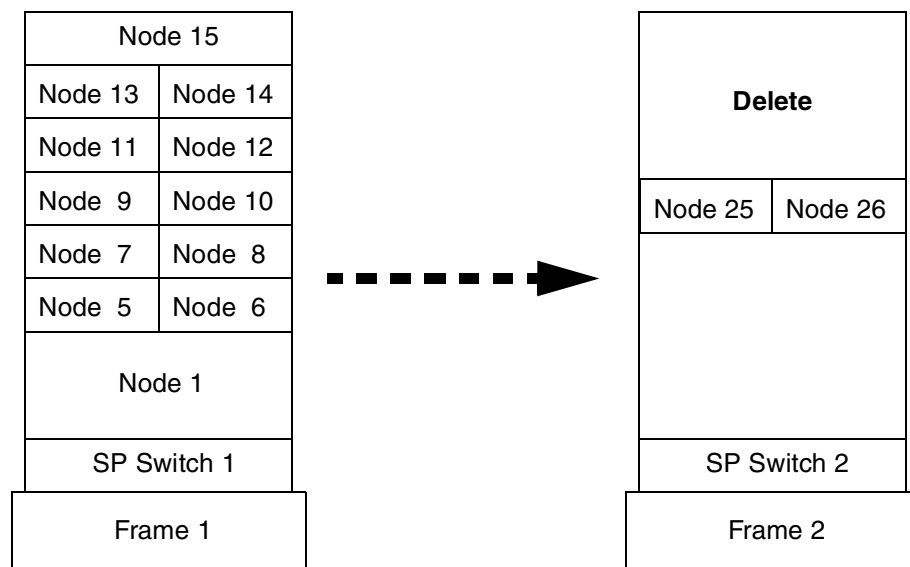


Figure 13. Deleting SP frame

To delete the SP frame, use the following steps:

Step 1: Archive the SDR

Before deleting the frame from your SP system, you should back up the SDR by issuing the `SDRArchive` command.

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note the location and the name of the file created after you issue this command. In this example, the backup.99064.1459.mybackup file is created in the /spdata/sys1/sdr/archives directory.

Step 2: Shut down the nodes

Shut down the nodes contained in the frame you are deleting from your SP system.

Step 3: Disconnect the frame

Your IBM Customer Engineer performs this step.

Step 4: Delete frame information from the SDR

To delete the frame information from the SDR, issue the `smitty delete_frame_dialog` fast path:

```

Delete Frame Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* Start Frame                        [2]                #
* Frame Count                        [1]                #

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell     F10=Exit      Enter=Do

```

Alternatively, issue the `spdelfram` command to perform the same operation:

```
# spdelfram 2 1
warning: 0042-140 m_mmmac: unable to remove the /etc/niminfo file on
"sp3n25"

delnimclient: Node 25 (sp3n25) unconfigured as a NIM client on
boot/install server node 0 (sp3en0).
warning: 0042-140 m_mmmac: unable to remove the /etc/niminfo file on
"sp3n26"

delnimclient: Node 26 (sp3n26) unconfigured as a NIM client on
boot/install server node 0 (sp3en0).
0513-044 The stop of the splogd Subsystem was completed successfully.
0513-044 The stop of the hardmon Subsystem was completed successfully.
0513-059 The hardmon Subsystem has been started. Subsystem PID is 42910.
0513-059 The splogd Subsystem has been started. Subsystem PID is 25682.
#
```

This step deletes the frame information and all of the node information contained in the SP frame.

Step 5: Verify that the frame is deleted

To check that the frame information is deleted, issue the `spmon` command:

```

# spmon -d -G
1. Checking server process
   Process 42910 has accumulated 0 minutes and 0 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   1 frame(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 Slot  Node Number  Node Type  Host/Switch Power  Responds  Key Switch  Env Fail  Front Panel LCD/LED  LCD/LED is Flashing
-----
   1      1      high  on yes yes  normal no  LCDs 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      thin  on yes yes  normal no  LEDs are blank no
   8      8      thin  on yes yes  normal no  LEDs are blank no
   9      9      thin  on yes yes  normal no  LEDs are blank no
  10     10     thin  on yes yes  normal no  LEDs are blank no
  11     11     thin  on yes yes  normal no  LEDs are blank no
  12     12     thin  on yes yes  normal no  LEDs are blank no
  13     13     thin  on yes yes  normal no  LEDs are blank no
  14     14     thin  on yes yes  normal no  LEDs are blank no
  15     15     wide  on yes yes  normal no  LEDs are blank no
#

```

In case of a problem, try tracing the `/usr/lpp/ssp/bin/delfram` script. It gives you the opportunity to see what is going wrong.

Step 6: Delete hardware addresses

If you use the `/etc/bootptab.info` file to keep the hardware Ethernet addresses of your nodes, delete the information about the nodes you deleted. This will avoid possible trouble when you install other nodes using the same node numbers used for the nodes you deleted.

Step 7: Refresh the system partition sensitive subsystems

To propagate the change in the configuration to the subsystems, issue the `syspar_ctrl` command:

```
# syspar_ctrl -G -r
0513-095 The request for subsystem refresh was completed successfully.
0513-095 The request for subsystem refresh was completed successfully.
#
```

Step 8: Annotate and store an SP Switch topology file

Annotate an SP Switch topology file and store it to the SDR. Update the SP Switch topology file's connection labels with their correct physical locations. Using an annotated file makes the debug of SP Switch easier because the SP Switch diagnostics information is based on physical locations.

Issue the `smitty` annotator fast path:

```
Topology File Annotator

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Input Topology File Name         [/etc/SP/expected.top.1> /
Output Topology File Name       [/etc/SP/expected.top.a> /
Save Output File to SDR         [yes] +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

Alternatively, issue the `Eannotator` command with the `-O yes` flag to store the SP Switch topology file in the SDR:

```
# Eannotator -F /etc/SP/expected.top.insb.0isb.0 \
> -f /etc/SP/expected.top.annotated -O yes
#
```

Step 9: Verify that the SP Switch information is deleted

To verify that the SP Switch information is deleted from the SDR correctly, issue the `splstdata` command:

```
# splstdata -s
List Node Switch Information

node# initial_hostname switch node# protocol number chip chip_port
-----
 1 sp3n01              0      IP      1      5      3
 5 sp3n05              4      IP      1      5      1
 6 sp3n06              5      IP      1      5      0
 7 sp3n07              6      IP      1      6      2
 8 sp3n08              7      IP      1      6      3
 9 sp3n09              8      IP      1      4      3
10 sp3n10              9      IP      1      4      2
11 sp3n11             10      IP      1      7      0
12 sp3n12             11      IP      1      7      1
13 sp3n13             12      IP      1      4      1
14 sp3n14             13      IP      1      4      0
15 sp3n15             14      IP      1      7      2

switch number frame number slot number switch partition number switch type clock input switch level
-----
      1      1      17      1      129      0

switch_part topology primary arp switch_node
number filename name enabled nos._used
-----
      1 expected.top.an sp3n01.msc.itso. yes no

#
```

1.3.3 Deleting SP nodes

This section uses the assumption that you have two SP frames in your SP system, which is the configuration of your SP system right after you completed the operation described in 1.2.2, “Adding SP nodes” on page 18. You are going to delete one SP node from the second SP frame.

Figure 14 on page 65 illustrates how you are going to delete the SP node from the second SP frame.

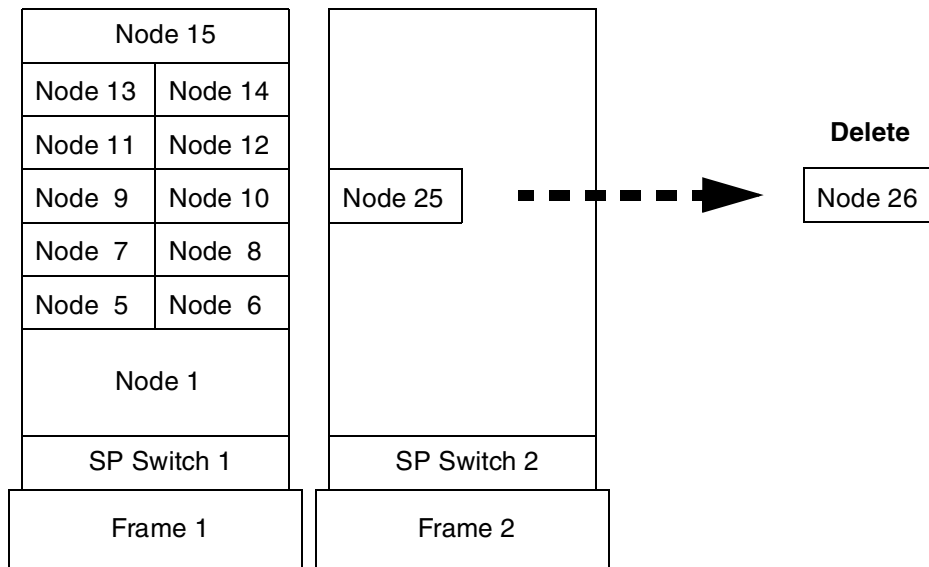


Figure 14. Deleting SP node

To delete the SP node, use the following steps:

Step 1: Archive the SDR

Before deleting the node from your SP system, you should back up the SDR by issuing the `SDRArchive` command.

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2: Shut down the node

Shut down the node that you are deleting from your SP system.

Step 3: Disconnect the node

Your IBM Customer Engineer performs this step.

Step 4: Delete node information from the SDR

To delete the node information from the SDR, issue the `smitty delete_node_dialog` fast path:

```

Delete Node Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                #
Start Slot                       [10]               #
Node Count                       [1]                #

OR

Node Group                       []                 +

OR

Node List                       []

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit      F8=Image
F9=Shell    F10=Exit       Enter=Do

```

Alternatively, issue the `spdelnode` command to perform the same operation:

```
# spdelnode 2 10 1
warning: 0042-140 m_mmac: unable to remove the /etc/niminfo file on
"sp3n26"

delnimclient: Node 26 (sp3n26) unconfigured as a NIM client on
boot/install server node 0 (sp3en0).
#
```

Step 5: Verify that the node is deleted

To verify that the node information is deleted from the SDR, issue the `spmon` command:


```

# spon -d -G
1. Checking server process
   Process 25670 has accumulated 3 minutes and 31 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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
   2     yes      yes    on     3        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    high    on  yes  yes  normal  no  LCDs 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    thin    on  yes  yes  normal  no  LEDs are blank  no
   8     8    thin    on  yes  yes  normal  no  LEDs are blank  no
   9     9    thin    on  yes  yes  normal  no  LEDs are blank  no
  10    10    thin    on  yes  yes  normal  no  LEDs are blank  no
  11    11    thin    on  yes  yes  normal  no  LEDs are blank  no
  12    12    thin    on  yes  yes  normal  no  LEDs are blank  no
  13    13    thin    on  yes  yes  normal  no  LEDs are blank  no
  14    14    thin    on  yes  yes  normal  no  LEDs are blank  no
  15    15    wide    on  yes  yes  normal  no  LEDs are blank  no

----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   9     25    thin    on  yes  yes  N/A     no  LCDs are blank  no
#

```

In case of a problem, try tracing the `/usr/lpp/ssp/bin/delnode` script. It gives you the opportunity to see what is going wrong.

Step 6: Delete hardware addresses

If you use the `/etc/bootptab.info` file to keep the hardware Ethernet addresses of your nodes, delete the information about the node you deleted. This will

avoid possible trouble when you install other nodes to the slot used for the node you deleted.

Step 7: Refresh the system partition sensitive subsystems

To propagate the change in the configuration to the subsystems, issue the `syspar_ctrl` command:

```
# syspar_ctrl -G -r
0513-095 The request for subsystem refresh was completed successfully.
0513-095 The request for subsystem refresh was completed successfully.
#
```

Step 8: Annotate and store an SP Switch topology file

Annotate an SP Switch topology file and store it to the SDR. Update the SP Switch topology file's connection labels with their correct physical locations. Using an annotated file makes the debug of SP Switch easier because the SP Switch diagnostics information is based on physical locations.

To do this, issue the `smitty` annotator fast path:

```
Topology File Annotator

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

Input Topology File Name      [Entry Fields]
Output Topology File Name    [/etc/SP/expected.top.1> /
Save Output File to SDR     [/etc/SP/expected.top.a> /
                             [yes] +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell    F10=Exit      Enter=Do
```

Alternatively, issue the `Eannotator` command with the `-O yes` flag to store the SP Switch topology file in the SDR:

```
# Eannotator -F /etc/SP/expected.top.2nsb.0isb.0 \
> -f /etc/SP/expected.top.annotated -O yes
#
```

Step 9: Verify that the SP Switch information is deleted

To verify that the SP Switch information is deleted from the SDR correctly, issue the `splstdata` command:

```
# splstdata -s
List Node Switch Information

switch switch switch switch switch
node# initial_hostname node# protocol number chip chip_port
-----
1 sp3n01 0 IP 1 5 3
5 sp3n05 4 IP 1 5 1
6 sp3n06 5 IP 1 5 0
7 sp3n07 6 IP 1 6 2
8 sp3n08 7 IP 1 6 3
9 sp3n09 8 IP 1 4 3
10 sp3n10 9 IP 1 4 2
11 sp3n11 10 IP 1 7 0
12 sp3n12 11 IP 1 7 1
13 sp3n13 12 IP 1 4 1
14 sp3n14 13 IP 1 4 0
15 sp3n15 14 IP 1 7 2
25 sp3n25.msc.itso. 24 IP 2 4 3

switch frame slot switch_partition switch clock switch
number number number number type input level
-----
1 1 17 1 129 0
2 2 17 1 129 3

switch_part topology primary arp switch_node
number filename name enabled nos._used
-----
1 expected.top.an sp3n01.msc.itso. yes no

#
```

1.3.4 Deleting SP-attached servers

This section uses the assumption that your SP system has one SP frame with a SP Switch and one SP-attached server connected to the SP frame. You are going to delete the SP-attached server from your SP system.

Figure 15 on page 70 illustrates how you are going to delete the SP-attached server from your SP system.

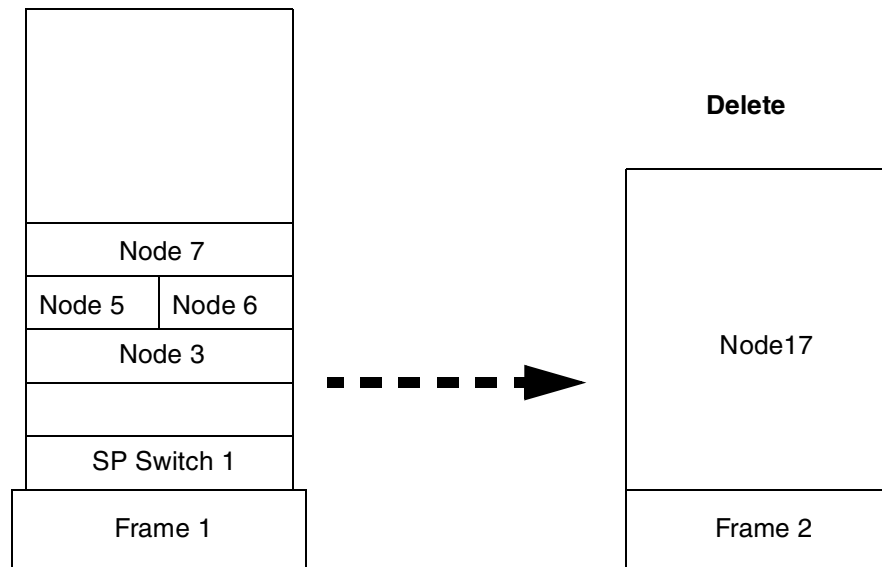


Figure 15. Deleting the SP-attached server

To delete the SP-attached server, use the following steps:

Step 1: Archive the SDR

Before deleting the SP-attached server from your SP system, you should back up the SDR by issuing the `SDRArchive` command.

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2: Shut down the SP-attached server

Shut down the SP-attached server.

Step 3: Disconnect the SP-attached server

Your IBM Customer Engineer performs this step.

Step 4: Delete the frame information from the SDR

To delete the frame information from the SDR, issue the `smitty delete_frame_dialog` fast path:

```
Delete Frame Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* Start Frame                        [2]                #
* Frame Count                        [1]                #

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Alternatively, issue the `spdelfram` command to perform the same operation:

```
# spdelfram 2 1
warning: 0042-140 m_mmac: unable to remove the /etc/niminfo file on
"f01n17"

delnimclient: Node 17 (f01n17) unconfigured as a NIM client on
boot/install server node 0 (cws1).
0513-044 The stop of the splogd Subsystem was completed successfully.
0513-059 The splogd Subsystem has been started. Subsystem PID is 11290.
#
```

This step deletes both the frame and the node information belonging to the SP-attached server.

Step 5: Verify that the SP-attached server is deleted

To check that the SP-attached server information is deleted, issue the `spmon` command:

```

# spmon -G -d
1. Checking server process
   Process 20390 has accumulated 0 minutes and 0 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   1 frame(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 N/A

5. Checking nodes
----- Frame 1 -----
Frame Slot  Node  Node  Host/Switch  Key  Env  Front Panel  LCD/LED is
      Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   3     3  wide  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
#

```

In case of a problem, try tracing the `/usr/lpp/ssp/bin/delfram` script. It gives you the opportunity to see what is going wrong.

Step 6: Delete hardware addresses

If you use the `/etc/bootptab.info` file to keep the hardware Ethernet addresses of your nodes, delete the information about the node you deleted. This will avoid possible trouble when you install other node to the same node number used for the SP-attached server you deleted.

Step 7: Refresh the system partition sensitive subsystems

To propagate the change in the configuration to the subsystems, issue the `syspar_ctrl` command:

```

# syspar_ctrl -G -r
0513-095 The request for subsystem refresh was completed successfully.
0513-095 The request for subsystem refresh was completed successfully.
#

```

Step 8: Annotate and store an SP Switch topology file

Annotate an SP Switch topology file and store it to the SDR. Update the SP Switch topology file's connection labels with their correct physical locations. Using an annotated file makes the debug of SP Switch easier because the SP Switch diagnostics information is based on physical locations.

To do this, issue the `smitty annotator fast` path:

```
Topology File Annotator

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Input Topology File Name         [ /etc/SP/expected.top.1> /
Output Topology File Name       [ /etc/SP/expected.top.a> /
Save Output File to SDR         [yes] +

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command   F7=Edit     F8=Image
F9=Shell     F10=Exit     Enter=Do
```

Alternatively, use the `Eannotator` command with the `-O yes` flag to store the SP Switch topology file in the SDR:

```
# Eannotator -F /etc/SP/expected.top.1nsb.0isb.0 \
> -f /etc/SP/expected.top.annotated -O yes
#
```

Step 9: Verify that the SP Switch information is deleted

To verify that the SP Switch information is deleted from the SDR correctly, issue the `splstdata` command:

```

# splstdata -s
List Node Switch Information

switch switch switch switch switch
node# initial_hostname node# protocol number chip chip_port
-----
 3 f01n03.itsc.aust      2      IP      1      6      0
 5 f01n05.itsc.aust      4      IP      1      5      1
 6 f01n06.itsc.aust      5      IP      1      5      0
 7 f01n07.itsc.aust      6      IP      1      6      2

switch frame slot switch_partition switch clock switch
number number number number type input level
-----
 1      1      17              1      129      0

switch_part topology primary arp switch_node
number filename name enabled nos._used
-----
 1 expected.top.an f01n03.itsc.aust yes no

#

```

1.4 Attaching/Detaching SP-attached servers

Attaching/Detaching SP-attached servers is similar to deleting them from the current location and then adding them to a different location. The difference is that with attaching/detaching you can preserve software resources in the moved SP-attached servers. In the case of attaching/detaching, you can use software resources, such as AIX or other applications, with only minor modification. This section presents the following two topics:

- Attaching SP-attached servers
- Detaching SP-attached servers

In section 1.4.1, “Attaching SP-attached servers” on page 75, the scenario is similar to the one described in 1.2.3, “Adding SP-attached servers” on page 38. The only difference is the SP-attached server is currently used as a stand-alone RS/6000 Enterprise Server, and you want to avoid to reinstalling AIX and some applications to the SP-attached server from scratch.

In section 1.4.2, “Detaching SP-attached servers” on page 81, assuming the scenario is similar to the one described in 1.3.4, “Deleting SP-attached servers” on page 69. The only difference is that you want to use the SP-attached server as a stand-alone RS/6000 Enterprise Server. In other words, you want to keep AIX and applications already installed on the

SP-attached server. Therefore, you need to delete PSSP-related software or configuration after deleting the SP-attached server from your SP system.

For more information, refer to Chapter 5, “Reconfiguring the RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347.

1.4.1 Attaching SP-attached servers

This section uses a scenario similar to the one described in 1.2.3, “Adding SP-attached servers” on page 38. The only difference is that for this example the SP-attached server is currently used as a stand-alone RS/6000 Enterprise Server. Therefore, you want to avoid reinstalling AIX, PSSP, and applications to the SP-attached server.

Figure 16 illustrates how you will attach the SP-attached server to your SP system.

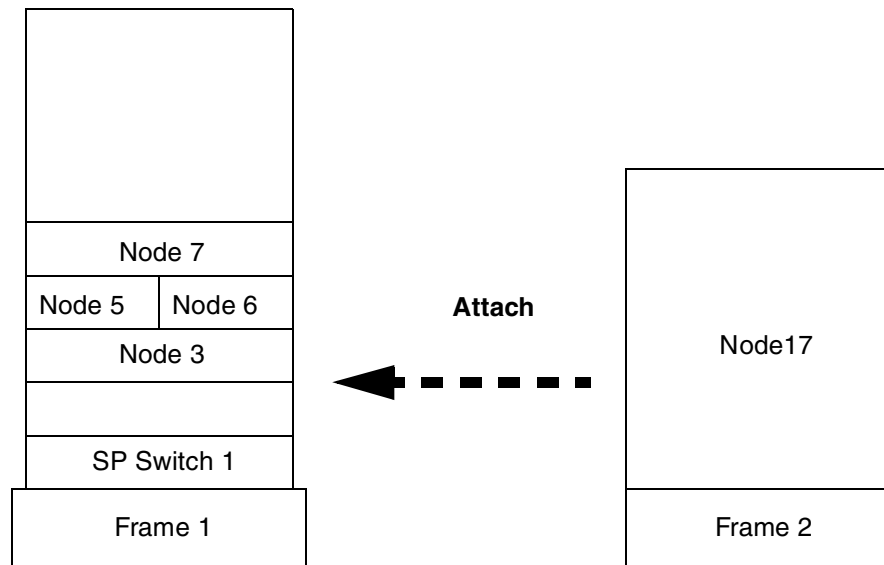


Figure 16. Attach SP-attached server

1.4.1.1 Preparation

To make the attach operation go smoothly, the following preparations are required. Assuming you can access the SP-attached server through a network other than the SP Ethernet, complete the following steps:

- Update host name resolution. Add the host name and IP address of the SP-attached server to the /etc/hosts file on the CWS. Add the host name and IP address of the CWS and all the nodes on your SP system to the /etc/hosts file on the SP-attached server. If you use DNS, update the DNS server.
- Update the /etc/bootptab.info file on the CWS so that you do not need to boot the SP-attached server just to acquire hardware Ethernet addresses. The following is the sample /etc/bootptab.info file:

```
# cat /etc/bootptab.info
17 02070123F8E1
#
```

Figure 17. The /etc/bootptab.info File

If you do not use the /etc/bootptab.info file currently, create it. To get the hardware Ethernet address, issue the `lscfg` command on the each nodes:

```
# lscfg -vl ent0
DEVICE          LOCATION      DESCRIPTION

ent0            00-03        Ethernet High-Performance LAN
                Adapter (8ef5)

Network Address.....02070123f8e1
ROS Level and ID.....0010
Displayable Message.....802.3/ETHERNET
Part Number.....071F1183
EC Level.....C26574
Device Driver Level.....00
Diagnostic Level.....00
FRU Number.....081F7913
Serial Number.....00054792
Manufacturer.....204491

#
```

- On the SP-attached server, upgrade AIX to the appropriate level.

1.4.1.2 The attachment process

To attach the SP-attached server to your SP system, use the following steps:

Step 1: Attach the SP-attached server to the frame

First, attach the SP-attached server to your SP frame by following the steps that are found in “Step 1: Archive the SDR” on page 38 to “Step 19: Verify initial host name is changed” on page 54.

Since you have an updated `/etc/bootptab.info` file on the CWS, in Step 10 you will see the following message when you issue the `sphrdwrad` command:

```
# sphrdwrad 2 1 1
Acquiring hardware ethernet address for node 17 from /etc/bootptab.info
#
```

When you have completed step 19, return to this page and continue with step 2, which follows:

Step 2: Issue the `spbootins` command

You must set the SP-attached server to customize instead of install. To set the SP-attached server to customize and run the `setup_server` command, issue the `smitty server_dialog` fast path:

```

                                Boot/Install Server Information
Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [2]                #
Start Slot                       [1]                #
Node Count                       [1]                #

OR

Node List                        []

Response from Server to bootp Request  customize      +
Volume Group Name                   []
Run setup_server?                    yes             +

F1=Help          F2=Refresh      F3=Cancel      F4=List
F5=Reset         F6=Command     F7=Edit        F8=Image
F9=Shell         F10=Exit       Enter=Do
```

Alternatively, issue the `spbootins` command:

```
# spbootins -r customize 2 1 1
```

Step 3: Configure SP Ethernet

Perform this step on the SP-attached server. Configure SP Ethernet by the `smitty mktcpip` fast path:

Minimum Configuration & Startup

To Delete existing configuration data, please use Further Configuration menus

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

```

                                     [Entry Fields]
* HOSTNAME                           [f01n17]
* Internet ADDRESS (dotted decimal) [9.3.187.219]
  Network MASK (dotted decimal)     [255.255.255.224]
* Network INTERFACE                   en0
  NAMESERVER
    Internet ADDRESS (dotted decimal) [9.3.187.202]
    DOMAIN Name                       [itsc.austin.ibm.com]
  Default GATEWAY Address             [9.3.187.202]
  (dotted decimal or symbolic name)
  Your CABLE Type                     N/A
  START Now                           yes
                                     +
                                     +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Step 4: Copy the /etc/SDR_dest_info file

During customization, certain information will be read from the SDR. The customization programs must know where the SDR resides. To specify the SDR location, copy the /etc/SDR_dest_info file from the CWS. Issue the `ftp` command to copy the file from the CWS to the SP-attached server. Check the mode and ownership of the file. It should be the same as the following:

```
# ls -al /etc/SDR_dest_info
-rw-r--r--  1 root  system    86 Feb 20 21:51 /etc/SDR_dest_info
#
```

Step 5: Verify perfagent.tools file set installed

Perform this step on the SP-attached server. Ensure that the `perfagent.tools 2.2.32.x` file set is installed on your SP-attached server. To check if it is installed, issue the `ls1pp` command:

```

# lspp -L perfagent.tools
Fileset                Level  State  Description
-----
perfagent.tools        2.2.32.0  C     Local Performance Analysis &
                               Control Commands

State Codes:
A -- Applied.
B -- Broken.
C -- Committed.
O -- Obsolete. (partially migrated to newer version)
? -- Inconsistent State...Run lppchk -v.
#

```

Step 6: Mount the psslpp directory

Perform this step on the SP-attached server. Mount the /spdata/sys1/install/psslpp directory on the CWS to the SP-attached server. To do this, issue the `mount` command. Use the /mnt mount point, for example:

```
# mount cws1:/spdata/sys1/install/psslpp /mnt
```

Step 7: Install ssp.basic file set

Perform this step on the SP-attached server. Install ssp.basic file set and its prerequisites onto the SP-attached server. To do this, issue the `smitty install_latest fast path:`

```

Install and Update from LATEST Available Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* INPUT device / directory for software      /mnt/PSSP-3.1
* SOFTWARE to install                       [+ 3.1.0.0 SP System S> +
PREVIEW only? (install operation will NOT occur) no +
COMMIT software updates?                       yes +
SAVE replaced files?                          no +
AUTOMATICALLY install requisite software?     yes +
EXTEND file systems if space needed?          yes +
OVERWRITE same or newer versions?            no +
VERIFY install and check file sizes?         no +
Include corresponding LANGUAGE filesets?     yes +
DETAILED output?                             no +
Process multiple volumes?                    yes +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Alternatively, issue the `installp` command to perform the same operation:

```
# installp -acgNQqwx -d /mnt/PSSP-3.1 -f ssp.basic
```

The following file sets will be installed:

- rsct.basic.rte
- rsct.basic.sp
- rsct.basic.hacmp
- rsct.clients.rte
- rsct.clients.sp
- rsct.clients.hacmp
- ssp.perlpkg
- ssp.clients
- ssp.basic

Step 8: Unmount the pssplpp directory

Perform this step on the SP-attached server. Unmount the `/spdata/sys1/install/pssplpp` directory mounted in Step 6. To do this, issue the `umount` command:

```
# umount /mnt
```

Step 9: Issue the pssp_script script

Perform this step on the SP-attached server. To customize the SP-attached server, issue the `pssp_script` script:

```
#/usr/lpp/ssp/install/bin/pssp_script

=====
pssp_script: = Making PSSP log directories... =
=====
+ /usr/bin/mkdir -p /var/adm/SPlogs/sysman
+ 1> /dev/null 2>& 1
+ exec
+ 3> /var/adm/SPlogs/sysman/NODE.config.log.8816

=====
pssp_script: = Switching output to log file... =
=====
#
```

Step 10: Reboot the SP-attached server

Perform this step on the SP-attached server. To reboot the SP-attached server, issue the `shutdown` command:

```
# shutdown -Fr
```

Step 11: Verify SP-attached server attaching

To verify SP-attached server attaching, perform the same steps described in 1.2.3, “Adding SP-attached servers” on page 38:

- Verify node installation. To check the `hostResponds` of nodes, issue the `spmon` command (“Step 22: Verify the `hostResponds`” on page 55).
- Start the SP Switch. The nodes are fenced; so, to start the SP Switch, issue the `Estart` command (“Step 23: Start the SP Switch” on page 56).
- Verify that the SP Switch is running. To check the `switchResponds` on the nodes, issue the `spmon` command (“Step 24: Verify the SP Switch is working” on page 56).

1.4.2 Detaching SP-attached servers

This section uses a scenario similar to the one described in 1.3.4, “Deleting SP-attached servers” on page 69. The only difference is that for this example you want to use the SP-attached server as a stand-alone RS/6000 Enterprise Server. In other words, you want to keep AIX and the applications already installed in the SP-attached server. Therefore, you need to delete PSSP-related software or configurations after deleting the SP-attached server from your SP system.

Figure 18 illustrates how you will detach the SP-attached server from your SP system:

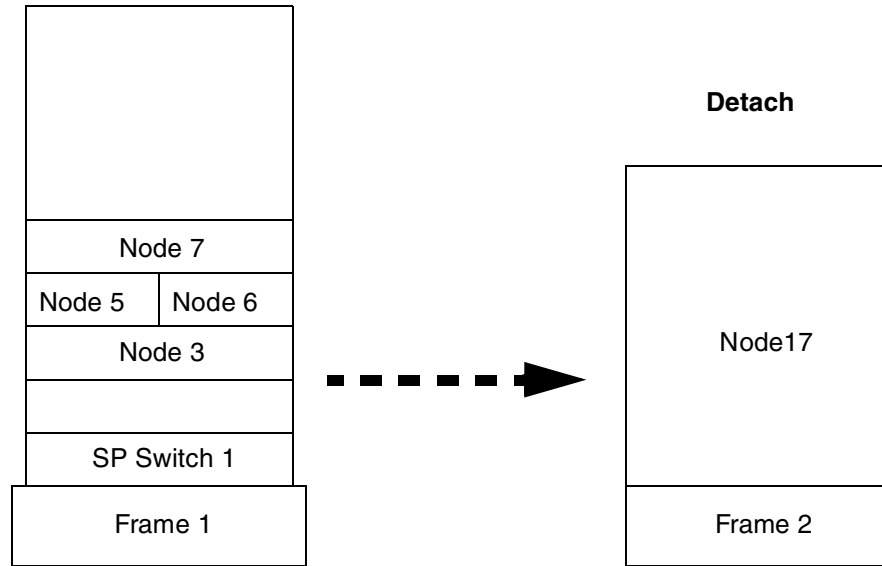


Figure 18. Detaching SP-Attached Servers

To detach the SP-attached server from your SP system, use the following steps:

Step 1: Modify the `/etc/inittab` file on the SP-attached server

To avoid starting the PSSP unique daemons on the next boot, delete the following entries from the `/etc/inittab` file on the SP-attached server:

- `start_net:2:wait:/usr/lpp/ssp/install/bin/start_net > /dev/console`
- `sp:2:wait:/etc/rc.sp > /dev/console 2>&1`
- `fsd:2:once:/usr/lpp/ssp/css/rc.switch`
- `sysctld:2:once:/usr/bin/startsrc -s sysctld`
- `st_sw_num:2:boot:/usr/lpp/ssp/bin/st_set_switch_number`
- `hats:2:once:/usr/bin/startsrc -g hats > /dev/console 2>&1`
- `hags:2:once:/usr/bin/startsrc -g hags > /dev/console 2>&1`
- `haem:2:once:/usr/bin/startsrc -g haem > /dev/console 2>&1`
- `pman:2:once:/usr/bin/startsrc -g pman >/dev/console 2>&1`
- `sp_configd:2:once:/usr/bin/startsrc -s sp_configd`
- `tty1:2:respawn:/usr/sbin/getty /dev/tty1`

Step 2: Delete the SP-attached server from your SP system

Delete the SP-attached server from your SP system. This step is the same operation as described in 1.3.4, “Deleting SP-attached servers” on page 69. Follow steps 1 through 9 described in that section; they are not reprinted here. When you have completed those steps, continue with Step 2, which follows. You perform the rest of the steps on the deleted SP-attached server, not on the SP system.

Step 3: Boot the RS/6000 Enterprise Server

Boot the RS/6000 Enterprise Server so that you can uninstall all the PSSP.

Step 4: Remove installed PSSP

Remove installed PSSP from the RS/6000 Enterprise Server. The following file sets are applied:

- perfagent.tools
- rsct.basic.hacmp
- rsct.basic.rte
- rsct.basic.sp
- rsct.clients.hacmp
- rsct.clients.rte
- rsct.clients.sp
- ssp.basic
- ssp.clients
- ssp.css
- ssp.ha_topsvcs.compat
- ssp.perlpkgx
- ssp.pman
- ssp.st
- ssp.sysctl
- ssp.sysman

To remove installed PSSP, issue the `smitty remove fast path`:

```

Remove Installed Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
SOFTWARE name                    [perfagent.tools rsct.b> +
PREVIEW only? (remove operation will NOT occur)    no      +
REMOVE dependent software?                       no      +
EXTEND file systems if space needed?              no      +
DETAILED output?                                 yes     +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Step 5: Delete directories

The following files and directories are PSSP-unique directories. You can delete them to save disk space:

- /spdata/*
- /var/adm/SPlogs/*
- /var/ha/*
- /var/sysman/*
- /usr/sbin/rsct/*
- /etc/SP/*
- /etc/ssp/*
- /etc/SDR_dest_info
- /etc/ha/*
- /etc/amd/*
- /etc/auto/*
- /etc/rc.sp
- /etc/ntp*

Step 6: Modify environment

Some PSSP configuration information still remains in AIX components. Check the following components:

cron file There are still entries for PSSP. To prevent starting unwanted commands automatically, delete them.

SMIT PSSP added its SMIT menu to the SMIT. You can delete it to save disk space.

/tftpboot/* AIX also uses this directory. You can delete PSSP-related files to save disk space.

security For security reasons, the Kerberos system should be deleted completely. The following files need to be deleted manually:

- /etc/krb.conf
- /etc/krb.realms
- /etc/krb-srvtab
- /.klogin

If you used AIX standard authentication method, delete the following file also:

- /.rhosts

Chapter 2. Software

Software is the basis for the functionality of your SP system, and it requires administration and maintenance to work properly. For an SP system, the AIX and IBM Parallel System Support Programs for AIX (PSSP) are the major pieces of software, but equally important are system firmware, microcode for devices, supervisor microcode, and program temporary fixes (PTF). These programs provide device support for your hardware devices, fix defects in code, provide error recovery, and produce new functionality. Software requires a lot of administration: Knowing what levels you are currently at, knowing how to upgrade and install new levels, and so forth. This chapter explains how to manage your software. Instructions and examples are used to make the information easier to understand.

Another important issue is software recovery. The most important part of recovering an SP system is to have a good backup. This chapter discusses how to make backups of your nodes and CWS as well as how to restore from your backups. Having a procedure to restore from a backup can save the day when you are recovering from a system crash, a failing hard disk, or accidentally removing an important file.

Finally, the SP system uses Public Domain Software (PDS) along with AIX and PSSP. If you like to understand every detail, and how everything functions together, there is source code available. This is also described in this chapter.

2.1 System firmware and microcode for devices

System firmware is required on SP nodes and CWS, and microcode for devices is required on devices to function properly. This system firmware and microcode is pre-installed on your system and on your devices and does not require any special installation procedures. Without this system firmware and microcode for devices, the SP node, CWS, and devices would not function. Occasionally, new versions of system firmware and microcode for devices are introduced for the following reasons:

- New features, connectivity, security, and resource support are included to improve system operation.
- Serviceability is also an issue, and new levels may include improved problem determination and fault isolation, which lead to accurate error codes.
- Performance enhancements in the system: Firmware may improve or resolve problems in the response times and throughput.

- Finally, new levels of system firmware may improve the user interface with easier to understand messages.

For these reasons, it is important to know your current level and be able to check this against the latest available version.

This section explains how to check what your installed version of system firmware is, what the latest available version is, and how to install it. It also explains how to check your microcode for devices on some of your devices. Note that firmware and microcode for devices, as well as their installation procedures, are different from machine to machine and device to device. The installation instructions must be followed exactly.

2.1.1 What is system firmware?

System firmware is a type of microcode for the system planar. It provides the system planar with information on how to handle and interact with devices, booting, error logs, and general behavior. System firmware does get updated, and your system might not be at the latest level. System firmware is available on symmetric multiprocessor (SMP) nodes.

2.1.2 Checking the system firmware level

System firmware requirements are dependent on the type of SP node that you have. Most SP nodes require firmware, but not all do. The following is a list of SP nodes that require firmware:

- SP-attached servers
- 332 MHz SMP wide nodes (F/C 2051)
- 332 MHz SMP thin nodes (F/C 2050)
- 200 MHz SMP high nodes (F/C 2009)

How you determine the firmware level your SP node should be at depends on your SP node type.

SP-attached servers

In the case of SP-attached servers, issue the `lscfg` command to check the system firmware level:

```
# lscfg -pv | grep alterable
ROM Level.(alterable).....19990121 (B) 19980825 (A)
ROM Level.(alterable).....19990122 (B) 19980825 (A)
#
```

The first line in the second column (before the (B)) shows the system firmware level. In this example, it is 19990121. The second line, same column shows the service processor firmware level, which, in this case, is 19990122.

332 MHz SMP wide/thin nodes

In the case of 332 MHz wide/thin nodes, issue the `lscfg` command to check the system firmware level:

```
#lscfg -pv | grep alterable
ROM Level .(alterable).....L99005
ROM Level .(non-alterable).....wc981228
ROM Level .(alterable).....wc981228
#
```

You are only concerned with the first and third line of output. The first line gives you the system firmware level, which, in this case, is 99005. On line 3, it gives you the output of the service processor firmware level, which is wc981228. The output from the second line can be ignored: It has no special meaning.

200 MHz SMP high nodes

In the case of 200 MHz SMP high nodes, you need to verify if your SP node has a 604 or 604e processor card. To check it, issue the `lscfg` command:

```
# lscfg -vl cpucard0
DEVICE          LOCATION      DESCRIPTION
cpucard0        00-0P        CPU card

EC Level.....E77287
FRU Number.....X4D
Device Specific.(MN).....IBM97N
Processor Component ID.....15010089000201890002020200070303000204
                                030002050300020603000201001600
Part Number.....93H9536
Serial Number.....156N358YZL
Size.....02
Machine Type and Model.....00A0
Device Specific.(Y0).....0000      20349   7015R40 20349
#
```

Look at the FRU Number and compare it to Table 1:

Table 1. FRU number and processor card

FRU Number	Processor Card
C1D	601
C4D	604
X4D	604e

In this example, the high node has a 604e processor card.

If your SP node has a 604 or 604e processor card, issue the `lscfg` command to determine the system firmware level:

```
#lscfg -vl ioplanar0 | grep RM
      Device Specific.(RM) .....3115G4833          09510902
#
```

In this case the system firmware level is the first four digits in the third column, or 0951.

If your SP node has FRU number C1D, in other words, your SP node has a 601 processor card, there is no system firmware available.

This will only tell you what system firmware level is currently installed on your SP node. To find out if your SP node is current, you can go to the following Web:

<http://www.rs6000.ibm.com/support/micro/>

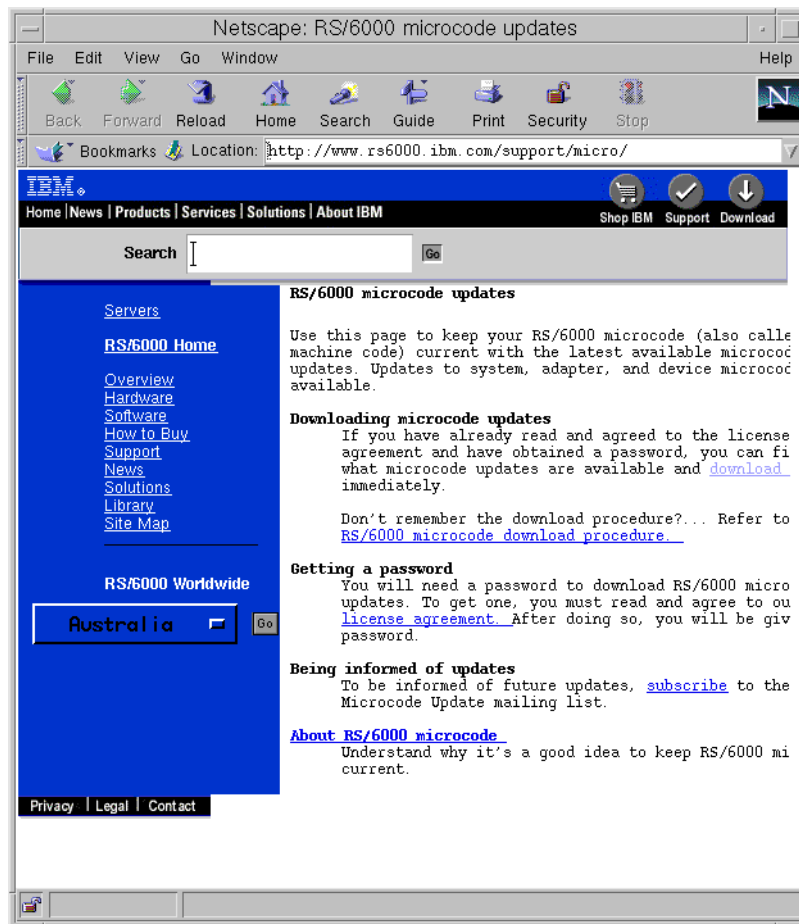


Figure 19. RS/6000 microcode updates

To view the current microcode levels and to download these microcodes, click on the download. This will bring up the Web page shown in Figure 20 on page 92:

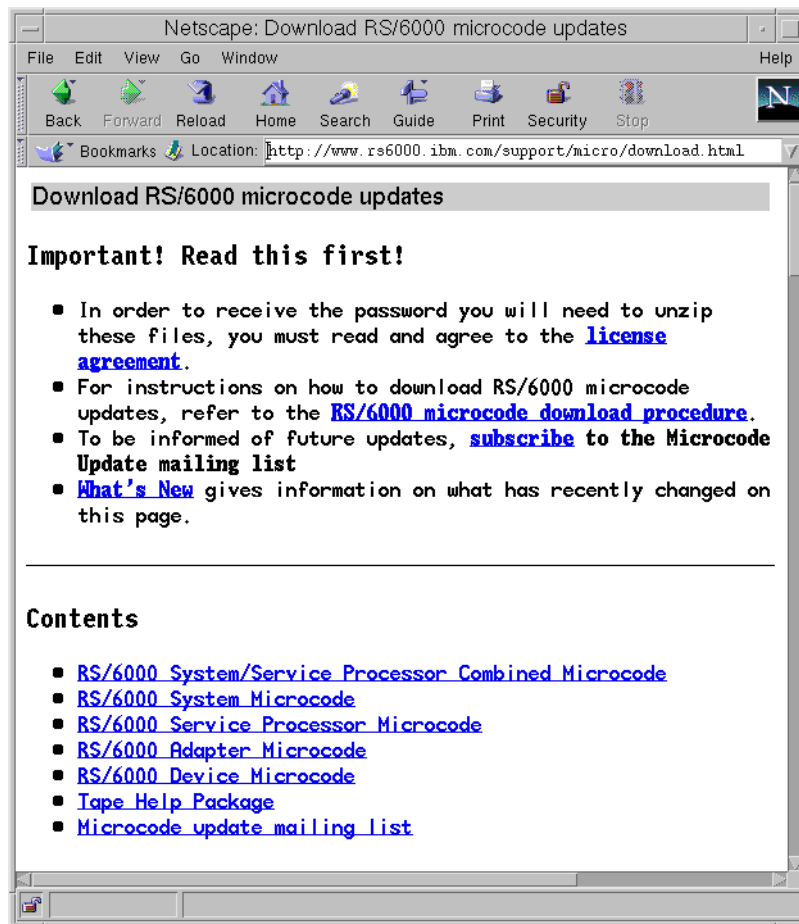


Figure 20. Download RS/6000 microcode updates

From this Web site, you can download the readme instructions as well as the microcode to install the system firmware. An alternative way to find out the current microcode levels is to call the IBM Support Center.

2.1.3 Upgrading system firmware

It is important to be at the current level of system firmware. A new version of system firmware can provide enhancements that make your SP node more reliable, efficient, and adaptable. For example, it can provide the following:

- Additional support for new hardware devices
- Diagnostic enhancements

- Improved data collection and reporting of errors
- Correction of specific bugs
- Additional error detection and isolation capabilities
- Added functionality

These are two options for upgrading your system firmware. The first is to contact your IBM Support Center and have the update done by your IBM Customer Engineer. Alternatively, you can download the instructions and the system firmware from the Web site shown in Figure 19 on page 91.

The actual steps to upgrade your system firmware vary with the machine type; so, it is important to read the instructions. Normally, they require the system firmware code to be copied onto the SP node in a specific directory. In some cases, an `install` command is run, and again, special attention must be paid to the syntax. Finally, a reboot is mandatory for the system firmware to function properly. If the installation instructions are not followed precisely, and the system firmware is not installed properly, it can cause your SP node to crash and render it useless.

2.1.4 What is the microcode for devices?

Microcode is used on devices, such as SSA disks, adapters, and tape drives. It is important to keep this microcode up to date. The IBM Support Center or the IBM Web site can keep you informed about what the latest microcode levels are.

2.1.5 Checking the microcode for devices

You should periodically check the devices in your system to ensure that they are running up-to-date versions of microcode. This section describes how to check specific devices.

Tape drives

For 7331 model 305, 7332, and 7336, issue the `lscfg` command to find out your current microcode level. The following is an example for a `rmt0` tape drive:

```

#lscfg -vl rmt0
  DEVICE          LOCATION          DESCRIPTION

  rmt0           00-00-0S-2,0    4.0 GB 4mm Tape Drive

  Manufacturer.....ARCHIVE
  Machine Type and Model.....IBM4326NP/RP !D
  Device Specific.(Z1).....4C00
  Serial Number.....          0011
  Device Specific.(LI).....0011
  Part Number.....87G4925
  FRU Number.....21H5172
  EC Level.....D48105
  Device Specific.(Z0).....0180020283000018
  Device Specific.(Z3).....L1
#

```

Your microcode level is shown in the Device Specific.(Z1) field. In this example, the microcode level is 4C00. It is also important to note the part number, the machine type and model, and in some cases the device specific(L1) field. This information is useful when you talk with the IBM Support Center.

SSA adapters and disks

Adapters and disks for SSA or RAID also have their own microcode. To find out the level of your SSA adapter microcode, ssa0, issue the `lscfg` command:

```

#lscfg -vl ssa0
  DEVICE          LOCATION          DESCRIPTION

  ssa0           00-06            SSA RAID Adapter

  Part Number.....084H9667
  Serial Number.....f6834019
  EC Level.....0000E48507
  Manufacturer.....IBM053
  ROS Level and ID.....2904
  Loadable Microcode Level...02
  Device Driver Level.....00
  Displayable Message.....SSA-ADAPTER
  Device Specific.(Z0).....DRAM=008
  Device Specific.(Z1).....CACHE=0
#

```

The microcode level is shown in the ROS Level and ID field. In this example the level is 2904.

For SSA disks, you will also want to make note of the Device Specific.(Z2) field. This field identifies the type of the SSA disk. This is important because different disk types have different microcode.

2.1.6 Upgrading the microcode for devices

To determine if the microcode level for your device is up-to-date, there are several Web sites that you can check. The first Web site for microcode for devices is the same as that for the system firmware shown in Figure 19 on page 91:

<http://www.rs6000.ibm.com/support/micro/>

From this Web site, you can download any of the following types of microcode:

- RS/6000 system/service processor combined microcode
- RS/6000 system microcode
- RS/6000 service processor microcode
- RS/6000 adapter microcode
- RS/6000 device microcode

This Web site shows you which level is current and includes step by step instructions on how to install the microcode.

The second Web site is specifically for SSA adapters and disks as shown in Figure 21 on page 96:

<http://www.hursley.ibm.com/~ssa/rs6k/>

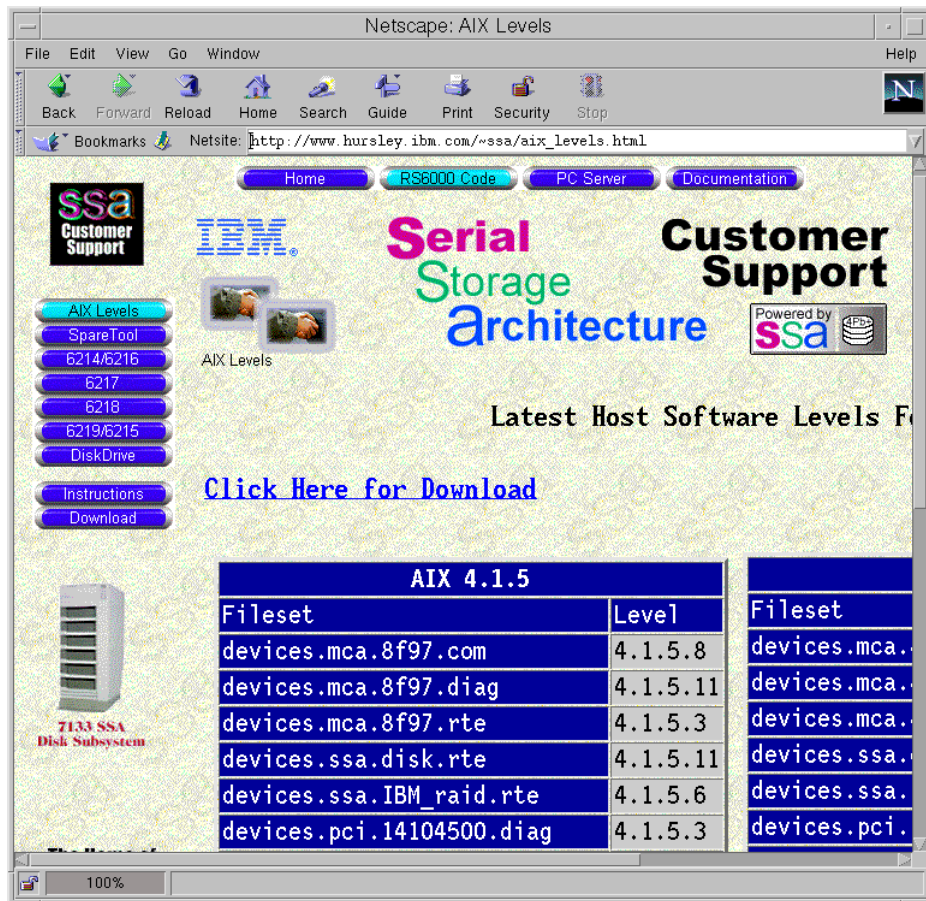


Figure 21. SSA customer support

This Web site allows you to select your SSA adapter, then it presents you with a download option. You can download the device drivers and microcode for both the adapter and SSA disks.

The microcode installation procedures vary depending on the adapter or disk. In some circumstances, the microcode is installed using the diagnostics on your system, and other times, it requires that the code be copied onto the system and then installed. Follow the instructions that come with the microcode. The alternative is to contact your IBM Support Center and arrange for an IBM Customer Engineer to come on site to install the microcode for you.

2.2 Supervisor microcode

Supervisor microcode is another piece of software that needs to be monitored and occasionally updated. Newer versions of supervisor microcode can fix defects, improve error logging, and enhance performance. Supervisor microcode can be downloaded to the frame supervisor card, the node supervisor card, and the switch supervisor card. Each of these microcodes are different because the cards to which they are downloaded have different jobs.

This section describes how to manage your supervisor microcode and gives you instructions on how to check your level and install, upgrade, and downgrade it.

2.2.1 What is a supervisor microcode?

To control a device attached to a computer, a device driver is necessary for the operating system on the computer side, and a supervisor microcode is necessary on the device side. Supervisor microcode is a small program that controls the device and that depends on instructions from the device driver. There are microcodes that you can upgrade, if you need to extend the device function. Therefore, you may need the latest version of microcodes downloaded on devices.

In the case of SP system, a CWS needs to communicate with a device. When you install an SP node with new architecture, you need to upgrade a frame supervisor microcode on the frame supervisor card so that the frame supervisor card can handle the new SP node.

2.2.2 Getting supervisor microcode

The ssp.unicode file set, which is a part of a PSSP, includes microcodes for the supervisor cards of SP system. When you install this file set, it places the microcodes in a /spdata/sys1/unicode directory. The following are the microcodes in the directory:

```
# cd /spdata/sys1/ucode
# ls
u_10.00.0709  u_10.1a.0615  u_10.36.0704  u_10.3c.070c  u_80.09.0609
u_10.00.070c  u_10.1c.0709  u_10.36.0706  u_10.3e.0704  u_80.09.060b
u_10.16.0704  u_10.1c.070c  u_10.3a.0614  u_10.3e.0706  u_80.11.060b
u_10.16.0706  u_10.1e.0704  u_10.3a.0615  u_80.01.0609  u_80.19.060b
u_10.1a.0614  u_10.1e.0706  u_10.3c.0709  u_80.01.060b
#
```

Figure 22. Supervisor microcodes in /spdata/sys1/ucode directory

Because ssp.ucode is a file set of PSSP, you can get a microcode as a PTF. To learn how to get a PTF, refer to 2.3.2, “Getting a PTF” on page 102.

2.2.3 Checking supervisor microcode levels

There are three types of supervisor microcode available for SP systems:

- Frame supervisor microcode
- Node supervisor microcode
- Switch supervisor microcode

To check if your current level of supervisor microcode is lower than what the ssp.ucode file set provides, you can issue the `smitty supervisor fast path`:

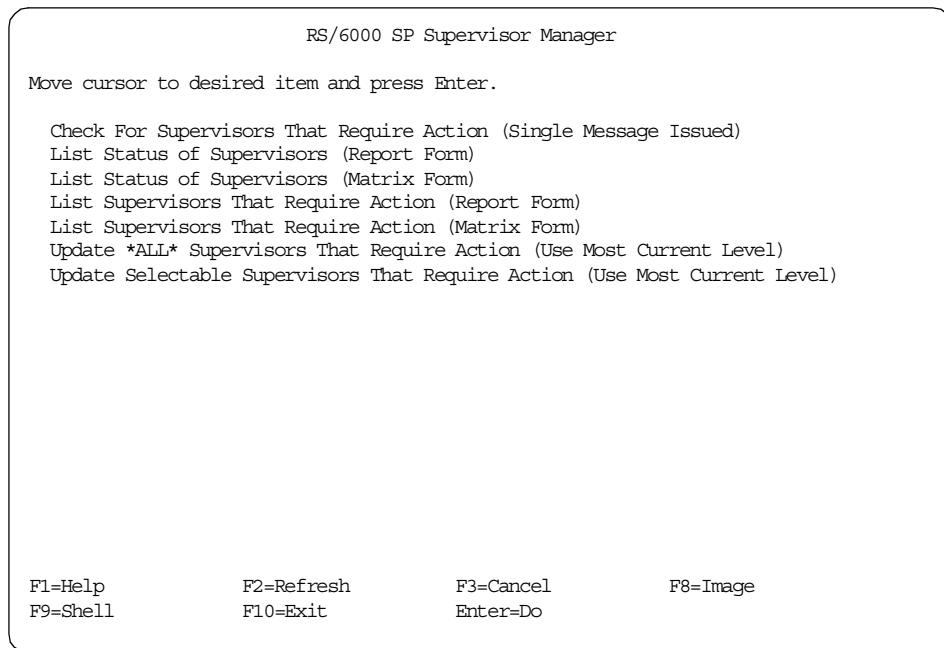


Figure 23. RS/6000 SP Supervisor Manager

To display the status of supervisor microcodes, select the **List Status of Supervisors (Report Form)**. You will see the following SMIT menu:

```

COMMAND STATUS

Command: OK          stdout: yes          stderr: no

Before command completion, additional instructions may appear below.

spsvrmgr: Frame Slot Supervisor State Media Versions Installed Version Required Action
-----
          1     0     Active      u_10.3c.0709 u_10.3c.0610 Upgrade
          |     |     |           u_10.3c.070c
          1     1     Active      u_10.3a.0614 u_10.3a.0615 None
          |     |     |           u_10.3a.0615
          9     9     Active      u_10.3e.0704 u_10.3e.0706 None
          |     |     |           u_10.3e.0706
          10    10    Active      u_10.3e.0704 u_10.3e.0706 None
          |     |     |           u_10.3e.0706
          17    17    Active      u_80.19.060b u_80.19.060b None

F1=Help          F2=Refresh      F3=Cancel      F6=Command
F8=Image        F9=Shell        F10=Exit       /=Find
n=Find Next

```

Figure 24. List status of supervisors (report form)

Frame 1, slot 0 represents a frame supervisor card. Frame 1, slot 17 represents a switch supervisor card. Other slots are node supervisor cards. In this example, a frame supervisor microcode should be upgraded.

You can issue the `spsvrmgr` command to perform the same operation:

```
# /usr/lpp/ssp/bin/spsvrmgr -G -r status all
```

Attention

In the case of old type frames, nodes, and switches, the `spsvrmgr` command does not show a status of supervisor microcodes because they don't have the ability to upgrade their supervisor microcode.

2.2.4 Upgrading supervisor microcode

To upgrade a supervisor microcode, issue the `smitty supervisor fast path`. You will see the SMIT menu as shown in Figure 23 on page 99.

To upgrade a supervisor microcode, select **Update *ALL* Supervisors That Require Action (Use Most Current Level)**. All supervisor microcodes that are marked for upgrade in Figure 24 on page 100 will be upgraded.

You can issue the `spsvrmgr` command to perform the same operation:

```
# /usr/lpp/ssp/bin/spsvrmgr -G -u all
```

When you select **Update Selectable Supervisors That Require Action (Use Most Current Level)** as shown in Figure 23 on page 99, you can specify the frame and slot number to be upgraded.

Attention

In most cases, the `-u` flag powers off the target slots for the duration of the update. It takes 5 to 10 minutes to upgrade supervisor microcodes. Do not interrupt the process, or you will lose a control line of the SP system.

2.2.5 Downgrading supervisor microcode

Supervisor microcode has backward compatibility. Generally, you will not be able to downgrade it because it is likely to cause problems. However, if the new supervisor microcode causes a problem, or downloading the new supervisor microcode fails, you might need to download a specific version of supervisor microcode. Use the following steps to do this:

Step 1: Activate basecode mode

To activate basecode mode, issue the `hmcmds` command:

```
# hmcmds -G basecode 1:0
```

This command performs a power off of the node, and switches the active frame, node, or switch supervisor microcode to basecode mode. This operation causes the active supervisor microcode to become non-active and the basecode supervisor microcode to become active. The `1:0` parameter indicates frame 1, slot 0, the frame supervisor card.

For more details about the `hmcmds` command, refer to 9.4.4, “Switching to basecode version” on page 245.

Step 2: Download specific microcode

Issue the `hmcmds` command to download a specific version of supervisor microcode:

```
# hmcmds -G -v -u /spdata/sys1/ucode/u_10.3c.0610 microcode 1:0
```

In this example, you are going to download u_10.3c.0610 to the frame supervisor card.

Be sure to note the current microcode level before you downgrade it. You may need to upgrade the microcode to the current level after downgrading it. In this case you need the file name for the current level of microcode.

For more details about the `hmcmds` command, refer to 9.4.6, “Downloading supervisor microcode” on page 247.

2.3 Program temporary fixes

Applications and operating systems are never perfect, and defects are inevitable. To resolve problems, there are Program Temporary Fixes (PTFs). PTFs contain fixes to code defects, but they sometimes also provide new device driver support, increased functionality, and other software changes. A PTF is a good thing, and by keeping your system up-to-date, you can relieve yourself from a lot of heartache. There is nothing worse than having your system crash and then discovering that if you had applied a PTF, you could have avoided the situation entirely. For this reason, it is a good idea to apply PTFs as a preventative measure.

This section explains how to download PTFs from the internet and install them. It also explains when you should apply PTFs and how to install the PTFs on your SP nodes.

2.3.1 What is a PTF?

A PTF is a temporary solution for a software defect in a current release of a licensed program product (LPP). When you encounter a software defect, you can call your IBM Support Center to report the problem. The problem is reported as an Authorized Program Analysis Report (APAR) to the third level supporters. Based on the report, they create a PTF. In general, one PTF fixes several problems; therefore, it corresponds to several APARs. A PTF is assigned an identification number that starts with the letter U, for example U461631. An APAR is assigned an identification number that starts with letters IX, for example, IX78929. If you apply a PTF, files that have defects are replaced by new ones.

2.3.2 Getting a PTF

The most common way to get a PTF is to contact your IBM Support Center, which can supply you with the PTF code. A quicker and more convenient

method to obtain PTFs is to download them from the IBM internet. You can do this using the `ftp` command or a special tool designed for this (FixDist).

Information for IBM employees only

You can check if a PTF that fixes the problem is already available. To check current PTF availability, visit the Web site:

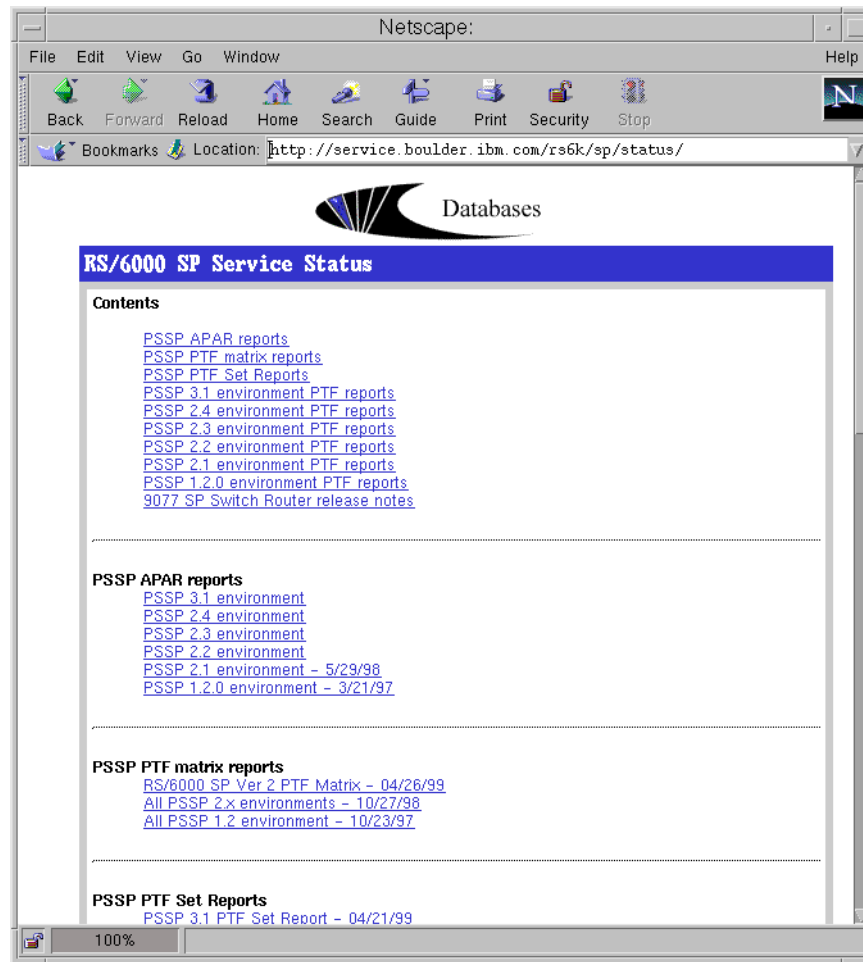
<http://w3.viewblue.ibm.com/>

In the case of a PTF for PSSP, several PTFs are grouped as a set. If you want to get the latest PTF set using the `ftp` command, perform the following steps:

Step 1: Get the file set level of the PTF set

To learn the contents of the latest PTF set, visit the following Web site:

<http://service.boulder.ibm.com/rs6k/sp/status/>



For example, you can find all PTF numbers, file set names, and file set levels that are included in the PTF set. Figure 25 on page 105 shows the PTF set 05 for PSSP 2.4.

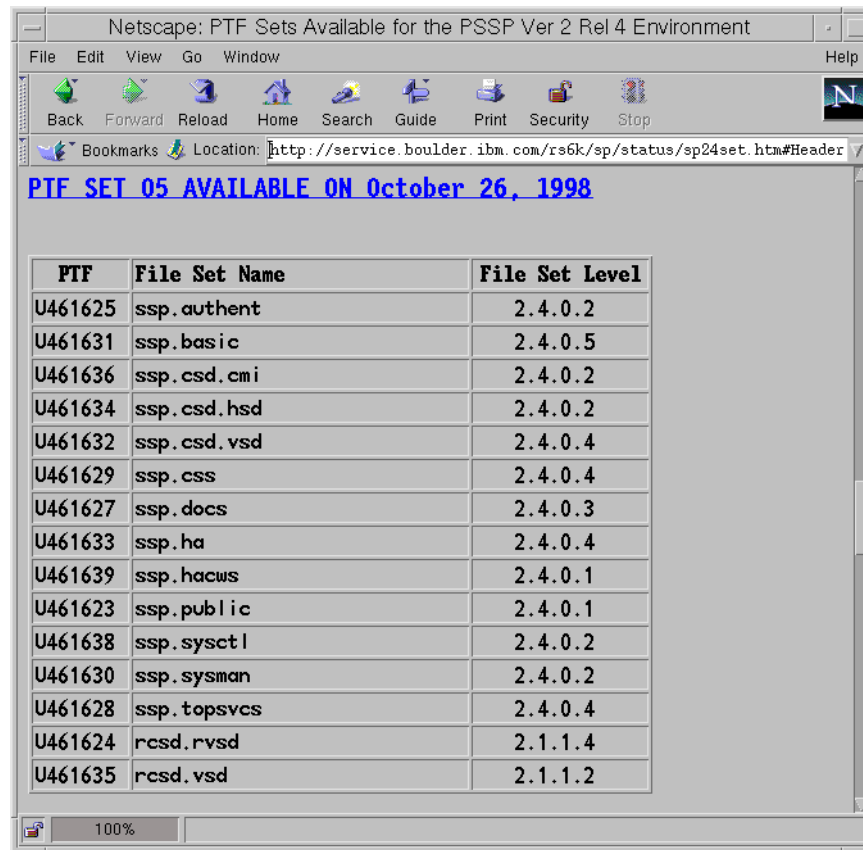


Figure 25. PTF Set 05 available on October 26, 1998

Step 2: Access to anonymous FTP servers

Currently, the following anonymous FTP servers are available:

Table 2. Anonymous FTP servers

Country	URL	IP Address
United States	service.software.ibm.com	197.17.57.66
United Kingdom	ftp.europe.ibm.com	193.129.186.2
Canada	rwww.aix.can.ibm.com	204.138.188.126
Germany	www.ibm.de	192.109.81.2
Japan	fixdist.yamato.ibm.co.jp	203.141.89.41

The following example (Figure 26 on page 106 through Figure 28 on page 108) shows how you can get PTF set 05 for PSSP 2.4 listed in Figure 25 on page 105. It uses the anonymous FTP server in the United Kingdom listed in Table 2 on page 105. First, you need to change your current directory to a file system that has enough space to receive the PTF.

```
# cd /usr/sys/space_for_ptf
# ftp ftp.europe.ibm.com
220-
220- Welcome to ftp.europe.ibm.com
220- *****
220-
220- You are connected to the FTP gateway for IBM Europe's
220- Electronic Customer Support Facility. This server is brought to you by
220- the IBM Global Services(UK) Web Server Team.
220-
220- WARNING: Access and usage of this computer system is monitored.
220- Unauthorized access is prohibited and is subject to criminal
220- and civil penalties.
220-
220- Login using the "anonymous" userid. Please supply your email address
220- when prompted for a password. If you have any odd problems, try logging
220- in with a minus sign (-) as the first character of your password. This
220- will turn off a feature that may be confusing your client program.
220-
220- Please send any questions, comments, or problem reports about this
220- server to web_farm@net.ibm.com.
220-
220 vi FTP server (Version wu-2.4(34) Wed May 13 11:37:43 BST 1998) ready.
ftp>
```

Figure 26. Get PTF from anonymous FTP server (1 of 3)

To log in as an anonymous user, use anonymous for a user name. You are prompted to enter your e-mail address as a password. You can input any text for the password.


```

ftp> user anonymous
331 Guest login ok, send your complete e-mail address as password.
Password:
230-
230-Welcome to IBM's European ftp server for electronic support.
230-Feel free to browse around and download anything you wish.
230-
230-To summarise the directories:
230-
230- aix/tools/fixdist..... Code to run AIX FixDist (fix distribution)
230- aix/tools/tapegen..... Code to run AIX Tapegen (create stacked tapes)
230- aix/fixes/v3..... AIX 3.2 PTF repository
230- aix/fixes/v4..... AIX 4.1 PTF repository
230- aix/fixes/cad..... AIX CATIA PTF repository
230-
230-Please ensure that all copyright and other proprietary notices are
230-retained in any download you make.
230-
230-Please send any questions, comments, or problem reports about this
230-server to web_fam@vnet.ibm.com
230-
230 Guest login ok, access restrictions apply.
ftp> cd aix/fixes/v4
250 CWD command successful.
ftp> ls -m
200 PORT command successful.
150 Opening ASCII mode data connection for /bin/ls.
X11, cics, comm, db2, devices, dsmit, essl, fixdistdb, graphics, lost+found,
ls_lR, monitors, netware, nv6000, os, other, printers, sapr3, sna, sp,
ums, xlc
226 Transfer complete.
ftp> cd sp
250 CWD command successful.
ftp> ls ssp.authent.2.4.0.2.*
200 PORT command successful.
150 Opening ASCII mode data connection for file list.
ssp.authent.2.4.0.2.bff
ssp.authent.2.4.0.2.info
226 Transfer complete.
ftp>

```

Figure 27. Get PTF from anonymous FTP server (2 of 3)

Issue the `bin` subcommand to transfer files by binary mode. If you are going to issue the `mget` subcommand, turn off the prompting for each transfer by issuing the `prompt` subcommand.

```

ftp> bin
200 Type set to I.
ftp> prompt
Interactive mode off.
ftp> mget ssp.authent.2.4.0.2.*
200 PORT command successful.
150 Opening BINARY mode data connection for ssp.authent.2.4.0.2.bff (317440 byte
s).
226 Transfer complete.
317440 bytes received in 160.7 seconds (1.929 Kbytes/s)
local: ssp.authent.2.4.0.2.bff remote: ssp.authent.2.4.0.2.bff
200 PORT command successful.
150 Opening BINARY mode data connection for ssp.authent.2.4.0.2.info (609 bytes)
.
226 Transfer complete.
609 bytes received in 0.3927 seconds (1.514 Kbytes/s)
local: ssp.authent.2.4.0.2.info remote: ssp.authent.2.4.0.2.info
ftp>

```

Figure 28. Get PTF from anonymous FTP server (3 of 3)

To get all the PTFs included in PTF set 05 for PSSP 2.4, repeat the `mget` subcommand operation.

Attention

If you want to get PTFs by using a PTF number or APAR number as the key, there is a FixDist utility available. Visit the following Web site for details:

<http://service.software.ibm.com/support/rs6000/>

2.3.3 Applying PTFs for AIX to CWS, nodes, and SPOT

When you manage your SP system, you can categorize FTP two ways. One is PTF for AIX itself, and the other is for Licensed Program Products (LPPs). When you apply PTFs to your SP system, be aware that AIX and LPPs require different methods of applying the PTFs. The following steps describe how you can apply PTFs to AIX:

Step 1: Put PTF in lppsource directory

Place all PTFs for AIX in the `/spdata/sys1/install/name/lppsource` directory on CWS. The directory name *name* is the `lpp_source` name for the nodes; commonly the AIX version number is used. This section uses `aix432` as an example. Issue the `inutoc` command to update the `.toc` file in the directory:

```
# cd /spdata/sys1/install/aix432/lppsource
# inutoc .
#
```

The .toc file contains all table of contents entries for every installation image in the directory and updates the Software Vital Product Data (SWVPD) database to indicate that the installation image is available.

Step 2: Apply PTF to CWS and nodes

We recommend that you do not apply the PTF to all nodes at once. Instead, verify that the PTF works correctly by applying it in stages. Begin with CWS, then one node, and finally the rest of the nodes. You can issue the `smitty installp` fast path:

```
# cd /spdata/sys1/install/aix432/lppsource
# smitty installp
```

To update all file sets with the file sets in the lppsource directory, select **Update Installed Software to Latest Level (Update All)**. You will see the following SMIT menu:

```
Update Installed Software to Latest Level (Update All)

Type or select a value for the entry field.
Press Enter AFTER making all desired changes.

* INPUT device / directory for software      [Entry Fields]
                                                [.]          +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset      F6=Command      F7=Edit        F8=Image
F9=Shell      F10=Exit       Enter=Do
```

Because you are in the right directory, you need only to specify [.] in the INPUT device / directory for software field. When you hit the **Enter** key, you will see the following SMIT menu:

```

Update Installed Software to Latest Level (Update All)

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[Entry Fields]
* INPUT device / directory for software      .
* SOFTWARE to update                          _update_all
PREVIEW only? (update operation will NOT occur)  no      +
COMMIT software updates?                      no      +
SAVE replaced files?                          yes     +
AUTOMATICALLY install requisite software?      yes     +
EXTEND file systems if space needed?          yes     +
VERIFY install and check file sizes?          no      +
DETAILED output?                              no      +
Process multiple volumes?                      yes     +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

You should not commit when you update the software for the first time. If you apply a software update, old files that are replaced by new files are still kept on the disks. You can reuse these old files easily if the new files cause trouble. If you commit software updates, old files will be removed. This is useful only if you want to save disk space. For the first time, you should specify no in the COMMIT software updates? field. Make sure that the SAVE replaced files? field is yes.

The following is the command that is executed from this SMIT menu:

```
/usr/lib/instl/sm_inst installp_cmd -a -d . -f _update_all -g -X
```

You can find this command in the /smit.script file.

The PTF may require you to reboot CWS. If you confirm that the PTF works correctly, you can also apply the PTF to the nodes.

To apply the PTF to the nodes, export the lppsource directory to the nodes. To do this, issue the `mkknfsexp` command:

```
# /usr/sbin/mkfs_xp -d /spdata/sys1/install/aix432/lppsource -t ro -N
/spdata/sys1/install/aix432/lppsource ro
Exported /spdata/sys1/install/aix432/lppsource
#
```

Then issue the `dsh` and `mount` command to mount this file system on the nodes:

```
# dsh -a mount cws:/spdata/sys1/install/aix432/lppsource /mnt
```

This example mounts `/spdata/sys1/install/aix432/lppsource` directory on the CWS to `/mnt` mount point on all the nodes.

To apply the PTF on the nodes, use scripts in the `/smit.script` file created by SMIT when you apply the PTF to the CWS. The following is the sample script, `sp_updateall.sh`, that can be used to apply the PTF:

```
#!/bin/ksh
#
# sp_updateall.sh
#
USAGE="Usage: sp_updateall.sh <name-of-input-directory>\n"
#
if (( $# < 1 ))
then
    print "$USAGE"
    exit 2
fi
/usr/lib/instr1/sm_inst installp_cmd -a -d $1 -f _update_all -g -X
```

Issue the `pcp` command to distribute this script to the nodes:

```
# cd /usr/local/bin
# ls
sp_updateall.sh
# dsh -a mkdir -p /usr/local/bin
# pcp -a /usr/local/bin/sp_updateall.sh
#
```

Then you can apply the PTF to one of the nodes by issuing the `dsh` command:

```
# dsh -w f01n01 /usr/local/bin/sp_updateall.sh /mnt
```

This example applies the PTF to the node `f01n01`.

If you confirm the PTF works fine, apply the PTF to the rest of the nodes. Afterwards, unmount the unnecessary directory by using the `umount` command:

```
# dsh -a umount /mnt
```

Step 3: Apply PTF to SPOT

The SP system has one more place to apply PTFs: The Shared Product Object Tree (SPOT). It is located in the `/spdata/sys1/install/aix432/spot/spot_aix432` directory and is managed by Network Installation Management (NIM). SPOT is equivalent to the `/usr` file system and is made from images in the `/spdata/sys1/install/aix432/lppsource` directory. When a `mksysb` image is installed to a node, SPOT is mounted, and some AIX commands are executed from there. Therefore, any PTF for AIX applied to that `mksysb` image must be placed in the `lppsource` directory and must be applied to SPOT also. This is documented in the "Read this First" memo that is shipped with PSSP installation media.

For more information about `lpp_source`, refer to 3.1.5, "What is the `lpp_source` object?" on page 138. For more information about SPOT, refer to 3.1.8, "What is the spot object?" on page 141.

In order to apply PTF for AIX to SPOT, perform the following steps on the CWS and boot/install servers (BIS):

1. To de-allocate SPOT from all clients, issue the `spbootins` command on CWS. For example:

```
# spbootins -r disk 1 1 10
```

2. For BIS nodes, add the BIS hostname to the `/.rhosts` file on the CWS.
3. Check the status of the `lppsource` directory by issuing the `nim` command:

```
# nim -o check
```

If there is no problem with the `lppsource` directory, no messages are displayed. If there are some error messages, refer to *AIX Version 4.3 Network Installation Management Guide and Reference*, SC23-4113

The name of the `lppsource` resource can be checked by issuing the `lsnim` command:

```

# lsnim
master          machines      master
boot           resources    boot
nim_script     resources    nim_script
spnet_en0      networks     ent
noprompt       resources    bosinst_data
lppsource_aix432 resources    lpp_source
migrate        resources    bosinst_data
psspscript     resources    script
mksysb_1       resources    mksysb
spnet          networks     ent
spot_aix432    resources    spot
sp4n01         machines    standalone
sp4n05         machines    standalone
sp4n06         machines    standalone
sp4n07         machines    standalone
sp4n08         machines    standalone
sp4n11         machines    standalone
sp4n13         machines    standalone
sp4n15         machines    standalone
sp4n10         machines    standalone
prompt         resources    bosinst_data
# nim -o check -F lppsource_aix432
#

```

In this example, spot_aix432 object is used for the SPOT (spot object type), and lppsource_aix432 object is used for the lppsource directory (lpp_source object type). They have no problem.

4. Issue the `smitty nim_res_op` fast path to apply PTF to the SPOT:


```

+-----+
                        Resource Name
+-----+
Move cursor to desired item and press Enter.

boot             resources    boot
migrate          resources    bosinst_data
noprompt         resources    bosinst_data
prompt           resources    bosinst_data
lppsource_aix432 resources    lpp_source
mksysb_1         resources    mksysb
nim_script       resources    nim_script
psspscript       resources    script
spot_aix432     resources    spot

F1=Help          F2=Refresh          F3=Cancel
Esc+8=Image      Esc+0=Exit          Enter=Do
/=Find           n=Find Next
+-----+

```

5. Select the spot_aix432 object for the SPOT (spot object type) from the list. You will see the following SMIT menu:

```

+-----+
                        Network Install Operation to Perform
+-----+
Move cursor to desired item and press Enter.

reset           = reset an object's NIM state
cust            = perform software customization
sync_roots     = synchronize roots for all clients using specified SPOT
showres        = show contents of a resource
maint          = perform software maintenance
lslpp          = list LPP information about an object
fix_query      = perform queries on installed fixes
showlog        = display a log in the NIM environment
check          = check the status of a NIM object
lppchk         = verify installed filesets
update_all     = update all currently installed filesets

F1=Help          F2=Refresh          F3=Cancel
Esc+8=Image      Esc+0=Exit          Enter=Do
/=Find           n=Find Next
+-----+

```

6. Select **update_all**. You will see the following SMIT menu:

```
Customize a SPOT

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* Resource Name                    spot_aix421
  Fixes (Keywords)                  update_all
* Source of Install Images         [lppsource_aix432]  +
  Expand file systems if space needed?  yes          +
  Force                               no           +

installp Flags
PREVIEW only? (install operation will NOT occur)  no          +
COMMIT software updates?                          no          +
SAVE replaced files?                               yes         +
AUTOMATICALLY install requisite software?         yes         +
OVERWRITE same or newer versions?                  no          +
VERIFY install and check file sizes?               no          +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit         Enter=Do
```

Press the **F4** key in the Source of Install Images field. Then select the `lppsource_aix432` object for the `lppsource` directory (`lpp_source` object type) . To apply PTF instead of commit, the `COMMIT software updates?` field should be `no`, and the `SAVE replaced files?` field should be `yes`.

The following `nim` command is equivalent to this SMIT operation:

```
# nim -o cust -a fixes=update_all -a lpp_source=lppsource_aix432 \
> -a installp_flags="g" spot_aix432
```

Attention

When you install AIX to a node, all PTFs for AIX placed in the `lppsource` directory are automatically installed on nodes if SPOT is updated with those PTFs.

2.3.4 Applying PTFs for LPPs to CWS and nodes

PTFs for LPPs (including for PSSP) are not managed by NIM. Therefore, you do not need to take care of the SPOT. The following steps shows how you can apply PTFs for LPPs:

Step 1: Put PTF on CWS

Make a directory under /spdata file system to keep the PTF. This example uses the /spdata/sys1/install/aix432/ptf4lp directory for this purpose.

To make the directory, issue the `mkdir` command:

```
# mkdir -p /spdata/sys1/install/aix432/ptf4lp
```

Then copy the PTF to this directory.

Step 2: Apply PTF to CWS and nodes

We recommend that you not apply the PTF to all nodes at once. Instead, verify that the PTF works correctly by applying it in stages. Begin with the CWS, then one node, and finally the rest of the nodes. You can issue the `smitty installp first path` or `installp` command.

Refer to “Step 2: Apply PTF to CWS and nodes” on page 109 to do this. Use the /spdata/sys1/install/aix432/ptf4lp directory instead of /spdata/sys1/install/aix432/lppsource.

2.3.5 When should you apply PTF?

Once your SP system becomes stable, you may hesitate to apply any PTFs, since there is a possibility that applying a PTF will cause unexpected problems. But, PTFs related to AIX, PSSP, or other LPPs are produced frequently, and several PTFs will be released in a year. If you fail to keep up with them, it might cause more difficulty applying PTFs in the future.

We recommend that you perform software maintenance at regular intervals even if there is no problem. Four times per year is an ideal schedule, and twice per year is the minimum.

2.4 System backup

Backing up your system is one of the most important things that a system administrator should do. Keeping a current and functional backup is always handy. The use of the backup can vary from restoring a file that was accidentally removed to restoring an entire system after the system crashes.

This section explains how to back up your CWS and SP nodes. It explains how to back up your AIX and the PSSP directories. This information is especially useful for new system administrators who are looking for ideas on how to maintain good backups.

2.4.1 Backing up rootvg on CWS

From the viewpoint of backing up and restoring the data, the CWS can be treated as a stand-alone RS/6000 machine. Therefore, the `mksysb` and `savevg` commands can be used to back up the data. These commands are included in the `bos.sysmgt.syspr` file set.

To back up rootvg, issue the `smitty mksysb` fast path:

```
Back Up the System

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]

WARNING: Execution of the mksysb command will
         result in the loss of all material
         previously stored on the selected
         output medium. This command backs
         up only rootvg volume group.

* Backup DEVICE or FILE                [/dev/rmt0]      +/
Create MAP files?                       no              +
EXCLUDE files?                          no              +
List files as they are backed up?       no              +
Generate new /image.data file?          yes             +
EXPAND /tmp if needed?                  no              +
Disable software packing of backup?     no              +
Number of BLOCKS to write in a single output []          #
(Leave blank to use a system default)

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

This example uses the tape device (`/dev/rmt0`) as a backup device.

You can issue the `mksysb` command to perform the same operation:

```
# mksysb -i /dev/rmt0
```

2.4.2 Restoring rootvg on CWS

To restore the data, boot CWS from the tape or AIX install CD-ROM. We recommend that you boot from the CD-ROM. If you restore the `mksysb` image on a machine other than the one on which the image is backed up, device drivers that are needed on the current machine are automatically installed from CD-ROM (AIX 4.2 or above).

After restoring your mksysb image, issue the `install_cw` command. All SP nodes and the CWS have a `node_number` entry in their Object Data Manager (ODM) Customized Attribute (CuAt). This entry gets lost when you restore the mksysb image to the CWS. The `install_cw` command will create the proper `node_number` entry in the ODM on the CWS.

If you have problems with the partition-sensitive subsystems, issue the `syspar_ctrl` command to recreate partition-sensitive subsystems.

For the CWS:

```
# syspar_ctrl -D -G
# syspar_ctrl -A -G
```

For the nodes:

```
# dsh -a syspar_ctrl -D
# dsh -a syspar_ctrl -A
```

If you have a problem with the SDR, issue the `SDR_config` command to reconfigure the SDR:

```
# SDR_config
```

This command queries the existing hardware on the SP system and updates the SDR as necessary.

2.4.3 Backing up /spdata on CWS

In general, we recommend that you create the `/spdata` file system that contains vital information of SP outside of the `rootvg`. To back up the `/spdata` file system, issue the `smitty savevg fast path`:

```

                                Back Up a Volume Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]

WARNING: Execution of the savevg command will
         result in the loss of all material
         previously stored on the selected
         output medium.

* Backup DEVICE or FILE          [/dev/rmt0]      +/
* VOLUME GROUP to back up       [spdatavg]        +
List files as they are backed up?  no          +
Generate new vg.data file?       yes          +
Create MAP files?                no          +
EXCLUDE files?                  no          +
EXPAND /tmp if needed?          no          +
Disable software packing of backup? no          +
Number of BLOCKS to write in a single output []      #
(Leave blank to use a system default)

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

You can issue the `savevg` command to perform the same operation:

```
# savevg -i -f /dev/rmt0 spdatavg
```

2.4.4 Restoring /spdata on CWS

To restore the `savevg` image, issue the `smitty restvg` fast path:

```

Remake a Volume Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* Restore DEVICE or FILE          [/dev/rmt0]      +/
  SHRINK the filesystems?         no              +
  PHYSICAL VOLUME names           []              +
  (Leave blank to use the PHYSICAL VOLUMES listed
   in the vname.data file in the backup image)
  Number of BLOCKS to read in a single input  []      #
  (Leave blank to use a system default)

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

You can issue the `restvg` command to perform the same operation:

```
# restvg -q -f /dev/rmt0
```

2.4.5 Backing up/restoring strategy for SP nodes

From the viewpoint of backing up and restoring the data, an SP node can be treated as a stand-alone RS/6000 machine. However, it is not practical to attach a tape device to each node. Besides, each node has a lot of data for applications, such as databases or Lotus Notes, in attached large hard disks.

In general, on large SP system, some backup/restore server nodes that are attached to a tape library, such as 3590, provide data backup/restore service to the other client nodes by using Adstar Distributed Storage Management (ADSM). We recommend that you use the `mksysb` command for rootvg backup, and use ADSM for other volume group or data backup.

To know the basic configuration of ADSM, refer to *Using ADSM to Back Up Database*, SG24-4335.

2.4.6 Backing up rootvg on SP node

In most cases, the `/spdata/sys1/install/images` directory on CWS is used to locate a rootvg backup image created by the `mksysb` command. You can NFS mount this directory from the node and make a `mksysb` image on it.

On the CWS, issue the following commands:

```
# /usr/sbin/mknfsxp -d /spdata/sys1/install/images -t ro -N
# dsh -w node-name mount CWS-name:/spdata/sys1/install/images /mnt
# dsh -w node-name mksysb -i /mnt/backup_node-name_YYMMDD
```

After completing the backup, remove unnecessary NFS export.

```
# dsh -w node-name umount /mnt
# exportfs -u /spdata/sys1/install/images
```

Attention

Write performance of NFS is improved from NFS Version 3 (AIX 4.2.1). But still, it is not suitable for the writing of enormous amounts of data. If your site has enough disk space on a node, you could take the `mksysb` image on it and transfer it to CWS later by issuing the `ftp` command.

To simplify the procedure of backup of rootvg on a node, it is recommended that you avoid varying the contents of rootvg between nodes. If you can make one common rootvg image, you save disk space on CWS on which it is kept. To accomplish this, user data or application data should be located outside of rootvg. Also, you should make a script for each node to handle node-specific configurations.

If your SP system does not have many nodes, and ADSM is not used for backup, you can use the standard AIX commands to back up and restore the file system, such as the `savevg`, `tar`, or `backup` command.

2.4.7 Restoring rootvg on SP node

To restore the rootvg backup image, you can use the combination of the `smitty changevg_dialog` and `smitty server_dialog` fast path. The following is the SMIT menu of the `smitty changevg_dialog` fast path:


```

Change Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
Start Frame                          []
Start Slot                            []
Node Count                            []

OR

Node List                             [1]

Volume Group Name                     [rootvg]
Physical Volume List                  [hdisk0]
Number of Copies of Volume Group      1
Set Quorum on the Node
Boot/Install Server Node              [0]
Network Install Image Name            [backup_node1_981101]
LPP Source Name                       [aix432]
PSSP Code Version                     PSSP-3.1

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell        F10=Exit             Enter=Do

```

This example installs backup image backup_node1_981101 to node 1.

You can issue the `spchvgobj` command to perform the same operation:

```

# spchvgobj -l 3 -r rootvg -h hdisk0 -c 1 -n 0 \
> -i backup_node1_981101 -v aix432 -p PSSP-3.1

```

To set the node to install mode and run the `setup_server` command, issue the `smitty server_dialog fast path`:

```

                                Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      []
Start Slot                       []
Node Count                       []

OR

Node List                        [1]

Response from Server to bootp Request    install
Volume Group Name                       [rootvg]
Run setup_server?                       yes

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit      F8=Image
F9=Shell     F10=Exit       Enter=Do

```

You can issue the `spbootins` command to perform the same operation:

```
# spbootins -l 1 -r install -c rootvg
```

You can verify the current setting of the node by issuing the `splstdata` command:

```

# splstdata -b -l 1
                                List Node Boot/Install Information

node#      hostname  hdw_enet_addr  svr      response      install_disk
last_install_image  last_install_time  next_install_image  lppsource_name
pssp_ver          selected_vg
-----
1 sp4n01.msc.itso.  02608C2E86CA  0          install      hdisk0
bos.obj.ssp.432 Sat_Aug_29_01:50:07 backup_node1_981101  aix432
PSSP-3.1          rootvg
#

```

Then, network boot the node by issuing the `nodecond` command:

```
# nodecond frame_number node_number &
```

Alternatively, use the Hardware Perspective. To use the Hardware Perspective, refer to 9.2.2.5, “Network booting a node” on page 241.

To check the LCD and LED display issue the `spmon` command:

```
# spmon -d
```

Prior to PSSP 3.1

Prior to PSSP 3.1, the `smitty changevg_dialog fast path` and `spchvgobj` command are not available. You can specify Network Install Image Name or Destination Hard Disk(s) by issuing the `smitty server_dialog fast path` or `spbootins` command.

2.4.8 Restoring rootvg mirroring node

PSSP 3.1 uses the new Volume_Group SDR class to store mirroring information. The Volume_Group object contains all the information about volume groups on each node. The following is an example of issuing the `SDRGetObjects` command to get information:

```
# SDRGetObjects Volume_Group node_number==1
node_number  vg_name      pv_list      rvg          quorum      copies      ma
pping      install_image code_version lppsource_name boot_server last_install_t
ime last_install_image last_bootdisk
1 rootvg      hdisk0       true         true         1 fa
lse        default      PSSP-3.1     aix432       0 Tue_Mar_16_14:25:
21_EST_1999 default      hdisk0
#
```

You can specify the number of copies for rootvg or quorum settings in the SDR by issuing the `spmkvgobj` or `spchvgobj` command. PSSP 3.1 installation will turn on AIX mirroring during installation on the node based on it even if the backup image is taken from a node on which AIX mirroring is turned off. For more details, refer to 6.3.1, “Configuring root volume group mirroring” on page 190.

Prior to PSSP 3.1, if you take a backup of a node on which AIX mirroring is turned off, PSSP installation with the image can not turn on AIX mirroring automatically. But, if you take a backup of a node on which AIX mirroring is turned on, the `/image.data` file, which is created with the `mksysb -i` command and is included in the `mksysb` image itself, contains mirrored logical volume information. The mirroring configuration is automatically recovered by the normal node restoring operation based on this information . Of course, you

need to specify multiple hard disks that correspond to the number of mirror copies that was specified on the target node.

2.5 Public domain software

Public domain software (PDS) is used by PSSP to help manage the SP system. This section describes which software included in PSSP is in fact PDS. Also, it explains where to locate the source code. This section is purely informational and is only relevant to those who want to know every detail about PSSP and what it uses.

2.5.1 Included PDS

PSSP 3.1 includes the following PDS:

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

All copyright notices in the documentation must be respected. You can find version and distribution information for each of these products in the `/usr/lpp/ssp/README/ssp.public.README` file.

2.5.2 Where is the source code?

PSSP uses several PDSs. Some of them are not well documented in IBM publications. The `ssp.public` file set provides all PDS source code that is used in PSSP. The following is the contents of the file set:

```

# lslpp -cf ssp.public
#Path:Fileset:File
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/tk3.6pl.patch
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/expect.5.7.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/sup.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/perl5.003.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/optlevel1.ssp.public
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/tk3.6.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/perl-4.036.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/README/ssp.public.README
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/tclX7.3a-p2.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/tcl7.3.tar.Z
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/README
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/clock.txt
/usr/lib/objrepos:ssp.public 3.1.0.0:/usr/lpp/ssp/public/ntp.tar.Z
#

```

Files with names with suffix Z are compressed by the `compress` command. You can expand them by issuing the `zcat` command:

```

# cd /usr/lpp/ssp/public
# zcat ntp.tar.Z | tar -xvf-

```

You can get Internet locations of these PDS codes in the `/usr/lpp/ssp/README/ssp.public.README` file.

The following is the contents of the `/usr/lpp/ssp/README/ssp.public.README` file at the time of this writing:

```

*****
                                ssp.public
*****
"@(#)76  1.15  src/ssp/public/ssp.public.README, public, ssp_rtro, rtrot3dg"

DESCRIPTION

The ssp.public option contains tar files of full public distributions
of expect, PERL, NTP, sup, Tcl, TclX, and Tk.

DOCUMENTATION UPDATES AND INFORMATION

Internet Locations of Public Code

The SP software includes and uses publicly available code which
is available via anonymous ftp off the internet. The SP software
includes, as a service to the customer, the tar file for each
application as it is on the internet.

This file is intended to document where to get the source code
and where the documentation is for each.

The applications are Perl(programming language), NTP (Network Time
Protocol), SUP(Software Update Protocol), Tcl(tool command language),
TclX(Extended Tcl), Tk(X11 toolkit based on Tcl), and expect(Tcl
application designed to help automate interactive programs).

The source for these applications are installed in
/usr/lpp/ssp/public
in the form of compressed tar files.

There may be newer versions of the applications at the ftp sites than is
shipped with PSSP.

```

Figure 29. /usr/lpp/ssp/README/ssp.public.README (1 of 3)

```

App:  ftp site      directory
      documentation
      (blank line)
      copyright
      misc. info
-----
Perl:  ftp.cis.ufl.edu /pub/perl/CPAN/src/unsupported/4.036/perl-4.036.tar.gz
      PSSP version:4.036

      ftp.cis.ufl.edu /pub/perl/CPAN/src/5.0
      PSSP version:5.003
      "Programming perl" by Larry Wall and Randal L. Schwartz
      O'Reilly & Associates, Inc. (A Nutshell Handbook). Also see

      http://www.cise.ufl.edu/ftp/perl/CPAN/CPAN.html for more information.

      Copyright (c) 1989,1990,1991, Larry Wall
      Copyright (C) 1989 Free Software Foundation, Inc.
      675 Mass Ave, Cambridge, MA 02139, USA
      Note - the Perl 5 version shipped with PSSP is no longer available
      though the ftp site.

NTP:  ftp.udel.edu   /pub/ntp
      PSSP version: 3.3c
      In tar file; uncompress and tar in ./ntp/doc
      clock.txt (/usr/lpp/spp/public) contains information
      about other NTP time servers... This file can be used
      to determine a server which is closest to your location.

      Copyright (c) David L. Mills 1992, 1993, 1994
      Note - the version shipped with PSSP is no longer available though the
      ftp site.

SUP:  mach.cs.cmu.edu /usr0/anon/project/mach/sup
      version: 7.24
      In tar file; uncompress and tar in ./sup
      Also see http://www.cs.cmu.edu/afs/cs.cmu.edu/project/mach/public/
      www/mach.html for more information.

      n/a

Tcl:  ftp.scriptics.com /pub/tcl/tcl_old
      PSSP version 7.3
      In tar file. Also in "Tcl and the Tk Toolkit" by John K.
      Ousterhout, Addison-Wesley Publishing Company. Also see
      http://www.scriptics.com/ for more information.

      n/a

```

Figure 30. /usr/lpp/spp/README/spp.public.README (2 of 3)

```
TclX:  ftp.neosoft.com      /pub/tcl/tclx-distrib
        version 7.3a-p2
        In tar file. See http://www.neosoft.com/tcl/ for more information.

        n/a
        Note - the version shipped with PSSP is no longer available though the
        ftp site.

Tk:    ftp.scriptics.com   /pub/tcl/tcl_old
        version 3.6
        In tar file. Also in "Tcl and the Tk Toolkit" by John K.
        Ousterhout, Addison-Wesley Publishing Company. Also see
        http://www.scriptics.com/ for more information.

        n/a

expect: ftp.cme.nist.gov   /pub/expect
        version 5.7
        In tar file. Also see http://expect.nist.gov/ for more information.

        n/a
```

Figure 31. `/usr/lpp/ssp/README/ssp.public.README` (3 of 3)

Part 2. Managing installation, customization, and configuration

Chapter 3. Network installation management

AIX Network Installation Management (NIM) is a tool used by an RS/6000 to install and maintain other RS/6000s over the network. This means that a system administrator does not have to physically go to each RS/6000, manually boot from an install CD-ROM, and install the operating system. All of this can be done easily over the network.

IBM Parallel System Support Programs for AIX (PSSP) uses NIM in its system administration. It allows nodes to be installed and customized using the control workstation (CWS) as the single point of control. If you have multiple nodes, you may have a node specified as a boot/install server (BIS). Because SP nodes do not have CD-ROM devices attached to them, they cannot be installed in the normal fashion by booting from a CD-ROM. This is why PSSP takes advantage of the service provided by NIM and customizes it so that it is easier for administration.

There are numerous documents and manuals explaining how NIM works and how it must be set up. This chapter focuses on using NIM on an SP system environment. In particular, standalone, lpp_source, and spot are three very important NIM object types used on SP system environment. These are discussed from the viewpoint of administrative operation.

Isolating the NIM problems is also useful in understanding how to manage the NIM tool. This includes looking at log files and checking the configuration files

If you are interested in further reading about NIM, refer to *AIX Version 4.3 Network Installation Management Guide and Reference*, SC23-4113.

3.1 Operations

As a system administrator, there are certain NIM operations that you have to oversee and manage. This section provides you with a brief overview of how NIM works in an SP system environment. This includes describing the different Perl scripts, called wrappers, that are used to make the SP NIM operation easier.

After this brief overview, NIM client is discussed focusing on how to create and how to delete it. Finally, different types of NIM objects, such as the lpp_source and the spot, are explained. Instructions on how to manage them and keep them up-to-date are given.

3.1.1 NIM in SP system environment

The NIM is used in heterogeneous RS/6000 environments to consolidate AIX installation and maintenance tasks on the machines belonging to it.

The NIM uses a TCP/IP based client/server network environment for its work. A client machine, called *NIM client*, registered to the server machine, called *NIM master*, can install AIX using the network by push and pull installation. This implies software maintenance goals, such as customizing, installing additional file sets, and executing scripts on the NIM clients.

The design of NIM takes an object-oriented approach. The objects belonging to NIM are categorized into the following three classes:

- Machines class
- Networks class
- Resources class

There is only one NIM master and one or more NIM clients for each NIM environment. The NIM master and NIM client belong to the *machines class*. In the SP system environment, the role of the NIM master is assumed by the CWS and SP nodes performing the function of a BIS for other nodes within the SP system. The rest of the SP nodes are NIM clients.

The network used by the NIM operation is the SP Ethernet. The SP Ethernet belongs to the *networks class*.

The resources used by the NIM operation include AIX and PSSP at least. The resources also include the Licensed Program Products (LPPs) for the AIX, customizing scripts, system images, and the Shared Product Object Tree (SPOT). The SPOT is responsible for providing the installation environment for the various types of RS/6000 systems. These resources belong to the *resources class*.

PSSP integrates the NIM to make network installation of SP nodes possible. The PSSP provides and uses interfaces to NIM that fit into its special design. This means handling NIM in the SP system environment is a little bit different from what you may know from outside the SP system environment. However, NIM, as such, has not changed; the way you manage it is different.

To list the NIM objects used in the SP system environment, issue the `lsnim` command:

```

# lsnim
master          machines      master
boot           resources    boot
nim_script     resources    nim_script
spnet_en0      networks     ent
psspscript     resources    script
prompt         resources    bosinst_data
noprompt       resources    bosinst_data
migrate        resources    bosinst_data
lppsource_aix432 resources    lpp_source
mksysb_1       resources    mksysb
spot_aix432    resources    spot
sp3n01         machines     standalone
sp3n17         machines     standalone
#

```

Reading from left to right, the output shows NIM object names, NIM class names, and NIM type names. This chapter focuses on some of the machines class and resources class NIM objects critical to the SP nodes installation. These are:

- standalone type NIM object
- lpp_source type NIM object
- spot type NIM object

Table 3 shows the relationship between NIM objects and SP system resources where *name* is replaced by the lppsource_name assigned with the `spchvgobj` command (for example, lppsource_aix432). If you do not assign the name, PSSP uses a word, *default* where *hostname* is replaced by the host name of the SP node (for example, sp3n01).

Table 3. NIM objects in the SP system resource

NIM object name	NIM object class	NIM object type	SP system resource
<i>hostname</i>	machines	standalone	SP nodes
<i>lppsource_name</i>	resources	lpp_source	/spdata/sys1/install/ <i>name</i> /lppsource
<i>spot_name</i>	resources	spot	/spdata/sys1/install/ <i>name</i> /spot/spot_ <i>name</i>

Attention

This chapter uses the following terms:

- NIM client for standalone type NIM object
- lpp_source object for lpp_source type NIM object
- spot object for spot type NIM object

3.1.2 What are wrappers?

The PSSP uses a set of Perl scripts, called *wrappers*, to perform dedicated NIM configuration tasks. This has the advantage that you do not have to perform each single step during object definitions known from the original NIM operation. Instead, NIM can easily be configured because most of the information it needs is already contained in the System Data Repository (SDR) on the CWS. Even the existence of wrappers is shielded most of the time when you are using the PSSP installation and maintenance mechanisms.

Use of these wrappers is optional; the `installation` and `configuration` commands use them to do the necessary work when you perform the high-level functions.

NIM administration within the SP system means referring to the following wrappers located in the `/usr/lpp/ssp/bin` directory:

allnimres	Allocates NIM resources from a NIM master to a NIM client.
delnimclient	Deletes a NIM client definition from a NIM master.
delnimmast	Unconfigures a node as a NIM master.
mknimclient	Makes a node a NIM client of its BIS.
mknimint	Creates the necessary NIM interfaces on a NIM master.
mknimmast	Configures a node as a NIM master.
mknimres	Creates the necessary NIM resources on a NIM master.
unallnimres	Deallocates NIM resources from a NIM master to one or more NIM clients.

3.1.3 Deleting a NIM client

In the SP environment, SP nodes are the NIM clients for the NIM master that is the CWS or BISs. Normally, the NIM clients are managed by the

`setup_server` command. If a NIM client's environment changes, the changes are reflected in the SDR. The `setup_server` command will notice these changes and re-create the NIM client referring to the SDR.

If you want the NIM client node to be deleted from a NIM master, use the `delnimclient` wrapper. The following example deletes a NIM client definition for node 1 from a NIM master.

Step 1: Delete the NIM client

To delete a NIM client definition from a NIM master, issue the `delnimclient` wrapper:

```
# delnimclient -l 1
warning: 0042-140 m_mmac: unable to remove the /etc/niminfo file on
"sp4n01"
rshd: 0826-813 Permission is denied.

delnimclient: Node 1 (sp4n01) unconfigured as a NIM client on
boot/install server node 0 (sp4en0new).
#
```

Ignore the warning message from the system. It is generated if the `niminfo` file does not exist. You can specify deleting more than one NIM client in a comma-separated list.

Step 2: Check if the NIM client is deleted

You can check if the NIM client definition for node 1 was deleted from the NIM master. The node 1 used to have a NIM object name `sp4n01`. To do this, issue the `lsnim` command:

```
# lsnim -l sp4n01
0042-053 lsnim: there is no NIM object named "sp4n01"
#
```

3.1.4 Creating a NIM client

A NIM client can be created by issuing the `mknimclient` wrapper. The following example creates the NIM client named `sp4n01`, which was deleted in 3.1.3, "Deleting a NIM client" on page 136. This example assumes that the NIM client's information in the SDR is correct.

Step 1: Create the NIM client

To create the NIM client, issue the `mknimclient` wrapper:

```
# mknimclient -l 1
mknimclient: Client node 1 (sp4n01.msc.itso.ibm.com) defined as NIM client on se
rver node (NIM master) 0 (sp4en0new).
#
```

Step 2: Check the NIM client information

You can check the NIM information generated by executing the `mknimclient` wrapper. To do this, issue the `lsnim` command:

```
# lsnim -l sp4en01
sp4n01:
class          = machines
type           = standalone
platform       = rs6k
netboot_kernel = mp
if1            = spnet_en0 sp4n01 02608C2E86CA ent
cable_type1    = bnc
Cstate         = ready for a NIM operation
prev_state     = ready for a NIM operation
Mstate         = currently running
#
```

3.1.5 What is the lpp_source object?

The `lpp_source` object is critical for NIM configuration. It contains all AIX file sets used for NIM client installation. These file sets use the AIX Backup File Format (BFF). Note, that PSSP file sets are not included in the AIX file sets. The `lpp_source` object should only contain file sets belonging to the AIX itself. On the NIM master, the `lpp_source` object file sets reside in the `/spdata/sys1/install/name/lppsource` directory.

It is possible to have multiple `lpp_source` objects. In other words, you can have `lpp_source` objects according to the AIX version, for example, AIX 4.1, 4.2, and 4.3. To distinguish these `lpp_source` objects, NIM uses a special naming convention. It follows the scheme:

`lppsource_name`

Where *name* is replaced by the `lppsource_name` assigned with the `spchvgobj` command.

Table 4 shows commonly used lpp_source object names.

Table 4. The naming convention for lpp_source objects

AIX version	lpp_source object name	lpp_source object directory
4.1.5	lppsource_aix415	/spdata/sys1/install/aix415/lppsource
4.2.1	lppsource_aix421	/spdata/sys1/install/aix421/lppsource
4.3.2	lppsource_aix432	/spdata/sys1/install/aix432/lppsource

All file sets in the lpp_source object directory are registered in a table of contents file named .toc that resides in the same directory. This file is managed by the `inutoc` command. If additional file sets are put in the directory, the .toc file has to be updated by issuing the `inutoc` command:

```
# inutoc .
```

This command example assumes that the lpp_source object directory is the current working directory.

NIM requires that a minimum group of AIX file sets be located in the lpp_source object directory. In the case of AIX 4.3.2, the following file sets are required:

- bos
- bos.64bit
- bos.up
- bos.mp
- bos.net
- bos.diag
- bos.sysmgmt
- bos.terminfo
- bos.terminfo.all.data
- devices.base.all
- devices.buc.all
- devices.common.all
- devices.graphics.all
- devices.mca.all
- devices.rs6ksmp.base
- devices.scsi.all
- devices.sio.all
- devices.sys.all
- devices.tty.all
- xIC.rte

Refer to 3.2.3, “The c_sh_lib file” on page 146 for detailed information about which file sets are required and which file sets are optional to NIM for various AIX versions.

3.1.6 Checking lpp_source object

The lpp_source object directory is created and filled during the PSSP installation process. Make sure the file sets required are contained in this directory.

To list NIM information about the lpp_source object, issue the `lsnim` command. The `-l` flag provides you with detailed information. To list NIM information for the object named `lppsource_aix432`, issue the `lsnim` command:

```
# lsnim -l lppsource_aix432
lppsource_aix432:
  class      = resources
  type       = lpp_source
  Rstate     = ready for use
  prev_state = unavailable for use
  location   = /spdata/sys1/install/aix432/lppsource
  simages    = yes
  alloc_count = 0
  server     = master
#
```

Output shows the `lppsource_aix432` located in `/spdata/sys1/install/aix432/lppsource` on the NIM master. During installation, customization, and maintenance operations, the NIM client refers to this directory. The `alloc_count` variable indicates how many NIM clients are currently using the `lppsource_aix432` object.

3.1.7 Updating lpp_source object

There exist a scenario in which the lpp_source object needs to be updated. When you need to install additional device drivers or Program Temporary Fixes (PTFs) on the SP system, updating the lpp_source object should not be forgotten.

To learn the exact operation for this task, refer to 2.3.3, “Applying PTFs for AIX to CWS, nodes, and SPOT” on page 108.

3.1.8 What is the spot object?

The spot object is critical for NIM configuration. It contains the SPOT that is the essential NIM component involved in the network installation process of the SP nodes. The SPOT is built using the file sets located in the `lpp_source` object directory for the specific AIX version. On the NIM master, the spot object files reside in the `/spdata/sys1/install/name/spot/spot_name` directory.

As with the `lpp_source` object, the NIM master can maintain multiple spot objects. Each spot object corresponds to a different AIX version. To distinguish these spot objects, NIM uses a special naming convention. It follows the scheme:

`spot_name`

Where *name* is replaced by the `lppsource_name` assigned with the `spchvgobj` command.

Table 5 shows commonly used spot object names.

Table 5. The naming convention for spot objects

AIX version	spot object name	spot object directory
4.1.5	spot_aix415	/spdata/sys1/install/aix415/spot/spot_aix415
4.2.1	spot_aix421	/spdata/sys1/install/aix421/spot/spot_aix421
4.3.2	spot_aix432	/spdata/sys1/install/aix432/spot/spot_aix432

The spot object is responsible for the creation of the network boot images (boot kernel) located in the `/tftpboot` directory of the NIM master. These network boot images follow the naming convention:

`spot_name.platform.processor_architecture.network_adapter`

When an SP node is installing, it transfers the network boot image matching its platform (*platform*), processor architecture (*processor_architecture*), and network boot adapter (*network_adapter*) from the NIM master using the Trivial File Transfer Protocol (TFTP). To give an example, a high node booting via SP Ethernet transfers the file `spot_aix432.rs6k.mp.ent`.

3.1.9 Checking spot object

To list NIMs information provided for a spot object named `spot_aix432`, issue the `lsnim` command:

```

# lsnm -l spot_aix432
spot_aix432:
class          = resources
type          = spot
Rstate        = ready for use
prev_state     = verification is being performed
location       = /spdata/sys1/install/aix432/spot/spot_aix432/usr
version        = 4
release        = 3
mod           = 2
alloc_count    = 0
server         = master
if_supported   = rs6k.mp ent
if_supported   = rs6k.up ent
Rstate_result = failure
mk_netboot    = yes
mk_netboot    = yes
plat_defined   = chrp
plat_defined   = rs6k
plat_defined   = rspc
#

```

Note, that the location attribute specified in the output refers to the /spdata/sys1/install/aix432/spot/spot_aix432/usr directory. This is different from /spdata/sys1/install/aix432/spot/spot_aix432. The reason for referring to the usr directory is that this directory is mounted by an SP node when it is installing. The SP node uses the directory specified in the location attribute value as its /usr file system during installation. However, there are other files belonging to the spot object. They are located in the directory /spdata/sys1/install/aix432/spot/spot_aix432.

3.1.10 Checking SPOT log file

During the PSSP installation process, the SPOT is created when the SP nodes are set to install and the `setup_server` command is executed. If you wonder why `setup_server` runs for so long, the SPOT build is the reason.

NIM logs the actions performed during SPOT creation in the /tmp directory. The file name is `spot.out.process_ID`. If problems occur, refer to the log file. Figure 32 on page 143 and Figure 33 on page 143 show an excerpt from it.

```

Creating SPOT in "/spdata/sys1/install/aix432/spot" on machine "master" from "l
ppsource_aix432" ...

    nim_realloc: total size = 11265
    nim_realloc: total size = 12289
    nim_realloc: total size = 13313
    nim_realloc: total size = 14337
    nim_realloc: total size = 15361
nim_realloc: total size = 14337
    nim_realloc: total size = 15361
    nim_realloc: total size = 16385
    nim_realloc: total size = 17409
    nim_realloc: total size = 18433
    nim_realloc: total size = 19457
Restoring files from BOS image. This may take several minutes ...

    nim_realloc: total size = 20481
    nim_realloc: total size = 21505
    nim_realloc: total size = 22529
    nim_realloc: total size = 23553
    nim_realloc: total size = 24577
Installing filesets ...

```

Figure 32. The log file for creating the SPOT (1 of 2)

```

SUCSESSES
-----
Filesets listed in this section passed pre-installation verification
and will be installed.

Selected Filesets
-----
bos.64bit 4.3.2.0           # Base Operating System 64 bit...
bos.diag.com 4.3.2.0       # Common Hardware Diagnostics
bos.diag.rte 4.3.2.0      # Hardware Diagnostics
bos.diag.util 4.3.2.0     # Hardware Diagnostics Utilities
bos.mp 4.3.2.0             # Base Operating System Multip...
bos.net.nfs.client 4.3.2.0 # Network File System Client
bos.net.tcp.client 4.3.2.0 # TCP/IP Client Support
bos.net.tcp.smit 4.3.2.0  # TCP/IP SMIT Support
bos.sysmgmt.nim.client 4.3.2.0 # Network Install Manager - Cl...
bos.sysmgmt.nim.spot 4.3.2.0 # Network Install Manager - SPOT
bos.sysmgmt.smit 4.3.2.0  # System Management Interface ...
bos.sysmgmt.sysbr 4.3.2.0 # System Backup and BOS Instal...
bos.terminfo.adds.data 4.3.0.0 # ADDS Terminal Definitions

```

Figure 33. The log file for creating the SPOT (2 of 2)

The log file is very lengthy; Look at it out on your SP system to get an idea. The important point about it is that you can see how the log keeps track of the AIX file sets installed into the SPOT. If problems occur during the creation of

the SPOT, you can check which AIX file sets are the cause, which helps to isolate the source. Reading the cryptic internal NIM information also contained in the log file can normally be avoided.

3.1.11 Updating the spot object

After you update the lpp_source object of a specific AIX version as described in 3.1.7, “Updating lpp_source object” on page 140, the spot object that corresponds to that AIX version should be updated too.

To learn the exact operation for this task, refer to 2.3.3, “Applying PTFs for AIX to CWS, nodes, and SPOT” on page 108.

3.2 Isolating NIM problems

This section covers specific details about NIM that are useful when dealing with problem determination and understanding how NIM works. A brief view of the NIM log file and configuration files is given so that you know where the files are contained. Also, there is a description of how you read the c_sh_lib script to know which file sets are required and which are optional when building the lpp_source object and spot object. Finally, the relationship between NIM information and Object Data Manager (ODM) is discussed.

For more information, refer to Chapter 17, “Diagnosing NIM Problems” in *IBM Parallel System Support Programs for AIX: Diagnosis Guide, GA22-7350*.

3.2.1 Checking NIM log files

NIM uses several log files to record its activity. The following list gives a brief definition of each:

nimlog	This file logs NIM operation activity. It is located in the /var/adm/ras directory on the NIM client and NIM master.
nim.installp	This file records the installation of the NIM file sets on the local machine. It is located in the /var/adm/ras directory on the NIM client and NIM master.
spot.out.PID	This file is filled during the creation of the spot object on the NIM master. It is located in the /tmp directory on the NIM master. It uses the process ID (<i>PID</i>) as a file extension. For more details, refer to 3.1.10, “Checking SPOT log file” on page 142.

3.2.2 Checking NIM configuration files

All of the data needed by NIM to configure NIM master and NIM client entities within the SP system is contained in the SDR; so, NIM primarily refers to it as the first source of information. But, the following file is created during the systems installation and contains NIM configuration data as well:

niminfo This configuration file is located in the /etc directory on the NIM client and NIM master. It contains all the information essential to NIM for defining that particular node.

The following shows how the niminfo file looks on the NIM master:

```
# cd etc
# cat niminfo
# nimconfig
export NIM_NAME=master
export NIM_CONFIGURATION=master
export NIM_MASTER_PORT=1058
export NIM_REGISTRATION_PORT=1059
export NIM_MASTER_HOSTNAME=sp5en0.msc.itso.ibm.com
#
```

On the NIM client, the niminfo file contains more information as follows:

```

# cd etc
# cat niminfo
#----- Network Install Manager -----
# warning - this file contains NIM configuration information
#         and should only be updated by NIM
export NIM_NAME=sp5n13
export NIM_HOSTNAME=sp5n13.msc.itso.ibm.com
export NIM_CONFIGURATION=standalone
export NIM_MASTER_HOSTNAME=sp5nb01.msc.itso.ibm.com
export NIM_MASTER_PORT=1058
export NIM_REGISTRATION_PORT=1059
export RC_CONFIG=rc.bos_inst
export NIM_BOSINST_RECOVER="/../SPOT/usr/lpp/bos.sysmgt/nim/methods/c_bosinst_en
v -a hostname=sp5n13.msc.itso.ibm.com"
export SPOT=sp5nb01.msc.itso.ibm.com:/spdata/sys1/install/aix432/spot/spot_aix43
2/usr
export NIM_BOSINST_DATA=/NIM_BOSINST_DATA
export NIM_CUSTOM="/../SPOT/usr/lpp/bos.sysmgt/nim/methods/c_script -a location=
sp5nb01.msc.itso.ibm.com:/export/nim/scripts/sp5n13.script"
export NIM_BOS_IMAGE=/NIM_BOS_IMAGE
export NIM_BOS_FORMAT=mksysb
export NIM_HOSTS=" 192.168.6.13:sp5n13.msc.itso.ibm.com 192.168.6.1:sp5nb01.msc
.itso.ibm.com 192.168.5.150:sp5en0.msc.itso.ibm.com "
export NIM_MOUNTS=" sp5en0.msc.itso.ibm.com:/spdata/sys1/install/aix432/lppsourc
e:/SPOT/usr/sys/inst.images:dir sp5nb01.msc.itso.ibm.com:/spdata/sys1/install/
ssp/bosinst_data:/NIM_BOSINST_DATA:file sp5nb01.msc.itso.ibm.com:/spdata/sys1/i
ninstall/images/bos.obj.ssp.432:/NIM_BOS_IMAGE:file "
#

```

The contents of the file provides a lot of information. It is not necessary to understand every item. However, most of the items are self-explanatory and, therefore, understandable.

3.2.3 The `c_sh_lib` file

Located in the `/usr/lpp/bos.sysmgt/nim/methods` directory, this Korn shell script provides library functions used for NIM operations. During the build of the `lpp_source` object and spot object, NIM refers to it as well.

The script is very lengthy, and you do not need to understand the whole thing. But the script contains, especially in the header, definitions of constants and strings that help you understand the AIX file set requirements NIM imposes. To avoid a disaster, do not modify the script.

The information in the script file covers AIX 4.1, 4.2, and 4.3.

3.2.3.1 The `lpp_source` object requirements

NIM has certain file set requirements to create the `lpp_source` object. This information is listed in the script. The following shows the

REQUIRED_SIMAGES stanza. The AIX file sets that belong to this stanza are required for AIX 4.3.

```
REQUIRED_SIMAGES="\
bos \
bos.64bit \
bos.up \
bos.mp \
bos.net \
bos.diag \
bos.sysmgmt \
bos.terminfo \
bos.terminfo.all.data \
devices.base.all \
devices.buc.all \
devices.common.all \
devices.graphics.all \
devices.mca.all \
devices.rs6kmp.base \
devices.scsi.all \
devices.sio.all \
devices.sys.all \
devices.tty.all \
xlC.rte"
```

The set of AIX file sets NIM requires for successfully creating an lpp_source object is called SIMAGES, which stands for Supported Images.

This means, if you want to create an AIX 4.3 lpp_source object successfully, the file sets listed here must be contained in the /spdata/sys1/install/*name*/lppsource directory.

The required file sets for lpp_source object are given by the script for AIX 4.1 and 4.2 as well. Check the script and find the following stanzas for these AIX versions:

- REQUIRED_SIMAGES_41 (for AIX 4.1)
- REQUIRED_SIMAGES_42 (for AIX 4.2)

3.2.3.2 The lpp_source object options

The /spdata/sys1/install/*name*/lppsource directory can contain more than just the basic set of AIX file sets described in 3.2.3.1, “The lpp_source object requirements” on page 146. Therefore, other stanzas named SIMAGES_OPTIONS exist. NIM refers to the listed file sets during the creation of the lpp_source object and checks the items. It is not necessary to keep all these file sets in the /spdata/sys1/install/*name*/lppsource directory.

As long as the required file sets are there, you do not have to worry. The following shows the SIMAGES_OPTIONS stanza for AIX 4.3:

```
SIMAGES_OPTIONS="\
bos \
bos.64bit \
bos.up \
bos.mp \
bos.adt \
bos.html.en_US.topnav \
bos.iconv \
bos.net \
bos.diag \
bos.loc.iso \
bos.msg.en_US \
bos.powermgt \
bos.docregister \
bos.sysmgt \
bos.terminfo.all \
bos.txt \
devices.all \
ifor_ls.base \
ifor_ls.client \
printers.rte \
sysmgt.sgguide \
sysmgt.websm \
sysmgt.msg.en_US.websm \
Java.rte \
x1C.rte \
x1C.cpp \
x1C.msg.Ja_JP.cpp \
x1C.msg.en_US.cpp \
x1C.msg.ja_JP.cpp \
X11.apps \
X11.base \
X11.compat \
X11.Dt \
X11.fnt \
X11.loc.all \
X11.motif \
X11.msg.all \
X11.vsm \
_SPOT._.pre_i.usr.1.0.0.0 \
_SPOT._.post_i.usr.1.0.0.0"
```

The optional file sets for lpp_source object are given by the script for AIX 4.1 and 4.2 as well. Check the script and find the following stanzas for these AIX versions:

- SIMAGES_OPTIONS_41 (for AIX 4.1)
- SIMAGES_OPTIONS_42 (for AIX 4.2)

3.2.3.3 The spot object options

The `c_sh_lib` file contains information about the optional file sets for the spot object as well. The `DEFAULT_SPOT_OPTIONS` stanza is used to provide a list of file sets. As an example, the following shows this list for AIX 4.3:

```
DEFAULT_SPOT_OPTIONS="\
    ${NIM_CLIENT_PACKAGE} \
    ${NIM_SPOT_PACKAGE} \
    bos.64bit \
    bos.up \
    bos.mp \
    bos.net.nfs.client \
    bos.net.tcp.client \
    bos.net.tcp.smit \
    bos.diag \
    bos.sysmgt.sysbr \
    bos.sysmgt.smit \
    bos.terminfo \
    devices.all"
```

Looking at the first two lines, variables containing the file set names for the NIM client and SPOT file sets are specified. They refer to the following file sets respectively:

- `bos.sysmgt.nim.client`
- `bos.sysmgt.nim.spot`

The last line states that all of the device drivers can be included in the SPOT. Note, that this does not mean all the device driver file sets have to be copied to the `lppsource` directory. You can be selective and save file system space, copying only the device drivers really needed for your SP system to the `lppsource`. However, copying all of them puts you on the safe side.

The optional file sets for the spot object are given by the script for AIX 4.1 and 4.2 as well. Check the script and find the following stanzas for these AIX versions:

- `DEFAULT_SPOT_OPTIONS_41` (for AIX 4.1)
- `DEFAULT_SPOT_OPTIONS_42` (for AIX 4.2)

3.2.4 Getting NIM information from ODM

The NIM master uses the AIX infrastructure and stores configuration information in the ODM. The following list shows the NIM-specific files located in the `/etc/objrepos` directory of the NIM master:

- `nim_attr`

- nim_attr.vc
- nim_object
- nim_object.vc
- nim_pdatatr
- nim_pdatatr.vc

This structure suggests an object-oriented view, because the `nim_object` class contains all objects defined within the NIM master. The `nim_attr` class contains the attributes describing a particular object.

You can query information about objects using the `odmget` command. The following example queries information about the NIM network object named `spnet_en0`:

```
# odmget -q name=spnet_en0 nim_object

nim_object:
  id = 910735966
  name = "spnet_en0"
  class = 3
  type = "43"
  attrs = "910735966"

#
```

Note, that an ID is assigned to each object. If you want to find the attributes of the object, you have to query the attribute class using this ID. The following example shows how to do this.

```
# odmget -q id=910735966 nim_attr

nim_attr:
  id = 910735966
  value = "192.168.5.0"
  seqno = 0
  pdattr = "110"

nim_attr:
  id = 910735966
  value = "255.255.255.0"
  seqno = 0
  pdattr = "111"

nim_attr:
  id = 910735966
  value = "20"
  seqno = 0
  pdattr = "109"

nim_attr:
  id = 910735966
  value = "2"

#
```

Hopefully, this output will alarm you. The attributes for the SP Ethernet specify the net address (pdattr="110") and the subnet mask (pdattr="111'). This means, if you are changing the network address or the subnet mask, NIM is affected.

Chapter 4. Node installation

Node installation is one of the first things that you will have to do when you set up your SP system. Having the control workstation (CWS) as your boot/install server (BIS) and using Network Installation Management (NIM) to install your nodes is the normal procedure. Although NIM is customized with the use of Perl scripts to minimize the administration work, there are many places where the install can fail or hang.

There are two different ways that you can monitor the progress of your node installation. The first method is to watch the installation online using the `s1term` command. This shows you each step, and you can follow the progress of your install. The other method is to monitor the installation offline by reading the various log files that are created. These log files can be used to determine if a portion of the install failed or if a particular file set or directory was missing.

This chapter discusses how to set up your system so that you can monitor your node installation. It also describes the different log files that collect information about the installation procedure. Finally, hints and tips on how to isolate a problem during a node installation will be discussed.

4.1 Monitoring node installation online

This section describes how to monitor your node installation while the installation is taking place. This is useful for system administrators because it shows the progress of the installation. This means you know if the installation has hung or if it is close to being finished.

4.1.1 Using the `s1term` command

To monitor node installation online, you can use the `s1term` command. An open read-only terminal session during the installation process of the node provides most of the necessary information. Besides looking at the produced output and messages from the installing node, the information shown on the screen can be recorded in a file on the CWS as well.

To split the output sent to the terminal and record it in the file as well, prior to network booting the node, issue the `s1term` command:

```
# s1term 1 13 | tee /tmp/install.node13
```

In this example, the /tmp/install.node13 will contain the installation record for node 13. If you miss something on the screen, you can always refer to the file containing the same information.

After having done so, start the node installation.

Because the architecture of the SP nodes varies, the information displayed during the selection of the nodes network boot device varies as well. However, you can see if the system's Bootstrap Protocol (BOOTP) request will be answered by the CWS or boot/install server (BIS).

Figure 34 shows the BOOTP packets sent by the node. The CWS or BIS answers them. After BOOTP does its job, the node is starting to load the network boot image located in the /tftpboot directory using the Trivial File Transfer Protocol (TFTP). You can see if the nodes packets are answered in Figure 34 as well.

```
STARTING SYSTEM (BOOT)

Booting . . . Please wait.

Ethernet: Built-In
Hardware address ..... 10005AFA07DF

          Packets Sent      Packets Received
BOOTP           00002           00001
TFTP            07546           07545
```

Figure 34. Output during netboot phase

The files used for the network boot image are named `spot_name.platform.processor_architecture.network_adapter` and are built by the NIM. For more details about network boot image, refer to 3.1.8, "What is the spot object?" on page 141.

The node starts to initialize the kernel and builds its RAM file system. From this point on, the installation process looks like the one on stand-alone RS/6000 machine.

4.1.2 Using LED or LCD messages

During the whole installation process, the node uses three digit codes indicating its state and what it is doing. You can use the `spmon` command or Hardware Perspective to monitor the codes shown. If the node gets stuck, and the code displayed does not change over a long period of time, refer to the following IBM publications to check the cause of the problem:

- *IBM Diagnostic Information for Multiple Bus Systems*, SA38-0509
- *OEM- Diagnostic Information for Micro Channel Bus System*, SA23-2765

4.2 Monitoring node installation offline

If you are looking for more detail than what appears online during the install, you can refer to the node logs. The node logs part of its activity during the installation phase in the `/var/adm/ras` directory. This mechanism is used by AIX.

For information about the boot phase, refer to the following file:

- `bootlog`

For information about the base operating system (BOS) installation, refer to the following files:

- `bosinst.data`
- `bosinstlog`
- `devinst.log`

AIX uses its error log mechanism to record problems during installation. The following files belong to the AIX Error Log:

- `errlog`
- `errtmpl`

System installation information about logical volumes for the root volume group and the corresponding file systems is saved in the following file:

- `image.data`

4.2.1 The bootlog file

The `bootlog` file, located in the `/var/adm/ras` directory, is worth looking at if you want to know what the node did during the boot phase. This file contains AIX Configuration Manager output produced by the `cfgmgr` command and

status messages issued during the boot phase of the node. The following example shows an excerpt from it:

```
invoking top level program -- "/etc/methods/startlft"
return code = 0
***** no stdout *****
***** no stderr *****
-----
invoking top level program -- "/etc/methods/starttrcm"
return code = 0
***** no stdout *****
***** no stderr *****
-----
invoking top level program -- "/etc/methods/starttty"
return code = 0
***** no stdout *****
***** no stderr *****
-----
invoking top level program -- "/usr/lib/methods/cfgfan"
return code = 0
***** no stdout *****
***** stderr *****
Command not allowed to be executed on this model
-----
invoking top level program -- "/usr/lib/methods/defaio"
return code = 0
***** no stdout *****
***** no stderr *****
-----
calling savebase
return code = 0
***** no stdout *****
***** no stderr *****

Saving Base Customize Data to boot disk
Starting the sync daemon
Starting the error daemon
System initialization completed.
Starting Multi-user Initialization
Activating all paging spaces
swapon: Paging device /dev/hd6 activated.
/dev/rhd1 (/home): ** Unmounted cleanly - Check suppressed
Performing all automatic mounts
Multi-user initialization completed
TB3 device already configured.

Method error (/etc/methods/cfglft -l lft0):
    0514-032 Cannot perform the requested function because the
        specified device is dependent on another device which does
        not exist.
```

4.2.2 The bosinst.data file

The bosinst.data file, located in the /var/adm/ras directory, contains installation information about how AIX is set up. Stanzas are used to group certain types of installation information. The file is a kind of response file used to answer prompts that occur during AIX installation. This avoids user interaction.

The following is the description of the stanzas in this file:

control_flow This stanza includes information, such as the installation method chosen or the installation console selected.

target_disk_data This stanza tells where to install AIX.

locale This stanza is the language environment specified.

An excerpt from the bosinst.data file is as follows:

```
control_flow:
  CONSOLE = /dev/tty0
  INSTALL_METHOD = overwrite
  PROMPT = no
  EXISTING_SYSTEM_OVERWRITE = yes
  INSTALL_X_IF_ADAPTER = no
  RUN_STARTUP = no
  RM_INST_ROOTS = no
  ERROR_EXIT =
  CUSTOMIZATION_FILE =
  TCB = no
  INSTALL_TYPE = full
  BUNDLES =
  SWITCH_TO_PRODUCT_TAPE =
  RECOVER_DEVICES =
  BOSINST_DEBUG =

target_disk_data:
  PVID = 0000152900049e6c
  CONNECTION = vscsi0//0,0
  LOCATION = 00-00-00-0,0
  SIZE_MB = 995
  HDISKNAME = hdisk0

locale:
  BOSINST_LANG = en_US
  CULTURAL_CONVENTION = en_US
  MESSAGES = en_US
  KEYBOARD = en_US
```

4.2.3 The bosinstlog file

The bosinstlog file, located in the /var/adm/ras directory, is filled before the AIX switches from the RAM file system environment to the disk file system. The logical volumes, including paging space, boot logical volume, and Journaled File System (JFS) log, are created. The dump device is initialized, and the boot image created. All this is recorded in the file.

The following is an excerpt from it:

```
Preparing target disks.
rootvg
Making boot logical volume.
hd5
Making paging logical volumes.
hd6
Making logical volumes.
hd8
hd4
hd2
hd9var
hd3
hd1
Forming the jfs log.
Making file systems.
Mounting file systems.
Restoring base operating system.
Initializing disk environment.
Over mounting /.
opstst.pok.ibm.com
Copying Cu* to disk.
Installing additional software.
opstst.pok.ibm.com
lft0 changed
Initializing dump device.
primary          /dev/hd6
secondary        /dev/sysdumpnull
copy directory   /var/adm/ras
forced copy flag TRUE
always allow dump FALSE
Network Install Manager customization.
Creating boot image.

bosboot: Boot image is 5907 512 byte blocks.
```

4.2.4 The devinst.log file

During the installation of the various AIX file sets, information is written to the devinst.log file in the /var/adm/ras directory. This reference can easily be used to check if a special file set was installed on the system, if problems occurred, or if the file set was even installed without errors. The following is an excerpt from the devinst.log file:

```

SUCSESSES
-----
Filesets listed in this section passed pre-installation verification
and will be installed.

Selected Filesets
-----
devices.base.rte 4.3.2.0                # RISC System 6000 Base Device...
devices.rs6kmp.base.rte 4.3.2.0        # Multiprocessor Base System D...

<< End of Success Section >>

FILESET STATISTICS
-----
  2 Selected to be installed, of which:
    2 Passed pre-installation verification
----
  2 Total to be installed

```

4.2.5 The errlog file

The errlog file contains the error log entries. To read them, you must use template file containing error templates used by the AIX Error Log. This is why the errtmpl file is there. Issuing the `errpt` command will show you the contents of the log file:

```
# errpt -a -i /var/adm/ras/errlog -y /var/adm/ras/errtmpl | more
```

Note, that the `-i` flag specifies the error log file to read. The `-y` flag specifies the error template file. The `errpt` command will produce the same output without any of the file options. But, the important point is that you notice how other correctly formatted error files can be read with the `errpt` command as well.

4.2.6 The image.data file

System installation information about logical volumes for the root volume group and the corresponding file systems is saved in the image.data file. This information is referred to by AIX during node installation. The following is an excerpt from the file:

```

##Command used for vg_data; /usr/sbin/lsvg
##Command used for source_disk_data; /usr/sbin/bootinfo
##Command used for lv_data; /usr/sbin/lslv
##Commands used for fs_data; /usr/bin/df and /usr/sbin/lsjfs

image_data:
  IMAGE_TYPE = bff
  DATE_TIME = Wed Oct 28 16:03:52 EST 1998
  UNAME_INFO = AIX opstst 3 4 000000401800
  LICENSE_INFO =
  PRODUCT_TAPE = no
  OSLEVEL = 4.3.2.0

logical_volume_policy:
  SHRINK = no
  EXACT_FIT = no

ils_data:
  LANG=en_US

source_disk_data:
  PVID = 000000405ce23b27
  CONNECTION = scsi0//0,0

LOCATION = 00-06-00-0,0
SIZE_MB = 995
HDISKNAME = hdisk0

vg_data:
  VGNAME = rootvg
  PPSIZE = 4
  VARYON = yes
  VG_SOURCE_DISK_LIST = hdisk0

```

4.3 Isolating problems during node installation

It is important to know how to isolate problems that occur while you are performing an installation. This section provides some hints and tips to isolate problems.

4.3.1 Hints and tips

A general methodology would be to use a technique called bottom up. This means, if a problem occurs, start at the lowest level (hardware) and go up to the higher levels (software). Do the following:

1. Check the three digit code to see if the problem is named by the codes description. Refer to the following IBM publications to check the cause of the problem:
 - *IBM Diagnostic Information for Multiple Bus Systems, SA38-0509*

- *OEM- Diagnostic Information for Micro Channel Bus System, SA23-2765*

2. If the cause is still unknown, check the hardware involved.

If the installation problem you face is a network boot problem, check the SP Ethernet cable connections, terminators, and cable length.

3. If the hardware is fine, go to the next level: Check the software involved. Make sure configuration files are set up properly and that the requested services are running on the CWS.

If the CWS or BIS is not answering the client request, check the `/etc/bootptab.info` file, if used. Make sure the nodes' SP Ethernet hardware address information in the SDR is correct.

If the problem occurs during TFTP file transfer, check the `/etc/tftpaccess.ctl` file.

Another good source of information to refer to if you are having network boot problems is Chapter 11, "Network Installation Management Troubleshooting" in *AIX Version 4.3 Network Installation Management Guide and Reference*, SC23-4113.

4. Refer to the terminal log file mentioned in 4.1.1, "Using the `s1term` command" on page 153, if the node gets stuck during AIX installation.

If you can login to the node, refer to the AIX files described in 4.2, "Monitoring node installation offline" on page 155.

5. Make sure the node information contained in the SDR is correct (Node SDR class). Check the boot response for the node, volume group, and installation disk information.

Chapter 5. Node customization

Node customization is the second step for successfully installing an SP node. The customization process installs IBM Parallel System Support Programs for AIX (PSSP) and configures it so that the control workstation (CWS) can integrate the SP node. Without customization, the node is not under the control of CWS.

Normally, when a node has a fresh install, the customization process is executed right after AIX has been installed. However, there are other occasions when you need to customize the node on its own. The purpose of this chapter is to briefly outline what occurs during customization, describe conditions when customization is needed, and explain the customization scripts that are run. Once these basics are understood, the focus is changed to how to customize and how to do problem determination when your customization fails.

The hints on what to do if your customization fails are very important to system administrators. Most likely, during one of your installs, a node will fail in its customization. The suggestions in this chapter are just a starting point in your journey, but at least it gives you tips on where the problem may lie.

5.1 Node customization

Node customization is a very important part of the installation of a node. This section explains what is involved in the customization process and what occurs on the nodes. Also, it describes different situations in which you may want to customize a node and it explains the procedure. First-time system administrators will find this section useful because it gives you a brief overview of what to expect.

5.1.1 What is customization?

When a node boots with the `bootp_response` attribute in Node class in the System Data Repository (SDR) set to `customize`, customization tasks are executed on the node. The main purpose of customization is to configure network adapters on the node based on the Adapter SDR class. The main tasks done by customization are:

- Install PSSP and its requisite file sets
- Add SP Switch ODM entries
- Configure and set up Kerberos authentication

- Create dump logical volume
- Start mirroring/unmirroring the root volume group
- Run the `tunig.cust` and `script.cust` scripts
- Configure network adapters on the node

5.1.2 When do you need customization?

There are several situations when you need to customize a node. When a node is initially installed, it will automatically customize itself. This allows PSSP to be installed and configured on the node. There are other times that a node may require customization, such as when you:

- Add new adapters
- Set up NFS, NIS, or AFS on the SP system
- Change the SDR
- Change authentication services
- Add an SP Switchboard to an SP system

When something is changed in the SDR, the nodes can be informed/reset by a customization.

Also, there can be times when your node does not fully install, meaning the `bootp_response` attribute does not change back to disk mode from install mode. In this case, if the node does boot up but not have host response, and PSSP is not configured, it is wise to recustomize the node and find out when and where it is failing.

5.1.3 Customizing a node with or without rebooting

To customize a node, follow Step 1, then either Step 2a, 2b, or 2c, depending on the configuration of your system.

Step 1: Issue the `spbootins` command

Issue the `spbootins` command as follows:

```
# spbootins -r customize -l node-number
```

It runs the `setup_server` on the node's BIS, which may be the CWS, or it may be another node. The `bootp_response` attribute of Node SDR class is changed to `customize`.

Step 2a: Reboot the node

Reboot the node, but do not perform network boot. When the node boots, it looks up the `bootp_response` of Node SDR class. If `customize` is specified, it runs the `pssp_script` script. This script executes customization tasks.

The `pssp_script` script is also executed at node installation time. This is why PSSP is installed automatically on the node even if the `mksysb` image does not contain PSSP.

Step 2b: Issue the `pssp_script` script on the node

If you want to customize your node without reboot, perform this step instead of “Step 2a: Reboot the node”. It assumes that you have the `pssp_script` script on the node.

Issue the `pssp_script` script on the node as follows:

```
# cd /usr/lpp/ssp/install/bin/pssp_script

=====
pssp_script: = Making PSSP log directories... =
=====
+ /usr/bin/mkdir -p /var/adm/SPlogs/sysman
+ 1> /dev/null 2>& 1
+ exec
+ 3> /var/adm/SPlogs/sysman/NODE.config.log.9768

=====
pssp_script: = Switching output to log file... =
=====

#
```

Step 2c: Install PSSP and issue the `pssp_script` script

If you do not have the `pssp_script` script on the node, for example, AIX has been installed, but PSSP has not been installed on the node, substitute the following steps for “Step 2a: Reboot the node” or “Step 2b: Issue the `pssp_script` script on the node” described previously.

1. Copy the `/etc/SDR_dest_info` file from the CWS to the node using the `ftp` command.
2. Verify that the `perfagent.tools.2.2.32.x` file set is installed on the node.
3. Mount the `/spdata/sys1/install/pssp1pp` directory onto the node.
4. Install the `ssp.basic` file set onto the node (this file set includes the `pssp_script` script).
5. Issue the `pssp_script` script.

For more details, refer to 1.4.1, “Attaching SP-attached servers” on page 75.

The `pssp_script` script logs its progress in the file called `/var/adm/SPlogs/sysman/node-name.config.log.PID`. The variable `node-name` is the host name of the node, and `PID` is the process ID of the `pssp_script` script.

The network adapter configuration is done by the `psspfb_script` script, which is executed from the `pssp_script` script. The `psspfb_script` script has its own log called `/var/adm/SPlogs/sysman/node-name/configfb.log`.

These logs are helpful when doing problem determination.

5.2 After node customization

When an SP node boots up, it must start special PSSP subsystems, which a stand-alone RS/6000 workstation does not run. These changes are what make the SP node part of an SP system complex. They handle communication, monitoring, security, and so on.

There are additions in the `/etc/inittab` file, both on the CWS and the SP nodes, after PSSP is customized. These additions allow daemons to start up, and they set up the security on the SP system.

This section describes the changes to the files and also the boot procedure.

5.2.1 `/etc/inittab` changes on the CWS

During the customization of PSSP on the CWS, additions are made to the `/etc/inittab` file. In the case of a CWS, the changes are shown in the following example:

```

nim:2:wait:/usr/bin/startsrc -g nim >/dev/console 2>&1
sdrd:2:once:/usr/bin/startsrc -g sdr
sp:2:wait:/etc/rc.sp > /dev/console 2>&1
hardmon:2:once:/usr/bin/startsrc -s hardmon
rvsd:2:once:/usr/lpp/csd/bin/ha_vsd > /dev/console 2>&1
st_sw_num:2:boot:/usr/lpp/ssp/bin/st_set_switch_number
spmgr:2:once:/usr/bin/startsrc -s spmgr
kerb:2:once:/usr/bin/startsrc -s kerberos
kadmind:2:once:/usr/bin/startsrc -s kadmind
sysctld:2:once:/usr/bin/startsrc -s sysctld
fsd:2:once:/usr/lpp/ssp/css/rc.switch
splogd:2:once:/usr/bin/startsrc -s splogd
hats:2:once:/usr/bin/startsrc -g hats > /dev/console 2>&1
hb:2:once:/usr/bin/startsrc -g hb > /dev/console 2>&1
hr:2:once:/usr/bin/startsrc -g hr
hags:2:once:/usr/bin/startsrc -g hags > /dev/console 2>&1
haem:2:once:/usr/bin/startsrc -g haem > /dev/console 2>&1
pman:2:once:/usr/bin/startsrc -g pman >/dev/console 2>&1
sp_configd:2:once:/usr/bin/startsrc -s sp_configd

```

Most of these additions are initializing an SP subsystem or group. For example, nim, sdr, hardmon, spmgr, kerberos, kadmind, sysctld, splogd, hats, hb, hr, hags, haem, pman, and sp_configd must all be started on bootup.

If the ssp.css file set was installed, it will add two entries to the /etc/inittab file. The st_set_switch_number script is run to determine the nodes' switch_node_number from the Object Data Manager (ODM). If the ODM is empty, it uses the SDR, then writes out the switch_node_number to the /spdata/sys1/st/switch_node_number data file. In the case of the CWS, this remains empty. The rc.switch script is also run.

If a specific PSSP component is installed, such as IBM Recoverable Virtual Shared Disk, it will also be started in the /etc/inittab file.

5.2.2 /etc/inittab changes on the SP node

There are some differences between the /etc/inittab file of a CWS and a node. The following is an example of the /etc/inittab file additions on a node:

```
start_net:2:wait:/usr/lpp/ssp/install/bin/start_net > /dev/console
sp:2:wait:/etc/rc.sp > /dev/console 2>&1
fsd:2:once:/usr/lpp/ssp/css/rc.switch
nimclient:2:once:/usr/sbin/nimclient -S running
sysctld:2:once:/usr/bin/startsrc -s sysctld
st_sw_num:2:boot:/usr/lpp/ssp/bin/st_set_switch_number
hats:2:once:/usr/bin/startsrc -g hats > /dev/console 2>&1
hags:2:once:/usr/bin/startsrc -g hags > /dev/console 2>&1
haem:2:once:/usr/bin/startsrc -g haem > /dev/console 2>&1
pman:2:once:/usr/bin/startsrc -g pman >/dev/console 2>&1
sp_configd:2:once:/usr/bin/startsrc -s sp_configd
```

The node does not require as many subsystems or groups to be started up as does the CWS. For example, `sysctld`, `hats`, `hags`, `haem`, `pman`, and `sp_configd` are only started on `bootup`.

The CWS is assigned as a Kerberos authentication server by default. Hence Kerberos only needs to be started on the CWS. For more information about Kerberos, refer to Chapter 16, “Security” on page 445.

NIM must also be started, but the node starts up `nimclient` since it is not the NIM master. For more information about NIM, refer to Chapter 3, “Network installation management” on page 133.

The RSCT daemons `hats`, `hags`, and `haem` are started. These daemons are added to the `/etc/inittab` when the node is customized. The `syspar_ctrl -A` command adds these entries into the `/etc/inittab`.

5.2.3 What does normal node bootup do?

During the normal bootup of a node from disk, the node will go through its normal hardware checks and tests. When it reaches the point where it starts `/etc/inittab`, the node changes from a stand-alone RS/6000 machine to an SP node. There are specific scripts and subsystems that are started up at this time; two scripts handle the SP side configuration.

/etc/rc.sp

This script is similar to a customization because it performs any reconfiguration that may be needed on the node since the last reboot. This script does the following:

- Gets node number from ODM
- Starts logging information in `/var/adm/SPlogs/sysman` directory
- Issues the `tuning.cust` script

- Restarts the inetd daemon
- Updates the /etc/SDR_dest_info file, if needed
- Sets up Kerberos security
- Updates the SDR if any information about the node is changed

This script also adjusts the system clock so that Kerberos works correctly.

/usr/lpp/ssp/css/rc.switch

This script handles the SP Switch. This script sets up the environment for the SP Switch to become active on the node. It does the following:

- Configures the SP Switch interface
- Starts the fault_service_Worm_RTG_SP daemon on the node
- Stores log information in /var/adm/SPlogs/css/rc.switch file

Once this configuration is done, then it is just a matter of starting the SP daemons, which is done in the /etc/inittab file.

5.3 Isolating problems during node customization

There are several different areas in which the customization can fail. To be able to recover from these errors, it is important to understand the scripts `pssp_script` and `psspfb_script`. This section describes the errors these scripts can encounter that can halt the customization process. It also describes what steps should be taken when you are trying to isolate a problem. This includes a list of commonly seen three digit codes that occur during the customization stage.

5.3.1 The customization script files

After AIX is installed on a node, the customizing phase starts. During this process, the PSSP components are installed on the node, and it becomes a functional SP node.

To achieve this, two scripts are executed on the node:

- `pssp_script`
- `psspfb_script`

The `pssp_script` script is called by Network Installation Management (NIM) before it reboots the node for the first time. The script is run under a single user environment with the RAM file system in place. It installs file sets required by PSSP and does post-PSSP installation setup.

Additional adapter configuration is performed after the node is rebooted. This is done by the `psspfb_script` script.

Both scripts use three digit codes to indicate the status of the node. This can be monitored using the `spmon` command or SP Perspectives.

5.3.2 Isolating problems by `pssp_script`

There are two methods to isolate problems using the `pssp_script` script.

The three digit codes

During the node customizing phase, the node uses three digit codes to show which state it is in and what it is doing. In case of problems, this information is very helpful in finding the cause.

The `pssp_script` script is located in the `/usr/lpp/ssp/install/bin` directory and is a Korn shell script. Therefore, you can read it and see how it works. The three digit codes are generated using the `showled` command located in the `/etc/methods` directory.

The script, as such, is divided into several functions performing special customize actions on the node. If you look at the following excerpt from the script, you can see how it works:

```
$shled 0xc33
install_ssp_sysctl          #-Install ssp.sysctl
[[ $? -ne 0 ]] && pssps_exit 1  #-Quit on error

$shled 0xc34                #-                               33542
install_ssp_pman           #-Install ssp.pman
[[ $? -ne 0 ]] && pssps_exit 1  #-Quit on error

$shled 0xc41                #-                               33331
config_switch             #-add switch odm entries          33331
[[ $? -ne 0 ]] && pssps_exit 1  #-Quit on error

$shled 0xc35                #-                               33542
install_ssp_css           #-Install ssp.css
[[ $? -ne 0 ]] && pssps_exit 1  #-Quit on error
```

The `$shled` string is an alias for the `showled` command. It is used to show the three digit code specified as argument (for example: `0xc33`). The values are specified in hexadecimal notation.

The function to be executed follows. The names of the functions are self-explanatory, and you can check out what they do by referring to the function code located in the section before the main code. Not all the code contained

in the `pssp_script` script is shown here; therefore, you can not see the function code.

After the function is executed, the system checks whether any bad return codes were given by the called function. If so, the script exits with return code 1. Otherwise it continues.

If the node gets stuck at a particular code during installation, you can refer to the `pssp_script` script and look up the function and the statements executed.

For an explanation of the three digit codes, refer to 5.3.4, “Meaning of the three digit codes” on page 173.

The log file

The script logs its activity in a file located in the `/var/adm/SPlogs/sysman` directory on the node. The file name looks like this:

`NODE.config.log.PID`

PID is the process ID used by the `pssp_script` script. During the execution of the script, the log file is renamed to:

`node_reliable_hostname.config.log.PID`

The variable `node_reliable_hostname` is the reliable host name that is set on the node’s SP Ethernet interface.

The following example shows an excerpt from the log file:

```

=====
pssp_script: = Beginning at Wed Aug 18 10:06:40 EDT 1999      =
=====

pssp_script: = Establishing node environment...              =
=====

+ [[ -z  ]]
+ mode=customize
+ [[ customize = install ]]
+ [[ customize = customize ]]
+ /usr/lpp/ssp/install/bin/update_dest_info
+ rc=1
+ [[ 1 -ne 0 ]]
+ [[ 1 -ne 1 ]]
+ + /usr/bin/grep Network Address
+ /usr/sbin/lscfg -l ent0 -v
hw_enet_addr=      Network Address.....006094E94F8F
+ rc=0
+ [[ 0 -ne 0 ]]
+ hw_enet_addr=006094E94F8F
+ hw_enet_addr=006094E94F8F
+ + /usr/lpp/ssp/bin/SDRGetObjects -G -x Node hw_enet_addr==006094E94F8F node_r
sdr_node_number=      5
+ rc=0
+ [[ 0 -ne 0 ]]
+ sdr_node_number=5
+ sdr_node_number=5
+ + /usr/lpp/ssp/bin/SDRGetObjects -G -x Node node_number==5 bootp_response
sdr_mode=customize
+ rc=0

```

The script is executed with the `set -x` Korn shell trace option. Referring to the log file should help you to quickly identify and fix the problem.

5.3.3 Isolating problems by `psspfb_script`

The `psspfb_script` script configures the node adapters. The script normally runs out of `/etc/inittab` (via `spfbcheck`) when the node is set to install or migrate, and the node is network booted. The script is also called from the `pssp_script` script when the node is set to customize.

There are two methods to isolate problems using the `psspfb_script` script.

The three digit codes

The structure of the script is similar to the `pssp_script` script described in the previous section. Three digit codes are used to indicate what the script is doing.

For an explanation of the meaning of the three digit codes, refer to 5.3.4, “Meaning of the three digit codes” on page 173.

The log file

Activity is logged in a file located in the `/var/adm/SPlogs/sysman` directory on the node. The file name looks like this:

`NODE.configfb.log.PID`

PID is the process ID used by the `psspfb_script` script. During the execution of the script, the log file is renamed to:

`node_reliable_hostname.configfb.log.PID`

The variable `node_reliable_hostname` is the reliable host name that is set on the node’s SP Ethernet interface.

5.3.4 Meaning of the three digit codes

During the customization phase, a progression of three digit codes will be displayed on your system. These codes are extremely helpful if your customization hangs. Table 6 presents a brief description of some of the codes you may see in the node customization phase.

Table 6. Customization phase codes

Code	Meaning
u20	Create PSSP directories
u21	Establish environment directories
u23	Create <code>/etc/ssp</code> files
u24	Update <code>/etc/hosts</code>
u25	Retrieve files from Server
u26	Kerberos activity
u27	Modify <code>/etc/inittab</code>
u28	up/mp selection
u29	Install pre-requisite lpps
u30	Install <code>ssp.client</code>
u31	Install <code>ssp.basic</code>
u32	Install <code>ssp.ha</code>
u33	Install <code>ssp.sysctl</code>

Code	Meaning
u34	Install ssp.pman
u35	Install ssp.css
u36	Install ssp.st
u37	Delete /.rhosts file
u38	Create dump logical volume
u39	Run tuning.cust script
u40	Run script.cust script
u41	Add switch ODM entries
u42	Run psspfb script
u43	Starting mirroring/unmirroring
u45	Start silver node surveillance
u50	Get tuning.cust file
u51	Determine processor type
u53	Install ssp_ha depending on PSSP version
u54	Get spfbcheck
u56	Get psspfb_script from nim_master
u57	Get <node>.config_info file
u58	Get psspfb_script from Control Workstation
u59	Get CuAt template
u60	Store identity information in /etc/ssp
u61	Get files from boot install server
u67	Remove and get new krb.conf file
u68	Remove and get new krb.realms file
u69	Get kerberos server key
u79	Get script.cust file
u80	Installing ssp.clients
u81	Installing ssp.basic

Code	Meaning
u82	Installing ssp.sysctl
u84	Installing ssp.css
u85	Installing ssp.st
u86	Creating dump logical volume
u87	Running script.cust

These codes are not in the sequence that they appear during a customization, and not all codes are shown during a customization. Refer to the `pssp_script` script directly to know the detailed structure and sequence of the code.

5.3.5 Hints and tips

There will be circumstances when a customization of one of your nodes may hang. This may occur during a new install, a migration, or a plain customization. There are several different things that you can do to try to understand why the process has failed.

1. Check the customization logs:
 - `/var/adm/SPlogs/sysman/node-name.config.log.PID`
 - `/var/adm/SPlogs/sysman/node-name.configfb.log.PID`
2. Check which three digit code it is hanging on, and then look in the `pssp_script` and `psspfb_script` script to find out what command it is doing at that time.
3. Check that the `initial_hostname` and `reliable_hostname` variables are correct in the SDR.
4. Check that the proper files existing on the CWS. They include:
 - `/tftpboot/node_hostname.config_info`
 - `/tftpboot/node_hostname.install_info`
 - `/etc/SDR_dest_info`
 - `/etc/krb.conf`
 - `/etc/krb.realms`
 - `/etc/krb-srvtab`
5. Check that you have all the PSSP in your `/spdata/sys1/install/name/lppsource` directory along with a current `.toc` file.

For more details, refer to 3.1.5, “What is the lpp_source object?” on page 138.

6. Check that you have an up-to-date SPOT in your `/spdata/sys1/install/name/spot/spot_name` directory.

For more details, refer to 3.1.8, “What is the spot object?” on page 141.

By following these steps, you should have a good idea of where the problem is. Then, it is just a matter of fixing it and redoing your customization.

Chapter 6. Disk configuration

Disk configuration and management is an integral part of system administration. Disks must be replaced when they fail, and new disks must be added to existing ones when systems run out of space. These types of jobs are done all the time, and every system administrator needs to understand what is involved.

This chapter is devoted to the different operations that are required for disk management. This includes adding disk drives to, and deleting them from, a volume group on an SP node. Some enhancements to IBM Parallel System Support Programs for AIX (PSSP) 3.1 regarding rootvg management are explained. Also there are sections that discuss creating, installing, and switching the boot system image, mirroring/unmirroring your volume groups, and booting your system from external disks. Many of the steps involved in disk management are similar to those for a standalone RS/6000 machine. But, not everything is identical, and this chapter informs you of the proper procedures to follow.

6.1 Physical configuration

This topic is not SP-specific, but you need to know about it when you manage your SP system. Therefore, this section describes two basic operations:

- Adding physical disks to an SP system
- Deleting physical disks from an SP system

In this case, an SP system includes the control workstation (CWS) and SP nodes.

It does not matter whether you are using Serial Storage Architecture (SSA) disks or Small Computer System Interface (SCSI) disks when discussing add and delete operations. SSA disks are more sophisticated because you have to check the microcode level of your SSA loop including SSA adapters. This should be done by you or by an IBM Customer Engineer while physically installing them. The following sections assume that your microcode level is fine and that the SSA/SCSI disks are recognized as hdisks by AIX.

To learn about updating your microcode for devices, refer to 2.1.6, "Upgrading the microcode for devices" on page 95.

6.1.1 Adding a physical disk to a volume group

Adding a physical disk to a volume group is a simple task. If you want to add `hdisk1` to the `growingvg` volume group, issue the `extendvg` command:

```
# extendvg growingvg hdisk1
```

However, when you add new disks to an existing volume group, there is one consideration: The physical partition (PP) size is a global parameter for the volume group, and it is already set. Once set, it can not be changed on the fly.

Before AIX 4.3.1, the number of PPs allowed per hard disk was restricted to 1016. It was important to stay within this boundary. But, AIX 4.3.1 and later releases have changed this limitation so that it is possible for a volume group to exceed 1016 PPs per physical volume. As a trade off, the number of disk drives supported in a volume group is decreased if the PPs exceeds 1016. The `chvg` and `mkvg` command have a new `-t` flag to specify the factor. The commands create or convert a volume group to allow multiples of 1016 PPs per disk. It allows $(1016 * \text{factor})$ PPs per physical disk.

For example, if you have a 2.0 GB disk drive and want to set the PP size to 1 MB, normally this would fail because it has exceeded the 1016 PPs per physical disk. You can either change your PP size or you can add a `-t` flag. The `mkvg` command can be used to create a volume group called `testvg` on a `hdisk1` with a factor of 2:

```
# mkvg -s 1 -t 2 -y testvg hdisk1
0516-1193 mkvg: WARNING, once this operation is completed, volume group testvg
cannot be imported into AIX 430 or lower versions. Continue (y/n) ?
y
testvg
#
```

The `-t` flag prompts you to confirm the factor because, once it is set, you cannot import this volume group to a lower version of AIX. If you take a look at the properties of the volume group, you can see what is the maximum number of PPs allowed and the maximum number of disks. To do this, issue the `lsvg` command:


```

# lsvg testvg
VOLUME GROUP:   testvg                VG IDENTIFIER: 00008984b87b0958
VG STATE:       active                 PP SIZE:       1 megabyte(s)
VG PERMISSION:  read/write            TOTAL PPs:    1918 (1918 megabytes)
MAX LVs:        256                   FREE PPs:     1918 (1918 megabytes)
LVs:            0                      USED PPs:     0 (0 megabytes)
OPEN LVs:       0                      QUORUM:       2
TOTAL PVs:      1                      VG DESCRIPTORS: 2
STALE PVs:      0                      STALE PPs:    0
ACTIVE PVs:     1                      AUTO ON:      yes
MAX PPs per PV: 2032                 MAX PVs:      16
#

```

Figure 35. Output of lsvg command

The MAX PPs per PV field in Figure 35 indicates the maximum number of PPs per physical volume (PV) is 2032. The MAX PVs field indicates the maximum number of PVs in this volume group is 16. The relationship between the factor and these values is shown in Table 7.

Table 7. Relationship between physical partitions and volumes

Factor	MAX PPs per PV	MAX PVs per VG
1	1016	32
2	2032	16
3	3048	10
4	4064	8

As the factor value increases, the maximum number of physical disks allowed is decreased. For further details on the changes, refer to Chapter 5, “Logical Volume Manager Enhancements” in *AIX Version 4.3 Differences Guide*, SG24-2014.

If you do not want to change the factor of a volume group, the alternative is to change the PP size. Here are the steps to follow:

1. Back up the volume group using the `smitty savevg` fast path.
2. Unmount all file systems in the volume group.
3. Delete the volume group using the `smitty reducevg` fast path.
4. Choose a PP size according to the size of the largest physical volume you are using within that volume group.
5. Recreate the volume group with the new PP size using the `smitty restvg` fast path and the backup tape.

6. Check your file systems.

6.1.2 Deleting a physical disk from a volume group

Reducing a volume group is surely more complicated than extending one. This depends on the current setup including factors, such as mirroring and logical volumes spanning several physical disks within the volume group.

This section uses the following scenario:

Node 1 has a volume group called:

- datavg

It consists of two physical disks:

- hdisk3
- hdisk4

On disk hdisk3 exist two logical volumes named:

- data31lv, containing the /data31 Journaled File System (JFS)
- data32lv, containing the /data32 JFS

On disk hdisk4 exists one logical volume named:

- data41lv, containing the /data41 JFS

The logical volume data32lv and the JFS log loglv00 are mirrored to hdisk4.

The configuration is shown by the `lsvg` command:

```
# lsvg -l datavg
datavg:
LV NAME          TYPE      LPs  PPs  PVs  LV STATE      MOUNT POINT
data31lv         jfs       20   20   1    open/syncd    /data31
data41lv         jfs       20   20   1    open/syncd    /data41
data32lv         jfs       40   80   2    open/syncd    /data32
loglv00          jfslog    1     2    2    open/syncd    N/A
#
```

You are going to delete the hdisk4 from the datavg volume group.

Attention

- Reducing volume groups does not move the logical volumes located on the deleted physical volumes to the remaining physical volumes of the volume group.
- Make sure there is enough space in the remaining physical volumes for the logical volumes to be moved.

Step 1: Check space requirements

Check if there will be enough space in the reduced volume group to contain the logical volumes from the physical volume to be deleted. If not, you may have to shrink the logical volumes or even delete the ones you no longer need.

This example assumes that the remaining physical volume `hdisk3` has enough disk space to keep all your logical volumes.

Step 2: Back up the volume group

Before reconfiguring it, save the volume group by issuing the `savevg` command:

```
# savevg -X -i -f/dev/rmt0 datavg
```

If you prefer to use SMIT, issue the `smitty savevg fast path`.

Step 3: Delete mirrors from the logical volumes

Make sure that there are no logical volume copies on the physical disk that you are going to delete from the volume group. To delete any copies of the logical volume `data32lv` from `hdisk4`, issue the `rmlvcopy` command:

```
# rmlvcopy data32lv 1 hdisk4
```

The physical volume to specify is the one from which the copy is to be removed. Alternatively, you can use SMIT by issuing the `smitty rmlvcopy fast path`.

If you use logical volume mirroring currently, do not forget to delete the JFS log mirror from the physical volume to be deleted. To delete the copy of the JFS log on `hdisk4`, issue the `rmlvcopy` command:

```
# rmlvcopy loglv00 1 hdisk4
```

After deleting all logical volume copies, check the volume group status by issuing the `lsvg` command:

```
# lsvg -l datavg
datavg:
LV NAME          TYPE      LPs  PPs  PVs  LV STATE    MOUNT POINT
data31lv         jfs       20   20   1    open/syncd  /data31
data41lv         jfs       20   20   1    open/syncd  /data41
data32lv         jfs       40   40   1    open/syncd  /data32
loglv00          jfslog    1     1    1    open/syncd  N/A
#
```

Step 4: Migrate logical volumes to different physical volumes

If you have available disk space on remaining physical disks within the same volume group, you can migrate the logical volumes from the disk to be deleted. To migrate the logical volume `data41lv` from physical volume `hdisk4` to `hdisk3`, issue the `migratepv` command:

```
# migratepv -l data41lv hdisk4 hdisk3
```

Alternatively, you can use SMIT by issuing the `smitty migratepv` fast path as well.

Attention

- You can not migrate logical volumes to spanning multiple physical volumes.
- You can not migrate striped logical volumes. As a work around, mirror them and delete the original.
- Issuing the `migratepv` command will lock your volume group. Furthermore, it will take some time.

Step 5: Save and close all remaining logical volumes

If there are still logical volumes left on the physical volume to be deleted, and you can not migrate them, save the files in the logical volumes with the `tar` or `backup` command. Then close all logical volumes on the physical volume to be deleted by issuing the `umount` command.

Step 6: Reduce the volume group

This example reduces volume group `datavg` by deleting physical volume `hdisk4`. To do this, issue the `reducevg` command:

```
# reducevg datavg hdisk4
```

Step 7: Restore the files

If you saved any files in Step 5, restore them if there is enough space.

6.2 Boot configuration

From PSSP 3.1, you can use the multiple boot system images. This allows you to boot SP nodes selectively with different configuration or different AIX versions. Also, PSSP 3.1 provides the function that allows you to boot an SP node from the external disks.

This section describes how you can define and install an alternative boot system image to your SP node and how you can boot your SP node from the external disk as well.

6.2.1 Using the alternate boot system image

If you want to use the alternate boot system image, in other words, root volume group, an SP node or the CWS must have the following hardware:

1. Two or more hard disks

Each root volume group has one or more hard disks.

2. A virtual battery on an SP node

Except for old type thin nodes, virtual battery is installed on all nodes. It keeps NVRAM powered. It keeps a boot list of the machine while the SP frame is plugged into a power outlet with power applied or until a node is removed from the SP frame even when the node is powered off. Power is supplied by the node supervisor card in each node.

6.2.2 Defining alternative boot system image

You can create an alternative boot system image using the `smitty createvg_dialog` fast path. This example is going to create an alternative boot system image named `altrootvg` in node 15.

```

                                Create Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
Start Frame                          []                #
Start Slot                            []                #
Node Count                            []                #

OR

Node List                             [15]

Volume Group Name                     [altrootvg]
Physical Volume List                   [00-00-0S-1,0]
Number of Copies of Volume Group      1                +
Boot/Install Server Node               [0]                #
Network Install Image Name             [bos.obj.ssp.432]
LPP Source Name                        [aix432]
PSSP Code Version                      PSSP-3.1          +
Set Quorum on the Node                 #

[BOTTOM]

F1=Help          F2=Refresh      F3=Cancel      F4=List
F5=Reset         F6=Command      F7=Edit        F8=Image
F9=Shell         F10=Exit       Enter=Do

```

Figure 36. Create volume group information

Alternatively, issue the `spmkvgobj` command:

```
# spmkvgobj -r altrootvg -h 00-00-0S-1,0 -n 0 -i bos.obj.ssp.432 \
> -v aix432 -P PSSP-3.1 -l 15
```

Use a different volume group name than ones that already exist.

You can use location code (00-00-0S-1,0, for example) or logical device name (hdisk1, for example) in the Physical Volume List field for SCSI disks. For more information, refer to 6.2.5, “Booting from external SCSI disks” on page 186.

If you are going to use SSA disks, refer to 6.2.6, “Booting from external SSA disks” on page 188.

To install an alternative boot system image to an SP node based on this setting, refer to 6.2.3, “Installing alternative boot system image” on page 185.

If you want to use external disks for an alternative boot system image, refer to 6.2.5, “Booting from external SCSI disks” on page 186 and 6.2.6, “Booting from external SSA disks” on page 188.

6.2.3 Installing alternative boot system image

After defining an alternative boot system image, you need to set the `bootp_response` attribute of the Node SDR class to install. Do this by issuing the `smitty server_dialog` fast path. Make sure that you specify the target boot system image name in the Volume Group Name field (`altrootvg`, in this case):

```

                                Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      []          #
Start Slot                       []          #
Node Count                       []          #

OR

Node List                        [15]

Response from Server to bootp Request  install      +
Volume Group Name                  [altrootvg]
Run setup_server?                   yes           +

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell    F10=Exit      Enter=Do

```

Figure 37. Boot/Install server information

Alternatively, you can issue the `spbootins` command:

```
# spbootins -l 15 -r install -c altrootvg
```

Then, issue the `nodecond` command to network boot the node:

```
# nodecond 1 15 &
```

This example uses `hdisk1` (00-00-0S-1,0) for the alternative boot system image, named `altrootvg`, in node 15. When the installation is complete, check the boot list by issuing the `bootlist` command:

```
# dsh -w sp4n15 bootlist -m normal -o
sp4n15: hdisk1
#
```

The bootlist for the normal mode should be set as hdisk1. If you boot node 15, it will use altrootvg as the boot system image.

6.2.4 Switching alternative boot system images

To alter the boot system image, use the `spbootins` command with the boot system image name that you want to switch to. Because execution of the `setup_server` is not needed for this purpose, suppress it with `-s no` flag. The following is a sample with node 15 as the target node. Currently, node 15 is using `altrootvg` as the boot system image. To switch it to `rootvg`, issue the `spbootins` command:

```
# spbootins -r disk -c rootvg -l 15 -s no
```

To use SMIT instead, refer to Figure 37 on page 185.

Then, change the bootlist on the target node using the `spbootlist` command:

```
# dsh -w sp4n15 bootlist -m normal -o
sp4n15: hdisk1
# spbootlist -l 15
spbootlist: total number of bootlists set successfully = 1
spbootlist: total number of bootlists which could not be set = 0
spbootlist: total number of invalid nodes selected = 0
spbootlist: total number of invalid volume groups selected = 0
# dsh -w sp4n15 bootlist -m normal -o
sp4n15: hdisk0
#
```

To use SMIT instead, issue the `smitty bootlist_dialog fast path`.

For the next boot, node 15 uses `rootvg` in `hdisk0` as system boot image.

6.2.5 Booting from external SCSI disks

From PSSP 3.1, you can boot an SP node from external disks. In the case of Small Computer Systems Interface (SCSI) disks, 7027-HSD SCSI-2 Fast/Wide disk can be used for this purpose.

Table 8 lists the supported combinations of SP node and SCSI adapter.

Table 8. Supported adapters for nodes with SCSI disk boot

Node feature code	Node type	Feature code numbers of supported SCSI adapters
2002	66 MHz Thin	#2412, #2416
2003	66 MHz Wide	
2004	66 MHz Thin2	
2005	77 MHz Wide	
2006	604 High	
2007	120 MHz Thin	
2008	135 MHz Wide	
2009	604e High	
2022	160 MHz Thin	
2050	332 MHz SMP Thin	
2051	332 MHz SMP Wide	

To use external disks as boot disks in addition to internal disks, you need to define an alternative boot system image as described in 6.2.2, “Defining alternative boot system image” on page 183, and install the alternative boot system image as described in 6.2.3, “Installing alternative boot system image” on page 185.

If you use external SCSI disks when you define an alternative boot system image, specify the location code in the Physical Volume List field as shown in Figure 36 on page 184.

The format is:

```
00-00-00-0,0
```

In a case where you specify multiple physical volumes:

```
00-00-00-0,0;00-00-00-1,0
```

The location code is displayed by the `lsdev -Cc disk` command on the target node. For interpretation of the location code, refer to Chapter 20. “Devices” in *AIX Version 4.3 System Management Guide: Operating System and Devices*, SC23-4126.

Attention

You can use a logical device name (hdisk1, for example) instead of location code (00-00-0S-1,0, for example) for an external SCSI disk. However, it is recommended that you use location code to indicate the target disk clearly.

6.2.6 Booting from external SSA disks

From PSSP 3.1, you can boot an SP node from external disks. In the case of Serial Storage Architecture (SSA) disks, 7133 IBM SSA Disk Subsystem Models 010, 020, 500, and 600 can be used for this purpose.

Table 9 lists the supported combinations of SP node and SSA adapter.

Table 9. Supported adapters for nodes with SSA disk boot

Node feature code	Node type	Feature code numbers of supported SSA adapters
2005	77 MHz Wide	#6214 SSA 4-Port Adapter #6216 Enhanced SSA 4-Port Adapter #6217 SSA RAID Adapter #6219 Enhanced SSA RAID Adapter
2006	604 High	
2007	120 MHz Thin	
2008	135 MHz Wide	
2009	604e High	
2022	160 MHz Thin	

To use external disks as boot disks in addition to internal disks, you need to define an alternative boot system image as described in 6.2.2, “Defining alternative boot system image” on page 183, and install the alternative boot system image as described in 6.2.3, “Installing alternative boot system image” on page 185.

If you use external SSA disks when you define an alternative boot system image, specify the SSA drive serial number. It is a 15-character unique identifier. You can find it as a connwhere attribute from the pdisk device. To do this, issue the `l.sdev` command:

```

# dsh -w sp3n13 lsdev -Cc pdisk
sp3n13: pdisk0 Available 00-02-P 2GB SSA C Physical Disk Drive
sp3n13: pdisk1 Available 00-02-P 2GB SSA C Physical Disk Drive
sp3n13: pdisk2 Available 00-02-P 4GB SSA C Physical Disk Drive
sp3n13: pdisk3 Available 00-02-P 2GB SSA C Physical Disk Drive
sp3n13: pdisk4 Available 00-02-P 4GB SSA C Physical Disk Drive
sp3n13: pdisk5 Available 00-02-P 4GB SSA C Physical Disk Drive
sp3n13: pdisk6 Available 00-02-P 4GB SSA C Physical Disk Drive
# dsh -w sp3n13 lsdev -Cc pdisk -r connwhere
sp3n13: 0004AC5052B500D
sp3n13: 0004AC50532100D
sp3n13: 0004AC50616A00D
sp3n13: 0004AC510D1E00D
sp3n13: 0004AC5150BA00D
sp3n13: 0004AC51535D00D
sp3n13: 0004AC51538200D
#

```

In this example, 0004AC51535D00D is the SSA drive serial number for pdisk5.

Issue the `smitty createvg_dialog` fast path to define an alternative boot system image:

```

                                Create Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]                                #
  Start Frame                        []                                #
  Start Slot                          []                                #
  Node Count                          []                                #

OR

Node List                             [13]

Volume Group Name                     [ssarootvg]
Physical Volume List                  [ssar//0004AC51535D00D]
Number of Copies of Volume Group      1                                +
Boot/Install Server Node              [0]                                #
Network Install Image Name            [bos.obj.ssp.432]
LPP Source Name                       [aix432]
PSSP Code Version                     PSSP-3.1                            +
Set Quorum on the Node                +

[BOTTOM]

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit            Enter=Do

```

In the Physical Volume List field, specify the SSA drive serial number with the string `ssar//` as a prefix.

This SMIT menu is equivalent to the `spmkgobj` command:

```
# spmkgobj -r ssarootvg -h ssar//0004AC51535D00D -n 0 \  
> -i bos.obj.ssp.432 -v aix432 -P PSSP-3.1 -l 13
```

6.3 Mirroring configuration

One of the disk-related enhancements in PSSP 3.1 is the possibility of mirroring the root volume group (boot system image) directly from the CWS. Mirroring writes simultaneous copies of the AIX logical volumes. Either two or three copies of logical volumes are allowed in AIX. It improves your SP system's availability.

6.3.1 Configuring root volume group mirroring

As you learned in previous sections, PSSP 3.1 provides commands and SMIT menus to maintain the root volume group information database. You can configure root volume group mirroring by issuing the `smitty changevg_dialog` fast path:

```

Change Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      []          #
Start Slot                       []          #
Node Count                       []          #

OR

Node List                        [13]

Volume Group Name                 [rootvg]
Physical Volume List              [hdisk0,hdisk1]
Number of Copies of Volume Group  2          +
Set Quorum on the Node            false       +
Boot/Install Server Node         [0]       #
Network Install Image Name       [bos.obj.ssp.432]
LPP Source Name                  [aix432]
PSSP Code Version                 PSSP-3.1    +

F1=Help          F2=Refresh      F3=Cancel      F4=List
F5=Reset         F6=Command     F7=Edit       F8=Image
F9=Shell        F10=Exit       Enter=Do

```

Figure 38. Configuring root volume group mirroring

In the Number of Copies of Volume Group field, specify 1, 2, or 3. If you want to have two copies of the root volume group, in other words, one mirroring, specify the following:

- In the Physical Volume List field, specify two physical volumes.
- In the Number of Copies of Volume Group field, specify 2.
- In the Set Quorum on the Node field, specify `false`.

This SMIT menu is equivalent to the `spchvgobj` command:

```
# spchvgobj -r rootvg -h hdisk0,hdisk1 -c 2 -q false -n 0 \
> -i bos.obj.ssp.432 -v aix432 -p PSSP-3.1 -l 13
```

This operation changes only the volume group information in the SDR. To initiate or discontinue root volume group mirroring, refer to the next two sections.

6.3.2 Initiating root volume group mirroring

To initiate root volume group mirroring, use the following steps. This section uses the example that configures root volume group mirroring on node 13 and then initiates mirroring.

Step 1: Check the current status

This step is optional, but it is a good idea to know the current root volume group status of your SP node. To do this, issue the `lspv`, `lsvg`, and `bootlist` commands:

```
# dsh -w sp4n13 lspv
sp4n13: hdisk0          0000174198d73070    rootvg
sp4n13: hdisk1          00001741eccdd80e    None
# dsh -w sp4n13 lsvg -l rootvg
sp4n13: rootvg:
sp4n13: LV NAME          TYPE      LPs    PPs    PVs    LV STATE    MOUNT POINT
sp4n13: hd5              boot      2      2      1      closed/syncd N/A
sp4n13: hd6              paging    80     80     1      open/syncd  N/A
sp4n13: hd8              jfslog    1      1      1      open/syncd  N/A
sp4n13: hd4              jfs       2      2      1      open/syncd  /
sp4n13: hd2              jfs       73     73     1      open/syncd  /usr
sp4n13: hd9var           jfs       8      8      1      open/syncd  /var
sp4n13: hd3              jfs       8      8      1      open/syncd  /tmp
sp4n13: hd1              jfs       1      1      1      open/syncd  /home
sp4n13: lv00             sysdump   7      7      1      open/syncd  N/A
# dsh -w sp4n13 lsvg rootvg | grep QUORUM
sp4n13: OPEN LVs:      8                      QUORUM:      2
# dsh -w sp4n13 bootlist -m normal -o
sp4n13: hdisk0
#
```

Figure 39. Check the current root volume group status

The output from the `lspv` and `lsvg` commands shows the root volume group resides in one disk (`hdisk0`) and is not mirrored. It also shows the quorum is active. The result of the `bootlist` command shows the current setting of the boot list contains only `hdisk0`.

Step 2: Configure mirroring

Change the volume group information as shown in Figure 38 on page 191.

Step 3: Verify the volume group information

Verify the volume group information in the SDR. To do this, issue the `splstdata` command:

```

# splstdata -v -l 13
                        List Volume Group Information

node# name                boot_server quorum copies  code_version lppsource_name
      last_install_image  last_install_time  last_bootdisk
      pv_list
-----
  13 rootvg                0                false    2          PSSP-3.1 aix432
      bos.obj.ssp.432      Tue_Nov_10_10:50:06_EST_1998 hdisk0
      hdisk0,hdisk1
#

```

This example shows the quorum, copies, and pv_list attributes are all correct.

Step 4: Initiate mirroring

To initiate root volume group mirroring, issue the `smitty start_mirroring` fast path:

```

                        Initiate Mirroring on a Node

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      []                #
Start Slot                       []                #
Node Count                       []                #

OR

Node List                        [13]
Force Extending the Volume Group? no                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

This SMIT menu is equivalent to the `spmirrorvg` command:

```

# spmirrorvg -l 13
sp4n13.msc.itso.ibm.com: spmirror: 0016-675 Physical volume hdisk0 is already
part of volume group rootvg.
sp4n13.msc.itso.ibm.com: spmirror: Successfully extended the volume group rootvg
by disk(s) hdisk1.
sp4n13.msc.itso.ibm.com: spmirror: This node must be rebooted to reflect the
change in quorum!
spmirrorvg: Volume group rootvg on node number 13 successfully mirrored.
spmirrorvg: The total number of volume groups mirrored successfully is 1.
spmirrorvg: The total number of volume groups which failed to mirror is 0.
spmirrorvg: total number of invalid nodes selected = 0
spmirrorvg: total number of invalid volume groups selected = 0
#

```

This task might take a while. As shown in the output, you need to reboot the node to reflect the change in quorum. This task issues the `mirrorvg` command.

Step 5: Verify the root volume group information

To verify that the mirroring is initiated correctly, issue the `lspv`, `lsvg`, and `bootlist` commands:

```

# dsh -w sp4n13 lspv
sp4n13: hdisk0          0000174198d73070    rootvg
sp4n13: hdisk1          00001741eccdd80e    rootvg
# dsh -w sp4n13 lsvg -l rootvg
sp4n13: rootvg:
sp4n13: LV NAME          TYPE      LPs   PPs   PVs   LV STATE    MOUNT POINT
sp4n13: hd5              boot      2     4     2     closed/syncd N/A
sp4n13: hd6              paging    80    160  2     open/syncd  N/A
sp4n13: hd8              jfslog    1     2     2     open/syncd  N/A
sp4n13: hd4              jfs       2     4     2     open/syncd  /
sp4n13: hd2              jfs       73    146  2     open/syncd  /usr
sp4n13: hd9var           jfs       8     16   2     open/syncd  /var
sp4n13: hd3              jfs       8     16   2     open/syncd  /tmp
sp4n13: hd1              jfs       1     2     2     open/syncd  /home
sp4n13: lv00             sysdump   7     7     1     open/syncd  N/A
# dsh -w sp4n13 lsvg rootvg | grep QUORUM
sp4n13: OPEN LVs:          8                                QUORUM:          1
# dsh -w sp4n13 bootlist -m normal -o
sp4n13: hdisk0
sp4n13: hdisk1
#

```

Compare the output with Figure 39 on page 192. Now, both disks (`hdisk0` and `hdisk1`) are contained in the `rootvg` root volume group. PPs are doubled by LPs by mirroring. Quorum is off, and the boot list is changed.

6.3.3 Discontinuing root volume group mirroring

If you want to discontinue the root volume group mirroring that was initiated in the previous section, use the following steps:

Step 1: Unconfigure mirroring

To unconfigure root volume group mirroring, issuing the `smitty changevg_dialog` fast path:

```
Change Volume Group Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      []          #
Start Slot                       []          #
Node Count                       []          #

OR

Node List                        [13]

Volume Group Name                [rootvg]
Physical Volume List             [hdisk0]
Number of Copies of Volume Group 1          +
Set Quorum on the Node          false       +
Boot/Install Server Node        [0]        #
Network Install Image Name      [bos.obj.ssp.432]
LPP Source Name                 [aix432]
PSSP Code Version               PSSP-3.1     +

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell     F10=Exit     Enter=Do
```

In this SMIT menu, fill in the following fields:

- In the Physical Volume List field, specify 1 physical volume.
- In the Number of Copies of Volume Group field, specify 1 copy.

By this operation, the rootvg will have only one physical volume (hdisk0) and one (1) copy of volume group.

This SMIT menu is equivalent to the `spchvgobj` command:

```
# spchvgobj -r rootvg -h hdisk0 -c 1 -n 0 \  
> -i bos.obj.ssp.432 -v aix432 -p PSSP-3.1 -l 13
```

Step 2: Discontinue mirroring

To discontinue root volume group mirroring, issue the `smitty stop_mirroring` fast path:

```
Discontinue Mirroring on a Node

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
Start Frame                          []          #
Start Slot                            []          #
Node Count                            []          #

OR

Node List                             [13]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit      F8=Image
F9=Shell     F10=Exit       Enter=Do
```

This SMIT menu is equivalent to the `spunmirrorvg` command:

```
# spunmirrorvg -l 13
sp4n13.msc.itso.ibm.com: spunmirror: Successfully reduced the volume group
rootvg by disk(s) hdisk1.
sp4n13.msc.itso.ibm.com: spunmirror: This node must be rebooted to reflect the
change in quorum!
spunmirrorvg: Volume Group rootvg on node number 13 successfully unmirrored.
spunmirrorvg: The total number of volume groups unmirrored successfully is 1.
spunmirrorvg: The total number of volume groups which failed to unmirror is 0.
spunmirrorvg: total number of invalid nodes selected = 0
spunmirrorvg: total number of invalid volume groups selected = 0
#
```

In this example, you do not need to reboot the node because you did not change the quorum.

Chapter 7. Network configuration

The network configuration on your SP system plays a vital role in the communication between the SP nodes and the control workstation (CWS). The SP Ethernet is the main channel of communication between the nodes and CWS for SP system management purposes. You may also need other networks to perform the application jobs in your SP system.

This chapter discusses how to manage your SP Ethernet adapter and other network adapters. It also discusses overall network configuration.

7.1 SP Ethernet

The SP Ethernet is a fundamental resource for the SP system because it is involved in most system management operations. It is needed when nodes are being installed and customized and when their software components are being controlled. Without it, you can not manage your nodes. Due to its importance, it is imperative that it remains functional and available.

7.1.1 Replacing an SP Ethernet adapter on a node

Replacing an SP Ethernet adapter on a node can have a serious impact on the availability of the SP Ethernet network. This section uses the following scenario: You are going to replace SP Ethernet adapter ent0 on node 8.

To perform this task, use the following steps:

Step 1: Archive the SDR

Before replacing the SP Ethernet adapter, back up the System Data Repository (SDR) by issuing the `SDRarchive` command:

```
# SDRarchive mybackup
SDRarchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note, the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2: Shut down the node

To shut down the node, issue the `cshutdown` command:

```
# cshutdown -F -N 8
```

Step 3: Replace the adapter card

Your IBM Customer Engineer performs this step.

Step 4: Delete the NIM client

Delete the NIM client definition for the node from a NIM master since the NIM database contains the hardware Ethernet address. To do this, issue the `delnimclient` wrapper:

```
# delnimclient -l 8
```

For more details, refer to 3.1.3, “Deleting a NIM client” on page 136.

Step 5: Acquire the hardware Ethernet address

To acquire the hardware Ethernet address, issue the `smitty hrdwrad_dialog` fast path:

Get Hardware Ethernet Addresses

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]	
Start Frame	[1]	#
Start Slot	[8]	#
Node Count	[1]	
OR		
Node Group	[]	+
OR		
Node List	[]	

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

Alternatively, issue the `sphrdwrad` command to perform the same operation:

```
# sphrdwrad 1 8 1
Acquiring hardware Ethernet address for node 8
Hardware ethernet address for node 8 is 10005AFA07DF
#
```

Note

If you use the `/etc/bootptab.info` file to acquire the hardware Ethernet address for the nodes, do not forget to delete the information for the Ethernet adapter to be replaced.

Step 6: Create the NIM client

To create the NIM client, issue the `mknimclient` wrapper:

```
# mknimclient -l 8
```

For more details, refer to 3.1.4, “Creating a NIM client” on page 137.

7.2 Other networks

The SP Ethernet and the SP Switch are not the only types of networks that are allowed on your SP system. The nodes and the CWS can have many different types of network adapters, such as Token-Ring or FDDI. This section will explain how to handle these additional adapters including:

- Adding an adapter
- Deleting an adapter
- Changing IP address and host name

This information will be helpful if you want to make a change on your adapter.

7.2.1 Adding a network adapter

When you add a network adapter, you can add it to a node or to the CWS. Except where noted, the following steps apply to adding a Token-Ring adapter in either situation.

Step 1: Check the device driver for the adapter

If it is the first time a Token-Ring adapter is installed to your SP system, in other words, if your SP system does not have a Token-Ring adapter, make sure the Token-Ring adapter device driver is located in the

/spdata/sys1/install/*name*/lppsource directory. You need the device driver in this directory for NIM operation.

If you do not have the device driver, do the following:

- a. Copy the device driver to the directory by using the `smitty bffcreate fast` path:

```
Copy Software to Hard Disk for Future Installation

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* INPUT device / directory for software      /dev/cd0
* SOFTWARE package to copy                  [devices.common.IBM.tok> +
* DIRECTORY for storing software package     [/spdata/sys1/install/a>
DIRECTORy for temporary storage during copying [/tmp]
EXTEND file systems if space needed?        yes +
Process multiple volumes?                   yes +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

- b. Update the NIM SPOT resource.

Refer to “Step 3: Apply PTF to SPOT” on page 113, for sample operation. A device driver can be treated the same as a PTF.

Step 2: Archive the SDR

Before adding a Token-Ring adapter, back up the SDR by issuing the `SDRArchive` command:

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note, the location and the name of the file created after you issue this command. In this example, the backup.99064.1459.mybackup file is created in the /spdata/sys1/sdr/archives directory.

Step 3: Physically install the network adapter

Your IBM Customer Engineer performs this step.

Step 4: Make the TCP/IP interface name resolution

If you use a plain /etc/hosts file for name resolution, simply add the new IP address of the Token-Ring adapter in it.

If you use Domain Name System (DNS), make sure that the name server can resolve the IP address of the Token-Ring adapter. For using DNS, refer to 7.3.2, “Using Domain Name System” on page 208.

Step 5 (for node): Enter the adapter information

All SP node adapter information must be kept in the SDR. Therefore, you need to provide any information about the Token-Ring adapter to the SDR. Issue the `smitty add_adapt_dialog` fast path:

```

Additional Adapter Database Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [1]                #
Start Slot                       [8]                #
Node Count                       [1]                #

OR

Node Group                       []                +

OR

Node List                        []

* Adapter Name                   [tr0]
* Starting Node's IP Address or Hostname [9.12.2.141]
* Netmask                        [255.255.255.0]
Additional IP Addresses          []
Ethernet Adapter Type           +
Duplex                          +
Ethernet Speed                  +
Token Ring Data Rate            +
Skip IP Addresses for Unused Slots? no                +
Enable ARP for the css0 Adapter? yes                 +
Use Switch Node Numbers for css0 IP Addresses? yes    +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit         Enter=Do

```

Alternatively, issue the `spadaptrs` command as follows:

```
# spadaptrs 1 8 1 tr0 9.12.2.141 255.255.255.0
```

This example uses slot numbers for assigning the IP address, and the Token-Ring adapter is installed on node 8.

Step 5 (for CWS): Enter the adapter information

If you are adding the Token-Ring adapter to the CWS, configure the adapter using the `smitty chinnet fast path`:


```

Change / Show a Token-Ring Network Interface

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
Network Interface Name                tr0
INTERNET ADDRESS (dotted decimal)    [9.12.0.4]
Network MASK (hexadecimal or dotted [255.255.255.0]
Current STATE                         up           +
Use Address Resolution Protocol (ARP)? yes         +
Enable Hardware LOOPBACK Mode?      no          +
BROADCAST ADDRESS (dotted decimal)  []
Confine BROADCAST to LOCAL Token-Ring? no          +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do

```

Enter the required adapter information for the Token-Ring adapter according to your IP network parameters.

Step 6: Run the *spbootins* command

To create a principal and a service key file for the Token-Ring adapter, set node 8 to customize mode and run the `setup_server` command. To do this, issue the `smitty server_dialog` fast path:

```

                                Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [1]                      #
Start Slot                       [8]                      #
Node Count                       [1]                      #

OR

Node List                        []

Response from Server to bootp Request  customize          +
Volume Group Name                  []
Run setup_server?                  yes                +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit           Enter=Do

```

Alternatively, issue the `spbootins` command:

```
# spbootins -r customize 1 8 1
```

In the case you add the adapter to CWS, issue the `setup_server` command.

Step 7 (for node): Reboot the node

Reboot the node. Refer to 5.1.3, “Customizing a node with or without rebooting” on page 164 for this operation.

Step 9: Distribute the Kerberos database

If you use secondary authentication servers, make sure that the database is propagated to them. Issue the `push-kprop` command to propagate the database.

Note

The following adapters can be added by this method:

- Ethernet (en)
- FDDI (fi)
- Token-Ring (tr)
- SP switch (css)

To add adapters, such as ESCON-adapter, you must add the adapter manually on each node using the `dsh` command or modify the `/tftpboot/firstboot.cust` script.

7.2.2 Deleting a network adapter

When you delete a network adapter, you can delete it from a node or from CWS. Except where noted, the following steps apply to deleting a Token-Ring adapter in either situation.

Step 1: Archive the SDR

Before deleting a Token-Ring adapter, back up the SDR by issuing the `SDRArchive` command:

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note, the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2 (for node): Delete the adapter information

To delete the Token-Ring adapter information from the SDR, issue the `spdeladap` command:

```
# spdeladap 1 8 1 tr0
```

Step 2 (for CWS): Delete the adapter information

If you are deleting the Token-Ring adapter from the CWS, issue the `rmdev` command to remove the Token-Ring adapter information from Object Data Manager (ODM):

```
# rmdev -l tr0 -d
```

Step 3: Run the `spbootins` command

To delete a principal and a service key file for the Token-Ring adapter, set node 8 to customize mode and run the `setup_server` command. To do this, issue the `smitty server_dialog` fast path:

Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]	
Start Frame	[1]	#
Start Slot	[8]	#
Node Count	[1]	#
OR		
Node List	[]	
Response from Server to bootp Request	customize	+
Volume Group Name	[]	
Run <code>setup_server</code> ?	yes	+

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

Alternatively, issue the `spbootins` command:

```
# spbootins -r customize 1 8 1
```

To delete the adapter from CWS, issue the `setup_server` command.

Step 4 (for node): Reboot the node

Reboot the node. Refer to 5.1.3, “Customizing a node with or without rebooting” on page 164 for this operation.

Step 7: Distribute the Kerberos database

If you use secondary authentication servers, make sure that the database is propagated to them. Issue the `push-kprop` command to propagate the database.

Step 8: Shut down and power off

To delete the Token-Ring adapter physically, power off the node or CWS. Issue the `shutdown` command:

```
# shutdown -F
```

Step 9: Delete the adapter physically

Your IBM Customer Engineer performs this step.

7.2.3 Changing an IP address and host name

The IP addresses and host names of the other network adapters can be changed the same way as adding adapters.

Follow the steps described in 7.2.1, “Adding a network adapter” on page 199, starting from Step 4.

7.3 Global configuration

This section describes network adapter-independent topics.

7.3.1 Changing an initial host name

For each node, you can choose network interface for the initial host name. To store this information to SDR, issue the `smitty hostname_dialog` fast path. In the following example, `tr0` is selected as an initial host name for node 9:

Hostname Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]													
Start Frame	[]	#												
Start Slot	[]	#												
Node Count	[]	#												
OR														
Node Group	[]	+												
OR														
Node List	[9]													
Adapter Name used for Initial Hostname	tr0	+												
Use Short or Long Hostnames	short	+												
<table style="width: 100%; border: none;"> <tr> <td style="width: 25%;">F1=Help</td> <td style="width: 25%;">F2=Refresh</td> <td style="width: 25%;">F3=Cancel</td> <td style="width: 25%;">F4=List</td> </tr> <tr> <td>F5=Reset</td> <td>F6=Command</td> <td>F7=Edit</td> <td>F8=Image</td> </tr> <tr> <td>F9=Shell</td> <td>F10=Exit</td> <td>Enter=Do</td> <td></td> </tr> </table>			F1=Help	F2=Refresh	F3=Cancel	F4=List	F5=Reset	F6=Command	F7=Edit	F8=Image	F9=Shell	F10=Exit	Enter=Do	
F1=Help	F2=Refresh	F3=Cancel	F4=List											
F5=Reset	F6=Command	F7=Edit	F8=Image											
F9=Shell	F10=Exit	Enter=Do												

In the Adapter Name used for Initial Hostname field, select one adapter available on the node.

This SMIT operation is equivalent to the `sphostnam` command:

```
# sphostnam -l 9 -a tr0 -f short
```

The information for the adapter must be stored in Adapter SDR class before this operation. If this is not the case, issue the `smitty add_adapt_dialog fast` path or `spadaptrs` command to do it. Refer to “Step 5 (for node): Enter the adapter information” on page 201, for a sample operation.

Then perform a customize boot of the target node. Refer to 5.1.3, “Customizing a node with or without rebooting” on page 164, for a sample operation.

7.3.2 Using Domain Name System

It is getting very common to use Domain Name System (DNS) server to resolve host names. Therefore, there is a strong demand to implement DNS for the SP system.

AIX 4 provides the `/etc/netsvc.conf` file to control the order for the name resolution method. You can specify using `/etc/hosts` first, and if the name cannot be resolved, then using DNS. The following is the contents of the `netsvc.conf` file for this:

```
hosts=local,bind
```

If you create the `/etc/resolv.conf` file on CWS and all nodes, you do not need to change any SP system configuration. It is recommended that you use this method instead of complete DNS integration.

Part 3. Controlling and monitoring the SP system

Chapter 8. Hardware information

Part of the job of an SP system administrator is to know your hardware. On an SP system, with multiple frames and numerous nodes, this information can seem overwhelming. But, knowing exactly what is on your SP system is important, particularly if you are planning on system expansions or if you need to talk to the IBM Support Center.

Getting the hardware information about your SP system means using two different sources of information:

- Vital Product Data (VPD), which is provided by AIX.
- System Data Repository (SDR), which is provided by IBM Parallel System Support Programs for AIX (PSSP).

VPD contains configuration information and can be accessed with the `lscfg` command. It tells you about all devices and components installed in the SP nodes or the control workstation (CWS) on which the command is executed.

SDR provides you with the information not known to AIX. This information is important when looking for SP frames, SP-attached servers, SP nodes, and SP Switches.

This chapter discusses how to use system commands to access the information desired. You can use Hardware Perspective as well.

8.1 Information in VPD

Vital Product Data (VPD) is the repository of your SP system that contains hardware information for all of your devices on nodes or CWS; it is provided by AIX. This section shows several ways that you can get access to VPD information. It also demonstrates how you can explore your hardware on your nodes and the CWS.

8.1.1 Accessing SP nodes information

AIX provides the hardware information to VPD. There are several ways to access VPD information:

- Using the `lscfg` command
- Reading the `node_number.lscfg` file
- Using the `diag` command

The *lscfg* command

To access hardware information, use the `lscfg` command. Issue the command on the node in which you are interested. The following is an excerpt of its informative output:

```
INSTALLED RESOURCE LIST WITH VPD

The following resources are installed on your machine.

Model Architecture: rs6k
Model Implementation: Uni-Processor, MCA bus

sysplanar0          00-00          CPU Planar

    Part Number.....08184781
    EC Level.....00D29534
    Processor Identification...00050930
    ROS Level and ID.....IPLVER1.0 LVL2.08,08184919
    Processor Component ID.....0100006700000014
    Device Specific.(Z0).....012048
    Device Specific.(Z1).....021047
    Device Specific.(Z2).....03204E
    Device Specific.(Z3).....041149
    Device Specific.(Z4).....CD234D
    Device Specific.(Z5).....FFFFFF
    Device Specific.(Z6).....0A114A
    Device Specific.(Z7).....2A114A
    Device Specific.(Z8).....FFFFFF
    Device Specific.(Z9).....FFFFFF
    ROS Level and ID.....OCS(07050000)
    ROS Level and ID.....SEEDS(AABAEAA2)

proc0              00-00          Processor
mem0               00-0B          128 MB Memory Card

    Device Specific.(Z3).....08
    EC Level.....35
    Device Specific.(Z0).....00
    Device Specific.(Z1).....00
    Device Specific.(Z2).....01
    Size.....128
```

The *node_number.lscfg* file

On the other hand, the `/var/adm/SPlogs/SPconfig` directory on the CWS already contains this kind of information about nodes in your SP system. The file name is `node_number.lscfg`:

```

# cd /var/adm/SPlogs/SPconfig
# ls -al *.lscfg
-rw-r--r-- 1 root system 5490 Mar 17 10:41 10.lscfg
-rw-r--r-- 1 root system 5490 Mar 17 11:13 11.lscfg
-rw-r--r-- 1 root system 5490 Mar 17 11:13 12.lscfg
-rw-r--r-- 1 root system 6031 Mar 17 11:13 13.lscfg
-rw-r--r-- 1 root system 5983 Mar 17 11:14 14.lscfg
-rw-r--r-- 1 root system 7371 Mar 17 15:43 15.lscfg
-rw-r--r-- 1 root system 5130 Mar 17 10:48 5.lscfg
-rw-r--r-- 1 root system 5130 Mar 17 14:51 6.lscfg
-rw-r--r-- 1 root system 5012 Mar 17 10:40 7.lscfg
-rw-r--r-- 1 root system 5838 Mar 17 10:42 8.lscfg
-rw-r--r-- 1 root system 6015 Mar 17 10:42 9.lscfg
#

```

These files include detailed information about the installed components, such as:

- System planar
- Memory
- Processor type
- Number of processors
- Bus architecture
- Adapter cards (for example: Network, multimedia, peripheral devices)
- Hard disks

Attributes describing the components even include:

- EC level
- Device driver level
- Part number
- Serial number
- Manufacturer
- Processor IDs

You will be provided with in-depth information from the VPD.

The diag command

Another way to see the VPD is through diagnostics. This is useful if you cannot boot your system, but you can get it up in diagnostic mode. Simply issue the `diag` command, and you will see the following menu:

LICENSED MATERIAL and LICENSED INTERNAL CODE - PROPERTY OF IBM
(C) COPYRIGHTS BY IBM AND BY OTHERS 1982, 1998.
ALL RIGHTS RESERVED.

These programs contain diagnostics, service aids, and tasks for the system. These procedures should be used whenever problems with the system occur which have not been corrected by any software application procedures available.

In general, the procedures will run automatically. However, sometimes you will be required to select options, inform the system when to continue, and do simple tasks.

Several keys are used to control the procedures:

- The Enter key continues the procedure or performs an action.
- The Backspace key allows keying errors to be corrected.
- The cursor keys are used to select an option.

Press the F3 key to exit or press Enter to continue.

Press **Enter** to continue. You will see the following menu:

Move cursor to selection, then press Enter.

Diagnostic Routines

This selection will test the machine hardware. Wrap plugs and other advanced functions will not be used.

Advanced Diagnostics Routines

This selection will test the machine hardware. Wrap plugs and other advanced functions will be used.

Task Selection(Diagnostics, Advanced Diagnostics, Service Aids, etc.)

This selection will list the tasks supported by these procedures. Once a task is selected, a resource menu may be presented showing all resources supported by the task.

Resource Selection

This selection will list the resources in the system that are supported by these procedures. Once a resource is selected, a task menu will be presented showing all tasks that can be run on the resource(s).

F1=Help

F10=Exit

F3=Previous Menu

Highlight **Task Selection** and press **Enter**. You will see the following menu:

From the list below, select a task by moving the cursor to the task and pressing 'Enter'.
To list the resources for the task highlighted, press 'List'.

[TOP]

Run Diagnostics
Display or Change Diagnostic Run Time Options
Display Service Hints
Display Previous Diagnostic Results
Display Hardware Error Report
Display Software Product Data
Display Configuration and Resource List
Display Hardware Vital Product Data
Display Resource Attributes
Change Hardware Vital Product Data
Format Media
Certify Media

[MORE...20]

F1=Help

F4=List

F10=Exit

Enter

F3=Previous Menu

Select **Display Hardware Vital Product Data** and press **Enter**. This will allow you to select a device and see the identical output of that from the `lscfg` command described at the beginning of this section.

The VPD is kept inside the Object Data Manager (ODM). The ODM files are located in the `/etc/objrepos` directory. It is strongly recommended that you do not modify or alter these files because it can cause major corruption on your system. However, it is OK to look at the information as long as it is not changed. To see how the VPD information is stored in the ODM, you can issue the `odmget` command:

```

#odmget CuVPD | pg
CuVPD:
    name = "sysplanar0"
    vpd_type = 0
    vpd = "**PN065G6518*EC00D18816*PI00000973*RLIPLVER1.0 LVL2.01,051G9929*PC
\n\
0100006700000014*Z0012048*Z1021047*Z203204E*Z3041149*Z4CD234D*Z5FFFFFF*Z60A114A*
Z72A114A*Z8FFFFFF*Z9FFFFFF*RL   OCS(07050000) *RL\n\
SEEDS(7334FC00) "

CuVPD:
    name = "ioplanar0"
    vpd_type = 0
    vpd = "**EC28"

CuVPD:
    name = "sio0"
    vpd_type = 0
    vpd = "**PN 43G2200*ECC74350*SN00001616*FN 43G2211*MFIBM97N*DSTANDARD I/O
*DD00*DG02*RL0000"

```

As you can tell from this output, the information in the ODM is not formatted in a user-friendly way. All the information is there, but it uses abbreviation and long single lines containing numerous bits of information. The `lscfg` command takes this information, reformats it, and presents it with a nice description of what each value means. This is why it is recommended that you use the `lscfg` command instead of going straight to the ODM for details.

8.1.2 Accessing control workstation information

To obtain CWS hardware information, you can apply nearly the same method you used with SP nodes.

However, there is one difference. The `/var/adm/SPlogs/SPconfig` directory does not contain VPD information for the CWS. Therefore, use the `lscfg` or `diag` command as described previously.

8.2 Information in SDR

PSSP uses the System Data Repository (SDR) as one of its main data depositories for an SP system. It also contains information about the hardware on the SP frames, the SP-attached servers, SP nodes, and SP Switches. There are several commands you can use to get specific data on a device. This section explains how to use these system commands, and Hardware Perspective, to get the information you are looking for.

For additional details about Hardware Perspective, refer to 9.2, “Using Hardware Perspective” on page 233.

For further information on the SDR class, refer to Appendix G, “The System Data Repository”, in the *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

8.2.1 Accessing SP frames information

The information about SP frames resides in the Frame SDR class. To get information about SP frames, issue the `SDRGetObjects` command:

```
# SDRGetObjects Frame
frame_number tty      frame_type  MACN      backup_MACN slots      fr
ame_in_config snn_index switch_config hardware_protocol s1_tty
          1 /dev/tty0  switch      sp3en0      ""          16
#           1          0          0 SP          ""
```

The attributes displayed tell you the following details about the frame:

- frame_type** This attribute indicates the type of frame. Possible values are: *switch* for switched frames, *noswitch* for switchless frames, and *allswitch* for frames consisting of SP Switch boards only.
- slots** This attribute indicates the number of slots in the frame. Possible values are: *16* for 1.93 m frames, *8* for 1.25 m frames and *1* for SP-attached servers.
- hardware_protocol** This attribute indicates the hardware protocols the frame uses. Possible values are: *SP* for SP frames and *SAMI* for SP-attached servers.
- s1_tty** This attribute indicates the tty port for s1term connection for SP-attached servers.

The `sp1stdata` command shows you similar information but in a more readable format:

```
# splstdata -f
List Frame Database Information

frame#          tty          s1_tty      frame_type  hardware_protocol
-----
      1         /dev/tty0          ""          switch                SP
#
```

The Hardware Perspective displays similar information as shown in Figure 40:

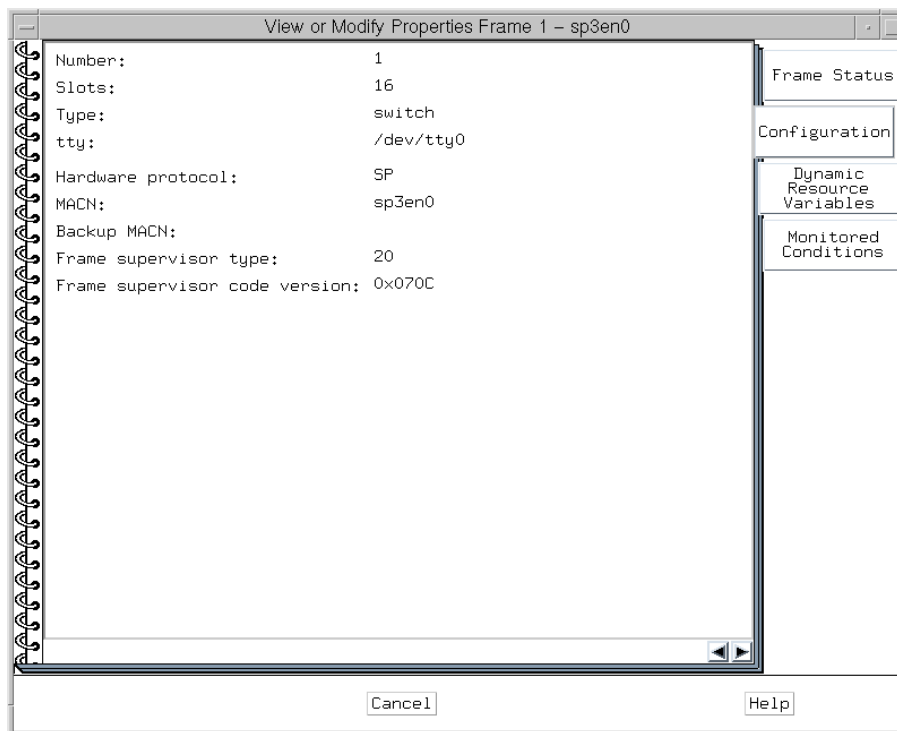


Figure 40. Hardware Perspective frame properties

8.2.2 Accessing SP-attached servers information

An SP-attached server has both frame and node personalities. In other words, you can find information about an SP-attached server in both the Frame and Node SDR class.

8.2.2.1 SP frame personality

Issue the `SDRGetObjects` command to explore an SP frame personality:

```
# SDRGetObjects Frame
frame_number tty      frame_type  MACN      backup_MACN  slots      fr
ame_in_config ssn_index  switch_config hardware_protocol s1_tty
      1 /dev/tty0  switch      s70test.ppd.pok.ibm.com ""
16      1      0      0 SP      ""
      2 /dev/tty2  ""          s70test.ppd.pok.ibm.com ""
1 ""      ""          ""          SAMI      /dev/tty3
#
```

The attributes displayed tell you the following details about the frame:

- frame_type** This attribute indicates the type of frame. Possible values are: *switch* for switched frames, *noswitch* for switchless frames, and *allswitch* for frames consisting of SP Switches only.
- slots** This attribute indicates the number of slots in the frame. Possible values are: *16* for 1.93 m frames, *8* for 1.25 frames, and *1* for SP-attached servers.
- hardware_protocol** This attribute indicates the hardware protocols the frame uses. Possible values are: *SP* for SP frames and *SAMI* for SP-attached servers.
- s1_tty** This attribute indicates the tty port for s1term connection for SP-attached servers.

The `splstdata` command shows you similar information but in a more readable format:

```
# splstdata -f
List Frame Database Information
frame#      tty      s1_tty      frame_type  hardware_protocol
-----
1      /dev/tty0      ""          switch      SP
2      /dev/tty2      /dev/tty3      ""          SAMI
#
```

The Hardware Perspective displays similar information as shown in Figure 41 on page 222:

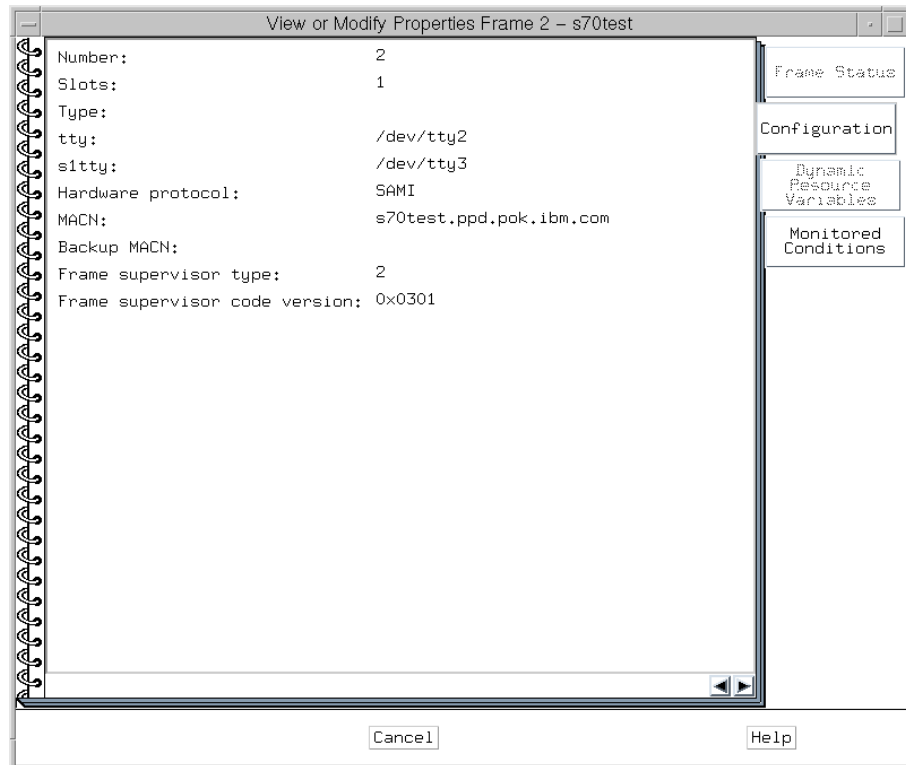


Figure 41. Hardware Perspective frame properties of an SP-attached server

8.2.2.2 SP node personality

Issue the `SDRGetObjects` command to explore an SP node personality:

```

# SDRGetObjects Node
node_number hdw_enet_addr frame_number slot_number slots_used switch_node_number
switch_chip_port switch_chip switch_number initial_hostname reliable_hostname
default_route boot_server bootdisk install_image install_disk last_inst
all_time last_install_image switch_protocol switch_partition_number bootp_respon
se_boot_device usr_maint JM_domain_id VSD_adapter VSD_max_buffer_count VSD_
request_blocks VSD_pbufs VSD_init_buffer_count VSD_min_buddy_buffer_size VSD_
max_buddy_buffer_size VSD_max_buddy_buffers VSD_do_ip_checksum cfs_adapter acct
_class_id acct_enable acct_job_charge acct_excluse_enable usr_server_ip usr_cli
ent_adapter has_usr_clients code_version usr_gateway_ip VSD_maxIPmsgsz lppsourc
e_name processors_installed processor_type description platform hardware_con
trol_type RVSD_version selected_vg dcehostname
1 08005A750830 1 1 2 0
3 5 1 ravtstn17.ppd.pok.ibm.com ravtstn17.ppd.pok
.ibm.com 9.114.49.125 0 hdisk0 default hdisk0 Wed_Ma
r_24_08:55:36_EST_1999 default IP 1 disk en0
false "" css0 256 256
48 64 4096 65536 4 false ""
default default 1.0 false local en0
false PSSP-3.1 0 24576 aix432 1
UP 135_MHz_P2SC_Wide rs6k 83 3010000 rootvg
""
17 02070123848B 2 1 1 1
2 5 1 ravtstn01.ppd.pok.ibm.com ravtstn01.ppd.pok
.ibm.com 9.114.49.125 0 hdisk0 default hdisk0 Wed_Ma
r_24_13:57:49_EST_1999 default IP 1 disk en0
false "" css0 256 256
48 64 4096 65536 4 false ""
default default 1.0 false local en0
false PSSP-3.1 0 24576 aix432 8
MP 7017-S70 chrp 10 3010000 rootvg ""
#

```

Figure 42. Node SDR class contents

The attributes displayed tell you the following details about the node:

- processor_type** This attribute indicates the type of processor. Possible values are: *MP* for multiprocessor nodes and *UP* for uniprocessor nodes.
- description** This attribute describes the hardware type of the node.
- platform** This attributes defines the architecture of the node. Possible values are: *chrp* for IBM PowerPC Common Hardware Reference Platform (CHRP), *rs6k* for RS/6000 architecture, and *rspc* for IBM PowerPC Personal Computer.

hardware_control_type This attribute defines the hardware used to control the node. If a supervisor card is used by the node, the attribute contains the card type.

From the description attribute, 7017-S70, you know that node 17 is an SP-attached server instead of an SP node. Referring to the NodeControl SDR class gives you information about the possible values for the different kinds of hardware control filling the hardware_control_type attribute in the Node class. To display this information, issue the `SDRGetObjects` command (the output is reformatted for readability):

```
# SDRGetObjects NodeControl
type      capabilities slots_used platform_type processor_type NC_timeout
115      power, reset, tty, keySwitch, LED, networkBoot      2 rs6k UP
65       power, reset, tty, keySwitch, LED, networkBoot      1 rs6k UP
11       power, reset, tty                                     1 "" "" ""
161      power, reset, tty, keySwitch, LCD, networkBoot        4 rs6k MP
33       power, reset, tty, keySwitch, LED, networkBoot      1 rs6k UP
10       power, tty, LCD, networkBoot                          1 chrp MP 3600
177      power, reset, tty, LCD, networkBoot                  1 chrp MP 1200
83       power, reset, tty, keySwitch, LED, networkBoot      2 rs6k UP
178      power, reset, tty, LCD, networkBoot                  2 chrp MP 1200
97       power, reset, tty, keySwitch, LED, networkBoot      1 rs6k UP
113      power, reset, tty, keySwitch, LED, networkBoot      2 rs6k UP
81       power, reset, tty, keySwitch, LED, networkBoot        2 rs6k UP
#
```

Figure 43. NodeControl SDR class contents

From the contents of Node SDR class shown in Figure 42 on page 223, you know that node 17 uses hardware_control_type 10. If you look this value up in Figure 43 on page 224, you see the following:

- You can control/monitor power, tty, LCD, or network boot for node 17, and SP-attached server.

The `splstdata` command shows you similar information but in a more readable format:

```

# splstdata -n
List Node Configuration Information

node# frame# slot# slots initial_hostname reliable_hostname dcehostname
default_route processor_type processors_installed description
-----
  1      1      1      2 rvtstn17.ppd.po rvtstn17.ppd.po ""
    9.114.49.125 UP 1 135_MHz_P2SC_Wide
17      2      1      1 rvtstn01.ppd.po rvtstn01.ppd.po ""
    9.114.49.125 MP 8 7017-S70
#

```

The Hardware Perspective displays similar information as shown in Figure 44:

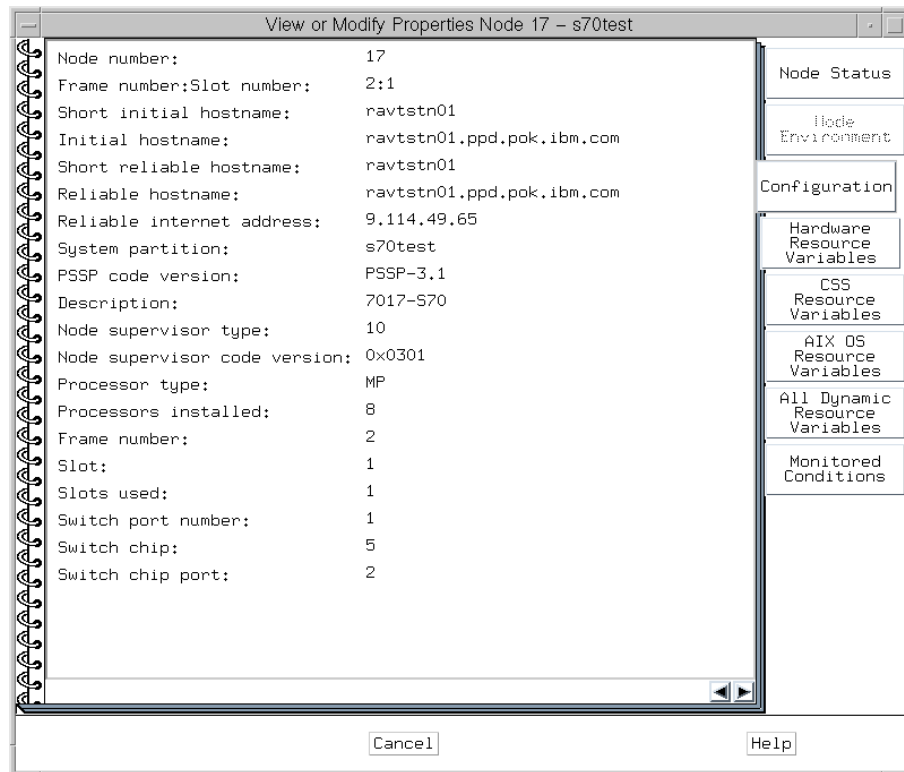


Figure 44. Hardware Perspective node properties of an SP-attached server

8.2.3 Accessing SP nodes information

Compared with SP frame or SP Switch, SP node has a lot of hardware information. Accessing the SDR for hardware information about a particular

node will give you a high level view of the node types used in your SP system. Issue the `SDRGetObjects` command to explore your SP node:

```
# SDRGetObjects Node
node_number hdw_enet_addr frame_number slot_number slots_used switch_node_nu
mber switch_chip_port switch_chip switch_number initial_hostname reliable_hostn
ame default_route boot_server bootdisk install_image install_disk last_inst
all_time last_install_image switch_protocol switch_partition_number bootp_respon
se boot_device usr_maint JM_domain_id VSD_adapter VSD_max_buffer_count VSD_
request_blocks VSD_pbufs VSD_init_buffer_count VSD_min_buddy_buffer_size VSD_
max_buddy_buffer_size VSD_max_buddy_buffers VSD_do_ip_checksum cfs_adapter acct
_class_id acct_enable acct_job_charge acct_exclude_enable usr_server_ip usr_cli
ent_adapter has_usr_clients code_version usr_gateway_ip VSD_maxIPmsgsz lppsouce
_name processors_installed processor_type description platform hardware_con
trol_type RVSD_version selected_vg dcehostname
1 02608CF534CC 1 1 4 0
3 5 1 sp3n01 sp3n01.msc.itso.ibm.com 192.16
8.3.130 0 hdisk0 default hdisk0 Tue_Mar_16_14:25:21
EST_1999 default IP 1 disk en0 false
"" css0 256 256 48
256 4096 262144 2 false "" default
default 1.0 false local en0 false
PSSP-3.1 0 61440 aix432 8 MP 11
2_MHz_SMP_High_rs6k 161 3010000 rootvg ""
15 02608CF53344 15 15 2 14
2 7 1 sp3n15 sp3n15.msc.itso.ibm.com 192.16
8.31.1 1 hdisk0 default hdisk0 Wed_Mar_17_15:42:18_E
ST_1999 default IP 1 disk en0 false
"" css0 256 256 48
256 4096 262144 2 false "" default
default 1.0 false local en0 false
PSSP-3.1 0 61440 aix432 1 UP 66
_MHz_PWR2_Wide_rs6k 81 3010000 rootvg ""
#
```

Figure 45. Node SDR class contents

The attributes displayed tell you the following details about the node:

- processor_type** This attribute indicates the type of processor. Possible Values are: *MP* for multiprocessor nodes and *UP* for uniprocessor nodes.
- description** This attribute describes the hardware type of the node.
- platform** This attribute defines the architecture of the node. Possible values are: *chrp* for IBM PowerPC Common Hardware Reference Platform (CHRP), *rs6k* for RS/6000 architecture, and *rspc* for IBM PowerPC Personal Computer.

hardware_control_type This attribute defines the hardware used to control the node. If a supervisor card is used by the node, the attribute contains the card type.

From the contents of Node SDR class shown in Figure 45 on page 226, you know that the node 1 uses hardware_control_type 161, and node 15 uses hardware_control_type 81. If you look up these values shown in Figure 43 on page 224, you see the following:

- You can control/monitor power, reset, tty, key switch, LCD, or network boot for node 1.
- You can control/monitor power, reset, tty, key switch, LED, or network boot for node 15.

The `splstdata` command shows you similar information but in a more readable format:

```
# splstdata -n
                                List Node Configuration Information

node# frame# slot# slots  initial_hostname  reliable_hostname  dcehostname
      default_route  processor_type  processors_installed  description
-----
   1     1     1     4  sp3n01                sp3n01.msc.itso.  ""
                        192.168.3.130    MP                8 112_MHz_SMP_High
  15     1    15     2  sp3n15                sp3n15.msc.itso.  ""
                        192.168.31.1      UP                1 66_MHz_PWR2_Wide

#
```

The Hardware Perspective displays similar information as shown in Figure 46 on page 228:

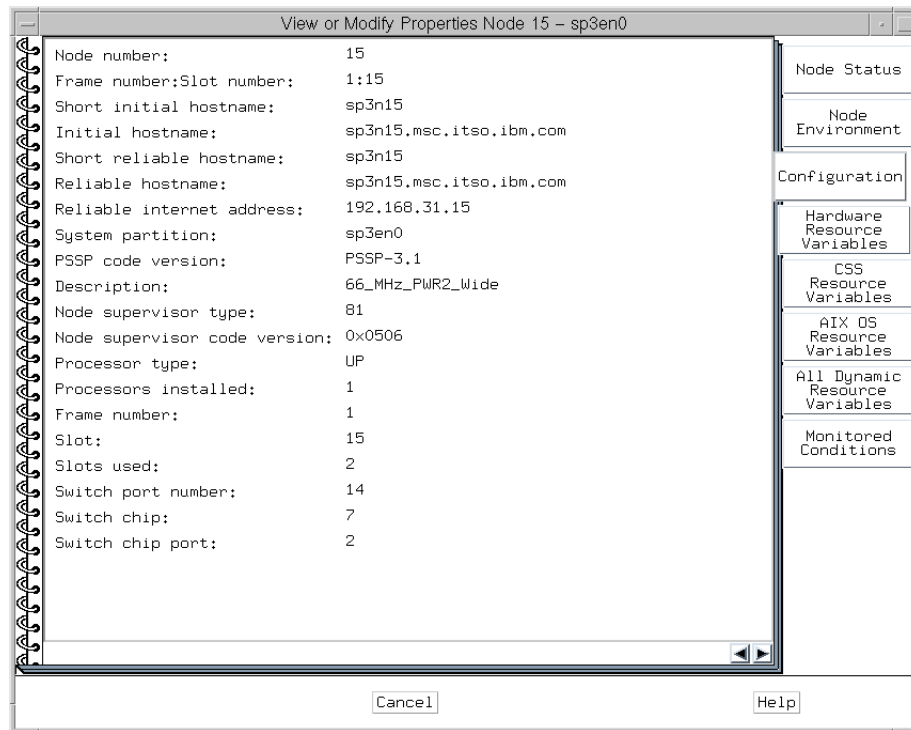


Figure 46. Hardware Perspective node properties

8.2.4 Accessing SP Switch information

Similar to the SP frame, the SP Switch is a component that does not have very much hardware information to offer because there are not many different types of SP Switches.

Issue the `SDRGetObjects` command to explore your SP Switch:

```
# SDRGetObjects Switch
switch_number frame_number slot_number switch_partition_number switch_type clo
ck_input switch_level switch_name clock_source clock_change
      1      1      1      17      1      129 0
      1 SP_Switch      ""      no
#
```

The attributes displayed tell you the following details about the switch:

- switch_level** This attribute indicates the level of the SP Switch supervisor microcode.
- switch_name** This attribute indicates the type of switch. If the switch is an SP Switch, the value is: *SP_Switch*.
- switch_type** This attribute indicates the type of supervisor card used for the switch board. The SP Switch board uses the value: 129.

The `splstdata` command shows you similar information but in a more readable format:

```
# splstdata -s
List Node Switch Information

node# initial_hostname  switch node# protocol number  switch chip chip_port
-----
   1 sp3n01              0      IP      1      5      3
  15 sp3n15             14      IP      1      7      2

switch frame slot switch_partition switch clock switch
number number number          number      type input level
-----
   1     1    17              1      129    0

switch_part topology primary arp switch_node
number filename name enabled nos._used
-----
   1 expected.top.an sp3n01.msc.itso. yes no

#
```

The Hardware Perspective displays similar information as shown in Figure 47 on page 230:

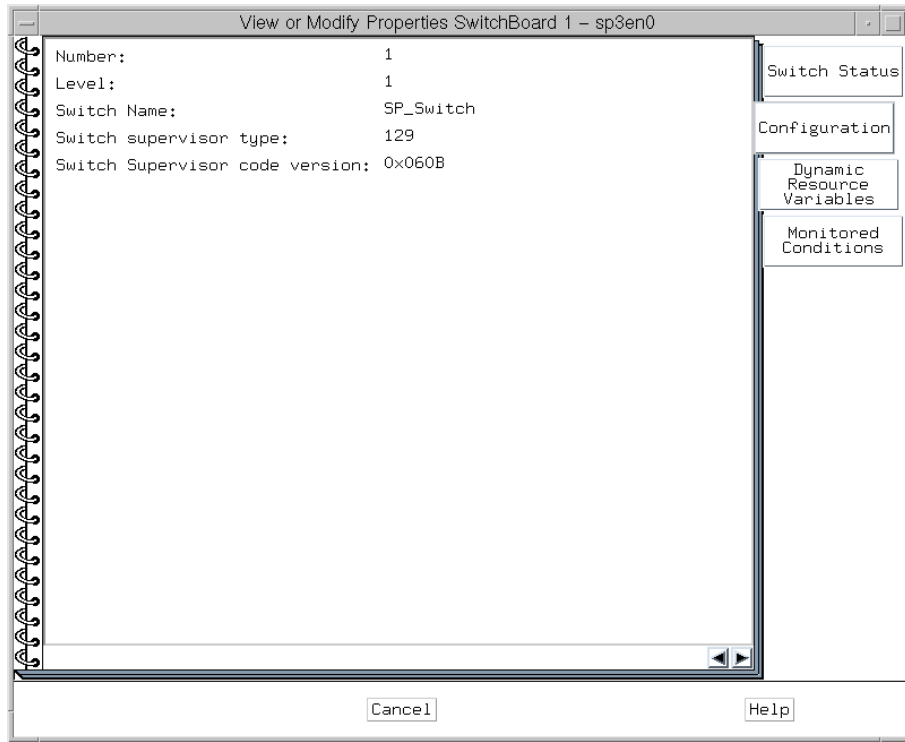


Figure 47. Hardware Perspective SP Switch board properties

Chapter 9. Controlling hardware

The control workstation (CWS) is a single point of control for the entire SP system because it can oversee the entire SP system and issue commands to a specific frame, node, or SP Switch if required. This means that the CWS can also control the hardware for the entire SP system.

This chapter explains how you can control your SP hardware from the CWS. The different tools that are available to you are discussed. The graphical user interface (GUI), called Hardware Perspective, is one option for you to use, if you prefer icons and buttons to handle your hardware. The alternative is to use command line tools. There are a couple of different command line tools at your disposal for controlling hardware. In this chapter, these different commands are explained and demonstrated by examples.

This chapter is useful for system administrators who would like to know various ways to control their hardware.

9.1 Controlling tools

There are two tools for taking control of your SP hardware:

- Hardware Perspective
- Command line tools

The command line tools are two commands providing the interface to the SP hardware:

- `spmon`
- `hmcmds`

The successor to the `spmon` GUI command, Hardware Perspective, gives you an easy look and feel when controlling the SP hardware. As the name suggests, this graphical tool looks at your SP hardware from different perspectives. The big advantage of the Hardware Perspective is that it enables you to see your hardware in different layers. Using an overview as a starting point, you can dig deeper into the SP hardware and look for the little details you are keen to find out.

On the other hand, command line tools provide you with detailed information; so, you will not miss a thing. This chapter concentrates on using the `spmon` command to take control of the SP hardware. The `spmon` command can be

used for monitoring the SP hardware as well. This is discussed in Chapter 10, “Monitoring hardware” on page 249.

Both approaches are good ones. Which one you prefer depends on the way you like to work with your SP hardware. However, a mixture of both can be especially useful.

9.1.1 Creating a Kerberos principal

To use the controlling tools, you must be successfully authenticated as a Kerberos principal. By default, there is a root.admin Kerberos principal available. If you need to have another Kerberos principal, you must add it.

To learn how you can add a Kerberos principal, refer to 16.4.3, “Adding a principal” on page 461.

9.1.2 Authorizing a Kerberos principal

Before you use the Hardware Perspective or command line tools, you must establish certain authorizations. Both Hardware Perspective and command line tools use the SP System Monitor. If your Kerberos principal is not registered in the SP System Monitor Access Control Lists (ACLs), you cannot use SP System Monitor.

Use the following procedures to authorize your Kerberos principal to use the SP System Monitor.

Step 1: Add a Kerberos principal

To authorize the Kerberos principal to use SP System Monitor, add the Kerberos principal to the hmacls file in the /spdata/sys1/spmon directory on the CWS. You can edit this file with your favorite text editor. In the case you want to add a Kerberos principal, UserA.admin, the file looks as follows:

```
# cat /spdata/sys1/spmon/hmacls
sp3en0 root.admin a
sp3en0 hardmon.sp3en0 a
sp3en0 UserA.admin a
1 root.admin vsm
1 hardmon.sp3en0 vsm
1 UserA.admin vsm
#
```

The third line, `sp3en0 UserA.admin a`, gives the following permission to the UserA.admin Kerberos principal:

- Administrative authority (by the last word *a*) to control SP System Monitor on the CWS (by the first word *sp3en0*. This is the CWS in your SP system).

The sixth line, `1 UserA.admin vsm`, gives the following permissions to the UserA.admin Kerberos principal:

- Virtual Front Operator Panel (VFOP) permission (by the character *v* in the last word *vsm*) to issue commands to the hardware in the first frame (by the first word *1*).
- Serial link permission (by the character *s* in the last word *vsm*) to read and write to a serial port in the first frame.
- Monitor permission (by the character *m* in the last word *vsm*) to receive state changes for the first frame.

Control is provided by the VFOP. VFOP is a set of commands that can be sent to the SP hardware components contained in one or more SP frames. Each SP frame consists of 18 slots, numbered 0 through 17, where slot 0 represents the SP frame supervisor card, slot 17 can represent an SP Switch supervisor card, and slots 1 through 16 can represent SP node supervisor cards.

Step 2: Activate permission

To activate these permissions, refresh the SP System Monitor. To do this, issue the `hmadm` command:

```
# hmadm setacl
```

By issuing this command, the SP System Monitor daemon reads the SP System Monitor ACL files to update the daemon's internal ACL tables. UserA.admin principal should now be able to use SP System Monitor.

For more information about SP System Monitor authorization, refer to Chapter 19, "Using the SP Command Line System Monitor" in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

9.2 Using Hardware Perspective

One of the advantages of Hardware Perspective is that it is a very consistently designed tool. This means that once you understand how to deal with one SP hardware component, you can manage the others as well. Since the scheme stays the same from one component to another, this section does not cover all components belonging to the SP hardware. Instead, it focuses

on SP nodes because they are a good example, and you will mostly deal with them.

To learn more about Hardware Perspective, refer to Chapter 3, “Using the Hardware Perspective Effectively” in *SP Perspectives: A New View of Your SP System*, SG24-5180.

9.2.1 Starting Hardware Perspective

Start the Hardware Perspective by issuing the `sphardware` command:

```
# sphardware
```

After a title screen, the Hardware Perspective window is displayed as shown in Figure 48.

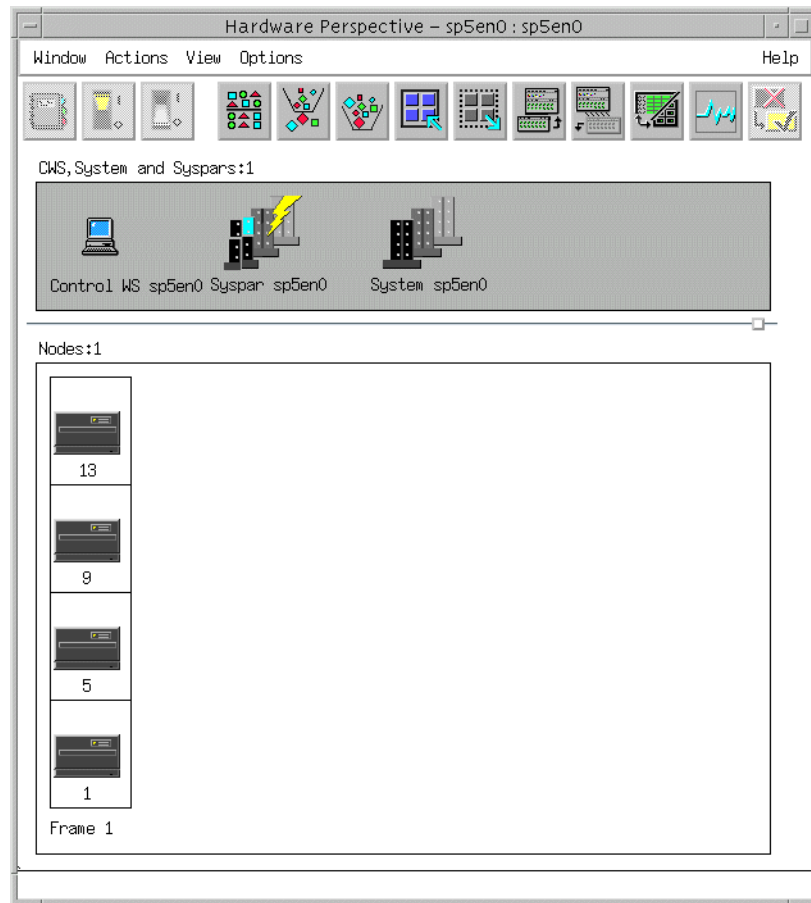


Figure 48. Hardware Perspective

As you notice, it only shows the SP nodes of the SP system. Since you will want to manage SP frames and switches a well, display these additional components using the following steps:

1. To show the Add Pane dialog box, click the **add pane** icon:



2. In the Pane type field, select **Frames and Switches** from the pull-down list:

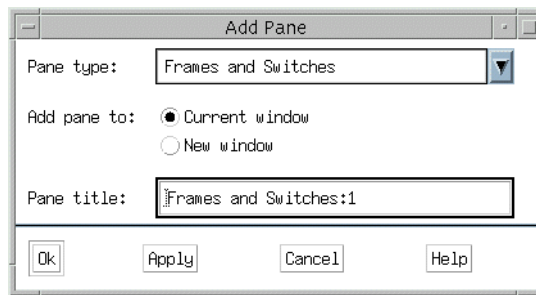


Figure 49. Add Pane dialog box

3. Click **OK**.

This will add the Frames and Switches pane to your Hardware Perspective window. Now, you can manage them together with the nodes.

4. Click the **down arrow** beside the SP frame object. The Frames and Switches pane shows the SP Switch, too. It should look similar to the screen shown in Figure 50 on page 237.

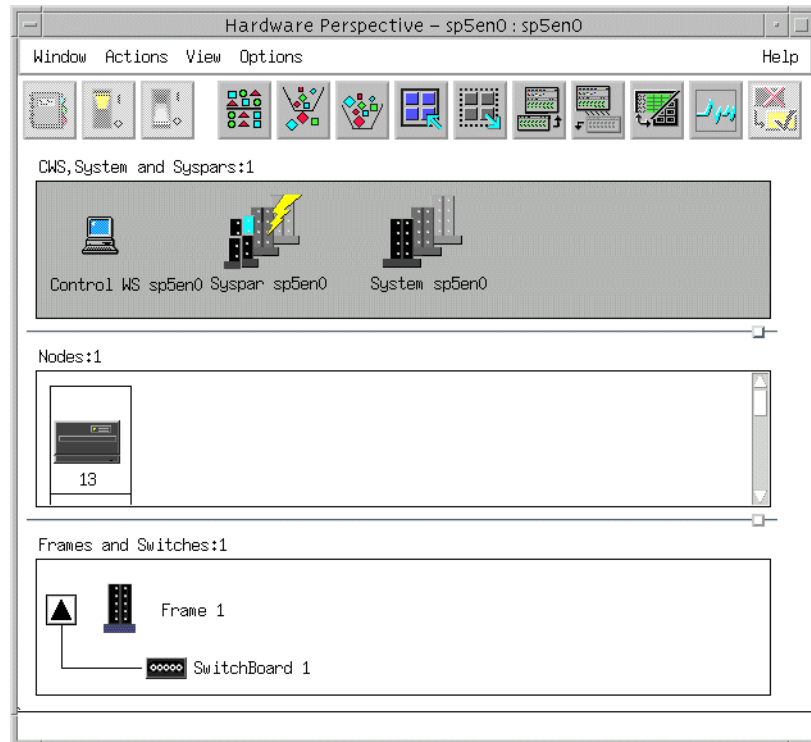


Figure 50. Hardware Perspective with Frames and Switches pane

9.2.2 Controlling SP nodes

If you double click one of the node objects in a Nodes pane, a notebook containing all necessary hardware information about the node will be displayed. The sample notebook shown in Figure 51 on page 238 contains the information for node 13.

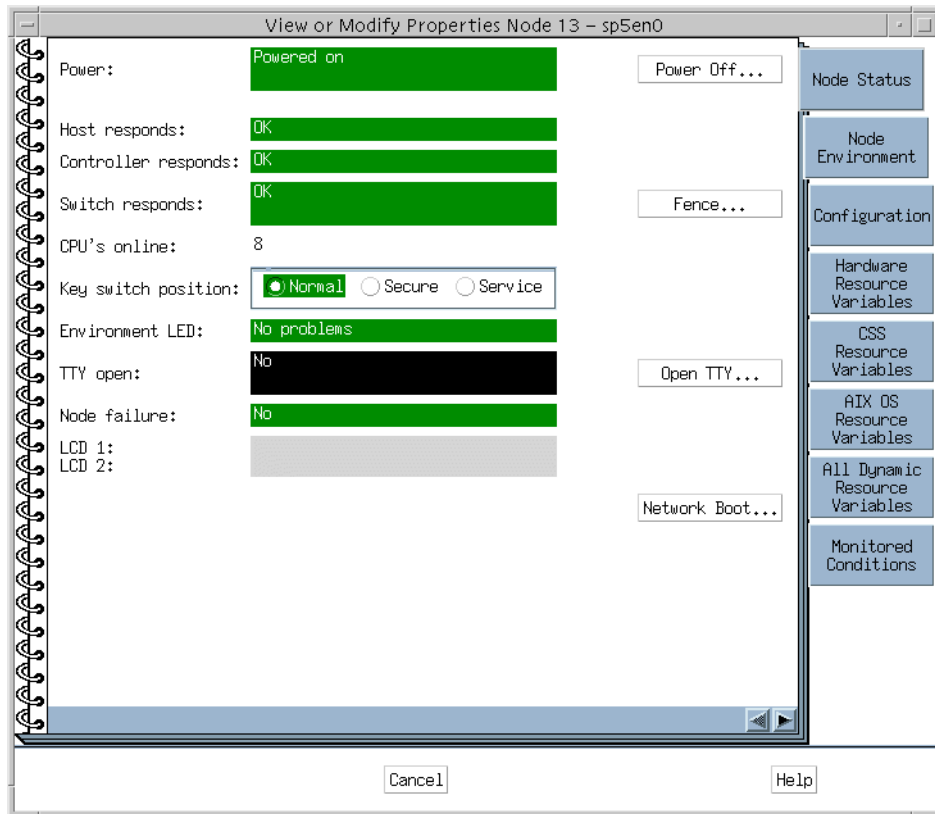


Figure 51. Hardware Perspective SP node notebook

The notebook contains several pages. Click the **Node Status** tab to access the Node Status page.

From this page, you can control the following operations:

- Changing the key switch position
- Powering off/on a node
- Fencing/unfencing a node
- Opening TTY
- Network booting a node

9.2.2.1 Changing the key switch position

To change the key switch position of a node, click the appropriate radio button in the Key switch position field. The choices for key switch position are **Normal**, **Secure**, or **Service**.

9.2.2.2 Powering off a node

To power off a node, click the **Power Off...** button. The display box shown in Figure 52 will be displayed:

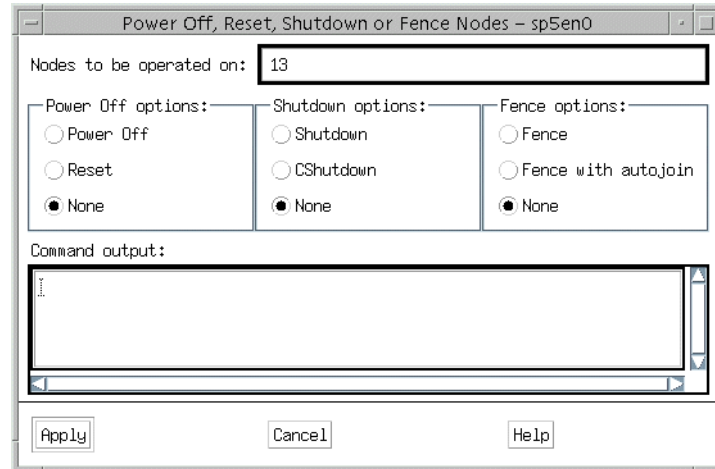


Figure 52. Power Off, Reset, Shutdown or Fence Nodes dialog box

After selecting your options in the Power Off options, Shutdown options, or Fence options field, click **Apply**. The output will be displayed in the Command output field.

An alternate way to access this dialog box is by clicking the **Power Off** icon on the tool bar, which appears as follows:



9.2.2.3 Fencing a node

To fence a node, click the **Fence...** button. The dialog box shown in Figure 53 on page 240 will be displayed:

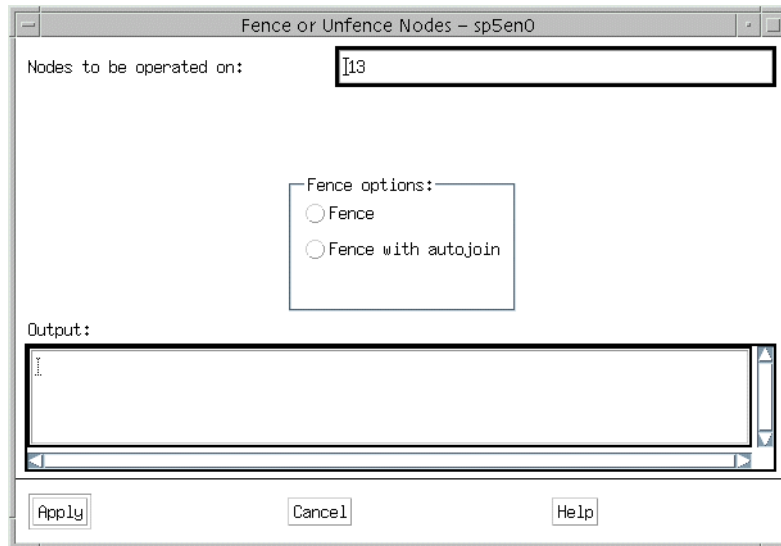


Figure 53. Fence or Unfence Nodes dialog box

After selecting your option in the Fence options field, click **Apply**. The output will be displayed in the Output field.

9.2.2.4 Opening tty

To open a tty console of a node, click the **Open TTY...** button. The Console window shown in Figure 54 will be displayed:

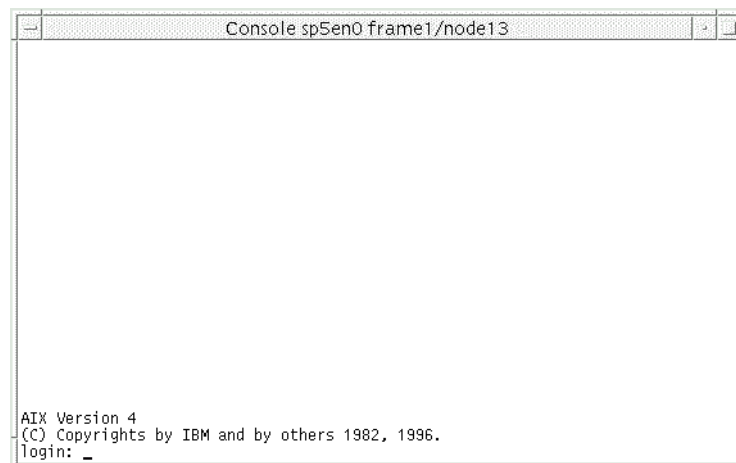


Figure 54. Console window

9.2.2.5 Network booting a node

To network boot a node, click the **Network Boot...** button. The Network Boot Nodes dialog box shown in Figure 55 will be displayed:

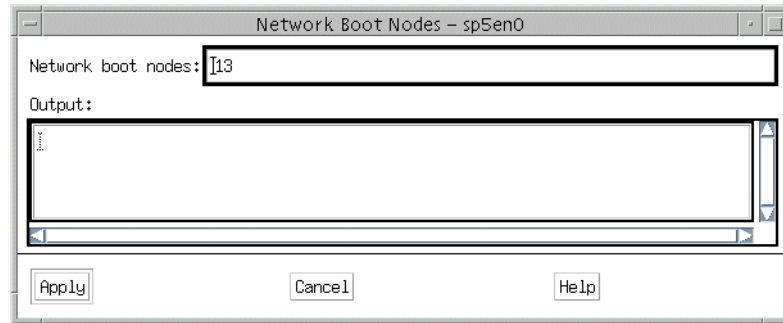


Figure 55. Network Boot Nodes dialog box

Click **Apply** to start network booting a node. The output will be displayed in the Output field.

9.3 Using the spmon command

If you prefer using command line tools rather than Hardware Perspective, you have a couple of options. One of the commands that offers a full range of hardware capabilities is the `spmon` command. In previous versions of IBM Parallel System Support Programs for AIX, there was a graphical option for the `spmon` command, but this option is no longer available. However, all of the functionality is still available from the command line. This section will explain how to use the `spmon` command to accomplish the following operations:

- Set the key switch position
- Query the key switch position
- Power on/off SP frames, SP nodes, and SP Switches
- Reset the nodes

This section provides examples of how to use the command so that the syntax is understandable.

For a detailed discussion about the `spmon` command, refer to *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351.

Attention

Setting/Querying the key switch position on the nodes is not applicable for the PCI bus architecture nodes:

9.3.1 Setting the key switch position

You can use the `spmon` command for thin, wide, and high nodes. Especially for the high nodes, it is very useful for accessing the Standby-Menu of the service processor. This is only possible if the nodes key is set to service mode.

To set the key switch position to *service* on node 13, issue the `spmon` command:

```
# spmon -G -key service node13
```

The general syntax for setting the key switch position on the node is:

```
spmon [-G] -key {normal | secure | service} nodenode#
```

9.3.2 Querying the key switch position

Use the `spmon` command to find out about the key switch position on the nodes.

To find out about the key switch position on node 1, issue the `spmon` command:

```
# spmon -G -Key node1
```

The general syntax for querying key switch position on the nodes is:

```
spmon [-G] -Key nodenode#
```

9.3.3 Powering on/off

Be careful with this task because the `spmon` command really does what you tell it to do. You can power on/off frames, nodes, and even the SP Switch boards with this command.

Attention

Powering off the SP frame means powering off *all the nodes* in it.

Powering off the SP Switch board means losing *all switch connections* on the nodes.

To power off frame 1 of your SP system, issue the `spmon` command:

```
# spmon -G -p off frame1
```

If you want to power off the SP Switch board in the first frame, issue the `spmon` command:

```
# spmon -G -p off frame1/switch
```

The general syntax for powering on/off the frames, nodes, and SP Switch boards with the `spmon` command is:

```
spmon [-G] -p {on | off} {frameframe#[/switch] | nodenode#}
```

9.3.4 Resetting the nodes

As with regular RS/6000 machines, you can reset SP nodes as well.

To reset node 9, issue the `spmon` command:

```
# spmon -G -r node9
```

The general syntax for resetting nodes is:

```
spmon -G -r nodenode#
```

9.4 Using the `hmcnds` command

The other command useful for controlling hardware is the `hmcnds` command. The `hmcnds` command offers more functions than the `spmon` command especially if you have to manage frame, node, or switch supervisor card. This section focuses on controlling the supervisor cards. It is necessary to know how to control the supervisor cards in the following SP system management tasks:

- Setting the frame ID into the frame supervisor card
- Initiating Power-On Self Tests in the frame supervisor card
- Booting the supervisor card
- Switching to basecode version
- Executing the basecode version
- Downloading supervisor microcode

Attention

Old type frames, nodes, and switches, cannot use the `hmcmds` command for these operations because they do not have the capability to handle supervisor microcode.

For more details about supervisor microcode, refer to 2.2, “Supervisor microcode” on page 97.

For a detailed discussion about the `hmcmds` command, refer to *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351.

9.4.1 Setting the frame ID into the frame supervisor card

If you have problems with your frame supervisor card, it can be helpful to set the correct SP frame ID. The SP frame ID in the frame supervisor card can be different from the SP frame ID that PSSP expected.

To set the SP frame ID for the first frame supervisor card, issue the `hmcmds` command:

```
# hmcmds -v -G setid 1:0
hmcmds: Sent VFOP command "setid" to 1 slots.
hmcmds: 100.00% complete.
hmcmds: Number of slots expected to be in state "setid": 1.
hmcmds: Number of slots currently in state "setid": 1.
#
```

The general syntax for setting the frame ID is:

```
hmcmds [-a | -v] [-f file_name] -G setid frame#:0
```

9.4.2 Initiating power-on self tests in the frame supervisor card

To find out if a frame supervisor card is malfunctioning or not, it can be useful to initiate Power-On Self Tests (POST).

To initiate the POST for the first frame supervisor card, issue the `hmcmds` command:

```
# hmcmds -v -G runpost 1:0
hmcmds: Sent VFOP command "runpost" to 1 slots.
hmcmds: 100.00% complete.
hmcmds: Number of slots expected to be in state "on": 1.
hmcmds: Number of slots currently in state "on": 1.
#
```

The general syntax for initiating POST is:

```
hmcmds [-a | -v] [-f file_name] -G runpost frame#:0
```

9.4.3 Booting a supervisor card

It is a good idea to boot a supervisor card when you have a problem with it. This action may solve the problem.

To boot the supervisor card on the first SP node in the first SP frame, issue the `hmcmds` command:

```
# hmcmds -v -G boot_supervisor 1:1
hmcmds: Sent VFOP command "boot_supervisor" to 1 slots.
hmcmds: 100.00% complete.
hmcmds: Number of slots expected to be in state "on": 1.
hmcmds: Number of slots currently in state "on": 1.
#
```

This command performs a boot of the frame, node, or switch basecode application and supervisor.

The general syntax for booting a supervisor card is:

```
hmcmds [-a | -v] [-f file_name] -G boot_supervisor frame#:slot#
```

9.4.4 Switching to basecode version

There may be exceptional situations that force you to back off the supervisor microcode to the base version. In preparation for this situation, the supervisor card keeps two microcode versions. One is the basecode version and the other is the version you are currently using.

Attention

Switching the base code level of the supervisor card can affect the usability of the component of which it is a part. Be careful with this command.

To switch the node supervisor microcode on node 13 to basecode version, issue the `hmcmds` command:

```
# hmcmds -v -G basecode 1:13
hmcmds: Sent VFOP command "basecode" to 1 slots.
hmcmds: 0.00% complete.
hmcmds: 100.00% complete.
hmcmds: Number of slots expected to be in state "basecode": 1.
hmcmds: Number of slots currently in state "basecode": 1.
#
```

This command performs a power off of the node and switches the active frame, node, or switch supervisor to basecode mode, causing the active supervisor to become non-active and the basecode supervisor to become active.

The general syntax for switching to basecode version is:

```
hmcmds [-a | -v] [-f file_name] -G basecode frame#:slot#
```

Attention

You must reboot the supervisor card or execute the basecode version after switching to basecode version.

If you applied this operation to the SP Switch board, do not forget to set the clock for the SP Switch board (refer to 13.1.11, “Reinitializing clock source” on page 351) after the operation.

9.4.5 Executing the basecode version

This step has to be performed after switching to the basecode version described in the previous section.

To execute the basecode version on node 13, issue the `hmcmds` command:

```
# hmcmds -v -G exec_supervisor 1:13
hmcmds: Sent VFOP command "exec_supervisor" to 1 slots.
hmcmds: 0.00% complete.
hmcmds: 100.00% complete.
hmcmds: Number of slots expected to be in state "on": 1.
hmcmds: Number of slots currently in state "on": 1.
#
```

This command causes the basecode to execute the non-active frame, node, or switch supervisor, thus making it active.

The general syntax for executing the basecode version is:

```
hmcnds [-a | -v] [-f file_name] -G exec_supervisor frame#:slot#
```

9.4.6 Downloading supervisor microcode

The reason for using this command is obvious. If you get a new microcode release for your supervisor cards, it has to be downloaded. Normally you do this using the `smitty supervisor fast path` or `spsvnmgr` command. To learn this method, refer to 2.2.4, “Upgrading supervisor microcode” on page 100.

To download the supervisor microcode to node 13, issue the `hmcnds` command:

```
# hmcnds -v -u /spdata/sys1/ucode/u_10.3a.0615 -G microcode 1:13
hmcnds: Application download begins to frame 1 node 13.
      Microcode file is "/spdata/sys1/ucode/u_10.3a.0615".
hmcnds: 25% complete. Elapsed time: 1 mins, 11 secs.
hmcnds: 50% complete. Elapsed time: 2 mins, 20 secs.
hmcnds: 75% complete. Elapsed time: 3 mins, 29 secs.
hmcnds: 100% complete. Elapsed time: 4 mins, 39 secs.
hmcnds: Sent VFOP command "boot_supervisor" to 1 slots.
hmcnds: 0.00% complete.
hmcnds: 100.00% complete.
hmcnds: Number of slots expected to be in state "microcode": 1.
hmcnds: Number of slots currently in state "microcode": 1.
hmcnds: Sent VFOP command "microcode" to 1 slots.
#
```

The command performs a download of supervisor microcode to the frame, node, or switch supervisor.

The general syntax for downloading supervisor microcode is:

```
hmcnds [-a | -v] [-f file_name] -u mcode_file -G microcode frame#:slot#
```

Attention

Prior to applying this command, you must switch the supervisor microcode on the supervisor card that you are planning to work on to the basecode version.

Chapter 10. Monitoring hardware

The control workstation (CWS) allows you to monitor your entire SP system. As an SP system administrator, you will want to make sure that all of your SP system components are behaving properly and not experiencing any problems. IBM Parallel System Support Programs for AIX includes several tools with which to monitor your SP system.

This chapter discusses the various monitoring tools that are available. One of the tools is the graphical user interface (GUI) Hardware Perspective. The command line tools `smon` and `hmon` are also discussed in detail. These two methods show you different ways in which you can monitor your SP system.

10.1 Monitoring tools

When you monitor a target, there are two different approaches:

- You monitor a target constantly. Therefore, you know all the changes that occurred, but it may be difficult to know all the detailed factors that belong to it. This method is generally called *real time monitor*.
- You monitor a target at specific points in time. Therefore, you will know the status of all the factors belonging to the target, but maybe not all the changes that occurred to get to this point. This method is generally called *snapshot*.

Monitoring the SP system in a real-time fashion means that you get informed by the SP system as soon as possible if something is changed. Monitoring the SP system in a snapshot fashion means that you trigger the SP system to give you some status information on demand. By combining and optimizing these two approaches, you should be able to keep an eagle's eye view on your SP system.

There are two tools available to monitor an SP system:

- Hardware Perspective
- Command line tools

The command line tools refer to two commands providing the interface to the SP system:

- `smon`
- `hmon`

The commands are intertwined, and the `smon` command refers to the `hmon` command. The `smon` command is also used for controlling the SP hardware in addition to monitoring. This was described in the previous chapter.

Both Hardware Perspective and command line tools provide you with both real-time and snapshot fashion monitoring.

10.1.1 Creating a Kerberos principal

To use the monitoring tools, you must be successfully authenticated as a Kerberos principal. By default, there is a `root.admin` Kerberos principal available. If you need to have another Kerberos principal, you must add it.

To learn how you can add Kerberos principal, refer to 16.4.3, “Adding a principal” on page 461.

10.1.2 Authorizing a Kerberos principal

Before you use the Hardware Perspective or command line tools, you must establish certain authorizations. Both Hardware Perspective and command line tools use the SP System Monitor. If your Kerberos principal is not registered in the SP System Monitor Access Control Lists (ACLs), you cannot use SP System Monitor.

Use the procedure described in 9.1.2, “Authorizing a Kerberos principal” on page 232 to establish the necessary authorization.

10.2 Using Hardware Perspective

Hardware Perspective is a simple and easy to understand GUI that lets you monitor your SP hardware. It is color coded, to catch your eye if something goes down, or if a problem occurs. This tool allows you to view your entire SP system with all its frames, or you can change the field of view and focus on a single frame or even a single node. Hardware Perspective is a powerful tool, and is easy to use because all the screens have a similar layout. This section explains how you can use Hardware Perspective to monitor your SP hardware.

To know more details about Hardware Perspective, refer to Chapter 3, “Using the Hardware Perspective Effectively” in *SP Perspectives: A New View of Your SP System*, SG24-5180.

10.2.1 Starting Hardware Perspective

To learn how to start Hardware Perspective, refer to 9.2.1, “Starting Hardware Perspective” on page 234.

10.2.2 Monitoring SP frames or SP-attached servers

You can monitor SP frames or SP-attached servers in both real-time and snapshot fashion.

10.2.2.1 Real-time monitor

This section describes how to use Hardware Perspective as a real-time monitor.

Start monitoring

To monitor conditions on your SP frames or SP-attached servers dynamically, perform the following steps:

1. Click one of the **frame** icons in the Frames and Switches pane to give it focus.
2. Click the **Monitor** icon on the tool bar:



3. You will see the Set Monitoring for Frames and Switches notebook as shown in Figure 56 on page 252.

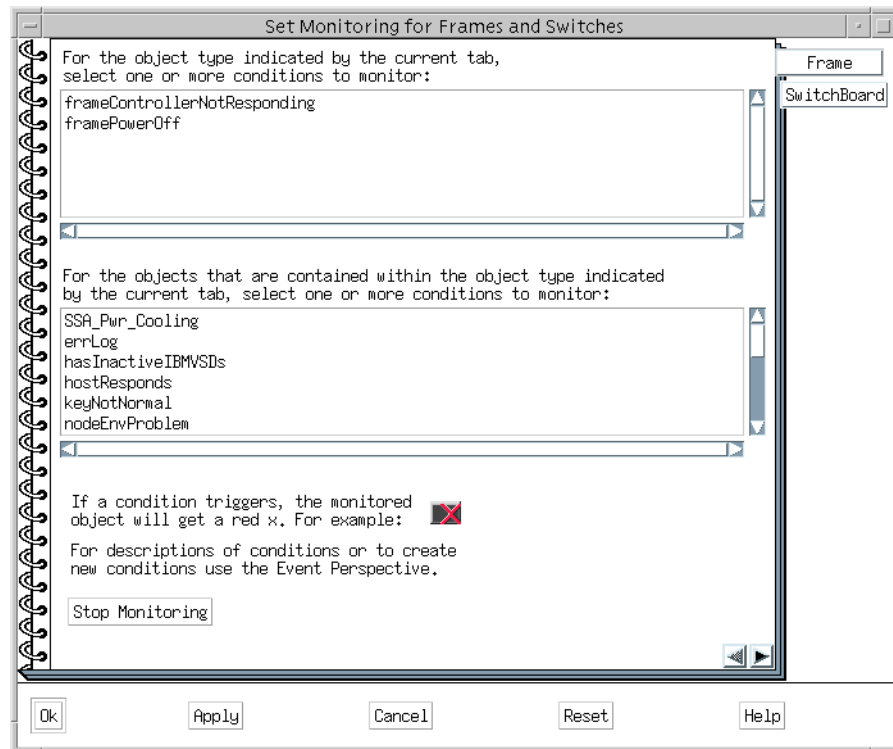


Figure 56. Frame page of Set Monitoring notebook

4. To see the conditions for the frame, in other words, to open the Frame page, click the **Frame** tab.
5. Select the conditions you want to monitor from the list boxes. For multiple selections, press the **Ctrl** key while clicking the items. Select the following conditions, for example:
 - **frameControllerNotResponding**
 - **framePowerOff**
6. Click **Ok**.

The color of the frame icon in the Frames and Switches pane will be changed. If it is green, that frame has no problem. If it is black with a red x, that frame has something wrong in the conditions you selected. The color will change dynamically based on the changes in the status of the conditions you selected.

Checking monitored conditions

To check the status history of the conditions you selected, perform the following steps:

1. Double click the frame icon on which you want to check the monitored conditions. You will see the View or Modify Properties Frame notebook as shown in Figure 57.

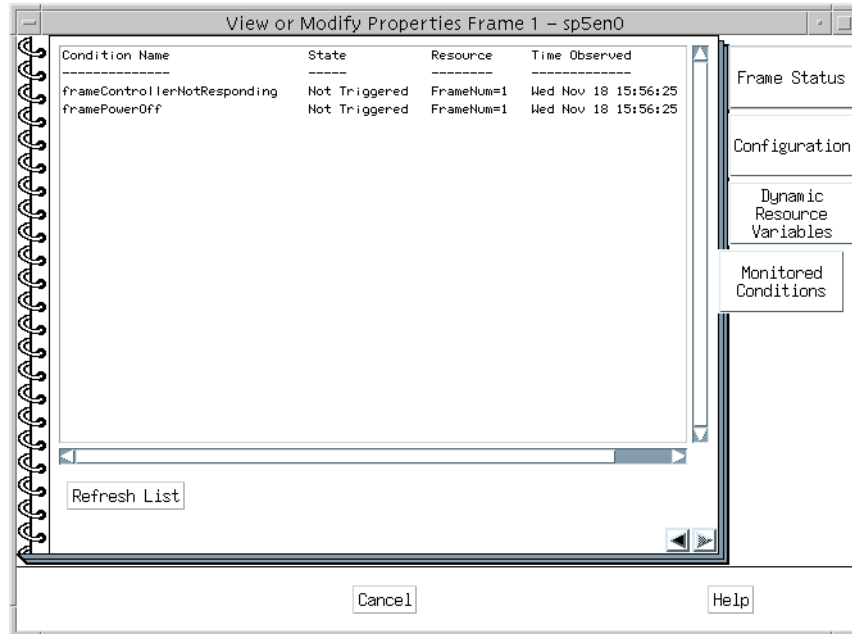


Figure 57. Monitored conditions page of frame notebook

2. To see the monitored conditions, click **Monitored Conditions**.
3. The information on the notebook will not be updated dynamically. If you need the latest status, click **Refresh List**.

Stop monitoring

To stop monitoring, click **Stop Monitoring** in the Set Monitoring for Frames and Switches notebook as shown in Figure 56 on page 252 or simply terminate the Hardware Perspective.

10.2.2.2 Snapshot monitoring

This section describes how to use Hardware Perspective as a snapshot monitoring tool.

If you want to check the status of a frame precisely at a certain time, there is a way. Hardware Perspective shows you the value of all the frame-related resource variables. To learn more about resource variables, refer to 11.1.1, “What is a resource variable?” on page 279.

To check all the frame-related resource variables, open the Dynamic Resource Variables page in the View or Modify Properties Frame notebook by clicking the appropriate tab. The page shown in Figure 58 will be displayed.

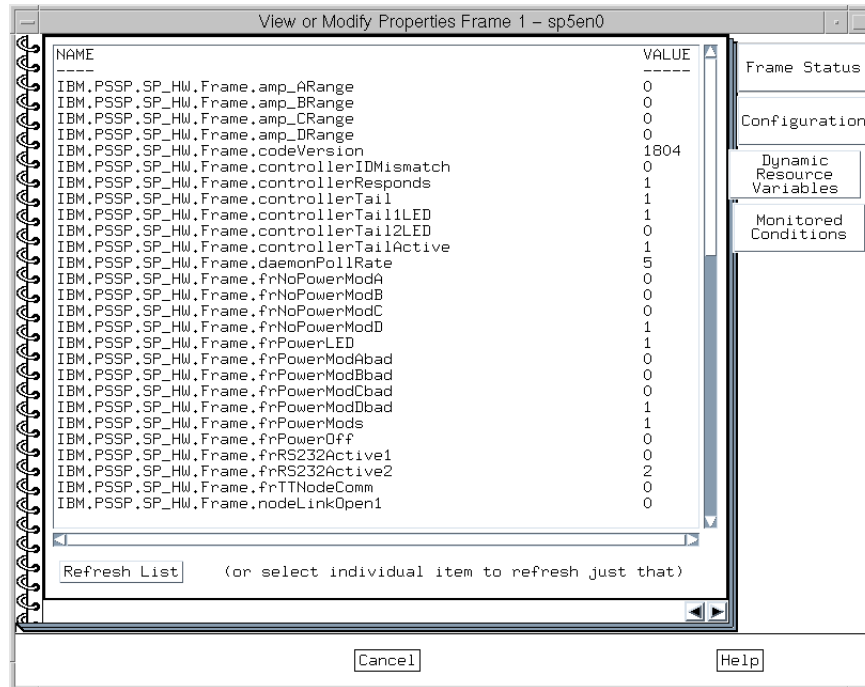


Figure 58. Dynamic Resource Variables page of frame notebook

This page shows a list of resource variables and their values. To learn the meaning of the resource variables, refer to 11.5.2, “Getting an explanation of resource variable” on page 302.

The information on the notebook page will not be changed dynamically. If you need the latest status, click **Refresh List**.

10.2.3 Monitoring SP nodes or SP-attached servers

You can monitor SP nodes or SP-attached servers in both the real-time and snapshot fashion.

10.2.3.1 Real-time monitoring

This section describes how to use Hardware Perspective as a real-time monitor.

Start monitoring

To monitor conditions on your SP nodes or SP-attached server dynamically, perform the following steps:

1. Click one of the **node** icons in the Nodes pane to give it focus.
2. Click the **Monitor** icon on the tool bar:



3. You will see the Set Monitoring for Nodes notebook as shown in Figure 59.

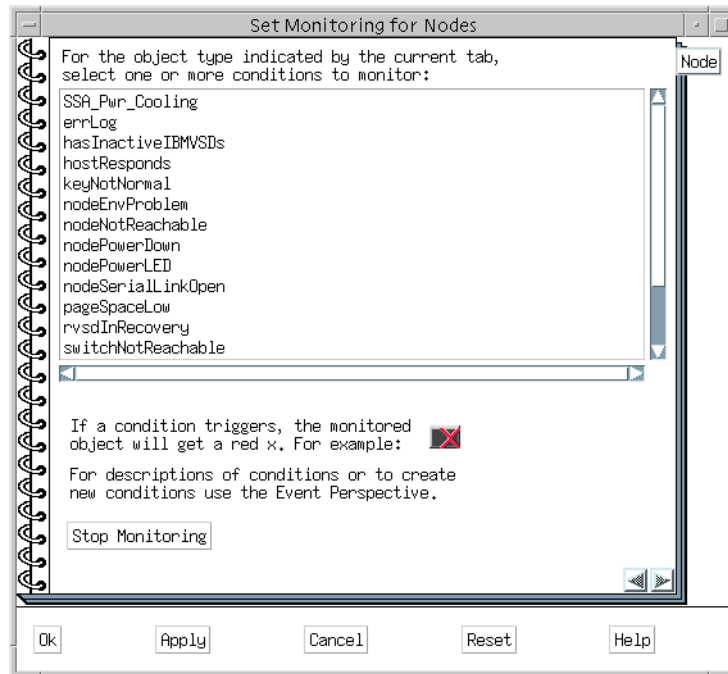


Figure 59. Node page of set monitoring notebook

4. Select the conditions you want to monitor from the list box. For multiple selections, press the **Ctrl** key while clicking the items. Select the following conditions, for example:
 - **hostResponds**

- **nodeEnvProblem**
- **nodePowerDown**
- **nodePowerLED**
- **pageSpaceLow**
- **switchResponds**
- **tmpFull**
- **varFull**

5. Click **Ok**.

The color of the node icon in the Nodes pane will be changed. If it is green, that node has no problem. If it is black with a red x, that node has something wrong in the conditions you selected. The color will change dynamically based on changes in the status of the conditions you selected.

Checking monitored conditions

To check the status history of the conditions you selected, perform the following steps:

1. Double click the **node** icon on which you want to check the monitored conditions. You will see the View or Modify Properties Node notebook as shown in Figure 60 on page 257.

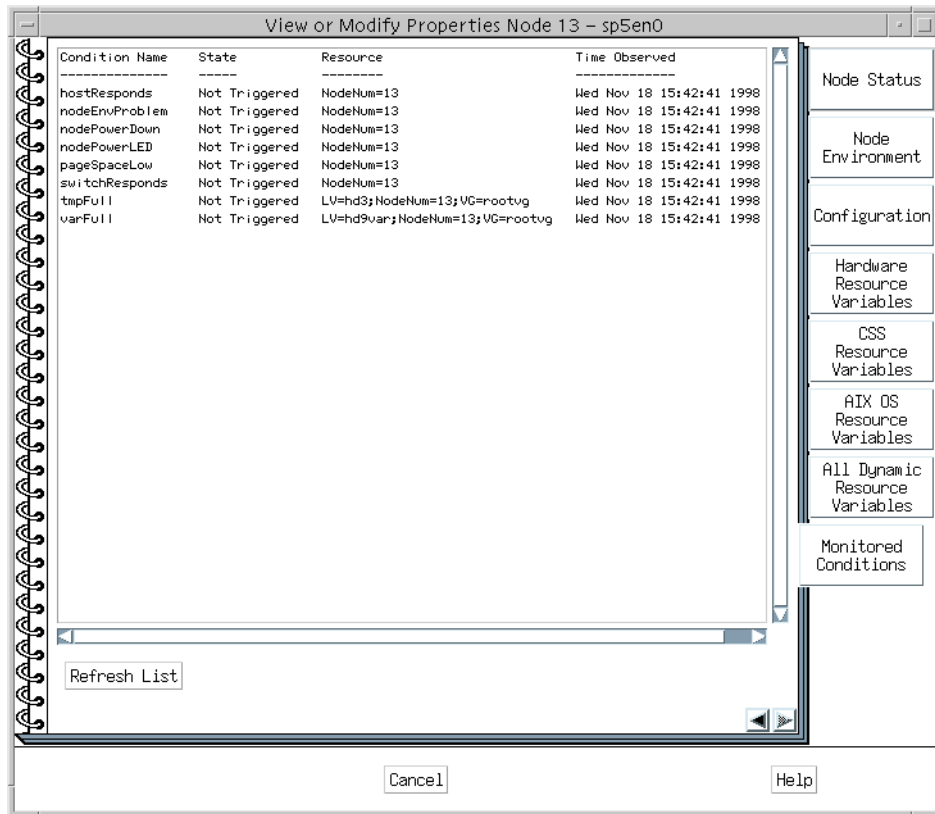


Figure 60. Monitored conditions page of node notebook

2. To see the monitored conditions, click **Monitored Conditions**.
3. The information on the notebook will not be updated dynamically. If you need the latest status, click **Refresh List**.

Stop monitoring

To stop monitoring, click **Stop Monitoring** in Set Monitoring for Nodes notebook as shown in Figure 59 on page 255 or simply terminate the Hardware Perspective.

10.2.3.2 Snapshot monitoring

This section describes how to use Hardware Perspective as a snapshot monitoring tool.

If you want to check the status of a node precisely at a certain time, there is a way. Hardware Perspective shows you the value of all the node-related

resource variables. To learn more about resource variables, refer to 11.1.1, “What is a resource variable?” on page 279.

To check all the node-related resource variables, open the All Dynamic Resource Variables page in the View or Modify Properties Node notebook by clicking the appropriate tab. The page shown in Figure 61 will be displayed.

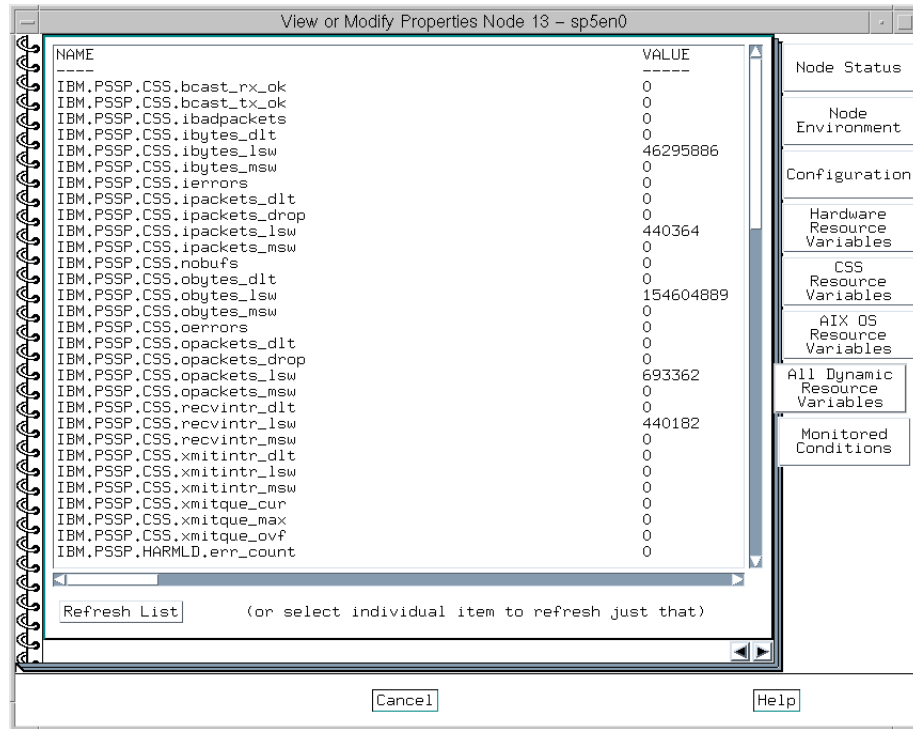


Figure 61. All Dynamic Resource Variables page of node notebook

This page shows a list of resource variables and their values. To learn the meaning of the resource variables, refer to 11.5.2, “Getting an explanation of resource variable” on page 302.

The information on the notebook page will not be changed dynamically. If you need the latest status, click **Refresh List**.

10.2.4 Monitoring SP Switch boards

You can monitor SP Switch boards in both the real-time and snapshot fashion.

10.2.4.1 Real-time monitoring

This section describes how to use Hardware Perspective as a real-time monitor.

Start monitoring

To monitor conditions on your SP Switch boards dynamically, perform the following steps:

1. Click one of the **switch board** icons in the Frames and Switches pane to give it focus.
2. Click the **Monitor** icon:



3. You will see the Set Monitoring for Frames and Switches notebook as shown in Figure 62.

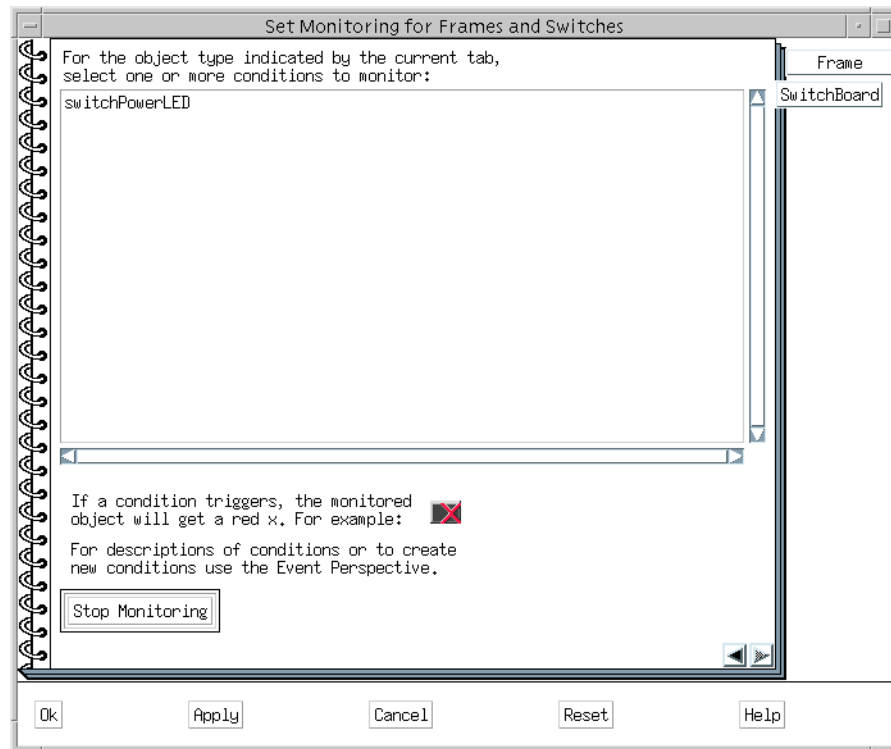


Figure 62. SwitchBoard page of set monitoring notebook

4. To see the conditions for the switch board, in other words, to open the SwitchBoard page, click **SwitchBoard**.
5. Select the **switchPowerLED** condition, which is the only one available for SP Switch boards.
6. Click **Ok**.

The color of the switch board icon in the Frames and Switches pane will be changed. If it is green, that switch board has no problem. If it is black with a red x, that switch board has something wrong in the switchPowerLED condition. The color will change dynamically based on the changes in the status of the condition.

Checking monitored conditions

To check the status history of the condition, perform the following steps:

1. Double click the **switch board** icon on which you want to check the monitored condition. You will see the View or Modify Properties SwitchBoard notebook as shown in Figure 63 on page 261.

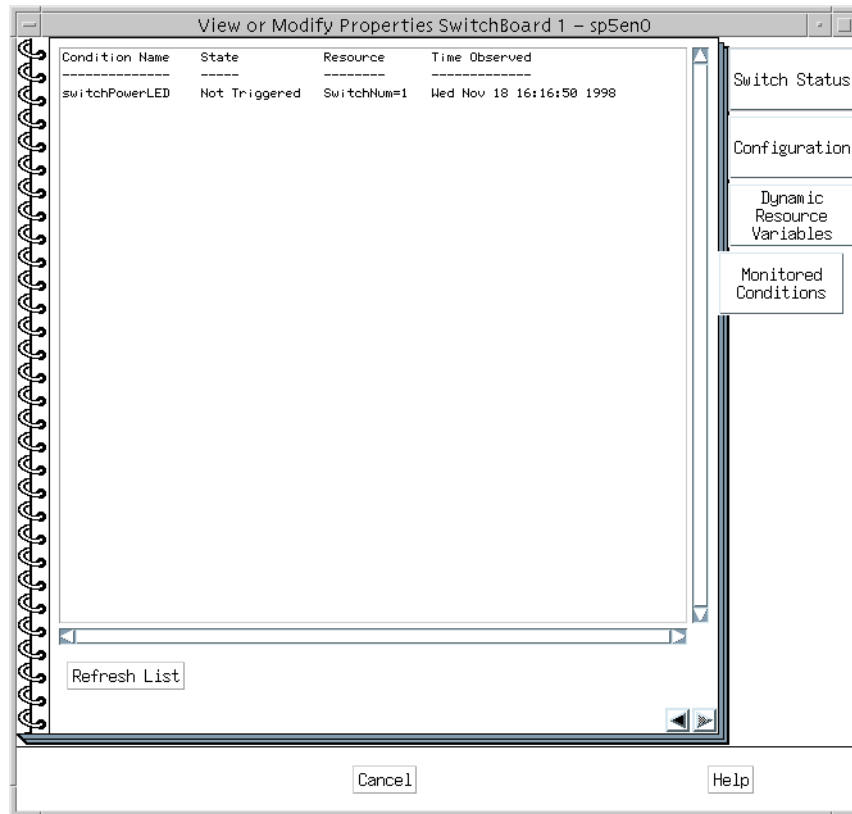


Figure 63. Monitored conditions page of SwitchBoard notebook

2. To see the monitored condition, click **Monitored Conditions**.
3. The information on the notebook will not be updated dynamically. If you need the latest status, click **Refresh List**.

Stop monitoring

To stop monitoring, click **Stop Monitoring** in the Set Monitoring for SwitchBoard notebook as shown in Figure 62 on page 259 or simply terminate the Hardware Perspective.

10.2.4.2 Snapshot monitoring

This section describes how to use Hardware Perspective as a snapshot monitoring tool.

If you want to check the status of a switch board precisely at a certain time, there is a way. Hardware Perspective shows you the value of all the switch

board-related resource variables. To learn more about resource variables, refer to 11.1.1, “What is a resource variable?” on page 279.

To check all the switch board-related resource variables, open the Dynamic Resource Variables page in the View or Modify Properties SwitchBoard notebook by clicking on the appropriate tab. The page shown in Figure 64 will be displayed.

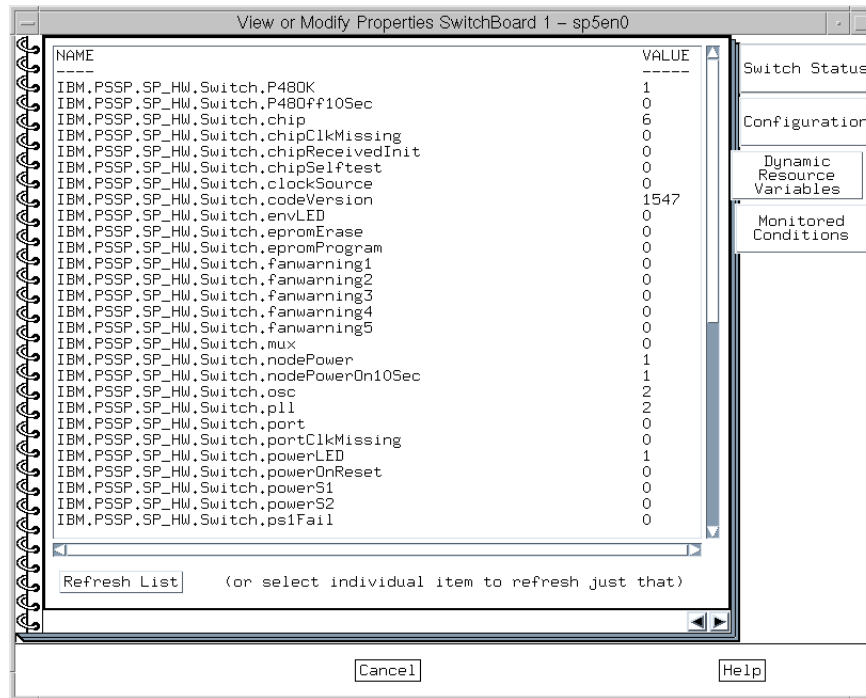


Figure 64. Dynamic Resource Variables page of SwitchBoard notebook

The page shows a list of resource variables and their values. To learn the meaning of the resource variables, refer to 11.5.2, “Getting an explanation of resource variable” on page 302.

The information on the notebook page will not be changed dynamically. If you need the latest status, click **Refresh List** button.

10.2.5 Monitoring your original conditions

The Set Monitoring notebooks for frames, nodes, and switch boards each provide you with the list of default conditions.

If the condition that you want to monitor is not on the list for the component of interest, Hardware Perspective allows you to define your own original condition. To do this, use Event Perspective to create the desired condition. Once it is created, it will be on the list in the Set Monitoring notebook for the specified component.

To learn how to create your original condition, refer to Chapter 4, “Using the Event Perspective Effectively” in the *SP Perspectives: A New View of Your SP System*, SG24-5180.

You may realize that Hardware Perspective can monitor resources other than hardware itself. For example, the switchResponds condition monitors availability of communication through the SP switch. It includes the switch daemon and other software resources too, not only SP switch hardware. Using a combination of Hardware Perspective and Event Perspective gives you more power to manage your SP system. For more details about Event Perspective, refer to Chapter 11, “Managing Events” on page 279.

10.3 Using the `smon` command

The alternative method to using Hardware Perspective is to use the command line tools. The `smon` command offers a straightforward and simple view of your system, while providing you with plenty of information. This section explains how you can use the `smon` command to monitor your system.

To learn the `smon` command syntax, refer to *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351.

10.3.1 Monitoring the SP system

One of the quickest and most widely used ways to have a look at the SP system is provided by the `smon` command. The `smon` command provides only snapshot monitoring.

To get the snapshot of your SP system, issue the `smon` command:

```

# spon -G -d
1. Checking server process
   Process 20388 has accumulated 4 minutes and 26 seconds.
   Check ok

2. Opening connection to server
   Connection opened
   Check ok

3. Querying frame(s)
   2 frame(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 N/A
   2    yes      no     N/A    N/A      N/A N/A N/A N/A

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
-----
   3    3    wide    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
----- Frame 2 -----
Frame Node  Node      Host/Switch  Key  Env  Front Panel  LCD/LED is
Slot  Number Type  Power  Responds  Switch  Fail  LCD/LED  Flashing
-----
   1    17   extrn   on  yes  yes  N/A    N/A  080C      N/A
                                   LCD2 is blank
#

```

This is a perfect way of gathering important information at a glance. From the output, you know information about frames, nodes, and switches. It includes the power status, host responds, switch responds, or LCD/LED status.

Note, that the status of the four power supplies in the SP frame are displayed as well. To display the checking result of SP frames, you need to specify the `-G` flag. In this example, node 17 on frame 2 is an SP-attached server. The information about an SP-attached server shows up both as a frame and a node.

10.4 Using the `hmmon` command

The other command line tool that you can use to monitor your SP system is the `hmmon` command. This command can monitor your frames, nodes, SP-attached servers, and SP Switches as well. This section describes how you can use this command to monitor your SP system.

To learn the `hmmon` command syntax, refer to *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351.

For more information about the SP System Monitor variables, refer to Appendix E, “System Monitor Variables” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

10.4.1 Monitoring SP frames

You can monitor SP frames in both a real-time and snapshot fashion.

10.4.1.1 Snapshot

To get the current information for an SP frame as a snapshot, issue the `hmmon` command:

```

# hmmon -G -Q 1:0

frame 001, slot 00:
  node 01 I2C not responding      FALSE node 02 I2C not responding      TRUE
  node 03 I2C not responding      TRUE  node 04 I2C not responding      TRUE
  node 05 I2C not responding      FALSE node 06 I2C not responding      TRUE
  node 07 I2C not responding      TRUE  node 08 I2C not responding      TRUE
  node 09 I2C not responding      FALSE node 10 I2C not responding      TRUE
  node 11 I2C not responding      TRUE  node 12 I2C not responding      TRUE
  node 13 I2C not responding      FALSE node 14 I2C not responding      TRUE
  node 15 I2C not responding      TRUE  node 16 I2C not responding      TRUE
  switch I2C not responding      FALSE controller tail is active      TRUE
  node 01 serial link open        FALSE node 02 serial link open        FALSE
  node 03 serial link open        FALSE node 04 serial link open        FALSE
  node 05 serial link open        FALSE node 06 serial link open        FALSE
  node 07 serial link open        FALSE node 08 serial link open        FALSE
  node 09 serial link open        FALSE node 10 serial link open        FALSE
  node 11 serial link open        FALSE node 12 serial link open        FALSE
  node 13 serial link open        FALSE node 14 serial link open        FALSE
  node 15 serial link open        FALSE node 16 serial link open        FALSE
  SEPBU frame LED 1 (green)      0x0001 SEPBU frame LED 2 (green)      0x0001
  frame LED 3 (green)            0x0000 frame LED 4 (green)            0x0001
  frame LED 6 (yellow)           0x0000 frame LED 7 (yellow)           0x0002
  frame LED 8 (yellow)           0x0000 module A power bad              FALSE
  module B power bad             FALSE module C power bad              FALSE
  module D power bad             TRUE  module A not present             FALSE
  module B not present           FALSE module C not present           FALSE
  module D not present           TRUE  SEPBU frame power off           FALSE
  mod A KW 0=3.5 1=5 3=n/a      0x0000 mod B KW 0=3.5 1=5 3=n/a      0x0000
  mod C KW 0=3.5 1=5 3=n/a      0x0000 mod D KW 0=3.5 1=5 3=n/a      0x0003
  supervisor timer ticks        0x02b5 diagnosis return code          0x0000
  +48 voltage                    0x0074 temperature                    0x0035
  section A current              0x0034 section B current                0x002b
  section C current              0x0015 section D current                0x0015
  +48 voltage out of range      FALSE temperature out of range    FALSE
  A current out of range        FALSE B current out of range          FALSE
  C current out of range        FALSE D current out of range        FALSE
  frame ID mismatch             FALSE supervisor serial number     0x1997
  supervisor type            0x0016 supervisor code version          0x070c
  hardware monitor poll rate    0x0005 active controller tail      0x0001

frame 001, slot 00:
  frame responding to polls      TRUE  RS232 link DCD is active        TRUE
  RS232 link CTS is active      TRUE
#

```

Figure 65. The hmmon command output for an SP frame

In this example, you requested the information for the first frame of your SP system. To be precise, you requested the information on the frame supervisor card (slot 0) of the first frame of your SP system.

Most of the information provided by this command is self-explanatory. For a more comprehensive interpretation, you need to know something about hardware variables on the SP System Monitor.

Issue the `hmmon` command in a similar way as shown in Figure 65 on page 266 but add an `-s` flag:

```
# hmmon -G -Q -s 1:0
1 0 nodefail1 FALSE 0x8802 node 01 I2C not responding
1 0 nodefail2 TRUE 0x8803 node 02 I2C not responding
1 0 nodefail3 TRUE 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 TRUE 0x8807 node 06 I2C not responding
1 0 nodefail7 TRUE 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 TRUE 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 controllerTailActive TRUE 0x8877 controller tail is active
1 0 nodeLinkOpen1 FALSE 0x8813 node 01 serial link open
1 0 nodeLinkOpen2 FALSE 0x8814 node 02 serial link open
1 0 nodeLinkOpen3 FALSE 0x8815 node 03 serial link open
1 0 nodeLinkOpen4 FALSE 0x8816 node 04 serial link open
1 0 nodeLinkOpen5 FALSE 0x8817 node 05 serial link open
1 0 nodeLinkOpen6 FALSE 0x8818 node 06 serial link open
1 0 nodeLinkOpen7 FALSE 0x8819 node 07 serial link open
1 0 nodeLinkOpen8 FALSE 0x881a node 08 serial link open
1 0 nodeLinkOpen9 FALSE 0x881b node 09 serial link open
1 0 nodeLinkOpen10 FALSE 0x881c node 10 serial link open
1 0 nodeLinkOpen11 FALSE 0x881d node 11 serial link open
1 0 nodeLinkOpen12 FALSE 0x881e node 12 serial link open
1 0 nodeLinkOpen13 FALSE 0x881f node 13 serial link open
1 0 nodeLinkOpen14 FALSE 0x8820 node 14 serial link open
1 0 nodeLinkOpen15 FALSE 0x8821 node 15 serial link open
1 0 nodeLinkOpen16 FALSE 0x8822 node 16 serial link open
1 0 frPowerMods 1 0x883c SEPBU frame LED 1 (green)
1 0 frPowerLED 1 0x883d SEPBU frame LED 2 (green)
1 0 controllerTail2LED 0 0x887a frame LED 3 (green)
1 0 controllerTail1LED 1 0x8879 frame LED 4 (green)
1 0 frTINodeComm 0 0x889a frame LED 6 (yellow)
1 0 frRS232Active2 2 0x8899 frame LED 7 (yellow)
1 0 frRS232Active1 0 0x8896 frame LED 8 (yellow)
```

This time, the command shows you hardware variable names in the third column. For a more detailed description for the `nodefail1` hardware variable,

see Appendix E, “System Monitor Variables” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

This publication provides you with many tables. You can find the `nodefail1` hardware variable in the table of Frame Hardware Variables Common to ALL Frame Supervisors. It is described as follows:

Loss of communication from frame supervisor to specified node supervisor (1 [True] = communication loss, 0 [False] = communication okay)

You may find some hardware variables in more than one table. If this is the case, look up the appropriate table using your frame supervisor card type. To find the frame supervisor card type, issue the `hmmmon` command:

```
# hmmmon -G -Q -s -v type 1:0
  1  0  type                22  0x883a  supervisor type
#
```

The fourth column indicates the decimal value of the variable. Now you know your frame supervisor card type is 22. Alternatively, you can find the hexadecimal value as shown in Figure 65 on page 266 for the supervisor type field.

10.4.1.2 Real-time monitor

You can also issue the `hmmmon` command as a real-time monitor:

```
# hmmmon -G -q 1:0
```

The difference in usage from that described in 10.4.1.1, “Snapshot” on page 265 is a flag. Use the `-q` flag instead of the `-Q` flag. This command will not return to the system prompt. Instead, it waits until the status of hardware components changes and reports the changed status as output to the console.

10.4.2 Monitoring SP-attached servers

You can monitor SP-attached servers in a both real-time and snapshot fashion.

10.4.2.1 Snapshot

To get the current information for an SP-attached server, issue the `hmmmon` command:

```

# hmmon -G -Q 2:1

frame 002, slot 01:
  DC-DC power on           TRUE  serial link is open           FALSE
  delayed power off active FALSE  SRC contains a message       FALSE
  SPCN contains a message  FALSE LED/LCD contains a message   TRUE
  System Reference Code    BLANK System Power Cntl Network    BLANK
  hardware status byte     0x0040 diagnosis return code       0x000f
  supervisor timer ticks   0xca55 supervisor type         0x000a
  supervisor code version  0x0301
  LCD line 1               080C
  LCD line 2               BLANK
#

```

Figure 66. The `hmmon` command output for an SP-attached server

In this example, you requested the information for node 17 of your SP system. To be precise, you requested the information on node supervisor card (slot 1) of the second frame of your SP system. An SP-attached server does not have a frame supervisor card, so the `s70d` daemon simulates it as if an SP-attached server has a frame supervisor card.

Most of the information provided by this command is self-explanatory. For a more comprehensive interpretation, you need to know something about hardware variables on the SP System Monitor because the `hmmon` command treats an SP-attached server as an SP node. For more details, refer to 10.4.3.1, “Snapshot” on page 270.

Since an SP-attached server does not have a node supervisor card, the `s70d` daemon simulates it as if an SP-attached server has a node supervisor card. To find the node supervisor card type, issue the `hmmon` command:

```

# hmmon -G -Q -s -v type 2:1
2 1 type           10 0x000a supervisor type
#

```

The fourth column indicates the decimal value of the variable. Now, you know your node supervisor card type is 10. Alternatively, you can find the hexadecimal value in Figure 66 for the supervisor type field.

10.4.2.2 Real-time monitor

Use the `-q` flag instead of `-Q` flag in conjunction with the `hmmon` command:

```

# hmmon -G -q 2:1

```

The effect is that the system prompt will not return, and the hardware status changes will be written to the console.

10.4.3 Monitoring SP nodes

You can monitor SP nodes in both a real-time and snapshot fashion.

10.4.3.1 Snapshot

To get the current information for an SP node, issue the `hmmon` command:

```
# hmmon -G -Q 1:13

frame 001, slot 13:
  SMP LED 1 off (green)      FALSE  SMP LED 2 off (yellow)    FALSE
  power switch status      TRUE   temperature warning       FALSE
  key: 0=norm 1=sec 2=serv  0x0000 serial link is open       FALSE
  hardware status byte     0x0000 LCD line 1 is flashing    FALSE
  LCD line 2 is flashing   FALSE  node/switch LED 1 (green) 0x0001
  node/switch LED 2 (yellow) 0x0000 DC-DC power on           TRUE
  +5 DC-DC output good     TRUE   +48 volt power on         TRUE
  hardware key id          0x003a board level              0x001a
  supervisor level         0x0001 LCD line 1                BLANK
  supervisor timer ticks   0xa3c7 supervisor type      0x00a1
  supervisor code version  0x0615 7 segment LED A         0x00ff
  7 segment LED B         0x00ff 7 segment LED C         0x00ff
  LED/LCD contains a message FALSE  7 segment display flashing FALSE
  LCD line 2              BLANK  temperature              0x0035
  temperature out of range FALSE

#
```

Figure 67. The `hmmon` command output for an SP node

In this example, you requested the information for node 13 of your SP system. To be precise, you requested the information on the node supervisor card (slot 13) of the first frame of your SP system.

Most of the information provided by this command is self-explanatory. For a more comprehensive interpretation, you need to know something about hardware variables on the SP System Monitor.

Issue the `hmmon` command in a similar way as shown in Figure 67 but add an `-s` flag:

```

# hmon -G -Q -s 1:13
1 13 smpPowerLEDOff FALSE 0x98f6 SMP LED 1 off (green)
1 13 smpDiagLEDOff FALSE 0x98f7 SMP LED 2 off (yellow)
1 13 powerSwitchStatus TRUE 0x98f0 power switch status
1 13 warningTemp FALSE 0x9858 temperature warning
1 13 keyModeSwitch 0 0x989b key: 0=norm 1=sec 2=serv
1 13 serialLinkOpen FALSE 0x989d serial link is open
1 13 hardwareStatus 0 0x98f3 hardware status byte
1 13 lcd1flash FALSE 0x98f1 LCD line 1 is flashing
1 13 lcd2flash FALSE 0x98f2 LCD line 2 is flashing
1 13 powerLED 1 0x9847 node/switch LED 1 (green)
1 13 envLED 0 0x9848 node/switch LED 2 (yellow)
1 13 nodePower TRUE 0x984a DC-DC power on
1 13 P5DCok TRUE 0x9897 +5 DC-DC output good
1 13 P48OK TRUE 0x9849 +48 volt power on
1 13 hardwareKeyID 58 0x98dc hardware key id
1 13 boardLevel 26 0x98e5 board level
1 13 supLevel 1 0x98e6 supervisor level
1 13 lcd1 BLANK 0x98f4 LCD line 1
1 13 timeTicks 40041 0x9830 supervisor timer ticks
1 13 type 161 0x983a supervisor type
1 13 codeVersion 1557 0x983b supervisor code version
1 13 LED7SegA 255 0x989f 7 segment LED A
1 13 LED7SegB 255 0x98a0 7 segment LED B
1 13 LED7SegC 255 0x98a1 7 segment LED C
1 13 LCDhasMessage FALSE 0x9906 LED/LCD contains a message
1 13 7segChanged FALSE 0x9898 7 segment display flashing
1 13 lcd2 BLANK 0x98f5 LCD line 2
1 13 temp 21.560 0x9834 temperature
1 13 tempRange FALSE 0x9884 temperature out of range
#

```

This time, the command shows you hardware variable names in the third column. For a more detailed description for the `smpPowerLEDOff` hardware variable, see Appendix E, “System Monitor Variables” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

This publication provides you with many tables. You can find the `smpPowerLEDOff` hardware variable in the table of Node-Specific Hardware Variables For Card Type 161 Node Supervisor. It is described as follows:

Indicates power status. LED 1 (0 [False] = DC power okay, 1 [True] = no DC power supplied)

You may find some hardware variables in more than one table. If this is the case, look up the appropriate table using your node supervisor card type. To find the node supervisor card type, issue the `hmon` command:

```
# hmmon -G -Q -s -v type 1:13
  1 13 type          161 0x983a supervisor type
#
```

The fourth column indicates the decimal value of the variable. Now you know your node supervisor card type is 161. Alternatively, you can find the hexadecimal value in Figure 67 on page 270 for the supervisor type field.

10.4.3.2 Real-time monitor

Issue the `-q` flag instead of `-Q` flag in conjunction with the `hmmon` command:

```
# hmmon -G -q 1:13
```

The effect is that the system prompt will not return, and the hardware status changes will be written to the console.

10.4.4 Monitoring SP switch boards

You can monitor SP Switch boards in both a real-time and snapshot fashion.

10.4.4.1 Snapshot

To get the current information for an SP Switch board as a snapshot, issue the `hmmon` command:

```

# hmon -G -Q 1:17

frame 001, slot 17:
  chip selftest active          FALSE waiting for init          FALSE
  chip clock missing           FALSE port clock missing          FALSE
  send port not tuned          TRUE  receive port not tuned        TRUE
  synchronous reset active     FALSE power on reset active        FALSE
  power supply #1 failure      FALSE power supply #2 failure      FALSE
  overcurrent warning          FALSE overcurrent shutdown          FALSE
  parallel power failure       FALSE program eeprom complete    FALSE
  erase eeprom complete        FALSE active clock source          0x0000
  active oscillator            0x0002 active phase lock loop      0x0002
  active chip                   0x0006 active port                      0x0000
  switch MUX setting           0x0000 node/switch LED 1 (green)      0x0001
  node/switch LED 2 (yellow)   0x0000 +48 volt power on              TRUE
  DC-DC power on               TRUE  DC-DC power on > 10 secs      TRUE
  power off > 10 secs          FALSE fan 1 warning                  FALSE
  fan 2 warning                 FALSE fan 3 warning                  FALSE
  fan 4 warning                 FALSE fan 5 warning                  FALSE
  +3.3 volt high warning       FALSE +3.3 volt high shutdown        FALSE
  +3.3 volt low warning        FALSE +3.3 volt low shutdown          FALSE
  temperature warning          FALSE temperature shutdown          FALSE
  hardware key id              0x0009 board level                    0x0001
  supervisor subtype           TRUE  supervisor level                0x0000
  power supply #1              0x0000 power supply #2                  0x0000
  diagnosis return code        0x0000 parallel voltage                0x0029
  ps1 power                     0x0027 ps2 power                        0x0027
  ps1 fuse                       0x0074 ps2 fuse                          0x0075
  temperature                    0x0032 +3.3 voltage                      0x00a9
  parallel out of range         FALSE ps1 power out of range          FALSE
  ps2 power out of range        FALSE ps1 fuse out of range            FALSE
  ps2 fuse out of range         FALSE temperature out of range      FALSE
  +3.3 voltage out of range     FALSE supervisor timer ticks       0xf4ed
  supervisor code version       0x060b supervisor type           0x0081
  supervisor serial number      0x1996

```

Figure 68. The hmon command output for an SP switch board

In this example, you requested the information for the first SP Switch board of your SP system. To be precise, you requested the information on the SP Switch supervisor card (slot 17) of the first frame of your SP system.

Most of the information provided by this command is self-explanatory. For a more comprehensive interpretation, you need to know something about hardware variables on the SP System Monitor.

Issue the hmon command in a similar way as shown in Figure 68 but add an -s flag:

```

# hmon -G -Q -s 1:17
1 17 chipSelftest FALSE 0x8cc0 chip selftest active
1 17 chipReceivedInit FALSE 0x8cc1 waiting for init
1 17 chipClkMissing FALSE 0x8cc2 chip clock missing
1 17 portClkMissing FALSE 0x8cc3 port clock missing
1 17 sendPortNotTuned TRUE 0x8cc4 send port not tuned
1 17 recPortNotTuned TRUE 0x8cc5 receive port not tuned
1 17 synchReset FALSE 0x8cc6 synchronous reset active
1 17 powerOnReset FALSE 0x8cc7 power on reset active
1 17 ps1Fail FALSE 0x8cc8 power supply #1 failure
1 17 ps2Fail FALSE 0x8cc9 power supply #2 failure
1 17 warningOC FALSE 0x8cca overcurrent warning
1 17 shutdownOC FALSE 0x8ccb overcurrent shutdown
1 17 psParallelFail FALSE 0x8ccc parallel power failure
1 17 epromProgram FALSE 0x8ccd program eprom complete
1 17 epromErase FALSE 0x8cce erase eprom complete
1 17 clockSource 0 0x8ce0 active clock source
1 17 osc 2 0x8ce1 active oscillator
1 17 pll 2 0x8ce2 active phase lock loop
1 17 chip 6 0x8ce3 active chip
1 17 port 0 0x8ce4 active port
1 17 mux 0 0x8c57 switch MUX setting
1 17 powerLED 1 0x8c47 node/switch LED 1 (green)
1 17 envLED 0 0x8c48 node/switch LED 2 (yellow)
1 17 P48OK TRUE 0x8c49 +48 volt power on
1 17 nodePower TRUE 0x8c4a DC-DC power on
1 17 nodePowerOn10Sec TRUE 0x8c4b DC-DC power on > 10 secs
1 17 P48Off10Sec FALSE 0x8c4c power off > 10 secs
1 17 fanwarning1 FALSE 0x8c4d fan 1 warning
1 17 fanwarning2 FALSE 0x8c4f fan 2 warning
1 17 fanwarning3 FALSE 0x8c51 fan 3 warning
1 17 fanwarning4 FALSE 0x8c53 fan 4 warning
1 17 fanwarning5 FALSE 0x8c55 fan 5 warning
1 17 warningP3_3High FALSE 0x8ccf +3.3 volt high warning
1 17 shutdownP3_3High FALSE 0x8cd0 +3.3 volt high shutdown
1 17 warningP3_3Low FALSE 0x8cd1 +3.3 volt low warning
1 17 shutdownP3_3Low FALSE 0x8cd2 +3.3 volt low shutdown
1 17 warningTemp FALSE 0x8c58 temperature warning
1 17 shutdownTemp FALSE 0x8c59 temperature shutdown

```

This time the command shows your hardware variable names in the third column. For a more detailed description for the chipSelftest hardware variable, see Appendix E, “System Monitor Variables” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*

This publication provides you with many tables. You can find the chipSelftest hardware variable in the table of SP Switch Variables for Card Type 129. It is described as follows:

Indicates whether the active chip, as defined by “chip,” is currently in selftest mode as a result of a power on reset, as defined by “powerOnReset” ([True] = in selftest mode, [False] = not in selftest mode)

You may find some hardware variables in more than one table. If this is the case, look up the appropriate table using your switch supervisor card type. To find the switch board supervisor card type, issue the `hmmmon` command:

```
# hmmmon -G -Q -s -v type 1:17
1 17 type 129 0x8c3a supervisor type
#
```

The fourth column indicates the decimal value of the variable. Now you know your switch supervisor card type is 129. Alternatively, you can find the hexadecimal value in Figure 68 on page 273 for the supervisor type field.

10.4.4.2 Real-time monitor

Issue the `-q` flag instead of `-Q` flag in conjunction with the `hmmmon` command:

```
# hmmmon -G -q 1:17
```

The effect is that the system prompt will not return, and the hardware status changes will be written to the console.

Part 4. Managing the SP system events

Chapter 11. Managing Events

An *event* is the result of a change in the state of a resource within a system partition of your SP system. Examples of resources include nodes, disk drives, memory, software applications, and file systems.

Managing events is taking actions automatically when resource changes that are important to you have occurred in your SP system. For example, you can be notified by a pop up window message when a node goes down or becomes unreachable, or when the system is close to running out of paging space, or when there is something wrong with the SP Switch.

To manage events, this chapter provides you with the following tools:

- Event Perspective
- Command line tools

The command line tools refer to two commands:

- `pmandef`
- `haemqvar`

Using both tools provides you with effective event management for your SP system.

11.1 Event Management subsystem concepts

The Event Management subsystem is a one of the components of RS/6000 Cluster Technology (RSCT). It is important to understand the concepts underlying Event Management subsystem prior to using it. This section describes some basic terminologies and its usage.

11.1.1 What is a resource variable?

The Event Management subsystem informs interested parties whenever the state of a resource changes. A resource is an entity in the system that provides a set of services. Examples of resources include hardware entities, such as processors, disk drives, and memory, as well as software entities, such as database applications, processes, and file systems. Each resource in the system has one or more attributes that define its state.

The resource state change is reflected by a change in the resource attribute. Each resource attribute is represented by a *resource variable*, which is the representation of the attribute in the Event Management subsystem.

Associated with each resource variable is a name, description, value type, data type, resource identifier, and, optionally, a location.

The resource variable name is a string that consists of a resource name, followed by a period, followed by the resource attribute.

The list below includes three examples of resource variable names. Their *resource attributes* are %totused, busy, and power, respectively. Everything to the left of the resource attribute represents the name of the *resource* (for example, IBM.PSSP.aixos.FS). Note, that the name represents a hierarchical organization, from the general to the specific, similar to an AIX file name.

- IBM.PSSP.aixos.FS.%totused
- IBM.PSSP.aixos.Disk.busy
- IBM.PSSP.HW.Frame.power

Although, in simple, terms an *event* is the change in the state of a resource. More correctly, an event is the notification that a relational expression, when applied to a resource variable, evaluates to TRUE.

If you like to know what kind of resource variables are available in your SP system, use the `haemqvar` command. For the command usage, refer to 11.5.1, “Listing resource variables” on page 302.

11.1.2 What is a resource ID?

Most systems include multiple instances of resources. For example, there is more than one logical volume per node, more than one processor per node (if the node is an SMP node), more than one pool of kernel memory buffers, more than one node in the system partition, and so on.

There are also multiple instances of the resource variables representing these resource states. To uniquely identify each instance of a resource and all of its resource variables, each resource in the system has its own resource ID.

The *resource ID* is specified as an *element name/element value* pair. The pairing consists of a element name, followed by an equal sign, followed by the element value. For example, node 5 is identified using the resource ID: NodeNum=5, where NodeNum is the element name, and 5 is the element value.

The element name is a string that identifies the resource. An element value identifies the particular instance of the resource. It can be:

- A single value
- A range of values
- A comma-separated list of single values
- A comma-separated list of ranges

As the preceding list suggests, an element name can have multiple element values. For example, the element name:

`NodeNum=1, 3, 5`

represents three separate resources: Nodes 1, 3, and 5. A range takes the form a-b and is valid only for integer values.

Note, that a resource ID can contain up to four element names. For example, a variable that represents the number of used blocks in the /tmp file system requires a resource ID with three element names: The node ID, the volume group name, and the logical volume name. On the other hand, a variable that represents the power state of a node requires only a single element, the node ID.

Table 10 shows some examples of resource variable names and related resource IDs:

Table 10. Related Resource ID for Resource Variable Name

Resource Variable Name	Resource ID
IBM.PSSP.aixos.CPU.gluser	NodeNum=5
IBM.PSSP.aixos.CPU.kern	NodeNum=5;Cpu=cpu0
IBM.PSSP.aixos.Mem.Kmem.inuse	NodeNum=5;Type=mbuf
IBM.PSSP.aixos.Disk.busy	NodeNum=5;Name=hdisk1
IBM.PSSP.aixos.FS.%totused	NodeNum=5;VG=rootvg;LV=lv100
IBM.PSSP.aixos.SP_HW.Frame.frACLED	FrameNum=1

Note, that although most resource IDs contain a node number, some variables are not associated with a particular node.

If you like to know what kind of resource IDs are available for the resource variable, use the `haemqvar` command. For the usage of the command, refer to 11.5.2, “Getting an explanation of resource variable” on page 302.

11.1.3 What is an event expression?

As previously mentioned, an event is the notification that a relational expression, when applied to a resource variable, evaluates to TRUE. The expression is specified as part of a condition.

The *expression* compares the resource variable to a constant or the previous value of the variable. When the expression is observed to be true, an event is generated. An example of an expression is $X < 10$, where X represents the resource variable IBM.AIX.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 percent, the Event Management subsystem generates an event. Note, that X does not always represent a percentage; it represents other types of values as well.

If you like to know how you can specify event expression for the resource variable, use the `haemqvar` command. For the usage of the command, refer to 11.5.2, “Getting an explanation of resource variable” on page 302.

11.1.3.1 Specifying an Expression

Use an uppercase letter *x* to represent the resource variable. Unmodified *x* represents the value of the latest observation of an instance of the resource variable.

You can also use other characters to modify the meaning of *x* as follows:

- Use *P* to indicate the value of the previous observation of the instance.
- Use *R* to indicate the raw value of the variable instance. This is useful only with variables of type Counter.
- Use the serial number of a specific structured field to indicate the value of the field in a structured byte string (SBS) variable.

The expression you specify can be arithmetic, logical, or a combination of both. The variable name may be repeated in the expression.

Here are some examples of valid expressions:

- | | |
|---------------------------------------|--|
| X == 0 | The value of the resource variable instance is equal to zero. |
| X &< 20 X > 80 | The value of the resource variable instance is less than 20 or greater than 80. |
| !(X &< 20 X > 80) | The value of the resource variable instance is neither less than 20 nor greater than 80. |

X@R > X@PR	The current raw value of the variable instance is greater than the raw value of the variable instance from its previous observation.
X >= X@P + 5	The current value of the variable instance is greater than or equal to the value of the variable instance from its previous observation plus 5.
X@2 != 100	For a resource variable instance that is defined as a structured byte string, the value of structured field number 2 is not equal to 100.

This expression is referred to as the *event expression* when differentiating it from the rearm expression.

11.1.4 What is a rearm expression?

A *rearm expression* generates an event that alternates with the event expression as follows:

1. The event expression is used until it is true, then
2. The rearm expression is used until it is true, then
3. The event expression is used, and so on.

The rearm expression can be used in a couple of different ways. Generally, the rearm expression is the inverse of the event expression, but this is not necessary. For example, if the event expression tests whether a resource variable value is on, the rearm expression tests whether it is off.

Here is another example. Suppose you want to know how much disk space is used in your system. You decide that normal disk space usage should not exceed 90 percent. If the amount of disk space used rises above 90 percent, you want to be notified. Likewise, you also want to be notified once it falls to an acceptable level again (90 percent or less).

In this case, you would define the event expression to be: The amount of disk space used is greater than 90 percent. The expression you would write to support this would be $X > 90$. This expression says that an event will occur when disk space usage (X) rises above 90 percent. You would then define the rearm expression to be: The amount of disk space used is less than 91 percent. The rearm expression you would write to support it would be $X < 91$. This says that a rearm event will occur when the disk space usage (X) falls to 90 percent or less.

The first time disk space usage rises above 90 percent, you are notified, and the Event Management subsystem switches to the rearm expression. You are not notified again until disk space usage drops to 90 percent or below (at which time the Event Management subsystem switches back to the event expression).

Use the same rules and syntax for the rearm expression as you used for the event expression.

If you like to know how you can specify rearm expression for the resource variable, use the `haemqvar` command. For the usage of the command, refer to 11.5.2, “Getting an explanation of resource variable” on page 302.

11.2 Security Considerations for the Event Perspective

To perform some Event Perspective tasks, there may be some special authorization that is required. This section will explain what type of restrictions may occur and what type of security considerations you should have when you define new conditions or events. Also, this section will include some common problems that occur with Event Perspective and suggestions on how to fix it.

For more information about problem management authorization, refer to Chapter 25, “Using the Problem Management Subsystem” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

11.2.1 To define new conditions

To define new conditions, your user name must have root authority.

If your user name does not have root authority:

- You can create event definitions using existing conditions, but you cannot create or modify conditions.
- Any buttons, tool bar icons, or menu bar options for creating or modifying a condition will not be selectable.

11.2.2 To define event or rearm event actions

To define actions to take in response to an event or rearm event, your user name must be associated with a Kerberos principal, and your Kerberos principal must be authorized to create problem management subscriptions.

If your Kerberos principal is not authorized to create problem management subscriptions:

- You can select notification options, but you cannot define actions in response to an event or rearm event.
- The Actions page of the Event Definition Notebook is not selectable.

To authorize the Kerberos principal to create problem management subscriptions:

1. Associate your user name with a Kerberos principal.
2. Add the Kerberos principal to the `/etc/sysctl.pman.acl` on the control workstation (CWS).
3. Copy the new file to all of the nodes. On the CWS, issue the `pcp` command:

```
# pcp -a /etc/sysctl.pman.acl /etc/sysctl.pman.acl
```

11.2.3 To take action

For any actions to succeed that you have defined, your Kerberos principal must also be authorized to take the requested actions.

If your Kerberos principal is not authorized to take the actions you define, the actions fail.

To authorize the Kerberos principal to take the actions you plan to define:

- To be able to execute a command, add the Kerberos principal to the `$HOME/.klogin` file of the user that will be used to run the command.
- To be able to write an entry in the AIX Error Log and BSD Syslog or to generate an SNMP trap, add the Kerberos principal to the root user's `$HOME/.klogin` file.

11.3 Using the Event Perspective

The Event Perspective, as part of the SP Perspectives, provides you with a graphical user interface (GUI) to manage events. There are many pre-defined event definitions available. If you register some of them to the Event Perspective, you can utilize them very easily.

This section implements the following scenario by using Event Perspective.

Scenario

When one of the switch adapters on the node stops responding, or starts responding, you want to know these events from the pop up event notification window on the CWS.

To implement this scenario by using the `pmandef` command, refer to 11.4, “Using the `pmandef` command” on page 296, or by using the `haemqvar` command, refer to 11.5, “Using the `haemqvar` command” on page 301.

For an easy understanding, this section uses only the pre-defined event definition. Event Perspective has a lot of flexibility and useful functions. For complete understanding, refer to Chapter 4, “Using the Event Perspective Effectively” in *SP Perspectives: A New View of Your SP System*, SG24-5180.

11.3.1 Starting Event Perspective

Start the Event Perspective by issuing the `spevent` command:

```
# spevent
```

After a title screen, the Event Perspective window is displayed as shown in Figure 69 on page 287.

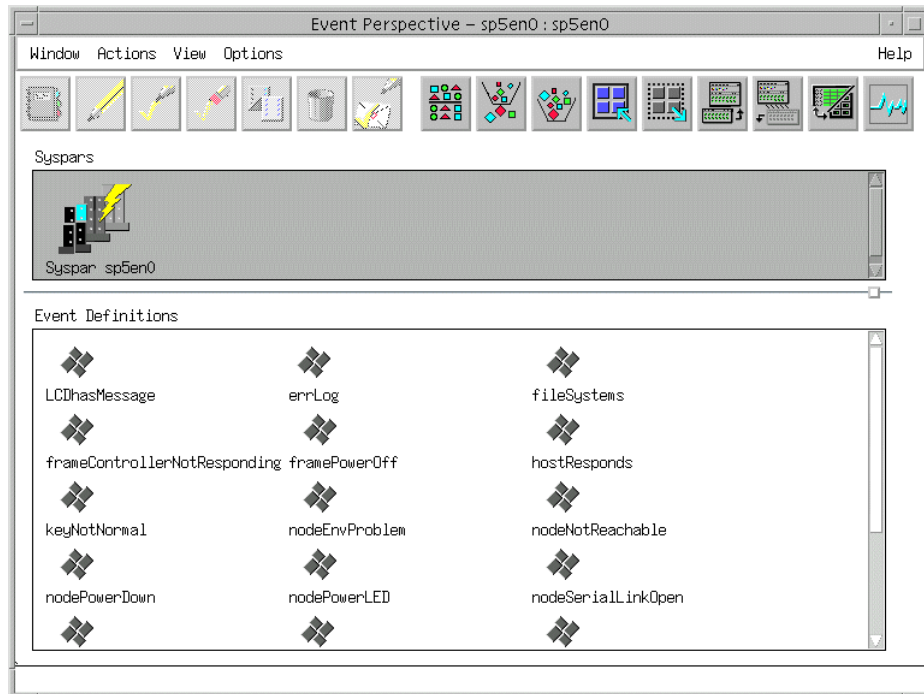


Figure 69. Event Perspective window

11.3.2 Viewing an event definition

There are 19 pre-defined event definitions available in the Event Definitions pane. To decide if your requirement fits one of these pre-defined event definitions, you need to know the contents of the event definitions.

To know the contents of the event definitions, use the following steps:

1. Viewing an event definition

At this point, you can decide if the event definition fits your requirement. But, if you would like to know more about this event definition, you can follow the two additional steps.

2. Viewing a condition
3. Viewing a resource variable

Step 1: Viewing an Event Definition

Double click the **switchResponds** icon, for example:



you will see the Event Definition notebook for switchResponds definition shown in Figure 70.

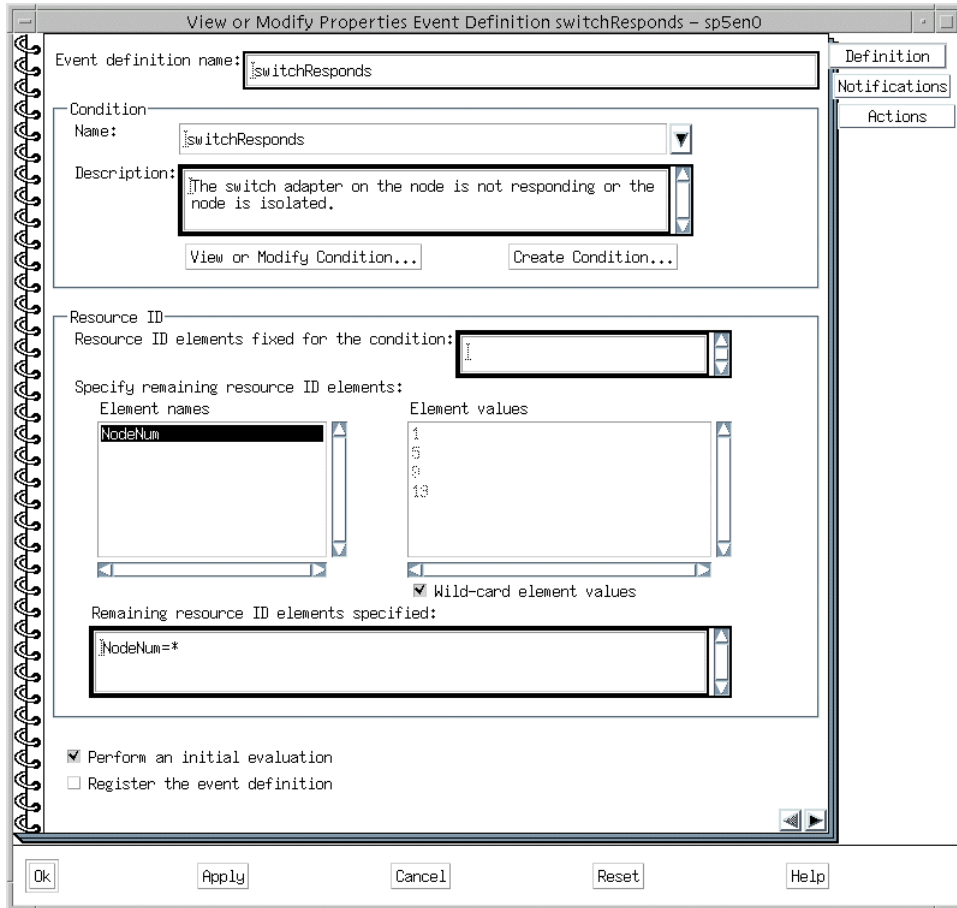


Figure 70. Event Definition Notebook

Some of the parameters are hard to read because the Event Perspective does not allow you to manipulate these parameters.

The following are the important fields that you need to pay attention to:

- Event definition name:

This field is the name of event definition that you choose. In this example, it is switchResponds.

- Name:

This field is the name of the condition that the switchResponds event definition uses. In this example, it is switchResponds. Do not get confused. Both the event definition and condition uses the same name, switchResponds, but they are different.

- Description:

This field is the description for the switchResponds condition. You can know what kind of event this condition manages.

- Remaining resource ID elements specified:

This field is the resource ID for the resource variable that the switchResponds condition uses. In this example, it is NodeNum=*. This means the switchResponds condition is applied for all the nodes in your SP system.

Now you know switchResponds event definition is the one you can use to implement the scenario described on page 286.

Attention

Actually, the scenario is created from the switchResponds pre-defined event definition. In the real case, you may not be able to find a pre-defined event definition that you can use. If this is the case, Event Perspective allows you to create your own event definition. To do this, refer to Chapter 4, “Using the Event Perspective Effectively” in *SP Perspectives: A New View of Your SP System*, SG24-5180.

Step 2: Viewing a condition

You can know about the condition when you read the Description: field shown in Figure 70 on page 288. There are more details about the switchResponds condition. To show it, click the **View or Modify Condition...** button underneath of Description: field. You will see the Condition notebook as shown in Figure 71 on page 290.

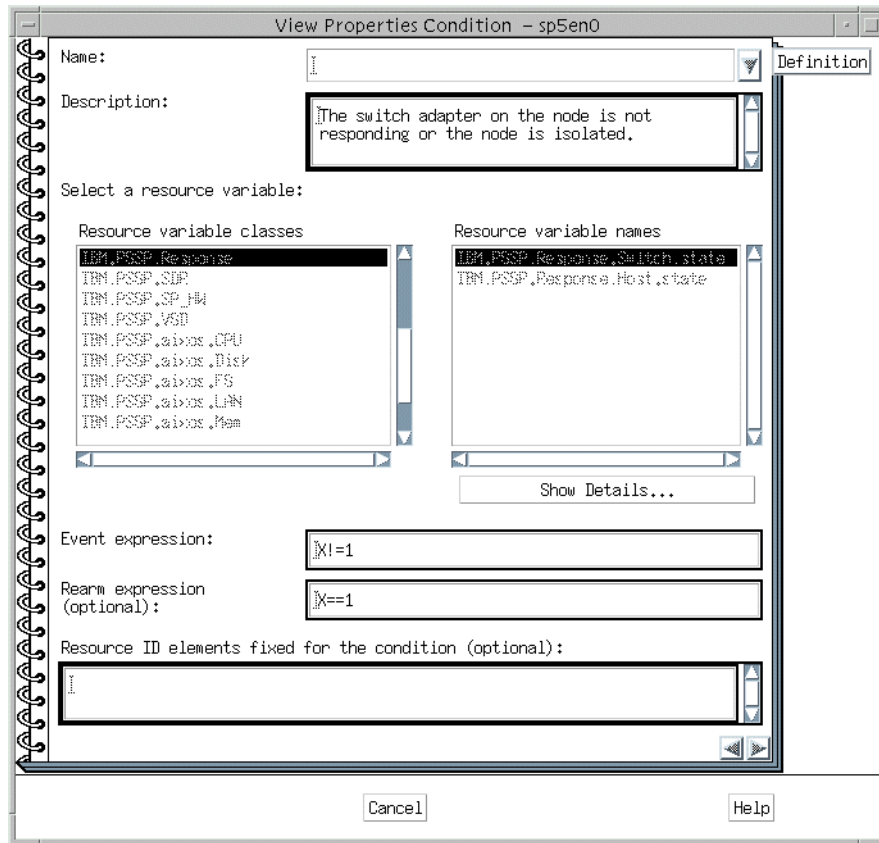


Figure 71. Condition notebook

The following fields are the important fields that you need to pay attention to:

- **Event expression:**
This field is an event expression for the resource variable that the switchResponds condition uses. It is $X \neq 1$. This means that if the value of resource variable becomes other than 1, the event is generated.
- **Rearm expression (optional):**
This field is a rearm event expression for the resource variable that the switchResponds condition uses. It is $X = 1$. This means that if the value of resource variable becomes 1, the rearm event is generated.

Step 3: Viewing a resource variable

To know what kind of resource variable the switchResponds condition uses, click the **Show Details...** button in the Condition notebook as shown in Figure

71. You will see the Show Resource Variable Details window as shown in Figure 72.

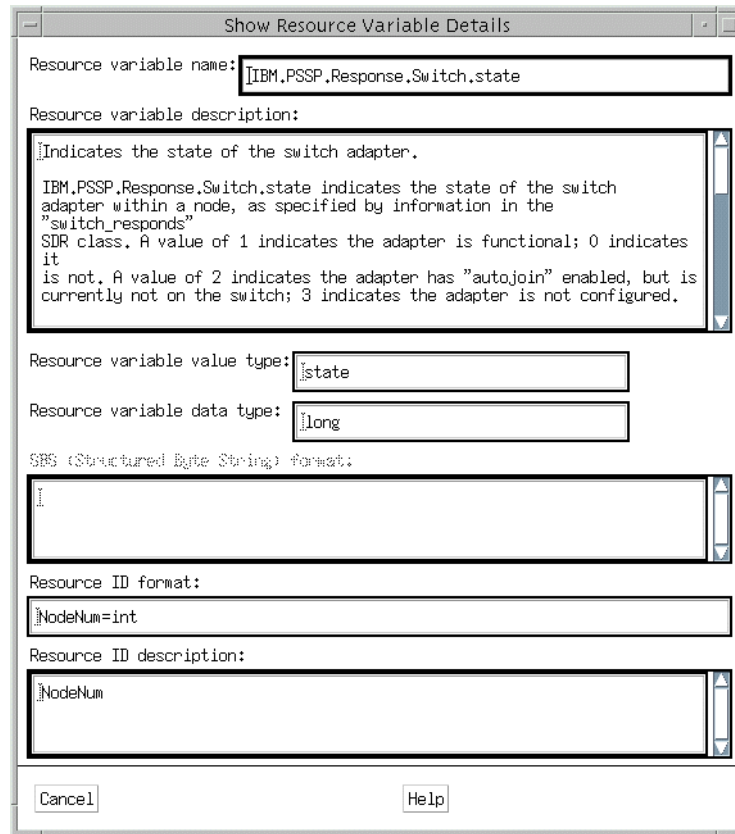


Figure 72. Resource Variable Details window

The following fields are the important fields that you need to pay attention to:

- Resource variable name:
This field is the name of resource variable that the switchResponds condition uses. It shows that the switchResponds condition uses the IBM.PSSP.Resource.Switch.state resource variable.
- Resource variable description:
This field is a complete description for the resource variable that the switchResponds condition uses. It describes the IBM.PSSP.Resource.Switch.state resource variable.

11.3.3 Registering an event definition

If you are satisfied with one of the 19 pre-defined event definitions, and want to use it, it is time to register it.

To register the event definition, click the **switchResponds** icon in the Event Definitions pane to implement the scenario described on page 286:



Then click the **register** icon on the tool bar:



The switchResponds icon turns into the four colors icon:



That is all for registering an event definition. Now, what you need to do is just wait for the events.

11.3.4 Checking event notification

When the event occurs the event definition icon turns to the envelope icon:



At the same time, the Event Notification Log window, as shown in Figure 73, will pop up.

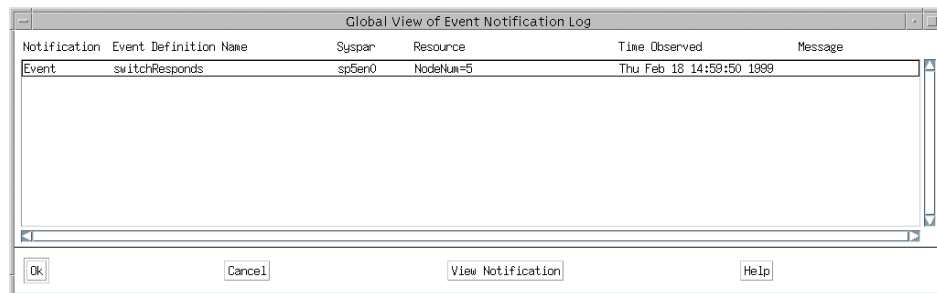


Figure 73. Event Notification Log Window

To know more details about the event, click the **Event** in the list, then click the **View Notification** button. You will see the View Event Notification window as shown in Figure 74.

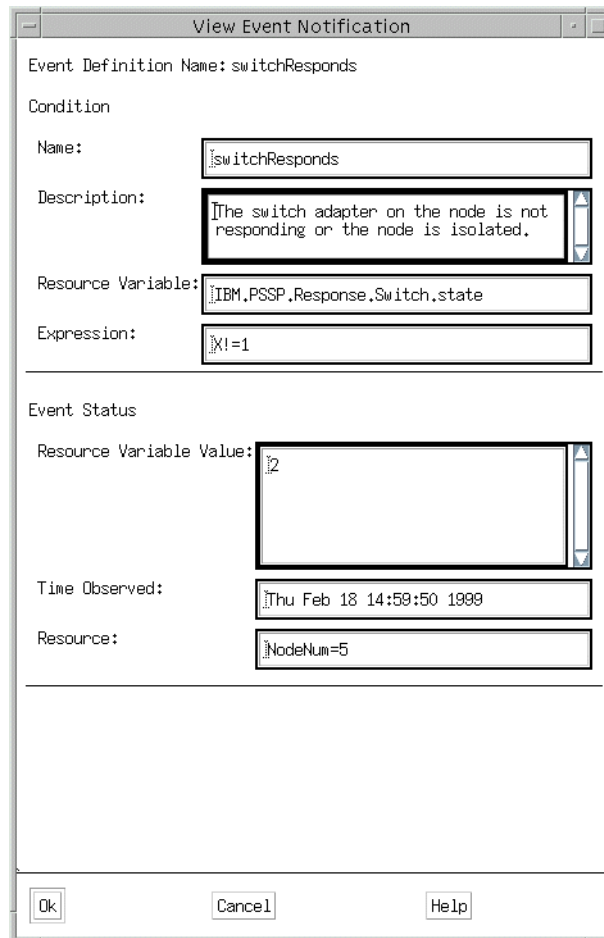


Figure 74. Event Notification window

11.3.5 Checking rearm event notification

When the rearm event occurs the event definition icon turns from the envelope icon to four colors icon:



At the same time, the Event Notification Log window, as shown in Figure 75, will pop up.

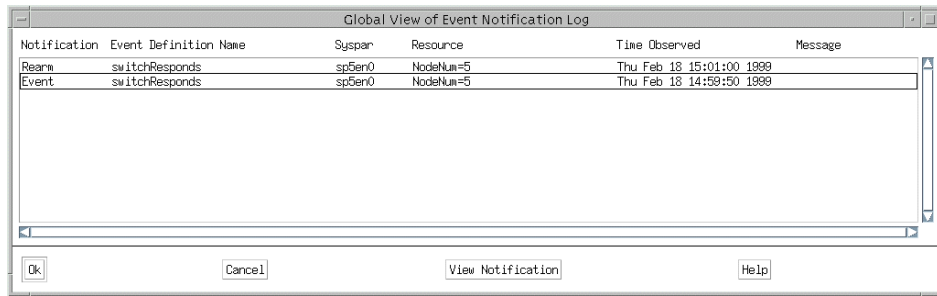


Figure 75. Event Notification log window

To know more details about the rearm event, click the **Rearm** in the list, then click the **View Notification** button. You will see the View Event Notification (Rearm) window as shown in Figure 76 on page 295.

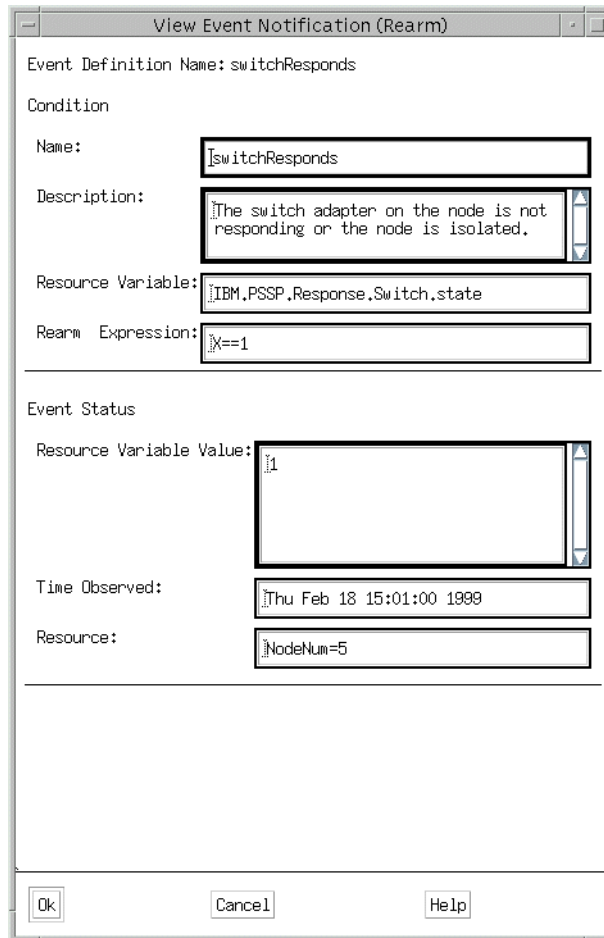


Figure 76. Event Notification (rearm) window

11.3.6 Unregistering event definition

When you do not need to manage the event definition that you choose, you can unregister the event definition. Click the **switchResponds** icon in the Event Definitions pane:



Then click the **unregister** icon on the tool bar:



The icon becomes the grey color icon:



You will not be notified of the event any more.

11.4 Using the pmandef command

The Problem Management subsystem (pman) provides you with a command interface to the Event Management subsystem. One of the commands of pman, the `pmandef` command, allows you to subscribe and unsubscribe events to the pman.

This section implements the following scenario by using the `pmandef` command.

Scenario

When one of the switch adapters on the node stops responding, or starts responding, you want to know these events from the pop up event notification window on the CWS.

To implement the scenario by using the Event Perspective, refer to 11.3, “Using the Event Perspective” on page 285, or by using the `haemqvar` command, refer to 11.5, “Using the haemqvar command” on page 301.

For more information on the syntax of the `pmandef` command, refer to *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351.

11.4.1 pmandefaults file

To subscribe the event, you can use the `pmandef` command. If you are not familiar with using the `pmandef` command, a good start is to have a look at the `pmandefaults` file in the `/usr/lpp/ssp/install/bin` directory on the CWS.

This file is a Korn shell script and defines several default events. The following excerpt gives you an idea of what it looks like.

```

#
# Watch /var space on each node in the partition
#
pmandef -s varFull \
-e 'IBM.PSSP.aixos.FS.%totused:NodeNum=*;VG=rootvg;LV=hd9var:X>95' \
-r 'X<70' \
-c /usr/lpp/ssp/bin/notify_event \
-C "/usr/lpp/ssp/bin/notify_event -r" \
-n 0 -U root -m varFull
[[ $? -ne 0 ]] && print -u2 "Problem with varFull" && exit 1

```

It subscribes an event named `varFull` using the `IBM.PSSP.aixos.FS.%totused` resource variable. If total usage of the `/var` file system becomes over 95 percent in any of your SP nodes, `pman` issues the `/usr/lpp/ssp/bin/notify_event` script. If total usage becomes below 70 percent, `pman` issues the `/usr/lpp/ssp/bin/notify_event -r` script. In the both case, `pman` issues the script on the CWS as a root user.

11.4.2 Subscribing an event

For using the `pmandef` command to subscribe an event to implement the scenario described on page 296, you need to set up the following flags:

- s** This flag sets up the handle name.
- e** This flag sets up the resource variable, resource ID, and expression.
- r** This flag sets up the rearm expression.
- c** This flag sets up the command for event.
- C** This flag sets up the command for rearm event.

The following sections describe how you can set up these flags.

11.4.2.1 Setting up handle name

You need a *handle name* when you subscribe and unsubscribe your event. The scenario uses `MySwitchResponds` as handle name to make it similar to the example in “Step 1: Viewing an Event Definition” on page 288.

Use `-s` flag to set up the handle name:

```
-s MySwitchResponds
```

11.4.2.2 Setting up resource variable, resource ID, and expression

To set up a *resource variable*, *resource ID*, and *expression*, use the following steps:

Step 1: Selecting a resource variable

To monitor the switch adapter on the node, you need to use the following resource variable:

```
IBM.PSSP.Response.Switch.state
```

This is the same resource variable that the example in “Step 3: Viewing a resource variable” on page 290 uses.

To know the details about this resource variable, use the `haemqvar` command described in 11.5.2, “Getting an explanation of resource variable” on page 302.

Step 2: Selecting resource ID

When you have selected the resource variable, you need to specify a resource ID for it. The scenario monitors all the node on your SP system; so, you need to specify a resource ID as follows:

```
NodeNum=*
```

This is the same resource ID that is used by the example in “Step 1: Viewing an Event Definition” on page 288.

Step 3: Specifying expression

To complete the parameters for `-e` flag, you need an event expression. The scenario wants to have the event when the switch adapter on the node stops responding; so, you need to specify event expression as follows:

```
X!=1
```

This is the same event expression that is used in the example in “Step 2: Viewing a condition” on page 289.

Now, you can set up `-e` flag as follows:

```
-e 'IBM.PSSP.Response.Switch.state:NodeNum=:X!=1'
```

11.4.2.3 Setting up rearm expression

The scenario wants to have the rearm event when the switch adapter on the node starts responding. To specify this rearm event, use the following flag and rearm expression:

```
-r 'X==1'
```

This is the same rearm expression that is used in the example in “Step 2: Viewing a condition” on page 289.

11.4.2.4 Setting up event command

When the event occurs, the scenario wants to pop up the event notification window. To pop up the window, you can use the `dterror.ds` command. Use this command with the following parameters:

```
# /usr/dt/bin/dterror.ds \  
> "The node stopped responding" "Event Notification" "OK"
```

The command brings up the window shown in Figure 77.

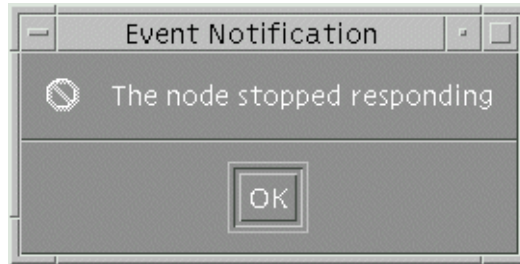


Figure 77. Event Notification Window

To execute this command from the `pmandef` command, use the `dterror.ds` command with the `aixterm` command:

```
-c "aixterm -display sp3en0:0.0 -i -e /usr/dt/bin/dterror.ds \  
  \"The node stopped responding\" \"Event Notification\" \"OK\""
```

11.4.2.5 Setting up rearm command

When the rearm event occurs, the scenario wants to pop up the rearm notification window. To pop up the window, you can use the `dterror.ds` command. Use this command with the following parameters:

```
# /usr/dt/bin/dterror.ds \  
> "The node start responding" "Rearm Notification" "OK"
```

The command brings up the window shown in Figure 78 on page 300:

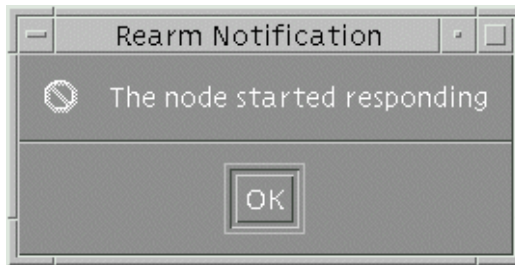


Figure 78. Rearm Notification window

To execute this command from the `pmandef` command, use the `dterror.ds` command with the `aixterm` command:

```
-C "aixterm -display sp3en0:0.0 -i -e /usr/dt/bin/dterror.ds \  
  \"The node started responding\" \"Rearm Notification\" \"OK\""
```

11.4.2.6 Subscribing an event

To implement the scenario described on page 296, you need to subscribe the event to `pman`.

To subscribe the event, issue the `pmandef` command:

```
# pmandef \  
-s MyswitchResponds \  
-e 'IBM.PSSP.Response.Switch.state:NodeNum=:X!=1' \  
-r 'X==1' \  
-c "aixterm -display sp3en0:0.0 -i -e /usr/dt/bin/dterror.ds \  
  \"The node stopped responding\" \"Event Notification\" \"OK\"\" \  
-C "aixterm -display sp3en0:0.0 -i -e /usr/dt/bin/dterror.ds \  
  \"The node started responding\" \"Rearm Notification\" \"OK\"\" \  
#
```

11.4.3 Listing events

You may want to know if your event is subscribed in `pman` correctly.

To list the events currently subscribed to `pman`, issue the `pmanquery` command:

```
# pmanquery -n all
pmActivated:pmHandle:pmRvar:pmIvec:pmPred:pmCommand:pmCommandTimeout:pmTrapId:pm
PPSLog:pmText:pmRearmPred:pmRearmCommand:pmRearmCommandTimeout:pmRearmTrapId:pmR
earmPPSLog:pmRearmText:pmUsername:pmPrincipal:pmHost:pmTargetType:pmTarget:pmUse
rLabel:pmInitEval
1:MySwitchResponds:IBM.PSSP.Response.Switch.state.NodeNum=*:X!=1:/tmp/event_noti
fy:0:-1:0: :X=1:/tmp/rearm_notify:0:-1:0: :root:root.admin@MSC.ITSO.IBM.COM:sp3
en0.msc.itso.ibm.com:NODE_LIST:sp3en0: :-1
#
```

In this example, you know the MySwitchResponds event is successfully subscribed.

11.4.4 unsubscribing the event

When you do not need to manage the MySwitchResponds event any more, you can unsubscribe the event.

To unsubscribe the event, issue the `pmandef` command:

```
# pmandef -u MySwitchResponds
sp3en0.msc.itso.ibm.com: event MySwitchResponds unsubscribed
#
```

11.5 Using the haemqvar command

The Event Management subsystem itself provides you with the command interface. You can ask the Event Management about the usage or status of resource variables. There is the `haemqvar` command available for this purpose.

This section focuses on the usage of the `haemqvar` command. The `haemqvar` command provides you with static information of resource variables. To use the command as a real time event management tool, this section gives you some hints and tips.

For more information on the syntax of the `haemqvar` command, refer to *IBM Parallel System Support Programs for AIX: Command and Technical Reference, SA22-7351*.

11.5.1 Listing resource variables

There are more than 400 resource variables available for the SP system. To list all the resource variables and a short explanation of them, you can issue the `haemqvar` command:

```
# haemqvar -d
```

This command displays more than 400 lines; so, you may need to use this command with the other commands, such as the `more` or `grep` command. The list is quite lengthy. Therefore, the following output uses the `haemqvar` command with the `grep` command to look for the resource variables that are related to the word *response*:

```
# haemqvar -d | grep -i response
IBM.PSSP.Response.Switch.state  Indicates the state of the switch adapter.
IBM.PSSP.Response.Host.state    Indicates if the node has connectivity over the e
n0 adapter.
IBM.PSSP.VSDdrv.rejected_responds  Rejected responses.
#
```

From the descriptions, you can easily guess what the resource variables are used for.

You can list out all the resource variables that belong to the specific resource class. To list up all the resource variables of `IBM.PSSP.Response` resource class, issue the `haemqvar` command:

```
# haemqvar -d "" IBM.PSSP.Response "*"
IBM.PSSP.Response.Switch.state  Indicates the state of the switch adapter.
IBM.PSSP.Response.Host.state    Indicates if the node has connectivity over the e
n0 adapter.
#
```

In the case of the `IBM.PSSP.Response` resource class, it has two resource variables.

11.5.2 Getting an explanation of resource variable

After having decided which resource variable you are interested in, list a detailed explanation about it. To display a detailed explanation of the `IBM.PSSP.Response.Switch.state` resource variable, issue the `haemqvar` command:

```

# haemqvar "" "IBM.PSSP.Response.Switch.state" "*"
Variable Name:  IBM.PSSP.Response.Switch.state
Value Type:    State
Data Type:     long
Initial Value:  0
Class:         IBM.PSSP.Response
Locator:
Variable Description:
    Indicates the state of the switch adapter.

    IBM.PSSP.Response.Switch.state indicates the state of the switch
    adapter within a node, as specified by information in the "switch_responds"
    SDR class. A value of 1 indicates the adapter is functional; 0 indicates it
    is not. A value of 2 indicates the adapter has "autojoin" enabled, but is
    currently not on the switch; 3 indicates the adapter is not configured.

    This variable is supplied by the "Response" resource monitor.

    The resource variable's resource ID specifies the number of the node
    containing the adapter. To register an event that indicates the switch
    adapter on node 5 is not functional, the variable, resource ID and expression
    would be:

        Resource Variable: IBM.PSSP.Response.Switch.state
        Resource ID:      NodeNum=5
        Expression:       X == 0

    Resource ID wildcarding:

    The resource ID element may be wildcarded.

    Related Resource Variables:

        IBM.PSSP.Response.Host.state
        IBM.PSSP.Membership.LANAdapter.state
        IBM.PSSP.Membership.Node.state

Resource ID:    NodeNum=int
                The number of the node.

#

```

The command displays a detailed explanation about the resource variable with a typical usage. You can refer to this when you use Event Perspective described in 11.3, "Using the Event Perspective" on page 285, or when you use the `pmundef` command described in 11.4, "Using the `pmundef` command" on page 296.

The explanation includes a list of related resource variables. You may find a more adequate resource variable from them.

11.5.3 Getting the value of a resource variable

Other than an explanation, the `haemqvar` command gives you the current value of a resource variable. To know the current value of `IBM.PSSP.Response.Switch.state`, for example, issue the `haemqvar` command:

```
# haemqvar -c "" IBM.PSSP.Response.Switch.state "*"
0 IBM.PSSP.Response.Switch.state NodeNum=15 1
0 IBM.PSSP.Response.Switch.state NodeNum=14 1
0 IBM.PSSP.Response.Switch.state NodeNum=13 1
0 IBM.PSSP.Response.Switch.state NodeNum=12 1
0 IBM.PSSP.Response.Switch.state NodeNum=11 1
0 IBM.PSSP.Response.Switch.state NodeNum=10 1
0 IBM.PSSP.Response.Switch.state NodeNum=9 1
0 IBM.PSSP.Response.Switch.state NodeNum=8 1
0 IBM.PSSP.Response.Switch.state NodeNum=7 1
0 IBM.PSSP.Response.Switch.state NodeNum=6 1
0 IBM.PSSP.Response.Switch.state NodeNum=5 1
0 IBM.PSSP.Response.Switch.state NodeNum=1 1
#
```

Each line contains the location (node number) of the resource variable instance, the resource variable name, the resource ID of the instance, and the resource variable instance value, from left to right.

In the case of the `IBM.PSSP.Response.Switch.state` resource variable, value 1 indicates that the switch adapter within a node is functional.

11.5.4 Managing events

As you can imagine, the `haemqvar` command does not provide you with a resource variable value dynamically. In other words, you need to execute this command periodically to check the status. To use the `haemqvar` command as a real time event management tool, this section provides some hints and tips.

This section implements the following scenario by using the `haemqvar` command.

Scenario

When one of the switch adapters on the node stops responding, or starts responding, you want to know these events through the pop up event notification window on the CWS.

To implement this scenario using the Event Perspective, refer to 11.3, “Using the Event Perspective” on page 285, or using the `pmandef` command, refer to 11.4, “Using the `pmandef` command” on page 296.

The following is the Perl script that partially implements the scenario:

```
#!/usr/lpp/ssp/perl5/bin/perl
#
# Event Notification Program
#
$haemqvar = "/usr/sbin/rsct/bin/haemqvar";
$dterror = "/usr/dt/bin/dterror.ds";
$data = "/tmp/data";
$title = "\"Event Notificatoin\"";
$message = "\"The node stopped responding\"";
$button = "\"OK\"";

for ($i = 1; $i <= 20; $i++) {
    system "$haemqvar -c \"\" IBM.PSSP.Response.Switch.state \"*\"> $data";
    open (DATA, "/tmp/data");
    while ($line = <DATA>) {
        $_ = $line;
        tr/ / /s;
        $line = $_;
        @fields = split(/ /, $line);
        if ($fields[3] != 1) {
            exec "$dterror $message $title $button";
        }
    }
    close (DATA);
    system "sleep 1";
    print "$i: No Event\n";
}
```

This script executes the `haemqvar` command every second to check the value of `IBM.PSSP.Response.Switch.state` resource variable:

```
system "$haemqvar -c \"\" IBM.PSSP.Response.Switch.state \"*\"> $data";
```

If the node stopped responding, in other words, the value of resource variable becomes other than 1:

```
if ($fields[3] != 1) {
```

It executes the `dterror.ds` command to pop up the event notification window:

```
exec "$dterror $message $title $button";
```

You will see a window as shown in Figure 79 on page 306 if the event occurs:

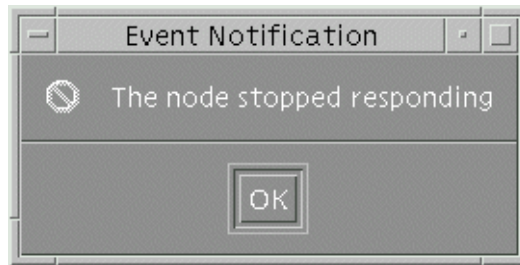


Figure 79. Event Notification window

The script itself is not practical and does not handle the rearm event. But it must provide you with the enough hints and tips so that you can create your own real time event management tool using the `haemqvar` command.

Chapter 12. Managing software resources

A large part of the system administration duties involves managing your resources. Initially, simple monitoring is done, but when an incident occurs, you have to understand how to manage it.

For software resources management on an SP system, you can use almost the same method used for a stand-alone RS/6000 machine. You need to take SP system unique points into consideration. Also, IBM Parallel System Support Programs for AIX (PSSP) provides convenient mechanisms to manage distributed software resources.

This chapter covers file systems, paging space, and log files.

12.1 File systems

Managing file systems on an SP system is no different than managing them on a regular RS/6000 machine. However, they are just as important, and maybe more so, on the SP system. To keep the system healthy, enough space should be allocated on the file system that contains working directory of the system or application. This section explains how to check your file systems and which file systems you should keep your eye on. Suggestions are also given in case a file system gets too large.

12.1.1 Considering file systems

The following is a list of usage and recommended free space for each file system when you manage AIX and PSSP:

- /** This file system is a home directory for the root user. Every time you use the SMIT as a root user on the control workstation (CWS) or nodes, the smit.log or smit.script files will be increased. Other than that, there is no special considerable point. Therefore, there is no recommended free space size.

In case of the CWS or boot/install server (BIS) nodes, you need to consider boot image files used by Network Installation Management (NIM). These files use the /tftpboot directory. If you do not define /tftpboot file system separately, an additional 25 MB of free space per lppsource version level is required.
- /usr** This file system is to hold Licensed Program Products (LPPs) or executable files. This means there should be no write operation to this file system. Therefore, there is no recommended free space size.

/var This file system keeps files that vary on each machine. It means that it keeps log files, temporary work files, and so on. They tend to grow; so, you need to pay attention to this directory. PSSP uses this directory in many ways. Especially, RS/6000 Cluster Technology (RSCT) daemons set their current directory in this file system and keep some configuration files also. And almost every daemon provided by PSSP keeps its log in this file system.

To keep the system healthy and to get log information when trouble occurs, it is recommended to keep 20 to 30 MB free space for the CWS and the nodes.

/tmp This file system keeps system-generated temporary files. Because Kerberos creates ticket cache files in this file system, it must not be full.

The recommended free space depends on the usage of the system. But it is often adjusted as the same size of /var file system.

/home This file system keeps a user's home directory. Because the size also depends on the usage of the system, there is no recommended size.

If you are using CWS as home directory server, you need to pay attention to CWS.

/spdata This file system contains, among other items, mksysb and installp file sets on the CWS. Therefore, it is recommended to create a separate volume group for it.

The minimum size of this file system is 2 GB. However, to keep backup files, PTFs, or other LPPs, it is not enough. 4 GB or more would be better although the System Data Repository (SDR) archive is stored in this file system, no other files tend to grow.

If you assign a node as a BIS node, three file systems are automatically created in the rootvg volume group:

- /dev/install_images
- /dev/spot_name
- /dev/install_pssplpp

Before you assign BIS to node, you need to check the rootvg volume group to see if there is enough space to keep these file systems.

The following is a sample file system configuration for a BIS node:

```

# df
Filesystem      512-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4         32768           5216   85%     1168   15% /
/dev/hd2        598016          75152   88%     9246   13% /usr
/dev/hd9var     65536           40384   39%      395    5% /var
/dev/hd3        65536           62424    5%        44    1% /tmp
/dev/hd1         8192            7840    5%        20    2% /home
/dev/install_images 237568        27360   89%        18    1% /spdata/sys1/inst
all/images
/dev/spot_aix432 434176         55104   88%    11724   22% /spdata/sys1/install
/aix432/spot
/dev/install_pssplpp 458752        248992   46%        22    1% /spdata/sys1/ins
tall/pssplpp
#

```

12.1.2 Getting more available file system space

There are some file systems that tend to grow and can eventually run out of space. This can cause major problems if it is the `/`, `/var`, or `/tmp` file system. Here are some suggestions on how you can get more available file system space when your file system is approaching its full size.

12.1.2.1 / file system

There are some techniques to get more available space for the `/` file system.

Checking the `/etc/security/failedlogin` file

If your `/` file system is close to full, then check the `/etc/security/failedlogin` file. This file contains a list of failed logins, and this can be caused by a `tty` that is respawning too rapidly. To read the contents of the file, use the `who` command:

```

# who /etc/security/failedlogin
root      dtlogin/_0  Mar 17 13:20
UNKNOWN_ pts/0       Mar 22 12:42 (TOT32.itso.ibm.c)
UNKNOWN_ pts/0       Mar 22 13:08 (TOT32.itso.ibm.c)
allnode   pts/1       Mar 25 15:45 (sp3en0)
allnode   pts/1       Mar 25 15:45 (sp3en0)
guest     PC-NFS      Mar 26 11:39 (JOHNDOE.itso.ibm)
UNKNOWN_  PC-NFS      Mar 26 11:41 (JOHNDOE.itso.ibm)
#

```

After reading the information in this file, you can empty it using the `cp` command:

```

# cp /dev/null /etc/security/failedlogin

```

Checking the /dev directory

The /dev directory should contain all of your devices and logical volumes. Look for devices that no longer exist on your system or devices that are not correct. You can remove files that are not correct, for example, bus0, instead of bus0. The devices listed here should have a major and a minor number. You can check what your major and minor numbers are by issuing the `ls` command:

```
# cd /dev
# ls -l | pg
drwxrwx--- 2 root system 512 Mar 29 00:00 .SRC-unix
crw-rw---- 1 root system 10, 0 Feb 18 18:09 IPL_rootvg
srwxrwxrwx 1 root system 0 Mar 09 20:47 SRC
crw----- 1 root system 10, 0 Mar 11 16:58 __vg10
crw----- 1 root system 40, 0 Mar 06 09:44 __vg40
cr--r----T 1 root system 8, 0 Feb 18 18:09 audit
crw----- 1 root system 3, 0 Feb 18 17:01 bus0
brw-rw---- 1 root system 10, 12 Mar 02 18:47 cache
br--r--r-- 1 root system 12, 0 Feb 18 17:01 cd0
crw-rw-rw- 1 root system 13, 0 Feb 18 17:01 clone
crw--w--w- 1 root system 4, 0 Feb 18 17:01 console
crw-rw-rw- 1 root system 39,4096 Feb 18 18:45 dserdasda0
crw-rw--w- 1 uucp uucp 2, 2 Mar 15 11:49 dtremote
```

A correct entry appears that has the following information:

```
crw----- 1 root system 3, 0 Feb 18 17:01 bus0
```

An incorrect entry would appear as a file, such as:

```
crw----- 1 root system 323420 Feb 18 17:01 bus0
```

There are some files in the /dev directory that are invalid; so, look for files that are larger than 500 bytes.

Checking large files

You can also check for large files on your / file systems by issuing the `find` and `sort` commands:

```
# find / -xdev -size +2048 -ls | sort -r +6
```

This will search for files greater than 1 MB in your / file system and sort them in decreasing order. With this list, you can look and see if anything is out of the ordinary.

12.1.2.2 /var file system

The /var file system is another file system that tends to get filled up easily. There are some techniques to get more available space for the /var file system.

Refer to 12.3.5, “Eliminating increased log files” on page 337 for reducing log files from /var file system.

Checking the /var/tmp directory

The first place to look is the /var/tmp directory. This directory may contain some old unused files that can be removed.

Checking the /var/adm/wtmp file

The next big space eater is the /var/adm/wtmp file. This log file contains information about logins, rlogins, and telnets done on the system. This file will grow indefinitely, and it must be monitored. To clean up this file, issue the `cp` command:

```
# cp /dev/null /var/adm/wtmp
```

Checking AIX error log

The AIX error report also is kept in the /var file system. It is located in the /var/adm/ras directory. It is not a file that you can just remove. To empty the file, issue the `errclear` command:

```
# errclear 0
```

This will clear out the error log completely. If you wanted to keep one day's errors, then replace the 0 with a 1.

Checking trace log

If you had a trace running at one point, the trace log would be recorded in the /var/adm/ras/trcfile file. If this file contains old information that is no longer relevant, you can remove it by issuing the `rm` command:

```
# rm /var/adm/ras/trcfile
```

Checking printer spool

If you have a printer attached to your system, you may find that the spool directories are large. They are located in the /var/spool directory. To clear the queueing system, issue the `stopsrc` command to stop the `qdaemon` daemon:

```
# stopsrc -s qdaemon
```

Then erase the files issuing the `rm` command:

```
# rm /var/spool/lpd/qdir/*  
# rm /var/spool/lpd/stat/*
```

```
# rm /var/spool/qdaemon/*
```

Start the qdaemon daemon with the `startsrc` command:

```
# startsrc -s qdaemon
```

This will clear up any old remnant files.

Checking user account

If you are running AIX accounting on your system, the `/var/adm/acct` directory may have some large accounting files. These files can be deleted by the `rm` command.

The `/var/adm/pacct` file is increased each time a process terminates. This file can be cleared by the `cp` command:

```
# cp /dev/null /var/adm/pacct
```

And if you want to stop AIX accounting, issue the `shutacct` command.

Checking the vi command

The `/var/preserve` directory contains old terminated vi sessions. These files can be removed by issuing the `rm` command:

```
# rm /var/preserve/*
```

Checking the su command

Every time you issue the `su` command, it adds a record in the `/var/adm/sulog` file. This file can grow indefinitely. This file can be modified by using a text editor so that it is smaller. If you do not need the log any more, you can simply erase it by issuing the `rm` command:

```
# rm /var/adm/sulog
```

Checking large files

You can also check for large files on your `/var` file system by issuing the `find` and `sort` commands:

```
# find /var -xdev -size +2048 -ls | sort -r +6
```

This will search for files greater than 1 MB in your `/var` file system and sort them in decreasing order. With this list, you can look and see if anything is out of the ordinary.

Using the skulker command

AIX provides the `skulker` command to clean up file systems by removing unwanted files. You can erase `smit.log` or core dump by issuing it.

12.1.3 Monitoring file systems

The free space size of the /var and /tmp is worth monitoring. You can get the currently available capacity of the /var and /tmp file systems by issuing the `pdf` command:

```
# pdf -a /var /tmp
Filesystem      Size-KB  Used-KB  Free-KB %Free  iUsed  iFree %iFree
-----
HOST: sp3n01.msc.itso.ibm.com
-----
/var            32768    5652     27116  83%    356    7836  96%
/tmp           32768    1528     31240  96%     43    8149  100%

HOST: sp3n05.msc.itso.ibm.com
-----
/var            32768    2528     30240  93%    330    7862  96%
/tmp           32768    1120     31648  97%     34    8158  100%

HOST: sp3n09.msc.itso.ibm.com
-----
/var            32768    2536     30232  93%    330    7862  96%
/tmp           32768    1120     31648  97%     34    8158  100%

HOST: sp3n13.msc.itso.ibm.com
-----
/var            32768    2564     30204  93%    334    7858  96%
/tmp           32768    1120     31648  97%     34    8158  100%
#
```

It is important to keep an eye on these file systems so that they do not become 100 percent full.

As long as you execute this command very frequently, you can know if the file system is full immediately. If you feel it is difficult to execute this command so often, there are two possible ways to know automatically if the file system is full immediately:

- Using the Event Perspective
- Using the `pmandef` command

12.1.3.1 Using the Event Perspective

There are three pre-defined event definitions available for Event Perspective:

1. Monitoring file system

Icon:



Name: fileSystemes
Description: Monitor the percentage of usage of the file systems.
Resource Variable: IBM.PSSP.aixos.FS.%totused

2. Monitoring /tmp file system

Icon:



Name: tmpFull
Description: The file system for the hd3 logical volume in the rootvg volume group is running out of space.
Resource Variable: IBM.PSSP.aixos.FS.%totused

3. Monitoring /var file system

Icon:



Name: varFull
Description: The file system for the hd9var logical volume in the rootvg volume group is running out of space.
Resource Variable: IBM.PSSP.aixos.FS.%totused

For more information about these pre-defined event definitions, refer to 11.3, “Using the Event Perspective” on page 285.

12.1.3.2 Using the pmandef command

There are two examples available in the `pmandefaults` script file for the `pmandef` command:

1. Monitoring /tmp file system

The following `pmandef` command example is an excerpt from the `pmandefaults` script file:

```

#
# Watch /tmp space on each node in the partition
#
pmandef -s tmpFull \
  -e 'IBM.PSSP.aixos.FS.%totused:NodeNum=*;VG=rootvg;LV=hd3:X>90' \
  -r 'X<80' \
  -c /usr/lpp/ssp/bin/notify_event \
  -C "/usr/lpp/ssp/bin/notify_event -r" \
  -n 0 -U root -m tmpFull
[[ $? -ne 0 ]] && print -u2 "Problem with tmpFull" && exit 1

```

The `pmandef` command subscribes the event named `tmpFull`. If the `/tmp` file system becomes more than 90 percent full on any node, the `notify_event` script is executed. After having this event, if the `/tmp` file system becomes less than 80 percent full, the `notify_event -r` script is executed. Both scripts are executed on the CWS by the root user.

2. Monitoring /var file system

The following `pmandef` command example is an excerpt from the `pmandefaults` script file:

```

#
# Watch /var space on each node in the partition
#
pmandef -s varFull \
  -e 'IBM.PSSP.aixos.FS.%totused:NodeNum=*;VG=rootvg;LV=hd9var:X>95' \
  -r 'X<70' \
  -c /usr/lpp/ssp/bin/notify_event \
  -C "/usr/lpp/ssp/bin/notify_event -r" \
  -n 0 -U root -m varFull
[[ $? -ne 0 ]] && print -u2 "Problem with varFull" && exit 1

```

The `pmandef` command subscribes the event named `varFull`. If the `/var` file system becomes more than 95 percent full on any node, the `notify_event` script is executed. After having this event, if the `/var` file system becomes less than 70 percent full, the `notify_event -r` script is executed. Both scripts are executed on the CWS by the root user.

For more information about these examples, refer to 11.4, “Using the `pmandef` command” on page 296.

12.2 Paging space

Paging space is a key issue for system administration. An adequate amount of paging space will improve the performance of your system. An insufficient

amount can cause your entire system to crash. This section explains the rule of thumb used for determining the proper amount of paging space required on a system. It also explains how to get paging space information and how to adjust your paging space if it is not sufficient. You can adjust paging space both during and after installation. It is important to monitor the paging space. If it gets full, it is not difficult to increase paging space. However, if you wait too long, your system can hang.

12.2.1 Sizing paging space

The install process creates paging space according to the following formulas:

1. If a system has less than 64 MB of memory:

$A \text{ paging space} = \text{memory size} \times 2$

2. If a system has more than or equal to 64 MB and less than 256 MB of memory:

$A \text{ paging space} = \text{memory size} + 16 \text{ MB}$

3. If a systems has more than or equal to 256 MB of memory:

$A \text{ paging space} = (\text{memory size} - 256 \text{ MB}) \times 1.25 + 512 \text{ MB}$

It is recommended to follow these formulas when you set up paging space on your SP system.

However, in the case a node has a large amount of memory, such as number of GBs, and the node is used by a few small programs and a large amount of data, you could start using paging space the same size as memory size.

12.2.2 Getting paging space information

AIX is installed to SP nodes by using mksysb image. The mksysb image includes the `./image.data` file. This file contains the paging space information. Therefore, if you use a given mksysb image for your SP node to install AIX, it might not follow the paging space allocation formula described in 12.2.1, “Sizing paging space” on page 318.

To check the paging space size indicated in `./image.data` file, issue the `restore` command:

```
# cd /spdata/sys1/install/images
# restore -xf bos.obj.ssp.421 ./image.data
Please mount volume 1 on bos.obj.ssp.421.
  Press the Enter key to continue.

x ./image.data
#
```

Take a look at the contents of ./image.data file. The stanza that has TYPE= paging is the information about paging space.

The following is an excerpt from the ./image.data file that indicates the paging space information:

```
lv_data:
  VOLUME_GROUP= rootvg
  LV_SOURCE_DISK_LIST= hdisk0
  LV_IDENTIFIER= 000119726b4b35b9.1
  LOGICAL_VOLUME= hd6
  VG_STAT= active/complete
  TYPE= paging
  MAX_LPS= 512
  COPIES= 1
  LPS= 18
  STALE_PPs= 0
  INTER_POLICY= minimum
  INTRA_POLICY= middle
  MOUNT_POINT=
  MIRROR_WRITE_CONSISTENCY= off
  LV_SEPARATE_PV= yes
  PERMISSION= read/write
  LV_STATE= opened/syncd
  WRITE_VERIFY= off
  PP_SIZE= 8
  SCHED_POLICY= parallel
  PP= 18
  BB_POLICY= non-relocatable
  RELOCATABLE= yes
  UPPER_BOUND= 32
  LABEL=
  MAPFILE=
  LV_MIN_LPS= 18
  STRIPE_WIDTH=
  STRIPE_SIZE=
```

The LOGICAL_VOLUME= hd6 indicates the logical volume (LV) name of the paging space. In this example, it is hd6. The PP_SIZE= 8 indicates the size of one physical partition (PP). In this example, it is 8. The PP= 18 indicates the number of PPs for this LV. In this example, it is 18.

Therefore, if you use this mksysb image for node installation, the logical volume name for the paging space will be:

hd6

and the size of paging space on the node will be:

$8 \text{ (PP_SIZE)} \times 18 \text{ (PP)} = 144 \text{ MB}$

To check the current size and usage of your paging space, you can issue the `lspvs` command with the `-a` flag:

```
#lspvs -a
Page Space   Physical Volume   Volume Group   Size   %Used   Active   Auto   Type
paging02     hdisk2             pagingvg       400MB   22      yes     no    lv
hd6          hdisk0             rootvg         56MB    100     yes     no    lv
#
```

This shows detailed information for each paging space. If you prefer to view a cumulative view, you can issue the `lspvs` command with `-s` flag:

```
#lspvs -s
Total Paging Space   Percent Used
456MB                32%
#
```

When the Percent Used field gets close to 100 percent, your system will soon have a major problem. Appropriate actions will be required.

12.2.3 Adjusting paging space size during installation

If you use the `/tftpboot/script.cust` file, you can adjust the paging space size during the node installation. You can make use of the `/usr/lpp/ssp/samples/script.cust` file to create `/tftpboot/script.cust` file.

The following (Figure 80 on page 321 and Figure 81 on page 322) shows an excerpt from the `/usr/lpp/ssp/samples/script.cust` file:

```

#-----#
# Modify page space.                                     #
#-----#
#
# To make a new logical volume mypage for page space with 5 logical
# partitions (4 MB per partition - 20 MB total) on the volume group vg04,
# issue the mklv command.
#
#   -t specifies the logical volume type (jfs, jfslog, or paging)
#   -y specifies the name of the logical volume to be used instead of
#       the system generated logical volume name (pagelvXX - where XX
#       starts at 00 and increases by one).
#
# Use lsps -a to list all available paging space and its status.
#
#mklv -t paging -y mypage vg04 5

#
# To cause the /dev/mypage device to become available, use the swapon
# command.
#
#swapon /dev/mypage

#
# To change page space to be active on the next and subsequent reboots
# us the chps command.
#
#   -a specifies to use page spae at next reboot
#   y specifies to use page space at subsequent reboots
#
#chps -a y mypage

```

Figure 80. /usr/lpp/ssp/samples/script.cust file (1 of 2)

```

#
# Change the paging space - to add 100 logical partitions (400 MB)
# to mypage page space.
#
# NOTE: The default logical partition size is 4 MB.
# NOTE: To use smit to change page space issue: smit chps.
#
# Sample check to see if we are greater than what we are asking for.
#
# NOTE: Read comments for increasing logical volume space below.
#
#num_pps_tox=100      # number of PPs to extend page space by

#
# Determine current size of mypage in PPs
#
#pps='/usr/sbin/lspcs -c mypage | tail +2 | cut -d":" -f4'
#if [[ $pps -lt $num_pps_tox ]]; then
#    chps -s$num_pps_tox mypage
#fi

#
# Create a 40 logical partition (default 4 MB partitions - 160 MB) page
# space on the root volume group.
#
# -a activates page space on subsequent reboots
# -n activates page space immediately
# -s is the size of the page space in partitions
#
#mkpvs -s'40' -n -a rootvg

```

Figure 81. /usr/lpp/ssp/samples/script.cust file (2 of 2)

If you customize this file and place it as /tftpboot/script.cust, PSSP changes paging space according to this information during node installation.

12.2.4 Adjusting paging space size after installation

It may not be suitable for some situations to adjust paging space during the node installation time as described in 12.2.3, “Adjusting paging space size during installation” on page 320. Only one /tftpboot/script.cust file can exist for the CWS or each BIS node. You can place conditional code in the file to perform selected operations based on PSSP level or even based on node_number. However, it might be difficult for you to modify the script.cust file that absorbs all the different configuration between nodes. Another situation is that AIX has been already installed to a node.

On the very first installation period of the node, PSSP itself is the only application running on the node. The PSSP does not require a lot of paging

space. So, the default paging space configuration, in other words, the mksysb image shipped with PSSP, will not cause any paging space problems.

Once installation is succeeded, you can adjust the paging space configuration with the following steps:

Step 1a: Check the memory size (for the MCA nodes)

The `lsdev` command can be used to know the memory size for the MCA nodes:

```
# dsh -w sp4n05 lsdev -Cc memory
sp4n05: mem0 Available 00-0C 128 MB Memory Card
sp4n05: mem1 Available 00-0B 128 MB Memory Card
#
```

You can see that node `sp4n05` has a total memory of 256 MB.

Step 1b: Check the memory size (for the PCI nodes)

The `lsdev` command can be used to know the available memory devices for the PCI nodes:

```
# dsh -w sp4n09 lsdev -Cc memory
sp4n09: mem0 Available 00-00 Memory
sp4n09: L2cache0 Available 00-00 L2 Cache
#
```

Because the node `sp4n09` is a PCI node, the `lsdev` command does not show the size of the memory. You need to issue the `lsattr` command to know this:

```
# dsh -w sp4n09 lsattr -E -l mem0
sp4n09: size 512 Total amount of physical memory in Mbytes False
sp4n09: goodsize 512 Amount of usable physical memory in Mbytes False
#
```

You know node `sp4n09` has 512 MB memory. In the last column, you see `False`. It does not mean this information is untrue. It indicates whether you can change this value by the command. You can change the value only when you add or remove memory physically; so, it is `False`.

Step 2: Check the paging space size

The `lspcs` command can be used to know the current paging space size:

```
# dsh -w sp4n05,sp4n09 lssps -a
sp4n05: Page Space  Physical Volume  Volume Group  Size  %Used  Active  .....
sp4n05: hd6        hdisk0        rootvg        144MB  1      yes     .....
sp4n09: Page Space  Physical Volume  Volume Group  Size  %Used  Active  .....
sp4n09: hd6        hdisk0        rootvg        144MB  1      yes     .....
#
```

You know both nodes have 144 MB paging space. According to the formulas described in 12.2.1, “Sizing paging space” on page 318, you need 272 MB paging space for node sp4n05 and 832 MB paging space for node sp4n09. Either node does not have enough paging space.

Step 3: Check the physical partition size

To adjust paging space, you need to know the physical partition size. It is used as incremental units. The `lsvg` command can be used to know the physical partition size of a volume group:

```
# dsh -w sp4n05,sp4n09 lsvg rootvg | grep "PP SIZE"
sp4n05: VG STATE:      active          PP SIZE:      4 megabyte(s)
sp4n09: VG STATE:      active          PP SIZE:      8 megabyte(s)
#
```

You can see that node sp4n05 uses 4 MB physical partition size (PP SIZE), and node sp4n09 uses 8 MB physical partition size (PP SIZE).

Step 4: Change the size of paging space

In this example, node sp4n05 has 256 MB memory and 144 MB paging space. Node sp5n05 has 512 MB memory and 144 MB paging space.

For the node sp4n05, you need to add 128 MB paging space. In other words, you need to add 32 physical partitions (using 4 MB physical partition size). For node sp4n09, you need to add 688 MB paging space. In other words, you need to add 86 physical partitions (using 8 MB physical partition size).

First, you need to check the available disk space to balance the paging space location. To do this, use a combination of the `lspv` and `lsvg` command. The following is a sample operation for node sp4n05:

```

# dsh -w sp4n05 lspv
sp4n05: hdisk0          000011830004ca74    rootvg
# dsh -w sp4n05 lsvg rootvg
sp4n05: VOLUME GROUP:  rootvg          VG IDENTIFIER: 000073490002c113
sp4n05: VG STATE:      active          PP SIZE:      4 megabyte(s)
sp4n05: VG PERMISSION: read/write     TOTAL PPs:    248 (992 megabytes)
sp4n05: MAX LVs:       256            FREE PPs:    110 (440 megabytes)
sp4n05: LVs:           9              USED PPs:    138 (552 megabytes)
sp4n05: OPEN LVs:     8              QUORUM:       2
sp4n05: TOTAL PVs:    1              VG DESCRIPTORS: 2
sp4n05: STALE PVs:    0              STALE PPs:    0
sp4n05: ACTIVE PVs:   1              AUTO ON:      yes
sp4n05: MAX PPs per PV: 1016         MAX PVs:      32
#

```

The rootvg volume group for node sp4n05 has only one physical volume, hdisk0. Therefore, you do not need to think about paging space location balance. The hdisk0 has 110 free physical partitions (FREE PPs), and it is enough for the 32 physical partitions requirement. Issue the `chps` command to add more paging space:

```

# dsh -w sp4n05 chps -s 32 hd6
#

```

The `-s` flag for the `chps` command specifies the number of logical partitions to add.

Issue the `lspvs` command to check that the paging space changed:

```

# dsh -w sp4n05 lspvs -a
sp4n05: Page Space  Physical Volume  Volume Group  Size  %Used  Active  ....
sp4n05: hd6        hdisk0          rootvg        272MB  1      yes     ....
#

```

The following is a sample operation for node sp4n09. Check the physical volume information first by issuing the `lspv` command:

```

# dsh -w sp4n09 lspv
sp4n09: hdisk0          0004368600024657    rootvg
sp4n09: hdisk1          none                 None
#

```

Node sp4n09 has two physical volumes, hdisk0 and hdisk1. Currently, the rootvg volume group uses only hdisk0, and the hdisk1 is not used. To balance the paging space location, in other words, to create the paging space on both hdisk0 and hdisk1, add hdisk1 to the rootvg. Issue the `extendvg` command to add the hdisk1 to the rootvg, then issue the `lsvg` command to check the free physical partitions:

```
# dsh -w sp4n09 extendvg -f rootvg hdisk1
# dsh -w sp4n09 lsvg rootvg
sp4n09: VOLUME GROUP:   rootvg           VG IDENTIFIER:  00005312393dab94
sp4n09: VG STATE:      active           PP SIZE:       8 megabyte(s)
sp4n09: VG PERMISSION: read/write        TOTAL PPs:    1074 (8592 megabytes)
sp4n09: MAX LVs:      256             FREE PPs:    1002 (8016 megabytes)
sp4n09: LVs:         9              USED PPs:     72 (576 megabytes)
sp4n09: OPEN LVs:    8              QUORUM:       2
sp4n09: TOTAL PVs:   1              VG DESCRIPTORS: 2
sp4n09: STALE PVs:   0              STALE PPs:    0
sp4n09: ACTIVE PVs:  1              AUTO ON:     yes
sp4n09: MAX PPs per PV: 1016        MAX PVs:     32
#
```

The rootvg has 1002 free physical partitions (FREE PPs). It is enough for the 86 physical partitions requirement. To create paging space on both hdisk0 and hdisk1 evenly, add 34 logical partitions to the paging space on hdisk0 and create a paging space with 52 logical partitions on hdisk1. Issue the `chpps` command for hdisk0 and issue the `mkps` command for hdisk1:

```
# dsh -w sp4n09 chpps -s 34 hd6
# dsh -w sp4n09 mkps -s 52 -n -a rootvg hdisk1
paging00
#
```

The `mkps` command uses `paging00` for the logical volume name of paging space.

To check the paging space information, issue the `lspss` command:

```
# dsh -w sp4n09 lspss -a
sp4n09: Page Space  Physical Volume  Volume Group  Size  %Used  Active.....
sp4n09: paging00   hdisk1         rootvg        416MB  1      yes.....
sp4n09: hd6        hdisk0         rootvg        416MB  1      yes.....
#
```

You see see the 416 MB paging space is available in both hdisk0 and hdisk1.

If you create a mksysb image with these paging space settings, you can use the mksysb image for the next installation without adjusting the paging space. However, you need to create and manage a mksysb image for each node if your SP nodes have a different hardware configuration from each other.

Note

If the rootvg consists of multiple disks, do not forget to specify the corresponds numbers of the disks by issuing the `spchvgobj` command. See 2.4.7, “Restoring rootvg on SP node” on page 122.

12.2.5 Monitoring page space

There is one pre-defined event definition available for the Event Perspective:

Monitoring paging space:

Icon:



Name: pageSpaceLow

Description: The paging space utilized on the node exceeds 85 percent.

Resource Variable: IBM.PSSP.aixos.PagSp.%totalused

For more information about this pre-defined event definition, refer to 11.3, “Using the Event Perspective” on page 285.

12.3 Log files

The PSSP uses both the AIX Error Log facility and the BSD syslog facility as well as a number of function-specific log facilities to record error events to files on each node. Some of these log files will grow indefinitely, and this can fill up your file system. Also, it makes looking at the logs tiresome because you have to sift through all the information to find what you are looking for. For these reasons, it is important to monitor the size of these log files. This section explains how to collect the information in these files, how to remove unwanted information, and how to monitor them.

For further information, refer to Chapter 27, “Managing Error Logs” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

12.3.1 Getting authorization

You can collect the necessary logs by issuing the combination of the `dsh` command and standard AIX commands, such as the `errpt`, `cat`, and so on. But for your convenience, several functions, or commands, are provided by PSSP to collect logs from nodes. They are based on `Sysctl`; so, the `rcmd` principal is added to the `/etc/logmgt.acls` file during installation. This file resides on both the CWS and nodes. The following is an example:

```
#acl#
# This sample acl file for log management commands contains a commented line for
# a principal
#_PRINCIPAL root.admin@HPSSL.KGN.IBM.COM
# Principal for trimming SPdaemon.log by cleanup.logs.ws
#_PRINCIPAL rcmd.sp3en0
```

The `root.admin` principal should be added to the file:

```
# echo _PRINCIPAL root.admin@SP4EN0 >> /etc/logmgt.acls
# dsh -a "echo _PRINCIPAL root.admin@SP4EN0 >> /etc/logmgt.acls"
#
```

12.3.2 Collecting AIX error logs

AIX Error Log keeps error logs in the `/var/adm/ras/errlog` file in binary format. The `errdemon` daemon manages this file. To convert the binary format to a text format, you need to issue the `errpt` command. To clear the contents, you need to issue the `errclear` command.

PSSP provides the `smitty perrpt` fast path to generate error report for the specified nodes:

Generate an Error Report

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]	[Entry Fields]	
Generate Report on all Nodes in Partition	no	+
Generate Report on Hosts	[sp4n01,sp4n05]	
SUMMARY or DETAILED error report	summary	+
Error CLASSES (default is all)	[]	+
Error TYPES (default is all)	[]	+
Error LABELS (default is all)	[]	+
Error ID's (default is all)	[]	+X
Resource CLASSES (default is all)	[]	
Resource TYPES (default is all)	[]	
Resource NAMES (default is all)	[]	
SEQUENCE numbers (default is all)	[]	
STARTING time interval	[]	
ENDING time interval	[]	
LOGFILE	[/var/adm/ras/errlog]	
TEMPLATE file	[/var/adm/ras/errtmpl]	
MESSAGE file	[/var/adm/ras/codepoint>	
FILENAME to send report to (default is stdout)	[/tmp/errpt.1-5]	

[BOTTOM]

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

If you specify the file name in the FILENAME to send report to (default is stdout) field, you can save a error report as a file instead of stdout.

You can also perform the trimming of error logs or creation of error notification objects, and so on. To perform these operations, issue the `smitty sperrlog` fast path:

```
AIX Error Log

Move cursor to desired item and press Enter.

Generate an Error Report
Show Characteristics of the Error Log
Change Characteristics of the Error Log
Clean the Error Log
Add a Notification Object
Remove a Notification Object
Show a Notification Object
Add an Error Template
Remove an Error Template
Show an Error Template

F1=Help      F2=Refresh   F3=Cancel    F8=Image
F9=Shell     F10=Exit    Enter=Do
```

12.3.3 Collecting BSD syslog logs

BSD syslog facility is maintained by the syslogd daemon. It writes log entries to the file that is specified in /etc/syslog.conf file. PSSP specifies the /var/adm/SPlogs/SPdaemon.log file for the destination in this file on each node.

BSD syslog log is a text; therefore, AIX does not provide any special command to manipulate the file. However, PSSP provides the `psyslrpt` command to generate reports of BSD syslog log files for the specified nodes:


```

# /usr/lpp/ssp/bin/psyslrpt -w sp4n01,sp4n05 \
> -s 11010000 -e 11011159
>> sp4n01.msc.itso.ibm.com
psyslrpt: /var/adm/SPlogs/SPdaemon.log=====
Nov 01 11:04:25 sp4n01 Worm[17142]: LPP=PSSP,Fn=TBSrecovery.c,SID=1.42,L#=1257,
Nov 01 11:05:05 sp4n01 Worm[17142]: LPP=PSSP,Fn=fsd_fsm.c,SID=1.56.5.2,L#=4172,
Nov 01 11:05:05 sp4n01 Worm[17142]: LPP=PSSP,Fn=fsd_fsm.c,SID=1.56.5.2,L#=4221,
Nov 01 11:05:05 sp4n01 css [17142]: LPP=PSSP,Fn=set_node_info.c,SID=1.9,L#=469,
Nov 01 11:06:28 sp4n01 css [17142]: LPP=PSSP,Fn=tbs_errlogger.c,SID=1.3,L#=101,
Nov 01 11:08:28 sp4n01 css [17142]: LPP=PSSP,Fn=tbs_errlogger.c,SID=1.3,L#=101,
Nov 01 11:10:27 sp4n01 css [17142]: LPP=PSSP,Fn=tbs_errlogger.c,SID=1.3,L#=101,
Nov 01 13:05:45 sp4n01 Worm[17142]: LPP=PSSP,Fn=TBSrecovery.c,SID=1.42,L#=1257,
Nov 01 17:27:55 sp4n01 Worm[17142]: LPP=PSSP,Fn=TBSrecovery.c,SID=1.42,L#=1257,
Nov 01 17:49:45 sp4n01 Worm[17142]: LPP=PSSP,Fn=TBSrecovery.c,SID=1.42,L#=1257,
<<
>> sp4n05.msc.itso.ibm.com
psyslrpt: /var/adm/SPlogs/SPdaemon.log=====
Nov 01 11:32:33 sp4n05 xntpd[9032]: ** adjust: STEP 192.168.4.130 offset ....
<<
#

```

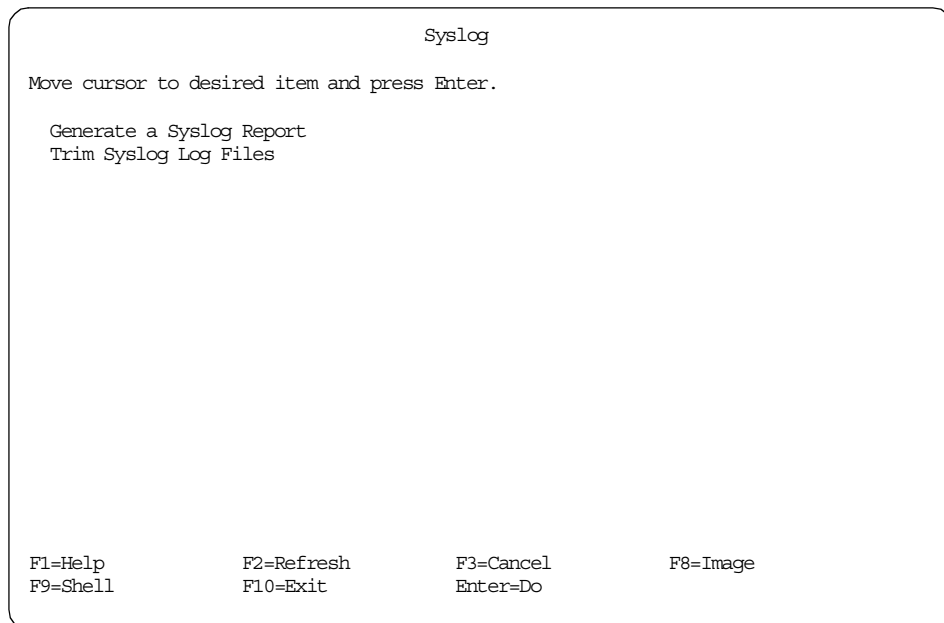
The `-s` and `-e` flag indicate start time and end time (MMDDhhmm format), respectively. In this example, it reports BSD syslog log entries from Nov. 1st, 00:00 to Nov. 1st, 11:59.

You can also trim the file by issuing the `psyslclr` command:

```
# psyslclr -a -y 7
```

In this example, it trims records more than seven days old from all nodes.

You can execute both commands from the `smitty spsyslog` fast path:



12.3.4 Collecting other logs

PSSP provides the `sp1m` command to collect specified log files from nodes to CWS. The command requires two stages, creating archives on each node, then gathering them to the CWS. To perform this task, perform the following steps:

Step 1: Edit the log table

There are sample log table files in the `/spdata/sys1/logtables` directory. Using a log table file, you can specify target nodes, directory to be archived, and directory to place archives. Figure 82 on page 333 shows the `sysman.tab` file.

```

# cd /spdata/sys1/logtables
# ls
amd.tab      filec.tab   sdr.tab     sysman.tab
css.tab      jm.tab      ssp.tab     sysmon.tab
# cat sysman.tab
#
# This is a sample service collection table for sysman problems.
# To use:
# 1. Uncomment lines to be collected.
# 2. Replace the target node list for each line with node names,
#    a file containing node names, or a ! to designate the
#    local node. (See the splm reference page)
# 3. Add to the node list or add additional collection commands
#    if needed.
#
# In this table the following target nodes need to edited:
# allnodes      - All nodes in the SP system

#allnodes: /var/adm/SPlogs/sysman/* /sysman/
sp3n01: /var/adm/SPlogs/sysman/* /sysman/
#

```

Figure 82. The sysman.tab log table file

In this example, the sysman.tab file specifies that all files in the /var/adm/SPlogs/sysman directory on node sp3n01 are archived to the /sysman/ directory on the node. You can specify multiple nodes.

Step 2: Create archives on the node

By issuing the smitty spcreate_archive fast path, the specified log files are archived on the node.

```

                                Create Archives

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* Archive Table                  [/spdata/sys1/logtables> +
Archive directory                [/tmp]
Append timestamp to directory    yes +
Maximum Fanout                  [] +#
Create compressed tar file      no +
Run on local node only          no +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

You can select one of the log table files in the Archive Table field. Use the **F4** key to list all the available log table files. The `sysman.tab` file, shown in Figure 82 on page 333, is selected in this example.

All files in the `/var/adm/SPlogs/sysman` directory are archived to the `/tmp/YYYYMMDD/arch_sysman.tab/sysman` directory on the node, `sp3n01`.

Or you can issue the `splm` command to perform the same operation.

```

# /usr/lpp/ssp/bin/splm -a archive \
> -t /spdata/sys1/logtables/sysman.tab \
> -d /tmp -y
sp3n01: splm: Found authorization file splm.allow.
sp3n01: splm: Created directory: /tmp/990529/arch_sysman.tab.
sp3n01: splm: Calculating space available in /tmp/990529/arch_sysman.tab.
sp3n01: splm: Space available is 31204 kb.
sp3n01: splm: Calculating space requirements.
sp3n01: splm: Space required is 23 kb.
sp3n01: splm: Created directory: /tmp/990529/arch_sysman.tab//sysman/.
sp3n01: splm: Processing archive actions...
sp3n01: splm: File /var/adm/SPlogs/sysman/spfbcheck.log copied to /tmp/990529/arch_sysman.tab/sysman/.
sp3n01: splm: File /var/adm/SPlogs/sysman/sp3n01.console.log.1 copied to /tmp/990529/arch_sysman.tab/sysman/.
sp3n01: splm: File /var/adm/SPlogs/sysman/sp3n01.console.log copied to /tmp/990529/arch_sysman.tab/sysman/.
sp3n01: splm: File /var/adm/SPlogs/sysman/sp3n01.console.log.2 copied to /tmp/990529/arch_sysman.tab/sysman/.
#

```

Check if the log files are archived correctly:

```

# dsh -w sp3n01 ls -al /var/adm/SPlogs/sysman
sp3n01: total 49
sp3n01: drwxr-xr-x  2 root    system    512 May 16 00:00 .
sp3n01: drwxr-xr-x 12 bin      bin      512 May 17 14:19 ..
sp3n01: -rw-r--r--  1 root    system   5301 May 24 09:27 sp3n01.console.log
sp3n01: -rw-r--r--  1 root    system    0 May 23 00:00 sp3n01.console.log.1
sp3n01: -rw-r--r--  1 root    system  17099 May 23 00:00 sp3n01.console.log.2
sp3n01: -rw-r--r--  1 root    system    556 Mar 16 14:00 spfbcheck.log
# dsh -w sp3n01 ls -alR /tmp/990529/*
sp3n01: total 24
sp3n01: drwxr-xr-x  3 root    system    512 May 29 11:47 .
sp3n01: drwxr-xr-x  3 root    system    512 May 29 11:47 ..
sp3n01: drwxr-xr-x  2 root    system    512 May 29 11:47 sysman
sp3n01: /tmp/990529/arch_sysman.tab/sysman:
sp3n01: total 80
sp3n01: drwxr-xr-x  2 root    system    512 May 29 11:47 .
sp3n01: drwxr-xr-x  3 root    system    512 May 29 11:47 ..
sp3n01: -rw-r--r--  1 root    system   5301 May 24 09:27 sp3n01.console.log
sp3n01: -rw-r--r--  1 root    system    0 May 23 00:00 sp3n01.console.log.1
sp3n01: -rw-r--r--  1 root    system  17099 May 23 00:00 sp3n01.console.log.2
sp3n01: -rw-r--r--  1 root    system    556 Mar 16 14:00 spfbcheck.log
#

```

You can see all the log files in the /var/adm/SPlogs/sysman directory are archived in the /tmp/990529/arch_sysman.tab/sysman directory.

Step 3: Gather archives from nodes to CWS

All the archives are available on the node. You can gather these archives from the node to the CWS. To do this, issue the `smitty spgather_archive` fast path:

```
Gather Archives

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Archive Table                    [/spdata/sys1/logtables> +
Node Archive Directory           [/tmp]
Append timestamp to directory    yes +
Maximum Fanout                   [] +#
Local Gather Directory           [/spdata/sys1/logholder]
Gather Output Destination        []
Remove Node Archives after Gather yes +
Remove Gathered Archives after Output no +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Based on the log table file indicated in the Archive Table field, the archives are gathered to the directory indicated in the Local Gather Directory field. If the directory does not exist, it is automatically created.

In this example, the archives in the `/tmp/YYYYMMDD/arch_sysman.tab/sysman` directory on node `sp3n01` are gathered to the `/spdata/sys1/logholder` directory on the CWS.

When archives are gathered, archives are packed to one file by the `tar` and `compress` command. Archives on the node are removed after gathering.

You can issue the `splm` command to perform the same operation.

```

# /usr/lpp/ssp/bin/splm -a gather -k archive \
> -t /spdata/sys1/logtables/sysman.tab \
> -d /tmp -y -l /spdata/sys1/logholder -r
splm: Space available in /spdata/sys1/logholder is 1133072 kb.
sp3n01: splm: Created file: /tmp/990529/arch_sysman.tab/sp3n01.tar.Z.
sp3n01: splm: Copied tar /tmp/990529/arch_sysman.tab/sp3n01.tar.Z to sp3en0:/spdata/sys1/logholder/sp3n01.tar.Z, tar size is 9588.
sp3n01: splm: Removing the following directories and files...
sp3n01: /tmp/990529/arch_sysman.tab
sp3n01: /tmp/990529/arch_sysman.tab/sysman
sp3n01: /tmp/990529/arch_sysman.tab/sysman/spfbcheck.log
sp3n01: /tmp/990529/arch_sysman.tab/sysman/sp3n01.console.log.1
sp3n01: /tmp/990529/arch_sysman.tab/sysman/sp3n01.console.log
sp3n01: /tmp/990529/arch_sysman.tab/sysman/sp3n01.console.log.2
sp3n01: /tmp/990529/arch_sysman.tab/sp3n01.tar.Z
#

```

Check if the archives on the node are gathered correctly:

```

# ls -al /spdata/sys1/logholder
total 40
drwxr-xr-x  2 root    system    512 May 29 11:31 .
drwxr-xr-x 15 bin     bin      512 May 29 11:31 ..
-rw-r--r--  1 root    system    9588 May 29 11:31 sp3n01.tar.Z
# dsh -w sp3n01 ls -al /tmp/990529
sp3n01: total 16
sp3n01: drwxr-xr-x  2 root    system    512 May 29 11:31 .
sp3n01: drwxrwxrwt  6 bin     bin      1024 May 29 11:31 ..
#

```

You have the archives in the /spdata/sys1/logholder directory on the CWS, and archives on the node are deleted.

12.3.5 Eliminating increased log files

PSSP provides two scripts to clean up logs:

- The `cleanup.logs.ws` script for CWS
- The `cleanup.logs.nodes` script for nodes

They are issued from cron at 00:00 every day.

Most of the PSSP daemons use the /var file system for logging and have some log elimination mechanism. But, /var file system is shared by many applications besides PSSP, and some elimination mechanism are only

effective when the daemon is running. To get more space on the /var file system, you might have to eliminate some log files from the /var file system.

Files in the /var/adm/SPlogs and /var/ha/log directory are all log files. Therefore, it would not cause any problems when you delete some or all of them. But log files provide important information in determining a problem when it occurs. You should keep the recently created or updated log files.

The `ls` command with `-lt` flag sorts the file list by the last modification time. You can use this command to find old log files.

The `find` command is also useful to find old files. The following example searches files that have not been accessed or modified for three days in the /var/adm/SPlogs directory, and deletes them by issuing the `rm` command:

```
# find /var/adm/SPlogs -type f -atime +3 -mtime +3 \  
> -exec \rm {} \; -print
```

If you need to erase the latest log files, you should issue the `cp` command as follows:

```
# cp /dev/null target_log_files
```

It clears the log file but does not delete the file itself.

12.3.6 Monitoring log files

There are two methods to monitor log files:

- Using the Event Perspective
- Using the `pmandef` command

12.3.6.1 Using the Event Perspective

There are two pre-defined event definitions available:

1. Monitoring permanent error

Icon:



errLog

Name: errLog

Description: A permanent error entry is added to the error log.

Resource Variable: IBM.PSSP.pm.Errlog

2. Monitoring off-line processor

Icon:

Name:	processorsOffline
Description:	One or more processors may have been taken off-line.
Resource Variable:	IBM.PSSP.pm.Errlog

For more information about these event definitions, refer to 11.3, “Using the Event Perspective” on page 285.

12.3.6.2 Using the pmandef command

There are two examples available in the `pmandefaults` script file for the `pmandef` command:

1. Monitoring permanent error

The following `pmandef` command example is an excerpt from the `pmandefaults` script file:

```
#
# Monitor the error log on each node for PERM events
#
pmandef -s errLog \
  -e 'IBM.PSSP.pm.Errlog:NodeNum=*:X@0!=X@P0 && X@3=="PERM"' \
  -c /usr/lpp/ssp/bin/notify_event \
  -n 0 -U root -m errLog
[[ $? -ne 0 ]] && print -u2 "Problem with errLog" && exit 1
```

The `pmandef` command subscribes the event named `errLog`. If the error log entry with `PERM` error type is added on any node, the `notify_event` script is executed on CWS by root user.

2. Monitor non-critical power loss or fan failure

The following `pmandef` command example is an excerpt from the `pmandefaults` script file:

```

#
# Monitor the error log on each node for non-critical power loss or
# fan failure. (These errors should only occur on SMP nodes.)
#
pmandef -s Non_critical_power_loss_or_fan_failure \
-e 'IBM.PSSP.pm.Errlog:NodeNum=*:X@5=="SYSIOS" && X@8=="EPOW_SUS"' \
-c '/bin/ksh -c "print \"Potential non-critical power loss or fan failure on
node ${PMAN_LOCATION}.\nCheck the AIX error log on node ${PMAN_LOCATION}.\n\"
| mail -s Non_critical_power_loss_or_fan_failure $LOGNAME" ' \
-n 0 -U root
[[ $? -ne 0 ]] && print -u2 "Problem with Non_critical_power_loss_or_fan_failure
" && exit 1

```

The `pmandef` command subscribes the event named `Non_critical_power_loss_or_fan_failue`. If the error log entry with `SYSIOS` resource name and `EPOW_SUS` error label is added, mail is sent to the root user on the CWS.

For more information about these examples, refer to 11.4, “Using the `pmandef` command” on page 296.

To understand these two examples, you need to know Structured Byte String (SBS). Table 11 shows the SBS definitions for the `IBM.PSSP.pm.Errlog` resource variable.

Table 11. Structured byte string definitions for `IBM.PSSP.pm.Errlog`

Field Name	Field Length	Field Type	Field Serial Number
sequenceNumber	variable	cstring	0
errorID	variable	cstring	1
errorClass	variable	cstring	2
errorType	variable	cstring	3
alertFlagsValue	variable	cstring	4
resourceName	variable	cstring	5
resourceType	variable	cstring	6
resourceClass	variable	cstring	7
errorLabel	variable	cstring	8

For more information about RS/6000 Cluster Technology, refer to the following IBM publications:

- *RS/6000 Cluster Technology: Group Service Programming Guide and Reference, SA22-7355*
- *RS/6000 Cluster Technology: Event Management Programming Guide and Reference, SA22-7354*

Chapter 13. Managing sets of nodes

The SP switch can be very complex and has its own group of commands. Using examples, this section explains how to manage your SP switch. For new system administrators, who are not familiar with the maintenance required for the SP switch, this chapter is very important.

System management is required for all the nodes, and IBM Parallel System Support Programs for AIX (PSSP) has some functions that make this easier. Nodes can be configured to belong to separate partitions so that it is easier for the administrator to handle them. You can assign one partition to all test nodes, another partition can contain database nodes, and another can be production nodes. These partitions can be customized to suit your needs on your SP system. System partitions are discussed in this chapter with tips on how to apply and delete them.

In the process of managing your SP system, you may be required to issue the same command over and over to each separate node. This can be tiresome if you have several frames full of nodes. To fix this, PSSP allows you to use node groups when you run some commands. This chapter explains what node groups are and how they can be used.

13.1 Managing SP switch

The SP switch is a way for the nodes to have fast and efficient communication between each other. In some ways, it can be considered to behave like a regular LAN; however, there are many characteristics and commands that are unique to the SP switch. This section explains how the SP switch works and how to set up a primary node and a primary backup node. It explains how to start the SP switch and what to do if you can not get the SP switch up on a specific node. Managing the SP switch requires quite a bit of knowledge, and this section attempts to deliver it and provides you with information on the overall maintenance of the SP switch and where to look when things are not working.

13.1.1 What are primary and primary backup nodes?

A primary backup node passively listens for activity from a primary node. When the primary backup node detects that it has not been contacted by the primary node for approximately seven minutes, it assumes the role of the primary node. This takeover involves nondisruptively reinitializing the switch fabric, selecting another primary backup, and updating the System Data Repository (SDR).

A primary node also watches over a primary backup node. If the primary node detects that the primary backup node can no longer be contacted on the switch fabric, it selects a new primary backup node. The criteria to select a new primary backup node is as follows:

1. First, select a node on a switch board other than the switch board to which the primary node is attached.
2. Second, if no other switch board exists, select a node attached to a switch chip other than the one to which the primary node is attached.
3. If no other switch chip exists, select any available node on the switch chip to which the primary node is attached.

For sample operation to assign primary node or primary backup node, refer to 13.1.3, “Changing primary and primary backup nodes” on page 344.

13.1.2 Showing primary and primary backup nodes

You can see the current setting for a primary and primary backup node by issuing the `Eprimary` command:

```
# Eprimary
none - primary
1 - oncoming primary
none - primary backup
15 - oncoming primary backup
#
```

The current primary and primary backup node are shown as primary and primary backup, respectively. In this example, SP switch has not started yet; so, they are shown as `none`. When you start the SP switch, the nodes assigned to the oncoming primary and oncoming primary backup will replace the primary and primary backup nodes, respectively.

13.1.3 Changing primary and primary backup nodes

The first thing you have to know is that you can not change a current primary node and a primary backup node directly. You first need to assign an oncoming primary node and an oncoming primary backup node. To assign a primary node and a primary backup node, issue the `smitty primary_node_dialog fast` path:

```

Set Primary/Primary Backup Node

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
Primary Node Identifier                [5]                +
Primary Backup Node Identifier         [7]                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

This example assigns node 5 to an oncoming primary node and node 7 to an oncoming primary backup node.

You can issue the `Eprimary` command instead:

```
# Eprimary 5 -backup 7
```

After assigning an oncoming primary node and an oncoming primary backup node, you can initialize the SP switch by issuing the `Estart` command. This operation will replace a primary node and a primary backup node to an oncoming primary node and an oncoming primary backup node, respectively.

13.1.4 Starting the SP switch

The SP switch can be initialized by issuing the `Estart` command:

```

# Estart
Estart: 0028-061 Estart is being issued to the primary node: sp4n01
Switch initialization started on sp4n01.
Initialized 10 node(s).
Switch initialization completed.
#

```

You can issue the `Estart` command on the control workstation (CWS) or any nodes.

13.1.5 Stopping the SP switch

There is no specific command that stops the SP switch. However, you can add/delete an SP node to/from the current SP switch network by using the `Efence/Euncence` command. In other words, you can shut down a node without special care of the SP switch.

If you shut down nodes while the primary node is up and running, the primary node fences the nodes including the primary backup node. If you shut down the primary node first and the primary backup node is still up and running, the primary backup node takes over the primary node. This recovery action is useful when the SP system is operating. But, when you shut down all nodes, this action may not be proceeded correctly. The primary node or the primary backup node can be shutting down during the take over process. To avoid unnecessary recovery action, you can issue the `Equiesce` command:

```
# Equiesce
```

You should issue this command before you shut down all nodes.

13.1.6 Joining switch fabric automatically

Prior to PSSP 3.1, the `autojoin` attribute in the `switch_respond` SDR class is normally turned off when a node is shut down. To join a node to switch fabric automatically when it is rebooted, you need to fence the node by issuing the `Efence` command with `-autojoin` flag before you shut down the node.

From PSSP 3.1, the `autojoin` attribute is turned on by default. Therefore, when a node is rebooted with the primary node active, it joins switch fabric automatically. To prevent it, issue the `Efence` command without flags when you fence the node.

13.1.7 Getting switch buffer pool information

SP switch uses two memory pools to transact data. One is for outgoing (send) data, and the other is for incoming (receive) data. When an IP datagram is passed to the switch interface, and if the size of the datagram is large enough, a buffer is allocated from the pool.

The current send pool size and receive pool size are shown as the device attribute of `css0`. To get the information about pool size, issue the `lsattr` command:


```

# dsh -w sp4n01 lsattr -E -l css0
sp4n01: bus_mem_addr 0x04000000 Bus memory address      False
sp4n01: int_level   0xb      Bus interrupt level  False
sp4n01: int_priority 3       Interrupt priority   False
sp4n01: dma_lvl     9       DMA arbitration level False
sp4n01: spoolsize  524288  Size of IP send buffer True
sp4n01: rpoolsize  524288  Size of IP receive buffer True
sp4n01: adapter_status css_ready Configuration status  False
#

```

The `spoolsize` field means send pool size, and `rpoolsize` means receive pool size. Each pool size is 524,288 bytes (512 KB) by default settings. This value is also the minimum size of the buffer pool.

13.1.8 Changing switch buffer pool size

Currently, the upper limit for the send pool and receive pool is 16 MB for each. But of course, it is limited by the free space of physical memory on node. To change the size of buffer pool, issue the `chgcass` command located in the `/usr/lpp/ssp/css` directory. For example:

```

# dsh -w sp4n01 /usr/lpp/ssp/css/chgcass -l css0 \
> -a spoolsize=2097152 -a rpoolsize=2097152

```

This example changes both pool sizes to 2 MB. This operation only changes the data in Object Database Manager (ODM). You need to reboot the node to configure the switch with new settings.

While the `lsattr` command shows the data in ODM, the `vdid13` command located in the `/usr/lpp/ssp/css` directory shows the values currently used:

```

# dsh -w sp4n01 /usr/lpp/ssp/css/vdidl3 -i
sp4n01: get ifbp info...
sp4n01:
sp4n01: send pool: size=524288 anchor@=0x50002e00 start@=0x50e50000 tags@=0x50001000
sp4n01: bkt   allocd   free  success   fail   split   comb   freed
sp4n01: 12      0         0      0         0      4       0      0
sp4n01: 13      0         0      0         0      0       0      0
sp4n01: 14      0         0      0         0      0       0      0
sp4n01: 15      0         0      0         0      0       0      0
sp4n01: 16      0         8      0         0      0       0      0
sp4n01:
sp4n01: rsvd pool: size=262144 anchor@=0x50d1f200 start@=0x50ed0000 tags@=0x50c98800
sp4n01: bkt   allocd   free  success   fail   split   comb   freed
sp4n01: 12      0         0      0         0      0       0      0
sp4n01: 13      0         0      0         0      0       0      0
sp4n01: 14      0         0      0         0      0       0      0
sp4n01: 15      0         0      0         0      0       0      0
sp4n01: 16      0         4      0         0      0       0      0
sp4n01:
sp4n01: recv pool: size=524288 anchor@=0x50002000 start@=0x50f10000 tags@=0x50d1e100
sp4n01: bkt   allocd   free  success   fail   split   comb   freed
sp4n01: 12      0         0      0         0      0       0      0
sp4n01: 13      0         0      0         0      0       0      0
sp4n01: 14      0         0      0         0      0       0      0
sp4n01: 15      0         0      0         0      0       0      0
sp4n01: 16      0         0      0         0      0       0      0
sp4n01:
#

```

The size field in the send pool line and recv pool line indicates the currently used pool size, respectively.

For more details about switch buffer pool, refer to Chapter 11, “SP Switch-Specific Application and Server Tuning” in *Understanding and Using the SP Switch*, SG24-5161.

13.1.9 Using the cssadm daemon

In PSSP 3.1, the cssadm daemon has been introduced to monitor the nodes and the switch adapters. It runs only on the CWS. If it is required, the daemon will issue the `Estart` command automatically. Previously, this was done by the system administrator.

Here is an example of how the cssadm daemon behaves. Assume that you have the following settings for primary and primary backup nodes:

```
# Eprimary
6      - primary
5      - oncoming primary
13     - primary backup
15     - oncoming primary backup
# date
Wed Mar 31 12:04:20 EST 1999
#
```

The `cssadm` daemon uses the following logic when it issues the `Estart` command:

1. If the primary node goes down, the daemon checks if the primary backup node has `switch_responds`. If it does, the daemon will not do anything. The primary backup node will take over the primary node.
2. If the primary backup node does not have `switch_responds`, the daemon will check if the oncoming primary node has `host_responds`. If it does not, the daemon will not do anything. You need to assign the primary node manually.
3. If the oncoming primary node has `host_responds`, the daemon issues the `Estart` command to assign primary node to the oncoming primary node.

To see what will happen, let the `switch_responds` down for the node 6 and 13, in other words, isolate the primary node and the primary backup node as follows:

```
# SDRGetObjects switch_responds
node_number  switch_responds autojoin    isolated    adapter_config_status
      1           1           1           0 css_ready
      5           1           1           0 css_ready
      6           0           1           1 css_ready
     11           1           1           0 css_ready
     13           0           1           1 css_ready
     15           1           1           0 css_ready
#
```

In this case, the `cssadm` daemon should issue the `Estart` command to place node 5, an oncoming primary node, to a primary node.

Check the `cssadm.debug` log file for the `cssadm` daemon located in the `/var/adm/SPlogs/css` directory. The following is an excerpt from this file:

```

-----
Processing Event:
-----
event_type      = node down
node number     = 6
time = Wed Mar 31 12:05:03 1999

complete       = 2
(i) cssadm: Primary node is down on switch responds in partition sp4en0.
   Checking primary backup.
(i) cssadm: Primary backup is not up on switch responds.  Checking
   if oncoming primary is up on host responds.
(i) cssadm: Oncoming primary up on host responds.  Going to Estart.

```

It gives a detailed explanation of what the daemon saw and did.

Once the `Estart` command is issued, the oncoming primary node becomes the primary node. In this case, the `Estart` command did not have to be issued manually; it was automatically issued by the `cssadm` daemon.

The `cssadm` daemon uses three files in `/var/adm/SPlogs/css` directory:

- cssadm.debug** This file contains entries for each event that occurs and the corresponding outcome.
- cssadm.stdout** This file contains output from the commands run.
- cssadm.stderr** This file contains the errors from the commands run. This information is useful if you need understand what the daemon is doing.

For more information on the `cssadm` daemon, refer to Chapter 14, “Using a Switch” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

Attention

The `/usr/lpp/ssp/css/cssadm` daemon is registered in System Resource Controller (SRC) as `swtadmd` subsystem and is started from `inittab`. This registration is done by the `install_cw` command that is executed at the very first phase of installation. But, if `ssp.css` was not installed on CWS when the script was executed, it was not registered. If you want to use this function afterward, install `ssp.css` on CWS and execute the `install_swtadmd` command:

```
# /usr/lpp/ssp/css/install_swtadmd
```

13.1.10 Reading the switch topology file

The switch topology file is the base information for the SP switch daemon, `fault_service_Worm_RTG_SP`, to create a routing table. If the physical connectivity of switch to switch or switch to node is different from the information in the switch topology file, the switch does not work properly. If there is a possibility of mis-connection, you can interpret the file and verify the physical connection.

To know which switch topology file your SP system uses currently, issue the `SDRGetObjects` command:

```
# SDRGetObjects Switch_partition topology_filename
topology_filename
expected.top.annotated.3
#
```

In this example, your SP system uses the `expected.top.annotated.3` for the switch topology file.

For more information on the switch topology file, refer to Chapter 6, “Installation of the SP Switch” in *Understanding and Using the SP Switch*, SG24-5161.

13.1.11 Reinitializing clock source

Once clock synchronization is lost, you can not start a switch without reinitialize the clock. To synchronize the clock, issue the `Eclock` command:

```
# Eclock -d
```

With the `-d` flag, the command detects the switch configuration, automatically selects the clock topology file, and initializes the switch clock inputs for all switches in the system. If you use `-r` flag, the command extracts the clock topology file information from the SDR and initializes the switch clock inputs for all switches in the system.

By this operation, the `swtch_responds` of all nodes become 0, and the `fault_service_Worm_RTG_SP` daemon on each node is refreshed. Therefore, when you failed to issue the `Estart` command with error, such as:

```
"0028-071 Switch initialization failed. The fault_service_Worm_RTG_SP
daemon exited due to errors. See/var/adm/SPlogs/css/flt file for more
information."
```

the `Eclock` command might solve the problem.

The clock topology file resides in the /etc/SP directory:

```
# ls /etc/SP | grep Eclock
Eclock.top.1nsb.0isb.0
Eclock.top.1nsb_8.0isb.0
Eclock.top.2nsb.0isb.0
Eclock.top.3nsb.0isb.0
Eclock.top.4nsb.0isb.0
Eclock.top.4nsb.2isb.0
Eclock.top.5nsb.0isb.0
Eclock.top.5nsb.4isb.0
Eclock.top.6nsb.4isb.0
Eclock.top.7nsb.4isb.0
Eclock.top.8nsb.4isb.0
#
```

13.1.12 Checking a switch log

The error information about a switch is logged in AIX Error log on each node. PSSP 3.1 provides the log file named `summlog` in the `/spdata/sys1/ha/css` directory on the CWS. The file is linked to `/var/adm/SPlogs/css/summlog`. The file contains summary records of switch-related entries in AIX error logs on all the nodes. It provides a single point to monitor system-wide switch activities.

When the `summlog` file size becomes greater than 3 MB, it is renamed to `summlog.old`, and a new `summlog` file is created.

The following is a sample entry in this file:

```
110115421998 sp4n01 N sp4en0 777 TB3_TRANSIENT_RE
```

Each column indicates the following from left to right:

- | | |
|-------------------|--|
| Time stamp | This column uses the MMDDhhmmYYYY format. |
| Node name | This column indicates the short reliable host name. It shows where the error occurred. |
| Snap | When the switch support code (for example, device driver, worm, fault-service, or diags) detects a serious error, the <code>css.snap</code> script is called to collect log and trace information into a compressed tar file, <code>hostname.YYMMDDhhmmss.css.snap.tar.Z</code> in the <code>/var/adm/SPlogs/css</code> directory. If this snap shot was taken, this column indicates Y. |

Note

The `css.snap` avoids filling up `/var` file system by following these rules:

- If less than 10 percent of `/var` file system is free, `css.snap` exits.
- If the `css` portion is more than 30 percent of the total space in `/var` file system, `css.snap` erases old snap files until the `css` portion becomes less than 30 percent. If it is successful, the snap proceeds. If not, it exits.

Partition	This column indicates system partition name or global.
Index	This column indicates the sequence number field in the AIX Error Log as shown in Figure 83 on page 354. You can see the corresponding error log entry by using this number with the <code>errpt</code> command:
Label	The label field in the AIX Error log. It corresponds to the LABEL field shown in Figure 83 on page 354.

```

# dsh -w sp4n01 errpt -a -l 777
sp4n01: -----
sp4n01: LABEL:          TB3 TRANSIENT_RE
sp4n01: IDENTIFIER:     06C2F1C9
sp4n01:
sp4n01: Date/Time:        Tue Nov 1 15:42:32
sp4n01: Sequence Number:  777
sp4n01: Machine Id:      00091141A400
sp4n01: Node Id:        sp4n01
sp4n01: Class:         H
sp4n01: Type:         TEMP
sp4n01: Resource Name:  css
sp4n01: Resource Class: NONE
sp4n01: Resource Type:  NONE
sp4n01: Location:     NONE
sp4n01:
sp4n01: Description
sp4n01: Switch adapter transient error
sp4n01:
sp4n01: Probable Causes
sp4n01: Loose, disconnected or bad switch cable
sp4n01:
sp4n01: User Causes
sp4n01: Switch cable loose or disconnected
sp4n01:
sp4n01:          Recommended Actions
sp4n01:          Check / reconnect / replace cable if problem persists
sp4n01:
sp4n01: Failure Causes
sp4n01: Switch cable faulty
sp4n01:
sp4n01:          Recommended Actions
sp4n01:          Check / reconnect / replace cable if problem persists
sp4n01:
sp4n01: Detail Data
sp4n01: Software ID String
sp4n01: LPP=PSSP,Fn=TB3recovery.c,SID=1.32,L#=576,
sp4n01: Interrupt source (ISR or MX CFG3)
sp4n01: TBIC intr
sp4n01: Bus Err (DMA CSR, MX MBA_ER, PCI C/S)
sp4n01: not applicable
sp4n01: Error Status Regs (INT_ERR and INT_ERR2)
sp4n01: 00051000 00000000
#

```

Figure 83. The errpt command output

For more information on diagnosing method of switch, refer to Chapter 11, “Diagnosing Switch Problems” in *IBM Parallel System Support Programs for AIX: Diagnosis Guide, GA22-7350*.

Attention

The summlog file is provided by the `/usr/lpp/ssp/css/css.summlog` daemon that is running on CWS. This daemon is registered in System Resource Controller (SRC) as a swtlog subsystem and is started from inittab. This registration is done by the `install_cw` command that is executed at the very first phase of installation. But, if `ssp.css` was not installed on CWS when the script was executed, it was not registered. If you want to use this function afterward, install `ssp.css` on CWS and execute the `install_swtlog` command:

```
# /usr/lpp/ssp/css/install_swtlog
```

13.2 System partition

System partition is a method for organizing the SP system into non-overlapping groups of nodes for various purposes, such as testing new software and creating multiple production environments. An SP system is divided into one or more logical system by this function. In most cases, an SP system is configured as one partition system. When multiple partitions are configured, a node in a partition can not communicate with a node in another partition through the SP switch. On the CWS, partition sensitive daemons are spawned partition by partition to manage each of them separately.

13.2.1 Applying a system partition

The following is a sample procedure to divide one partition system into two partitions:

Step 1: Archive the SDR

Before applying a system partition, you should back up the SDR by issuing the `SDRArchive` command.

```
# SDRArchive mybackup
SDRArchive: SDR archive file name is /spdata/sys1/sdr/archives/backup.99064.1459
.mybackup
#
```

Note, the location and the name of the file created after you issue this command. In this example, the `backup.99064.1459.mybackup` file is created in the `/spdata/sys1/sdr/archives` directory.

Step 2: Add alias IP address to CWS

CWS must have a different IP address for each partition. An alias IP address can be defined for the SP Ethernet for this purpose.

First, the alias host name and IP address should be added in the `/etc/hosts` file. In this example, the default partition uses `sp4en0` as host name and `192.168.4.130` as IP address for SP Ethernet. The other partition uses `sp4alias` as the host name, and `192.168.4.131` as the IP address. Distribute the `/etc/hosts` file that includes aliases to all the nodes by issuing the `pcp` command:

```
# pcp -a /etc/hosts
```

Then set an alias on CWS using the `ifconfig` command:

```
# /usr/sbin/ifconfig en0 alias 192.168.4.131 netmask 255.255.255.0
```

Because this setting will be gone when CWS is rebooted, add this command to the `/etc/rc.net` file.

Step 3: Select a system partition configuration

To display supported system partition configuration on your SP system, issue the `spdisplay_config` command:

```
# spdisplay_config
config.16
config.4_12
config.4_4_4_4
config.4_4_8
config.8_8
#
```

In this example, your SP system has only one frame. You know there are five system partition configurations available for your SP system. The numbers in a file name indicate the number of partitions and the number of nodes that belongs to the partition. For example, `4_12` indicates there are two partitions and one partition has 4 nodes and the other has 12 nodes. `4_4_4_4` indicates there are four partitions and each partition has 4 nodes.

You use `config.8_8` configuration for this example.

Step 4: Select a system partition layout

One configuration may have more than one layout. The layout indicates which node belongs to which partition. To know the layouts of the `config.8_8` configuration, issue the `spdisplay_config` command:

```

# spdisplay_config -R -n config.8_8
layout.1/syspar.1/nodelist:
switch node numbers:  0  1  4  5  8  9 12 13
node numbers:        1  2  5  6  9 10 13 14

layout.1/syspar.2/nodelist:
switch node numbers:  2  3  6  7 10 11 14 15
node numbers:        3  4  7  8 11 12 15 16

layout.2/syspar.1/nodelist:
switch node numbers:  0  1  4  5 10 11 14 15
node numbers:        1  2  5  6 11 12 15 16

layout.2/syspar.2/nodelist:
switch node numbers:  2  3  6  7  8  9 12 13
node numbers:        3  4  7  8  9 10 13 14

layout.3/syspar.1/nodelist:
switch node numbers:  0  1  2  3  4  5  6  7
node numbers:        1  2  3  4  5  6  7  8

layout.3/syspar.2/nodelist:
switch node numbers:  8  9 10 11 12 13 14 15
node numbers:        9 10 11 12 13 14 15 16
#

```

The config.8_8 configuration has three layouts, layout.1, layout.2, and layout.3. In the case of layout.3, one partition, syspar.1, has node 1, 2, 3, 4, 5, 6, 7, and 8. The other partition, syspar.2, has node 9, 10, 11, 12, 13, 14, 15, and 16.

You use the layout.3 layout for this example.

Step 5: Customize a system partition

This step assigns the host name discussed in “Step 2: Add alias IP address to CWS” on page 356 to the partitions. In this example, you assign sp4en0 to the partition syspar.1 and sp4alias to the syspar.2. To assign these host names, issue the `smitty syspar_cust` fast path. When you select **config.8_8/layout.3/syspar.1:** or **config.8_8/layout.3/syspar.2:**, the following SMIT menu is displayed:

```

Enter Customization Arguments for this System Partition

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* System Partition Name or IP Address [sp4en0]
* PSSP Code Level                     PSSP-3.1      +
Default Install Image                 [bos.obj.ssp.432] +
Primary Node                           [1]          +
Backup Primary Node                   [8]          +
Authorization for root rcmds           k4 std      +
Authentication Methods                 k4 std      +
System Partition Path                  config.8_8/layout.3/sy>

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

This SMIT menu is for syspar.1. You need to specify sp4en0 in the System Partition Name or IP Address field. You also need to select primary and primary backup node from the nodes that belong to this partition.

The following `spcustomize_syspar` command performs the same operation:

```

# /usr/lpp/ssp/bin/spcustomize_syspar -n sp4alias -l PSSP-3.1 \
> -d bos.obj.ssp.432 -e 1 -b 8 -r k4 std -m k4 std \
> config.8_8/layout.3/syspar.1

```

For syspar.2, make sure that you specify sp4alias in the System Partition Name or IP address field.

Step 6: Shut down all nodes in changing system partitions

Before you shut down the nodes, issue the `Eunpartition` command:

```
# Eunpartition
```

This operation prepares the SP Switch that belongs to the current system partition for a new system partition configuration.

You need to shut down only the nodes in changing system partition. In this example, there is only one system partition; so, shut down all the nodes by issuing the `cshutdown` command:

```
# cshutdown -F ALL
```

Step 7: Run the `setup_server` command on CWS

The `setup_server` command defines the new rcmd principal associated with the new host name alias:

```
# setup_server
```

Step 8: Apply a system partition configuration

Issue the `smitty syspar_apply` fast path to apply a system partition configuration. When you select the **config.8_8/layout.3**, the following SMIT menu is displayed:

```
Apply System Partition Configuration

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
System Partition Apply Option      Apply this config.      +
Correct VSD configuration?        No. Discontinue         +
System Partition Path              config.8_8/layout.3

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

If you want to verify the system partition configuration before you apply it, specify **Verify Only** in the System Partition Apply Option field.

You can issue the `spapply_config` command to perform the same operation:

```
# /usr/lpp/ssp/bin/spapply_config config.8_8/layout.3
```

For verify, issue the `spapply_config` command with `-v` flag:

```
# /usr/lpp/ssp/bin/spapply_config -v config.8_8/layout.3
```

Step 9: Configure authentication methods for all new partitions

To set the authentication methods for all new partitions, issue the `smitty spauth_rcmd` fast path:

```

Select Authorization Methods for Root access to Remote Commands

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* System Partition names                sp4en0                +
* Authorization Methods                  k4 std                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Specify host name, in this case, `sp4en0` or `sp4alias`, in the System Partition names field. PSSP 3.1 requires `k4` and `std` in the Authorization Methods field at least. This step should be executed for each system partition.

This SMIT operation is equivalent to the `spsetauth` command:

```
# spsetauth -d -p sp4en0 k4 std
```

To enable selected authentication methods for each partition, issue the `smitty spauth_methods` fast path:

```

                                Enable Authentication Methods

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Enable on Control Workstation Only      no          +
Force change on nodes                   no          +
* System Partition names                 sp4en0     +
* Authentication Methods                 k4 std     +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

This SMIT menu operation is equivalent to the `chauthpar` command:

```
# chauthpar -p sp4en0 k4 std
```

This step should be executed for each system partition.

For more details about authentication methods, refer to 16.1, “Authentication and authorization methods” on page 446.

Step 10: Reboot all nodes in changed partitions

Reboot all nodes by issuing the `cstartup` command:

```
# cstartup ALL
```

For further reading, refer to Chapter 15, “Managing System Partitions” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

13.2.2 Deleting a system partition

If a SP system has two partitions, and you want to reconfigure it to one partition, it seems you need to delete one partition. But what you actually need to do is to apply a new configuration that has one partition. Therefore, refer to 13.2.1, “Applying a system partition” on page 355 to see how to delete one system partition. When you reconfigure the system partition,

delete the unnecessary alias setting. You can check current alias settings by using the `ifconfig` command:

```
# ifconfig en0
en0: flags=e080863<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,
GROUPRT,64BIT>
    inet 192.168.4.130 netmask 0xffffffff0 broadcast 192.168.4.255
    inet 192.168.4.131 netmask 0xffffffff0 broadcast 192.168.4.255
#
```

To delete the IP address 192.168.4.131, issue the `ifconfig` command with the `delete` flag:

```
# ifconfig en0 delete 192.168.4.131
# ifconfig en0
en0: flags=e080863<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,
GROUPRT,64BIT>
    inet 192.168.4.130 netmask 0xffffffff0 broadcast 192.168.4.255
#
```

Do not forget to delete unnecessary entries from the `/etc/rc.net` or `/etc/hosts` file also.

13.2.3 Creating your original partition configuration/layout

You can create your original partition configuration/layout by System Partitioning Aid Perspective. There is a restriction when you do this. In the case of switched SP system, you can not create partitions beyond the switch chip boundary. PSSP provides all possible configuration/layout for one frame configuration by default. So, if you have only one switched SP frame in your SP system, you can not create your original partition configuration/layout. If the SP system is switchless SP system, you can create your original partition configuration/layout to whatever you want.

To start the System Partitioning Aid Perspective, issue the `spsyspar` command:

```
# spsyspar
```

You will see the System Partitioning Aid Perspective window as shown in Figure 84 on page 363:

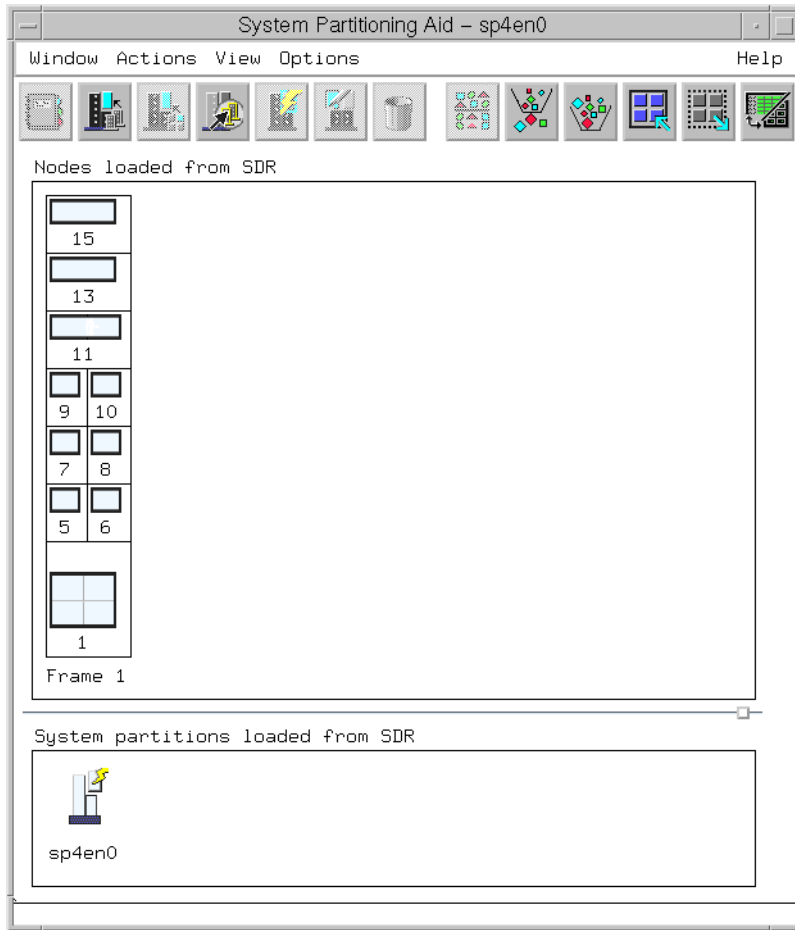









Figure 84. System Partitioning Aid Perspective window

Table 12 shows the tool bar icons and their names.

Table 12. System Partitioning Aid Perspective tool bar icons

Icon	Name of Icon
	View or modify properties of the selected object

Icon	Name of Icon
	Display defined and user generated system configurations
	Place selected nodes into an active partition
	Generate files used to define a system configuration
	Activate a system partition
	Define a new system partition
	Delete selected system partition

To add new configuration/layout, click the **Define a new system partition** tool bar icon:



The Define System Partition dialog box as shown in Figure 85 on page 365 will be displayed:

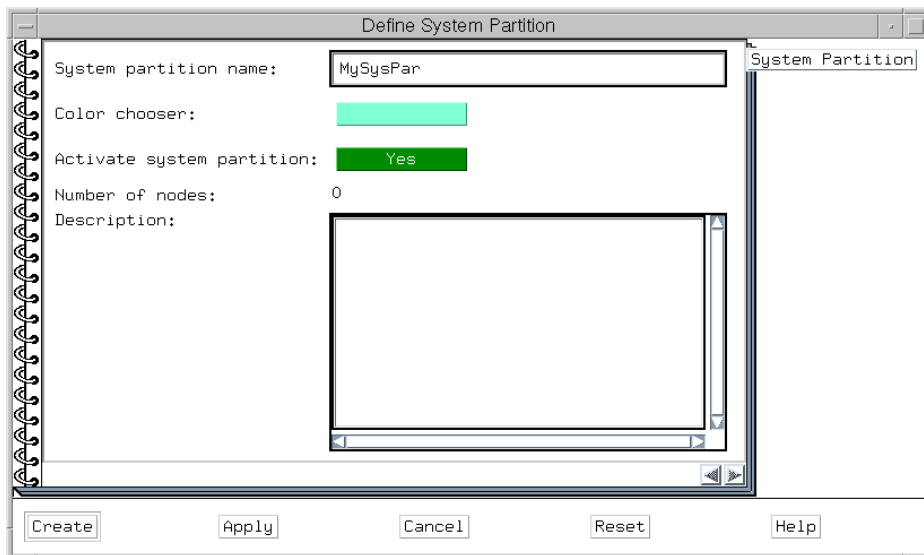


Figure 85. Define System Partition dialog box

Type your favorite system partition name in the System partition name: field. Then click the **Create** button. The new icon will be added in the System partitions loaded from SDR pane as shown in Figure 86 on page 366:

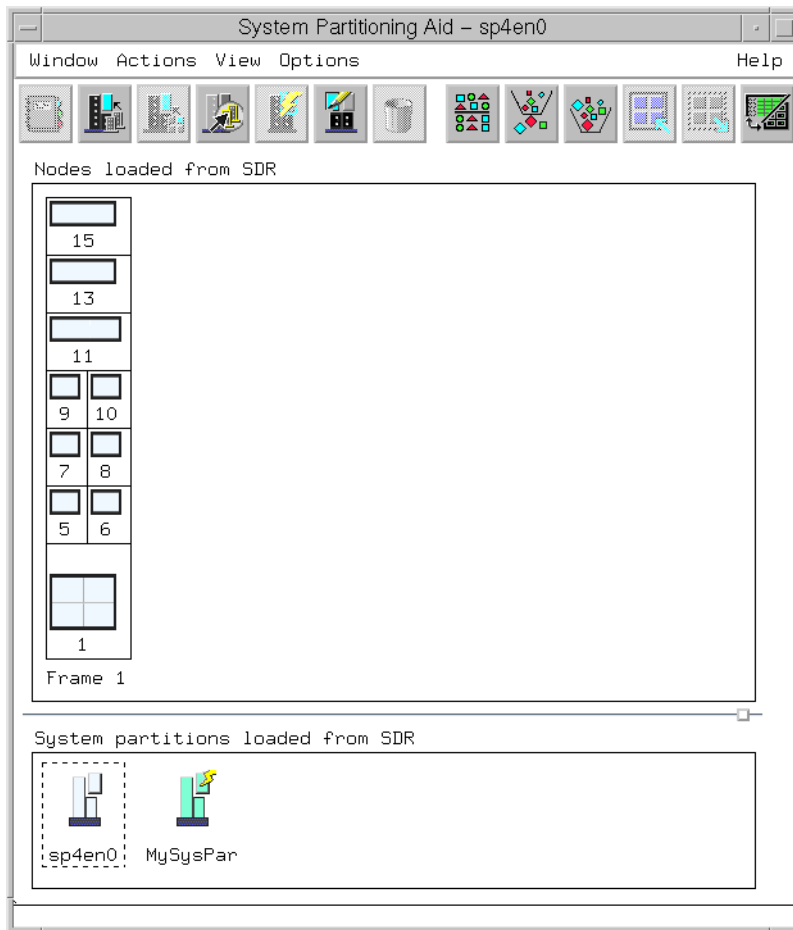


Figure 86. New system partition

Then, you can select nodes that you want to place in the MySysPar partition from the nodes loaded from SDR pane. To place the selected nodes in the MySysPar partition, click the **Place selected nodes into an active partition** tool bar icon:



If your SP system is a switched SP system, all nodes that belongs to the same switch chip to which your selected nodes belong are automatically placed into that partition.

If you want to place the selected nodes to the sp4en0 partition, select the **sp4en0** partition icon from System partition loaded from SDR pane and click the **Activate a system partition** tool bar icon:



It will set the sp4en0 partition as the active partition. Then, select a node from Nodes loaded from SDR pane and click the **Place selected nodes into an active partition** tool bar icon:



To generate system partition configuration/layout file based on your definition, click the **Generate files used to define a system configuration** tool bar icon:



If it is generated successfully, click the **Display defined and user generated system configurations** tool bar icon:



The Display Existing Configurations window shown in Figure 87 on page 368 will be displayed.

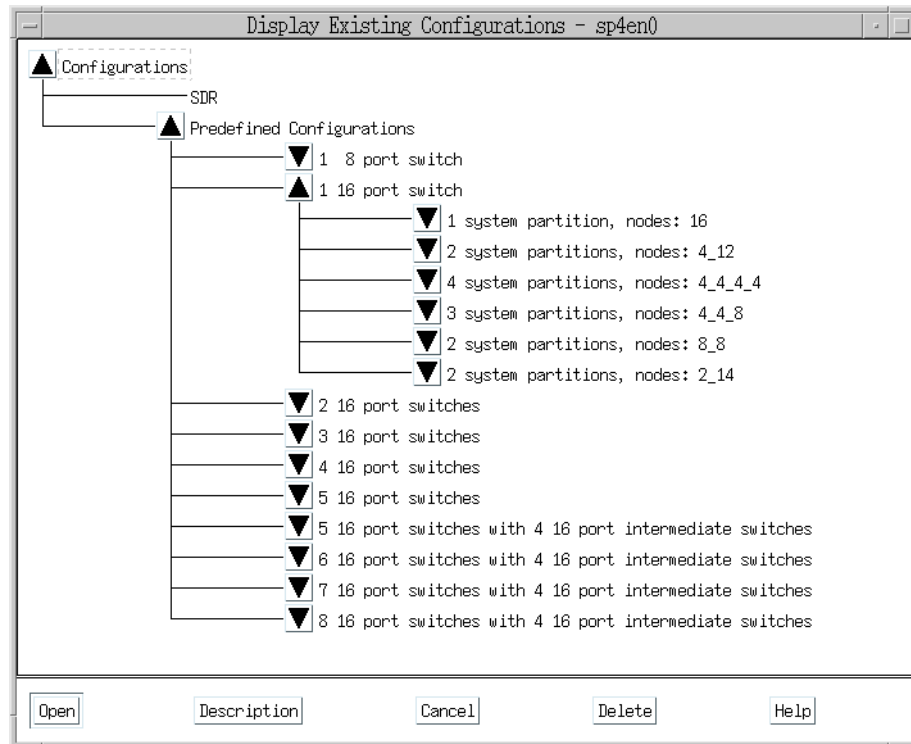


Figure 87. Display existing configurations

Because this example uses a switchless SP system, *2 system partitions, nodes: 2_14*, were defined.

Once it is generated correctly, it can be used the same as the system default configuration/layout that PSSP provides.

13.3 Node group

An SP system can be looked at as a consolidated stand-alone RS/6000 machine. You can log in to each node and manage it as a stand-alone RS/6000 machine. But, it is very verbose; so, there are some node grouping functions or commands available. They can handle several nodes as one object. This section provides this information.

13.3.1 Starting up nodes by group

On a SP System, there may be some client nodes that need services from the other server nodes. When you start up the SP system, you need to consider such dependencies. You need to start up the server node before you start up the client node, for example.

To solve this problem you can use the `cstartup` command. The command determines the dependencies of nodes from a special file `/etc/cstartSeq` and start up the nodes in proper sequence. The following is a sample `/etc/cstartSeq` file:

```
# cat /etc/cstartSeq
Server > Client1 > Client2
Server: sp4n01,sp4n09,sp4n10
Client1: sp4n05,sp4n06,sp4n07,sp4n08
Client2: sp4n11,sp4n13,sp4n15
#
```

This file defines three groups, Server, Client1, and Client2. You can name them as you like. Server group contains three nodes, sp4n01, sp4n09, and sp4n10. Client1 group contains four nodes, sp4n05, sp4n06, sp4n07, and sp4n08. Client2 group contains three nodes, sp4n11, sp4n13, and sp4n15. The first line, Server > Client1 > Client2, means that Server group is started first. When Server group is up, then Client1 group is started. When Client1 group is up, then Client 2 group is started.

To start up all the nodes using this sequence, issue the `cstartup` command:

```
# cstartup ALL
```

In the case that not all the nodes in Server group are available, and you try to start up one of the nodes in Client1 group, the command will fail:

```
# cstartup sp4n05
Progress recorded in /var/adm/SPlogs/cs/cstart.1106192208.34160.
cstartup: 0035-158 Node sp4n01 (1) of Predecessor Group Server is initially do
wn and is a non-target node.
cstartup: 0035-158 Node sp4n09 (9) of Predecessor Group Server is initially do
wn and is a non-target node.
cstartup: 0035-160 /usr/lpp/ssp/bin/csStart failed -- pre-existing sequence viol
ation.
cstartup: 0035-154 The initial state of nodes does not allow the
command to continue.
#
```

In this example, you know sp4n01 and sp4n09 are both down, so you can not start up sp4n05.

If you issue this with an `-x` flag, the command starts up the node even if there are non-target nodes gating the target node start up.

For more information on the `cstartup` command, refer to the following IBM publications:

- Chapter 2, “Starting Up and Shutting Down the SP System” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*
- *IBM Parallel System Support Programs for AIX: Command and Technical Reference, SA22-7351*

13.3.2 Shutting down nodes by group

On a SP System, there may be some client nodes that need services from the other server nodes. When you shut down the SP system, you need to consider such dependencies. You need to shut down the client node before you shut down the server node, for example.

To solve this problem, you can use the `cshutdown` command. The command determines the dependencies of nodes from a special file `/etc/cshutSeq` and shuts down the nodes in proper sequence. The following is a sample `/etc/cshutSeq` file:

```
# cat /etc/cshutSeq
Server > Client1 > Client2
Server: sp4n01,sp4n09,sp4n10
Client1: sp4n05,sp4n06,sp4n07,sp4n08
Client2: sp4n11,sp4n13,sp4n15
#
```

This file defines three groups: Server, Client1, and Client2. You can name them as you like. Server group contains three nodes: sp4n01, sp4n09, and sp4n10. Client1 group contains four nodes: sp4n05, sp4n06, sp4n07, and sp4n08. Client2 group contains three nodes: sp4n11, sp4n13, and sp4n15. The first line, `Server > Client1 > Client2`, means that Client2 group is shut down first. When Client2 group is down, Client1 group is shut down. When Client1 group is down, then Server is shut down.

To shut down all the nodes using this sequence, issue the `cshutdown` command:


```
# cshutdown ALL
```

You can also use similar flags to the `shutdown` command. If you want issue the `cshutdown` command to start the shut down immediately, without issuing warning messages, and restart the all the nodes:

```
# cshutdown -F -r ALL
```

In the case that not all the nodes in Client2 group are shut down, and you try to shut down one of the nodes in Client1 group, the command will fail:

```
# cshutdown sp4n05
Progress recorded in /var/adm/SPlogs/cs/cshut.1106181417.33254.
cshutdown: 0036-117 Some nodes are gated by others:
Target nodes (sp4n05.msc.itso.ibm.com ) are gated by non-target nodes (sp4n11.ms
c.itso.ibm.com sp4n15.msc.itso.ibm.com ).
#
```

In this example, you know sp4n11 and sp4n15 are both up; so, you can not shut down sp4n05.

If you issue this with an `-x` flag, the command shuts down the node even if there are non-target nodes gating the target node shut down.

For more information about the `cshutdown` command, refer to the following IBM publications:

- Chapter 2, “Starting Up and Shutting Down the SP System” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348
- *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351

13.3.3 Managing nodes using node group

There are a couple of ways to manage nodes by group. One of them is the node group management function that is provided by PSSP. You can use the `smitty png_create` fast path to use this function:

```

                                Create New Node Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* Node Group Name                [Group1]
  Nodes to Include                [13 15 ]      +
  Node Groups to Include          []           +

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell     F10=Exit     Enter=Do

```

You can name the group by the Node Group Name field. Specify the nodes you want to include in that group in the Nodes to Include field.

If you want to use commands, the following commands are available:

ngaddto	This command adds nodes and node groups to the definition list of the destination node group.
ngclean	This command cleans up a node group, removing references to nodes and node groups that are not in the current system partition.
ngcreate	This command creates and optionally populates a named node group.
ngdelete	This command removes node groups from persistent storage.
ngdelfrom	This command deletes nodes and node groups from the definition list of the destination node group.
ngfind	Returns a list of all node groups whose definition list contains the specified node or node group.
nglist	This command returns a list of all node groups in the current system partition.

<code>ngnew</code>	This command creates but does not populate new node groups in persistent storage.
<code>ngresolve</code>	This command returns a list of hosts in the specified node group.

The following is an example usage for the commands. It creates node group Group1, adds node 13 and 15 to this group, lists all the node groups, and lists all the nodes belongs to node group Group1:

```
# cd /usr/lpp/ssp/bin
# ngcreate Group1
# ngaddto Group1 13 15
# nglis
Group1
# ngresolve Group1
13
15
#
```

You can use group name when you issue remote commands.

- In the case of the `dsh` command:


```
# dsh -N Group1 date
```
- In the case of the `Efence` command:


```
# Efence Group1
```

For the complete list of commands that accepts a node group as an argument, refer to Chapter 16, “Managing Node Groups” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

13.3.4 Managing nodes using working collective

Another method to manage node by group is using a working collective. It can be specified by the `WCOLL` environment variable.

For example, if you want to issue remote commands only to node 13 and 15, you need to create a file that contains the host names of them. Then, you define the `WCOLL` environment variable that points the file:

```

# cat /.node1315
sp4n13
sp4n15
# dsh date
dsh: 5025-507 Working collective environment variable not set
# export WCOLL=/.node1315
# dsh date
sp4n13.msc.itso.ibm.com: Mon Nov  1 13:37:28 EST 1998
sp4n15.msc.itso.ibm.com: Mon Nov  1 13:37:28 EST 1998
#

```

To build working collective file easily, there is the `hostlist` command available.

For example, if you want to make working collective file that includes only responding nodes:

```

# hostlist -a -v
sp4n01
sp4n05
sp4n06
# cshutdown -F sp4n06 &
Progress recorded in /var/adm/SPlogs/cs/cshut.1101140020.12836.
# hostlist -a -v
sp4n01
sp4n05
# hostlist -a -v > /.wcoll
# export WCOLL=/.wcoll
# dsh date
sp4n01.msc.itso.ibm.com: Mon Nov  1 13:37:28 EST 1998
sp4n05.msc.itso.ibm.com: Mon Nov  1 13:37:28 EST 1998
#

```

The output of the `hostlist` command can be piped directly to the `dsh` command:

```

# hostlist -av | dsh -w - date
sp4n01.msc.itso.ibm.com: Mon Nov  1 13:37:28 EST 1998
sp4n05.msc.itso.ibm.com: Mon Nov  1 13:37:28 EST 1998
#

```

Attention

The path used resolving the `dsh` command on the target nodes is the path set by the user with the `DSHPATH` environment variable. If `DSHPATH` is not set, the path used is the `rsh` command default path, `/usr/ucb:/bin:/usr/bin`. The `DSHPATH` environment variable only works when the user's remote login shell is the Bourne or Korn shell. An example would be to set `DSHPATH` to the path set on the source machine (for example, `DSHPATH=$PATH`).

For the complete list of parallel commands that accept a working collective, refer to Chapter 3, "Parallel Management Commands" in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

Part 6. Managing administrative tasks

Chapter 14. Administration tools

IBM Parallel System Support Programs for AIX (PSSP) provides several tools to make system management easy. These tools can lessen the work load of a system administrator. This chapter covers the following tools:

- File collection technology
To keep specified files on each node identical, Software Update Protocol (SUP) is used.
- SP User Management
Several commands are provided for this task. It is closely related file collection technology to keep user management files identical on each node.
- Time synchronization
To synchronize the time all over the SP system, Network Time Protocol (NTP) is used.
- The Automounter
The Automounter is supported by AIX, and PSSP utilizes it. Users on each node can use their own home directory from the server on the request base.
- Accounting
SP Accounting provides you with usage information and daily records about specific users.

14.1 File collection technology

The SP system installs file collection technology by default to simplify the task of maintaining duplicate files on multiple nodes. In the delivered system, the files that are required on the control workstation (CWS), boot/install servers (BISs), and processor nodes, belong to file collections. Grouping these files into file collections, and using the provided tools to manage them, allows you to easily maintain their consistency and accuracy on multiple nodes in the SP system.

The file collections is closely related with SP User Management function. With default settings of PSSP, each node gets its user management files, such as `/etc/passwd`, `/etc/group`, `/etc/security/passwd`, and so on, from the CWS or BIS by the file collection technology. The `supper` command, that is executed from cron on each node, is responsible for it. It is based on Software

Update Protocol (SUP), a public domain software. If the files are updated on the server machine, the `supper` command gets them and replaces them on the node.

For further reading, refer to Chapter 4, “Managing File Collections” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

To get documentation updates and information about SUP, refer to the `/usr/lpp/spp/README/spp.public.README` file. There is SUP source code available on the CWS. Refer to 2.5.2, “Where is the source code?” on page 126, for more detail.

14.1.1 Getting information

When you install PSSP, file collection technology is configured by default. Information about file collection technology is stored in SP class of the System Data Repository (SDR). You can get this information by issuing the `splstdata` command:

```

# splstdata -e
      List Site Environment Database Information

attribute          value
-----
control_workstation  sp4en0
cw_ipaddrs          9.12.1.140:192.168.4.130:
install_image       bos.obj.ssp.432
remove_image        false
primary_node        1
ntp_config          consensus
ntp_server          ""
ntp_version         3
amd_config          true
print_config        false
print_id            ""
usermgmt_config     true
passwd_file         /etc/passwd
passwd_file_loc     sp4en0
homedir_server      sp4en0
homedir_path        /home/sp4en0
filecoll_config    true
supman_uid        102
supfilesrv_port   8431
spacct_enable       false
spacct_actnode_thresh  80
spacct_exclude_enable  false
acct_master         0
cw_has_usr_clients  false
code_version        PSSP-3.1
layout_dir          ""
authent_server      ssp
backup_cw           ""
ipaddrs_bucw       ""
active_cw           ""
sec_master          ""
ods_server          ""
cell_name           ""
cw_lppsource_name   aix432
cw_dcehostname      ""
#

```

In these attributes, file collection technology uses the following attributes:

- filecoll_config** This attribute indicates true or false if file collection technology code should be installed.
- supman_uid** This attribute indicates the user ID for supman.
- supfilesrv_port** This attribute indicates the file collection technology port number.

By default, the filecoll_config attribute is set true and configured to the SP system.

14.1.2 Checking status

If you want to know the name, resident status, and access point of all available file collections, plus the name and estimated size of their associated file systems, issue the `supper` command on a CWS, BIS, or node.

In the case you issue the `supper` command on the CWS, the output should look like the following:

```
# /var/sysman/upper status
```

Collection	Resident	Access Point	Filesystem	Size
node.root	-	/	-	-
power_system	-	/share/power/system	-	-
sup.admin	-	/var/sysman	-	-
user.admin	-	/	-	-

```
#
```

In the case you issue the `supper` command on a BIS, the output should look like the following:

```
# /var/sysman/supper status
```

Collection	Resident	Access Point	Filesystem	Size
node.root	Yes	/	-	-
power_system	Yes	/share/power/system	-	-
sup.admin	Yes	/var/sysman	-	-
user.admin	Yes	/	-	-

```
#
```

In the case you issue the `supper` command on a node, the output should look like the following:

```
# /var/sysman/supper status
```

Collection	Resident	Access Point	Filesystem	Size
node.root	Yes	/	-	-
power_system	-	/share/power/system	-	-
sup.admin	Yes	/var/sysman	-	-
user.admin	Yes	/	-	-

```
#
```

14.1.3 Checking served file collections

If you want to know all the served file collections on a BIS, issue the `supper` command on the BIS:

```
# /var/sysman/supper serve
sup.admin
user.admin
#
```

In this example, you know there are two file collections, `sup.admin` and `user.admin`, served on this BIS.

14.1.4 Checking resident files

If you want to know all the resident files resulting from the `supper update` or `supper install` command, issue the `supper` command on a BIS or node:

```

# supper files sup.admin
var/sysman
var/sysman/etc
var/sysman/file.collections
var/sysman/logs
var/sysman/sup
var/sysman/sup/lists
var/sysman/sup/lists/node.root
var/sysman/sup/lists/power_system
var/sysman/sup/lists/sup.admin
var/sysman/sup/lists/user.admin
var/sysman/sup/node.root/list
var/sysman/sup/power_system/list
var/sysman/sup/sup.admin/list
var/sysman/sup/user.admin/list
var/sysman/supper
# supper files user.admin
etc/group
etc/passwd
etc/passwd.id.idx
etc/passwd.nm.idx
etc/security/group
etc/security/passwd
etc/security/passwd.idx
# supper files node.root
#

```

From the output, you can find out what kind of files are handled by each file collection.

14.1.5 Checking file collection server

If you like to know the file collection server from which you receive the file collection, issue the `supper` command on a BIS or node:

```

# supper where
supper: Collection node.root would be updated from server sp4en0.msc.itso.ibm.com.
supper: Collection power_system would be updated from server sp4en0.msc.itso.ibm.com.
supper: Collection sup.admin would be updated from server sp4en0.msc.itso.ibm.com.
supper: Collection user.admin would be updated from server sp4en0.msc.itso.ibm.com.
#

```

In this example, all the file collections, `node.root`, `power_system`, `sup.admin`, and `user.admin`, use `sp4en0.msc.itso.ibm.com` as the file collection server.

14.1.6 Checking last updated time and date

If you want to know the last updated time and date when the `supper update` command executed, issue the `supper` command on a BIS or node:

```
# supper when

Collection                                Last Update
=====
node.root                                  Thu Mar 25 11:10:07 1999
sup.admin                                  Thu Mar 25 11:10:03 1999
user.admin                                 Thu Mar 25 11:10:05 1999
=====

#
```

In this example, all the file collections, node.root, sup.admin, and user.admin, are updated on Thu Mar 25 11:10 1999.

14.1.7 Updating files managed by file collection

PSSP provides system default for the file collections. The contents of the file collection is listed in the user.admin file located in the /var/sysman/sup/lists directory:

```
# cd /var/sysman/sup/lists
# ls
node.root    power_system  sup.admin    user.admin
# cat user.admin
symlinkall
upgrade ./etc/passwd
upgrade ./etc/passwd.nm.idx
upgrade ./etc/passwd.id.idx
upgrade ./etc/group
upgrade ./etc/security/group
upgrade ./etc/security/passwd
upgrade ./etc/security/passwd.idx
# Do not delete or change this comment. Automount configuration added.
upgrade ./etc/auto.master
upgrade ./etc/auto/maps/auto.*
execute /etc/amd/refresh_amd (./etc/auto/maps/auto.u)
upgrade ./etc/auto/cust/*
upgrade ./etc/amd/amd-maps/amd.*
execute /etc/amd/refresh_amd (./etc/amd/amd-maps/amd.u)
#
```

Figure 88. The user.admin file collection

The following is a explanation of the command keywords and operands:

symlinkall This keyword treats all symbolic links. They are transferred as links and not followed. By default, symbolic links are followed.

upgrade *filename* The specified files or directories are included in the list of files to be upgraded unless they are also specified by the `omit` or `omitany` command or are in the `refuse` file. If a directory name is given, it recursively includes all subdirectories and files within that directory.

execute *exec-command (filename)*

The command is specified in `executed` on the client process whenever any of the files listed in parentheses are upgraded.

On BISs and nodes, the `supper` command updates the file collections once an hour by default. So, if you add a user on the CWS, the user will be available on BISs and nodes in one hour. If you need the user available immediately, issue the `supper` command on BISs and nodes manually by using the `dsh` command:

```
# dsh -a /var/sysman/supper update
sp4n01: Updating collection node.root from server sp4en0.msc.itso.ibm.com.
sp4n01: File Changes: 0 updated, 0 removed, 0 errors.
sp4n01: Updating collection sup.admin from server sp4en0.msc.itso.ibm.com.
sp4n01: File Changes: 0 updated, 0 removed, 0 errors.
sp4n01: Updating collection user.admin from server sp4en0.msc.itso.ibm.com.
sp4n01: File Changes: 3 updated, 0 removed, 0 errors.
sp4n03: Updating collection node.root from server sp4en0.msc.itso.ibm.com.
sp4n03: File Changes: 0 updated, 0 removed, 0 errors.
sp4n03: Updating collection sup.admin from server sp4en0.msc.itso.ibm.com.
sp4n03: File Changes: 0 updated, 0 removed, 0 errors.
sp4n03: Updating collection user.admin from server sp4en0.msc.itso.ibm.com.
sp4n03: File Changes: 3 updated, 0 removed, 0 errors.
sp4n05: Updating collection node.root from server sp4en0.msc.itso.ibm.com.
sp4n05: File Changes: 0 updated, 0 removed, 0 errors.
sp4n05: Updating collection sup.admin from server sp4en0.msc.itso.ibm.com.
sp4n05: File Changes: 0 updated, 0 removed, 0 errors.
sp4n05: Updating collection user.admin from server sp4en0.msc.itso.ibm.com.
sp4n05: File Changes: 3 updated, 0 removed, 0 errors.
#
```

If you want to change the update cycle, refer to 14.1.8, “Changing update cycle” on page 386.

14.1.8 Changing update cycle

The `supper update` command is included in the `crontabs` file by default. The command is set to be issued hourly. You can modify the `crontabs` file to issue the `supper update` command more or less frequently.

Log in a node that you want to modify the update cycle, then issue the `crontab` command:

```
# crontab -e
```

You will enter the editor screen. Find the line for the `supper` command. It should look like the following:

```
10 * * * * /var/sysman/supper update sup.admin user.admin node.root
1>/dev/null 2>/dev/null
```

If you want to update more often, duplicate this line and change the first column as follows:

```
40 * * * * /var/sysman/supper update sup.admin user.admin node.root
1>/dev/null 2>/dev/null
```

Now the `supper update` command is issued every 30 minutes.

Attention

If you want to change update cycle for all the nodes, you have to do this operation on each node.

14.1.9 Update sequence

There are two things you need to be careful of when you issue the `supper update` command: The sequence within file collections and the sequence within machines.

Within file collections

Take a look at crontab entry for file collections by issuing the `crontab` command:

```
# dsh -w sp4n01 crontab -l | grep supper
sp4n01: 10 * * * * /var/sysman/supper update sup.admin user.admin node.root 1>/
dev/null 2>/dev/null
#
```

The `supper` command updates the `sup.admin` file collection first, then other file collections. The `sup.admin` file collection contains definitions about all file collections. The definition files are located in `/var/sysman/sup` directory. The `sup.admin` file collection must be updated first to reflect other file collections' change. They may be added, modified, or deleted. So, if you issue the `supper`

command manually, you need to pay attention the update sequence of file collections.

Within machines

If your SP system uses BIS, the file collections need to be propagated from the CWS to the BISs first, then BISs to the other nodes. So, if you issue `supper` command manually, you need to issue the command on the BISs first, then on the other nodes.

14.1.10 Checking log files

Logs for the file collection technology are saved in the `/var/adm/SPlogs/filec` directory on each node instead of the CWS. You can show these log files by using the `supper` command:

```
# supper log
                                     Supper Log
Host: sp4n01                          Date: 03/25/1999 13:10:02
                                     CHANGES TO SYSTEM
Collection      Server          Updated   Removed   Errors
-----
sup.admin       sp4en0.msc.its  0         0         0
user.admin      sp4en0.msc.its  0         0         0
node.root       sp4en0.msc.its  0         0         0
-----
Totals          0              0         0
Supper Finished 03/25/1999 13:10:08
#
```

The command shows the summary of the current or most recent command session.

If you want to know the raw output of the current or most recent command session log, issue the `supper` command:

```

# supper rlog
SUP 7.24 (4.3 BSD) for file /tmp/.sf21150 at Mar 25 13:10:03
SUP Upgrade of sup.admin at Thu Mar 25 13:10:03 1999
SUP Fileserver 7.12 (4.3 BSD) 44044 on sp4en0.msc.itso.ibm.com
SUP Locked collection sup.admin for exclusive access
SUP Requesting changes since Thu Mar 25 12:10:03 1999
SUP Upgrade of sup.admin completed at Thu Mar 25 13:10:03 1999
SUP Scan for sup.admin starting at Thu Mar 25 13:10:04 1999
SUP Scan for sup.admin completed at Thu Mar 25 13:10:04 1999
SUP 7.24 (4.3 BSD) for file /tmp/.sf21150 at Mar 25 13:10:05
SUP Upgrade of user.admin at Thu Mar 25 13:10:06 1999
SUP Fileserver 7.12 (4.3 BSD) 38778 on sp4en0.msc.itso.ibm.com
SUP Locked collection user.admin for exclusive access
SUP Requesting changes since Thu Mar 25 12:10:05 1999
SUP Upgrade of user.admin completed at Thu Mar 25 13:10:06 1999
SUP Scan for user.admin starting at Thu Mar 25 13:10:06 1999
SUP Scan for user.admin completed at Thu Mar 25 13:10:06 1999
SUP 7.24 (4.3 BSD) for file /tmp/.sf21150 at Mar 25 13:10:08
SUP Upgrade of node.root at Thu Mar 25 13:10:08 1999
SUP Fileserver 7.12 (4.3 BSD) 31372 on sp4en0.msc.itso.ibm.com
SUP Locked collection node.root for exclusive access
SUP Requesting changes since Thu Mar 25 12:10:07 1999
SUP Upgrade of node.root completed at Thu Mar 25 13:10:08 1999
#

```

If you find error messages in the log file, you should check the following:

- On the file collection server machine (CWS or BIS), the supfilesrv daemon must be running. You can check this by issuing the `lssrc` or `ps` command:

```

# lssrc -s supfilesrv
Subsystem      Group          PID    Status
supfilesrv     -              11218  active
# ps -e | grep sup
11218      -  0:10 supfilesrv
#

```

- On the file collection client machine, the `/etc/ssp/server_name` file must contain the right file collection server name:

```

# cd /etc/ssp
# cat server_name
192.168.31.1 sp3n01en1.msc.itso.ibm.com sp3n01en1
#

```

14.2 SP User Management

The SP system provides the ability for you to manage your user accounts by adding users, deleting users, or changing the user account information from a single point of control. SP User Management ensures that users have the same account, home directory, and environment across all the nodes in the SP system. SP User Management is optional, and it is designed to fit into your existing user management system.

For further reading, refer to Chapter 5, “Managing User Accounts” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

14.2.1 Getting information

When you install PSSP, SP User Management is configured by default. The configuration information is defined in the SP SDR class.

You can display this information by issuing the `splstdata` command:

```

# splstdata -e
      List Site Environment Database Information

attribute          value
-----
control_workstation  sp4en0
cw_ipaddrs          9.12.1.140:192.168.4.130:
install_image       bos.obj.ssp.432
remove_image        false
primary_node        1
ntp_config          consensus
ntp_server          ""
ntp_version         3
and_config         true
print_config        false
print_id            ""
usermgmt_config    true
passwd_file         /etc/passwd
passwd_file_loc     sp4en0
homedir_server      sp4en0
homedir_path        /home/sp4en0
filecoll_config    true
supman uid          102
supfilesrv_port     8431
spacct_enable       false
spacct_actnode_thresh 80
spacct_exclude_enable false
acct_master         0
cw_has_usr_clients  false
code_version        PSSP-3.1
layout_dir          ""
authent_server      ssp
backup_cw           ""
ipaddrs_bucw       ""
active_cw           ""
sec_master          ""
ods_server          ""
cell_name           ""
cw_lppsource_name   aix432
cw_dcehostname      ""
#

```

Figure 89. Site environment database information for SP user management

SP User Management is controlled by the following attribute:

usermgmt_config This attribute indicates true or false if the SP User Management code and SMIT menu is installed.

SP User Management refers to the following attributes:

filecoll_config This attribute indicates true or false if the file collection technology code is installed.

amd_config This attribute indicates true or false if the PSSP provides automounter support.

Because SP User Management works with file collection technology, the `filecoll_config` attribute needs to be set as true.

To learn about the file collections, refer to 14.1, “File collection technology” on page 379. To learn about the automounter, refer to 14.4, “The Automounter” on page 418.

14.2.2 Adding a user

To add a user, issue the `smitty spmkuser` fast path on the CWS. The following example adds `newuser` whose user ID is 2001:

```

                                Add a User

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* User NAME                       [newuser]
User ID                           [2001]          #
LOGIN user?                       true          +
PRIMARY group                      []           +
Secondary GROUPS                  []           +
HOME directory                    []
Initial PROGRAM                   []           /
User INFORMATION                  []

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell     F10=Exit     Enter=Do

```

Alternatively, issue the `spmkuser` command to perform the same operation:

```
# spmkuser id=2001 newuser
```

To reflect the user on the nodes immediately, you need to use the file collection technology. To do this, issue the `supper` command on the BISs then on the other nodes:

```
# /var/sysman/supper update
```

For more detail about updating files by file collection technology, refer to 14.1.7, “Updating files managed by file collection” on page 385.

The default values for primary group, secondary groups, and the initial program, are defined in the `/usr/lpp/ssp/bin/spmkuser.default` file:

```
# cat /usr/lpp/ssp/bin/spmkuser.default

user:
    group = staff
    groups = staff
    prog = /bin/ksh

#
```

The `spmkuser` command does some works besides what the `mkuser` AIX standard command does:

- Password generation

It generates a random password for the user and stores it in the `/usr/lpp/ssp/config/admin/newpass.log` file:

```
# cat /usr/lpp/ssp/config/admin/newpass.log
newuser      Kzu7MMbi
#
```

Gives the password to the user and ask him/her to change the password immediately. Deleting this file improves the security of your SP system.

- Keeps consistency with the automounter settings
 - If `amd_config` attribute is set as true:

An entry for the user is added to the `/etc/auto/maps/auto.u` file:

```
# The "netinst" entry is for the netinstall server. When automount is used,
# this entry must be defined for the netinstalls to work. Otherwise
# references to /u/netinst will not be resolved and netinstalls will fail.
#
netinst      $HOST:/home:&

newuser      cws1:/home/cws1:&
```

SP User Management uses the `homedir_server` and `homedir_path` attributes to create this entry. The user home directory is set as the `/u/username` directory.

- If `amd_config` attribute is set as `false`:

No entry is added to the `/etc/auto/maps/auto.u` file, and the user home directory is set using the `homedir_path` attribute.

For more detail about the automounter, refer to 14.4, “The Automounter” on page 418.

14.2.3 Deleting a user

To delete a user, issue the `smitty sprmuser` fast path on the CWS. You select the target user name on the first screen, then the following SMIT menu is displayed:

```

                                Remove a User

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Remove AUTHENTICATION information?      Yes          +
Remove HOME directory?                  Yes          +

* User NAME                             user1
User ID                                 204
PRIMARY group                           1
Secondary GROUPS                         staff
HOME directory                           /u/hiro on sp4sw09 (/hom>
Initial PROGRAM                           /bin/ksh
User INFORMATION

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit      Enter=Do

```

Alternatively, issue the `spruser` command to perform the same operation:


```
# spmuser -i -p -r user1
User Profile:
    User: user1
    User ID: 804400252
    Group ID: 804400256
    Groups: staff
    Information:
    Home: /u/user1 on sp4sw09:/home/sp4en0/user1
    Shell: /bin/ksh
    Login permitted: true

Verify Delete <y/n>:y
#
```

If the `amd_config` attribute of SP SDR class is set as true, the entry for the user is deleted from the `/etc/auto/maps/auto.u` file also. For more detail about the automounter, refer to 14.4, “The Automounter” on page 418.

14.2.4 Controlling log in to CWS or nodes

If your SP system uses the SP User Management function, users can not change their password on nodes. They have to log in to CWS to change their password. It likely will cause a security problem that all users can log in CWS anytime.

To protect CWS from unexpected operation, you can control the users who can log in to the CWS. There are several methods available for this requirement:

- Using `/etc/nologin` file
- Using `/etc/security/user` file
- Using `/etc/ftpusers` file
- Using `/usr/lpp/ssp/config/admin/cw_restrict_login` script
- Using the `spacs_cntrl` command

Attention

- All these files are not managed by the file collections by default.
- You can combine these methods to as many as you want.
- You can also utilize these methods on the nodes except the `cw_restrict_login` script.

14.2.4.1 Using the /etc/nologin file

If the /etc/nologin file exists, the system prevents all users except root user from logging in and displays the contents of the /etc/nologin file.

The /etc/nologin file will be removed when you reboot the system. This is a standard AIX function.

To block the log in from all the users except root user, create the /etc/nologin file on the CWS or node, for example:

```
#
#
# Log in to this node is prohibited
#
#
```

The following is the log in process when you use this file. Users except root user can not log in to the CWS or node:

```
AIX Version 4
(C) Copyrights by IBM and by others 1982, 1996.
login: user1
user1's Password:
#
#
# Log in to this node is prohibited
#
#

login:
```

Attention

This method is not effective to access by the `rsh`, `rexec`, `rcp`, or `ftp` command.

14.2.4.2 Using the /etc/security/user file

The /etc/security/user file is provided to manage authentication methods for users. You can control interactive access to the system user by user. This is a standard AIX function.

For example, the following stanza prevents user1 from access by the `login`, `rlogin`, `rsh`, `rexec`, and `rcp` command:

```

user1:
  login = false
  rlogin = false
  ttys = ALL, !RSH, !REXEC

```

Figure 90. The `/etc/security/user` file

You can edit the `/etc/security/user` file directly.

Alternatively, issue the `smitty chuser` fast path:

```

Change / Show Characteristics of a User

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]
* User NAME                               [Entry Fields]
User ID                                   user1
ADMINISTRATIVE USER?                     [202] #
Primary GROUP                             [staff] +
Group SET                                 [staff] +
ADMINISTRATIVE GROUPS                     [] +
ROLES                                      [] +
Another user can SU TO USER?              true +
SU GROUPS                                 [ALL] +
HOME directory                            [/home/sp4en0/user1]
Initial PROGRAM                            [/bin/ksh]
User INFORMATION                           []
EXPIRATION date (MDDhhmmy)                 [0]
Is this user ACCOUNT LOCKED?               false +
User can LOGIN?                            false +
User can LOGIN REMOTELY?                   false +
Allowed LOGIN TIMES                         []
Number of FAILED LOGINS before              [0] #
  user account is locked
Login AUTHENTICATION GRAMMAR                [compat]
Valid TTYS                                 [ALL, !RSH, !REXEC]
[MORE...29]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do

```

Alternatively, issue the `chuser` command:

```
# chuser login=false rlogin=false ttys=ALL, !RSH, !REXEC user1
```

The following is the log in process when you use the `/etc/security/user` file. Users controlled by this file are refused to log in to the CWS or node:

```
AIX Version 4
(C) Copyrights by IBM and by others 1982, 1996.
login: user1
user1's Password:
3004-306 Remote logins are not allowed for this account.

login:
```

Attention

This method is not effective to access by the `ftp` command.

14.2.4.3 Using the `/etc/ftpusers` file

The `/etc/ftpusers` file contains a list of local user names that the `ftpd` server does not allow remote File Transfer Protocol (FTP) clients to use. The format of the `/etc/ftpusers` file is a simple list of user names that also appear in the `/etc/passwd` file. For example:

```
user1
user2
```

To add a user `user1`, for example, issue the `smitty mkftpusers` fast path:

Add a Restricted User

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[Entry Fields]

* Name of Local USER ID [user1]

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

Alternatively, issue the `ruser` command:

```
# ruser -a -f user1
```

To delete a user `user1`, for example, issue the `smitty rmtfusers` fast path:

```

Remove a Restricted User

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

* Name of Local USER ID                                [Entry Fields]
                                                         [user1]          +

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell        F10=Exit           Enter=Do

```

Alternatively, issue the `ruser` command:

```
# ruser -d -f user1
```

This is the standard AIX command.

The following is the ftp process when you use the `/etc/ftpusers` file. Users listed in this file are refused to issue the `ftp` command to CWS, BIS, or node:

```

# ftp sp3cw0
Connected to sp3cw0.itso.ibm.com.
220 sp3en0 FTP server (Version 4.1 Tue Sep 8 17:35:59 CDT 1998) ready.
Name (sp3cw0:root): user1
530 User user1 access denied.
Login failed.
ftp> quit
221 Goodbye.
#

```

14.2.4.4 Using the `cw_restrict_login` script

The `/usr/lpp/ssp/config/admin/cw_restrict_login` script controls the users log in to the CWS. Only the users you designate are allowed to fully log in. Other users can log in to the CWS only for changing their password. When they

change the password, they are logged out. To use this script, perform the following steps:

Step 1: Edit /usr/lpp/ssp/config/admin/cw_allowed file

The /usr/lpp/ssp/config/admin/cw_allowed file has a definition that users are allowed to fully log in to the CWS. You do not need to include the root user in this file.

Add users line by line and do not place any comments in the file. If you want to permit only admin1, admin2, and admin3 to fully log in to the CWS, this file should look like the following:

```
# cd /usr/lpp/ssp/config/admin
# cat cw_allowed
admin1
admin2
admin3
#
```

Step 2: Edit /etc/profile file

To integrate the cw_restrict_login script into the login process, add the following lines to the beginning (or at the most appropriate place) of the /etc/profile file on CWS.

```
# Allow general users to login to control workstation to only
# change their password then log them out.
/usr/lpp/ssp/config/admin/cw_restrict_login
```

Attention

If you are using the AIX Common Desktop Environment (CDE) on the CWS, you will also need to make a link from the cde_cw_restrict_login script for CDE to the appropriate CDE directory:

```
# ln -s /usr/lpp/ssp/config/admin/cde_cw_restrict_login \
>/etc/dt/config/Xsession.d/cde_cw_restrict_login
```

The following is the log in process when you use this script. Users not listed in cw_allowed file will be logged out from the CWS after changing their password:

```

AIX Version 4
(C) Copyrights by IBM and by others 1982, 1996.
login: admin4
admin4's Password:
*****
*
* Welcome to AIX Version 4.3!
*
*****
Last login: Wed Mar 24 16:49:45 EST 1999 on /dev/pts/0 from ps82.itso.ibm.com

You are only permitted to change your password.

Your password has recently been changed.
Type y to change your password again, any other key to exit: y

Changing password for "admin4"
admin4's Old password:
admin4's New password:
Enter the new password again:

Password changed.
Password will take up to an hour to propagate to the rest of the SP system.
Logging out...

```

Attention

You can use this method only for the CWS .

These scripts are provided by PSSP.

14.2.4.5 Using the spacs_cntrl command

The `spacs_cntrl` command controls the user log in by changing the stanza in the `/etc/security/user` file as shown in Figure 90 on page 397.

For example, to block a user, `user1`, issue the `spacs_cntrl` command:

```
# spacs_cntrl block user1
```

To unblock the user, `user1`, issue the `spacs_cntrl` command:

```
# spacs_cntrl unblock user1
```


This command is provided by PSSP.

14.2.5 Managing users, groups, or password on specific nodes

As long as you use SP User Management with default setting, the users added on CWS are propagated on all nodes by the file collection technology. You can not create a user for only specific nodes.

If you want to block some users from logging in specific nodes, you can control their access rights in several ways. If this is the case, refer to 14.2.4, “Controlling log in to CWS or nodes” on page 395.

If you want to manage security information, such as users, groups, and passwords, on specific nodes uniquely, you need to modify the file collection technology configuration. Nevertheless, all the security information will be over written by the file collection technology.

By default, the security information is controlled by the user.admin file collection. The following is the list of security files under control of the user.admin file collection:

- /etc/group
- /etc/passwd
- /etc/passwd.id.idx
- /etc/passwd.nm.idx
- /etc/security/group
- /etc/security/passwd
- /etc/security/passwd.idx

The target of file collections, when you issue the `supper update` command, are defined in the `/var/sysman/sup/.resident` file. By default, the contents of this file should look as follows:

```
# cat /var/sysman/sup/.resident
node.root 0
sup.admin 0
user.admin 0
#
```

To manage security information for the specific node in its own way, delete the following line from this file:

```
user.admin 0
```

By default, when you issue the `supper update` command, you will see the following messages:

```
# /var/sysman/supper update
Updating collection node.root from server sp3n01en1.msc.itso.ibm.com.
File Changes: 0 updated, 0 removed, 0 errors.
Updating collection sup.admin from server sp3n01en1.msc.itso.ibm.com.
File Changes: 0 updated, 0 removed, 0 errors.
Updating collection user.admin from server sp3n01en1.msc.itso.ibm.com.
File Changes: 8 updated, 0 removed, 0 errors.
#
```

If you modify `/var/sysman/sup/.resident` file, you will see the following messages instead:

```
# /var/sysman/supper update
Updating collection node.root from server sp3n01en1.msc.itso.ibm.com.
File Changes: 0 updated, 0 removed, 0 errors.
Updating collection sup.admin from server sp3n01en1.msc.itso.ibm.com.
File Changes: 0 updated, 0 removed, 0 errors.
#
```

The `user.admin` file collection was not updated. Now, you can add/delete/change users, groups, or passwords without being under the control of SP User Management.

Be careful of the following points when you manage security information this way:

- A user home directory will be created under `/home` directory instead of `/home/CWS_hostname`. This means that the home directory will not be managed by automounter. If you want use automounter for managing the home directory, you need to set up automounter manually.
- The `passwd` command in `/usr/bin` directory is linked to the `sp_passwd` command in the `/usr/lpp/ssp/config` directory. The standard AIX `passwd` command was renamed to `passwd.aix`. To change the password, issue the `passwd.aix` command instead of the `passwd` command. Alternatively, you can rename the name of the commands.

14.2.6 Stop using SP User Management

You can stop using the SP User Management function to manage users node by node. To do this, you need to change the `usermgmt_config` attribute of SP SDR class as shown in Figure 89 on page 391.

To stop using SP User Management, issue the `smitty site_env_dialog fast` path and set the User Administration Interface field to false:

```
Site Environment Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                     [Entry Fields]
Default Network Install Image           [bos.obj.ssp.432]
Remove Install Image after Installs     false +

NTP Installation                         consensus +
NTP Server Hostname(s)                  ["" ] +
NTP Version                              3 +

Automounter Configuration               true +

Print Management Configuration          false +
Print system secure mode login name     ["" ]

User Administration Interface           false +
[MORE...15]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Alternatively, issue the `spsitenv` command:

```
# spsitenv usermgmt_config=false
```

This command issues the `services_config` script on all nodes. The `supper update` command is invoked from the script; therefore, all nodes get a new `user.admin` file collection that does not include `/etc/passwd`, `/etc/security/passwd`, and so on. Now, you can manage users on a node by node basis.

14.2.7 Using Network Information Service

The user space provided by SP User Management is effective only inside the SP system. If you want to include machines outside of the SP system, you may need to use the Network Information Service (NIS).

The following are prerequisites for using NIS:

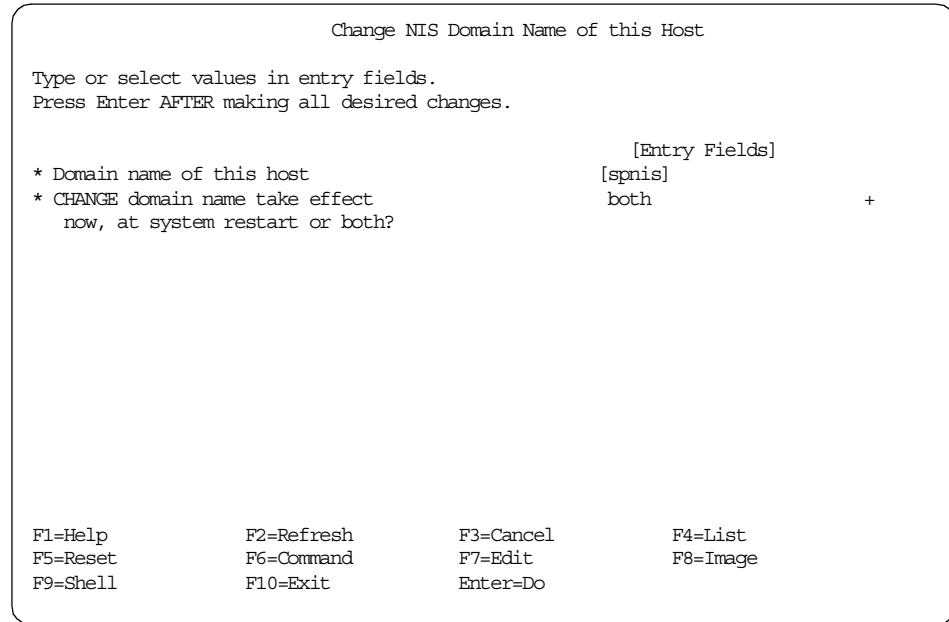
- SP User Management must be inactivated. To do this, refer to 14.2.6, “Stop using SP User Management” on page 404.
- The network routing to the NIS master or slave server machine should be established on the whole SP system.

- The `bos.net.nis.server` file set must be installed in the NIS server machine, and the `bos.net.nis.client` file set must be installed in the NIS client machine.

To configure NIS on the SP system, perform the following steps:

Step 1: Set NIS domain name

To set the NIS domain name, issue the `smitty chypdom` fast path. For example, set the domain name as `spnis`:



Alternatively, issue the `chypdom` command:

```
# /usr/sbin/chypdom -B spnis
```

The `-B` flag indicates that the domain name should be changed now, and the `/etc/rc.nfs` file should be updated to reflect the change.

Issue the command on all machines that uses NIS.

Step 2: Configure NIS master server

To configure NIS master server, issue the `smitty mkmaster` fast path on the machine that will be NIS master:

```

                                Configure this Host as a NIS Master Server

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
HOSTS that will be slave servers [slavesrv]
* Can existing MAPS for the domain be overwritten? yes +
* EXIT on errors, when creating master server? yes +
* START the yppasswdd daemon? yes +
* START the ypubdated daemon? yes +
* START the ypbind daemon? yes +
* START the master server now, both +
  at system restart, or both?

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell     F10=Exit      Enter=Do

```

Alternatively, issue the `mkmaster` command on the machine that will be NIS master:

```
# /usr/sbin/mkmaster -s slavesrv -O -E -P -U -C -B
```

One or more slave servers should be specified in the HOSTS that will be the slave servers field. If NIS client (the ypbind daemon) could not access server or slave server (the ypserv daemon), the system is hung up.

Step 3: Configure NIS slave server

The NIS master machine is only one in the domain. On the other hand, the NIS slave machine can be one or more. Having no slave server configuration is also possible, but it is not recommended because if master is down, no one can log in to the system.

To configure the NIS slave server, issue the `smitty mkslave` fast path on each slave machine:

```

Configure this Host as a NIS Slave Server

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* HOSTNAME of the master server      [mastersrv]      +
* Can existing MAPS for the domain be overwritten?  yes            +
* START the slave server now,        both            +
  at system restart, or both?
* Quit if errors are encountered?    yes            +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Alternatively, issue the `mkslave` command on each slave machine:

```
# /usr/sbin/mkslave -O -B -c sp4en0
```

Step 4: Configure NIS client

To use NIS, the `ybind` daemon must run on all the NIS machines including the master server, slave server, and client. If the `ybind` daemon is not running, they can not get information from the `ypserv` daemon even if it is running locally.

To configure the NIS client, issue the `smitty mkclient` fast path on all the machines:

```

                                Configure this Host as a NIS Client

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
* START the NIS client now,                                both                +
  at system restart, or both?
NIS server - required if there are no NIS servers on this subnet  []                +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Alternatively, issue the `mkclient` command on all the machines:

```
# /usr/sbin/mkclient -B
```

Under the NIS environment, on all client machines, the `/etc/passwd` file should have the following line called *escape sequence*:

```
+:0:0:::
```

For all user entries above the escape sequence, the user password information is not inquired to the NIS master server or slave server. The local password file is referred to. For the user entries below the escape sequence, the user password information is gotten from the NIS master server or slave server.

Users can change their password on any client by using the `yppasswd` command. Adding/deleting user is performed on master server only. Once you have changed `/etc/passwd`, `/etc/group`, and so on, on the master server, issue the `smitty mkmaps` fast path to recreate the map and transfer them to slave server:

```
Build / Rebuild NIS Maps for this Master Server

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

* MAPS that are to be built                                [Entry Fields]
                                                           [all]

F1=Help           F2=Refresh           F3=Cancel           F4=List
F5=Reset          F6=Command           F7=Edit            F8=Image
F9=Shell         F10=Exit            Enter=Do
```

Alternatively, issue the `make` command:

```
# cd /etc/yp
# ./make all
```

14.3 Time synchronization

Kerberos or RS/6000 Cluster Technology (RSCT) require that all the system clocks in the SP system are synchronized. For example, if there is more than a five minute time shift between CWS and a node, Kerberos does not allow you to issue a remote command.

Network Time Protocol (NTP) is provided to synchronize system time over all the SP system. Though it is originally a Public Domain Software (PDS), it is included in AIX as a standard protocol. PSSP uses its own NTP (`/usr/lpp/ssp/bin/xntpd`) instead of the standard AIX NTP (`/usr/sbin/xntpd`). However, PSSP NTP can co-work withan AIX NTP that is running on regular RS/6000 outside of the SP system.

For further reading, refer to Chapter 7, “Managing Time Synchronization” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

To get documentation updates and information about NTP, refer to the `/usr/lpp/ssp/README/ssp.public.README` file. There is NTP source code available on the CWS. Refer to 2.5.2, “Where is the source code?” on page 126, for more detail.

14.3.1 How does it work?

With the default setting, CWS becomes a *time server* for all other nodes. This means that the system time of all nodes are synchronized to the CWS. When a node boots, the `setclock` command is issued from the `rc.sp` script to set its time to the same time as the CWS. The `rc.sp` script issues the `/usr/lpp/ssp/install/bin/ntp_config` script indirectly. The `ntp_config` script creates `/etc/ntp.conf` file by using information in SDR. The following is a sample of the `/etc/ntp.conf` file:

```
# cat /etc/ntp.conf
#
#5799-FBX (C) Copyright IBM Corporation 1993
#Licensed Materials - Property of IBM
#All rights reserved.
#US Government Users Restricted Rights -
#Use, duplication or disclosure restricted by
#GSA ADP Schedule Contract with IBM Corp.
#
#"@(#)86 1.1 src/ssp/config/ntp.conf.base, sysman, ssp_rtro, rtrot3dg 1/27/"
#
# call this ntp.conf on 6000s

#
# Drift file. Put this in a directory which the daemon can write to.
# No symbolic links allowed, either, since the daemon updates the file
# by creating a temporary in the same directory and then rename()'ing
# it to the file.
#
driftfile /etc/ntp.drift

#
# Authentication delay. If you use, or plan to use someday, the
# authentication facility you should make the programs in the auth_stuff
# directory and figure out what this number should be on your machine.
#
authdelay 0.000047

#
# The following entries were added by ntp_config during sp
# configuration. They are (generally) the ip addresses of
# the time servers for the system and any peers that may be
# present as well... The configuration will differ based on
# the environment variable NTP_TYPE (set in the CMI).
#
server 127.127.1.10 version 3
#
```

The `/etc/ntp.conf` file contains the time server's IP address. Based on this file, the `xntpd` daemon starts time synchronization. It adjusts time using the following rules:

- If the time difference between the server and client is bigger than 1,000 seconds, it abandons synchronization.
- If the difference is less than or equal to 128 milliseconds, the `xntpd` daemon adjusts the client's time to the server. Adjust means that adding or subtracting a fixed amount at each timer interrupt for a calculated number of ticks. By this method, time shift does not occur on the client.
- If the difference is more than 128 milliseconds and less than or equal to 1,000 seconds, about 900 seconds after, system time on the client is shifted to the server time.

If the time shifted to earlier, Topology Services sends a heartbeat with the earlier time stamp. Therefore, the nodes are determined as down. Some applications that use RSCT, such as IBM Recoverable Virtual Shared Disk, might be affected by this. You should not change the system time on the time server without the nodes shut down.

14.3.2 Getting information

By default, NTP is configured as consensus mode. This means the CWS is a time server. This information is kept as an `ntp_config` attribute in SP SDR class. You can get this kind of information by issuing the `sp1stdata` command:

```

# splstdata -e
      List Site Environment Database Information

attribute          value
-----
control_workstation  sp4en0
cw_ipaddrs          9.12.1.140:192.168.4.130:
install_image       bos.obj.ssp.432
remove_image        false
primary_node        1
ntp_config          consensus
ntp_server          ""
ntp_version         3
amd_config          true
print_config        false
print_id            ""
usermgmt_config     true
passwd_file         /etc/passwd
passwd_file_loc     sp4en0
homedir_server      sp4en0
homedir_path        /home/sp4en0
filecoll_config     true
supman uid          102
supfilesrv_port     8431
spacct_enable       false
spacct_actnode_thresh  80
spacct_exclude_enable  false
acct_master         0
cw_has_usr_clients  false
code_version        PSSP-3.1
layout_dir          ""
authent_server      ssp
backup_cw           ""
ipaddrs_bucw       ""
active_cw           ""
sec_master          ""
ods_server          ""
cell_name           ""
cw_lppsource_name   aix432
cw_dcehostname      ""
#

```

Time synchronization is controlled by the following attributes:

- | | |
|-------------------|--|
| ntp_config | This attribute indicates the NTP installation mode: |
| none | Do not use NTP on the SP. |
| consensus | Set up NTP to configure CWS as the NTP server and BISs as NTP peers. |
| timemaster | Site has an existing NTP server. Configure NTP to use this. The ntp_server attribute contains the NTP server hostname. |

internet	The CWS has access to the Internet. Configure the CWS to be an NTP server using the Internet time server defined in <code>ntp_server</code> attribute.
ntp_server	This attribute indicates the hostname of NTP server.
ntp_version	This attribute indicates the NTP version number.

14.3.3 Changing NTP time server

You can specify the machine outside of SP system as a NTP time server. You need to set TCP/IP routing between the server and all nodes in advance.

To change the `ntp_config` attribute, issue the `smitty site_env_dialog fast` path:

```

Site Environment Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                     [Entry Fields]
Default Network Install Image             [bos.obj.ssp.432]
Remove Install Image after Installs       false +
                                           +
NTP Installation                           timemaster +
NTP Server Hostname(s)                     [ntpserver]
NTP Version                               3 +
                                           +
Automounter Configuration                  true +
                                           +
Print Management Configuration             false +
Print system secure mode login name       [""]
                                           +
User Administration Interface              true +
[MORE...15]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do

```

Figure 91. Change NTP installation mode (`smitty site_env_dialog`)

To use your site's existing NTP time server to synchronize the SP system clocks, put `timemaster` in the NTP Installation field and the host name in the NTP Server Hostname(s) field.

Alternatively, issue the `spsitenv` command:

```
# spsitenv ntp_config=timemaster ntp_server=ntpserver
```

This command updates the `/etc/ntp.conf` file on each node based on the new setting but does not refresh the `xntpd` daemon. To refresh the daemon, you need to kill the `xntpd` daemon on CWS and all nodes and restart it. The following is a sample script showing how to do this. Put it on CWS and all the nodes to be issued:

```
#!/bin/ksh
#
# refreshntp.sh
#
# USAGE="Usage: refreshntp.sh"
#
PID=$(ps -ef | grep /usr/lpp/ssp/bin/xntpd | egrep -v grep | awk '{print $2}')
kill $PID
/etc/rc.ntp
```

14.3.4 Monitoring NTP

You can monitor current NTP status by using the `xntpd` command. The following sample shows the status of the node `sp4n01`:

```
# dsh -w sp4n01 xntpd -c help
sp4n01: Commands available:
sp4n01: addpeer      addrefclock  addserver    addtrap      authinfo
sp4n01: broadcast   clkbug       clockstat    clrtrap      controlkey
sp4n01: ctlstats    debug        delay        delrestrict  disable
sp4n01: dmpeers     enable       fudge        help          host
sp4n01: hostnames   iostats     kerninfo     keyid         keytype
sp4n01: leapinfo    listpeers   loopinfo     memstats     monitor
sp4n01: monlist     passwd      peers        preset       pstats
sp4n01: quit        readkeys    requestkey   reset        reslist
sp4n01: restrict    setprecision showpeer     sysinfo     sysstats
sp4n01: timeout     timerstats  traps        trustkey     unconfig
sp4n01: unrestrict  untrustkey  version
# dsh -w sp4n01 xntpd -c peers
sp4n01:      remote      local      st poll reach  delay  offset  disp
sp4n01: =====
sp4n01: *sp4en0      192.168.4.1  11  64  377  0.00206  0.001825  0.00009
#
```

Using the `-c help` subcommand shows all available subcommands. Not all of the subcommands are useful because the meaning of the output is not well documented. You can type the `-p` flag (`xntpd -p`) instead of the `-c peers` subcommand (`xntpd -c peers`) for the `xntpd` command.

From the output, you can verify if the client is currently synchronizing to the server. The character in the left margin indicates the mode this peer entry is operating in:

- + This character denotes symmetric active.
- | This character indicates symmetric passive.
- = This character means the remote server is being polled in client mode.
- ^ This character indicates that the server is broadcasting to this address.
- ~ This character denotes that the remote peer is sending broadcasts.
- * This character marks the peer the server is currently synchronizing to.

Therefore, if there is *, the node is currently synchronizing to the server.

To know about the `xntpd` command, issue the `man` command. If the `man` command is not installed properly, refer to 15.1, “Man page” on page 429.

14.3.5 Changing system time

You may have a situation that you need to change system time, for example, the test for the year 2000 issue. You can set the system clock of SP system forward beyond year 2000 and back to the current time. But, you can not change the time on just some of the nodes to do the testing. The reason behind this is that many of the subsystems require a single clock value, such as Kerberos and the Network File System (NFS). If some nodes are at year 2000, and others are not, the Time of Day Clocks will not be the same. This will result in commands, such as `rsh`, `dsh`, or `make` failing. Nor can you do testing on a partition because time services are not partition sensitive. The only supported method to do your year 2000 testing is on a separate SP system with its own CWS.

The following is the procedure for changing your date for year 2000 testing on an SP system and then resetting your date back to the present day. Restoring your system back to the present is not as simple as just changing the date back. There will be files with the wrong date on them, and this could lead to unreliable behavior, and hence the proper method is to restore from the system from a backup.

To change the system time, perform the following steps:

Step 1: Make a mksysb image of your nodes

Depending on what is running on your nodes, you may need to make a separate image for each node, or use the same image. Save these images on the CWS.

Step 2: Make a mksysb image of your CWS

Create a mksysb image of your CWS and make sure that you have a good backup of the /spdata file system as well. If you have /spdata in its own volume group, issue the `savevg` command on it. Your /spdata file system should contain your good mksysb images for your nodes.

Step 3: Shut down all nodes

Avoid the Topology Services time stamp problem, described in 14.3.1, “How does it work?” on page 411. It is recommended to shut down all nodes. For example, issue the `cshutdn` command:

```
# cshutdn -F all
```

Step 4: Change the system time on time server

You can change system time on CWS by the `date` command. For example, to set it as 31 Dec 1999 11h 59m 00s:

```
# date 12311159.0099
```

If CWS is not the time server, synchronize its time to timeserver. The `setclock` command can be used:

```
# setclock timeserver_hostname
```

Step 5: Re-issue k4init

The Kerberos authentication ticket should be refreshed. To make it sure that it has not expired, reissue the `k4init` command:

```
# k4init root.admin
```

Step 6: Reboot the nodes

Then time synchronization starts with a new time setting.

After testing, you need to reset the system time as original. To do this, perform the follow steps:

Step 1: Shut down all the nodes

To do this, issue the `cshutdn` command:

```
# cshutdn -F ALL
```

Step 2: Shutdown the CWS

To do this, issue the `shutdown` command:

```
# shutdown -F
```

Step 3: Restore CWS from mksysb

Put the mksysb tape in the tape drive and reboot your CWS. Follow your normal mksysb restore procedures. Set the proper date and time on your CWS. If your /spdata file system is on its own volume group, also restore this as well.

Step 4: Set nodes to install

You can either issue the `smitty server_dialog` fast path:

```
# smitty server_dialog
```

Alternatively, you can issue the `spbootins` command. For example:

```
# spbootins -p PSSP-3.1 -r install -v AIX432 -i bos.obj.ssp.432 -l 1,2
```

This will install the mksysb image `bos.obj.ssp.432` onto nodes 1 and 2.

Step 5: Verify installation settings

Make sure that correct settings are in the SDR for the installation. You can issue the `splstdata -G -b` command to see what they are currently set to.

Step 6: Network boot the node

Network boot the nodes that you are installing with the `nodecond` command or by using Hardware Perspective. When the install is finished, verify that the `bootp_response` has been set back to disk by issuing the `splstdata` command.

For further information about Year 2000 testing on an SP system, please refer to the following web site:

<http://www.software.ibm.com/year2000/papers/aixy2k.html>

AIX and PSSP are Year 2000 compliant products; however, some versions of AIX require PTFs. The Web site includes this information also.

14.4 The Automounter

The basic function of automounter is merely to do NFS mounts dynamically. When you access a file or directory under automounter control, it transparently mounts the required file system. When there has been no activity to that file system for some pre-determined amount of time, the automounter unmounts the file system.

PSSP uses AIX automounter for home directory management. When a user is added to SP system by issuing the `spmuser` command, the entry for the user is automatically added to automounter configuration file. This configuration file will be put on nodes automatically in one hour, by the file

collection technology. For more details about SP User Management, refer to 14.2, “SP User Management” on page 390. For more details about file collection technology, refer to 14.1, “File collection technology” on page 379.

For further reading, refer to Chapter 8, “Managing the Automounter” in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

14.4.1 Getting information

The current setting of the automounter is stored in the SP SDR class. The `splstdata` command shows the attributes:

```
# splstdata -e
      List Site Environment Database Information

attribute          value
-----
control_workstation  sp4en0
cw_ipaddrs          9.12.1.140:192.168.4.130:
install_image       bos.obj.ssp.432
remove_image        false
primary_node         1
ntp_config           consensus
ntp_server           ""
ntp_version          3
amd_config          true
print_config         false
print_id             ""
usemgmt_config       true
passwd_file          /etc/passwd
passwd_file_loc      sp4en0
homedir_server     sp4en0
homedir_path       /home/sp4en0
filecoll_config      true
supman_uid           102
supfilesrv_port      8431
spacct_enable        false
spacct_actnode_thresh 80
spacct_exclude_enable false
acct_master           0
cw_has_usr_clients   false
code_version         PSSP-3.1
layout_dir           ""
authent_server        ssp
backup_cw             ""
ipaddrs_bucw         ""
active_cw            ""
sec_master            ""
ods_server           ""
cell_name            ""
cw_lppsource_name    aix432
cw_dcehostname        ""
#
```

Automounter is controlled by the following attribute:

amd_config This attribute indicates true or false as to whether the SP provides automounter support.

Automounter uses the following attributes:

homedir_server This attribute indicates the host name of the default user directory server. By default, automounter uses CWS as the home directory server.

homedir_path This attribute indicates the default path to user home directories. By default, automounter uses `/home/$homedir_server` for the home directory path.

By default, `amd_config` attribute is set true and configured to the SP system.

14.4.2 Changing home directory server and path

The CWS is a home directory server by default settings. However, you may need to change the home directory server to something other than CWS. For example, in the case your SP system has many users, and they access their home directory very frequently, it uses a lot of resources of CWS. To avoid decreasing the CWS availability, CWS should be used for SP system management purposes only.

In the case you want to use the node `sp3n01` as a home directory server, and its `/home/filesvr` directory as a home directory path, issue the `smitty site_env_dialog fast path`:

```

Site Environment Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[MORE...6]                                     [Entry Fields]

Automounter Configuration                       true          +
Print Management Configuration                 false         +
Print system secure mode login name           [""]
User Administration Interface                 true          +
Password File Server Hostname                 [sp3en0]
Password File                                 [/etc/passwd]
Home Directory Server Hostname                [sp3n01]
Home Directory Path                           [/home/filesvr]

File Collection Management                     true          +
[MORE...9]

F1=Help           F2=Refresh       F3=Cancel        F4=List
F5=Reset          F6=Command       F7=Edit          F8=Image
F9=Shell          F10=Exit         Enter=Do

```

Specify sp3n01 in the Home Directory Server Hostname field and /home/filesvr in the Home Directory Path field.

Alternatively, issue the `spsitenv` command:

```
# spsitenv homedir_server=sp3n01 homedir_path=/home/filesvr
```

The home directory, /home/filesvr, must be created on the home directory server sp3n01. The home directory server must NFS-export the home directory path by issuing the `mknfsexp` command:

```
# /usr/sbin/mknfsexp -d /home/filesvr -t rw -B
```

When a new user is added by the `spmuser` command, the new entry will be created in the /etc/auto/maps/auto.u file. For example, if you add a user user1, the following entry will be added in the auto.u file:

```
user1      sp3n01:/home/filesvr:&
```

This file will be distributed to all the nodes by file collection technology. You can find this file name in /var/sysman/sup/lists/user.admin file (refer to Figure 88 on page 385). The automountd daemon reads this file and knows which directory should be mounted from which machine. On SP, automounter manages /u directory. Therefore, this entry means that if user1 touch the

/u/user1 directory, the automountd daemon mounts /home/filesvr/user1 on /u/user1 from the node sp3n01.

Attention

- If your SP system is used as a server consolidated system, in other words, there is not so many login users who work on their home directory, the default home directory server settings does not cause any problem.
- You can have multiple home directory servers in your SP system. So, even if you change home directory server and path as described in this section, existing users' home directories continuously use the CWS. Newly created users' home directories will use the node sp3n01.

14.4.3 Using an SP switch for Automounter

In the case you assign a home directory server to one of the nodes, you can use the SP switch for automounter. To do this issue the `smitty site_env_dialog` fast path:

```
Site Environment Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[MORE...6]                                     [Entry Fields]

Automounter Configuration                       true          +
Print Management Configuration                  false         +
Print system secure mode login name            [""]
User Administration Interface                  true          +
Password File Server Hostname                  [sp3en0]
Password File                                  [/etc/passwd]
Home Directory Server Hostname                 [sp3sw01]
Home Directory Path                            [/home/filesvr]

File Collection Management                       true          +
[MORE...9]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do
```

Specify the SP switch interface name, in this example sp3sw01, in the Home Directory Server Hostname field.

Alternatively, issue the `spsitenv` command:

```
# spsitenv homedir_server=sp3cw01
```

With default settings, CWS does not have a network route to SP Switch. To add network route to the SP Switch, issue the `route` command on the CWS. For example, if the node IP address of SP Ethernet is 192.168.4.9 and SP Switch is 192.168.14.9:

```
# route add -host 192.168.14.9 192.168.4.9
```

If there is a router node between the CWS and the node, you need to specify the router node IP address instead of the SP Ethernet.

On the node, you need to change the `ipforwarding` variable to 1. To do this, issue the `no` command:

```
# no ipforwarding = 1
```

14.4.4 Stop using Automounter

To stop using the automounter, issue the `smitty site_env_dialog fast path`:

Site Environment Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[MORE...6]	[Entry Fields]	
Automounter Configuration	true	+
Print Management Configuration	false	+
Print system secure mode login name	[""]	
Automounter Configuration	false	+
Password File Server Hostname	[sp3en0]	
Password File	[/etc/passwd]	
Home Directory Server Hostname	[sp3en0]	
Home Directory Path	[/home/sp3en0]	
File Collection Management	true	+
[MORE...9]		

F1=Help F2=Refresh F3=Cancel F4=List
F5=Reset F6=Command F7=Edit F8=Image
F9=Shell F10=Exit Enter=Do

Specify false in the Automounter Configuration field.

Alternatively, issue the `spsitenv` command:

```
# spsitenv amd_config=false
```

Issuing the `stopsrc -s automountd` command is not enough to release the `/u` directory from automounter; so, you need to reboot the whole SP system.

To change the SP system to a simple NFS server/client system, you need to NFS-mount the home directory on the client machine. To do this, issue the `smitty mknfsmnt` fast path:

```

                                Add a File System for Mounting

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
PATHNAME of mount point                [/u]                                /
PATHNAME of remote directory            [/home/filesvr]
HOST where remote directory resides     [sp3en0]
Mount type NAME                          []
Use SECURE mount option?                 no                                  +
MOUNT now, add entry to /etc/filesystems or both? both                      +
/etc/filesystems entry will mount the directory no                          +
on system RESTART.
MODE for this NFS file system            read-write                          +
ATTEMPT mount in foreground or background background                       +
NUMBER of times to attempt mount         []                                  #
Buffer SIZE for read                     []                                  #
Buffer SIZE for writes                   []                                  #
[MORE...26]

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit       F8=Image
F9=Shell     F10=Exit       Enter=Do
```

Specify `/u` in the `PATHNAME` of mount point field. Specify the previous home directory path name, in this example `/home/filesvr`, in the `PATHNAME` of remote directory field, and specify the previous home directory server name, in this case `sp3en0`, in the `HOST` where remote directory resides field.

On the server machine, users are going to use `/u` directory for their home directory instead of `/home/filesvr` directory. So, you need to make a symbolic link. To do this, rename the `/u` directory by issuing the `mv` command:

```
# mv /u /u.back
```

Then issue the `ln` command:

```
# ln -s /home/filesvr /u
```

14.4.5 Checking logs

The automountd daemon keeps its log in the `/var/adm/SPlogs/auto/auto.log` file. All the output written to stdout or stderr by the daemon is recorded.

If there are some problems, you should check the status of the automountd daemon and the `/etc/auto/amd.map/auto.u` file on the client machine. On the server machine, check the NFS-exported file system.

14.5 SP Accounting

SP accounting is similar to AIX accounting, only its functionality has been increased to handle groups of nodes, and it allows exclusive use of resource accounting. With SP accounting enabled, you can collect various records on users usage and work loads. If you want to charge a usage fee to SP users, this section explains how you can collect the information that you need.

For further reading, refer to Chapter 10, “Accounting” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

14.5.1 Before using SP accounting

The AIX system accounting provides a function that records daily system resource usage per user. SP accounting uses this function; so, you need to install `bos.acct` file set on nodes in advance.

Attention

System accounting requires a large amount of space in the `/var` file system to keep its data. When `/var` becomes full, it can cause problems to AIX or PSSP. Keep 20-30 MB free space for the file system at anytime. It keeps the SP system healthy.

14.5.2 Getting information

The current setting of the SP accounting is stored in the SP SDR class. To show the attributes, issue the `splstdata` command:

```

# splstdata -e
      List Site Environment Database Information

attribute          value
-----
control_workstation  sp4en0
cw_ipaddrs          9.12.1.140:192.168.4.130:
install_image       bos.obj.ssp.432
remove_image        false
primary_node        1
ntp_config           consensus
ntp_server           ""
ntp_version          3
amd_config           true
print_config         false
print_id             ""
usermgmt_config      true
passwd_file          /etc/passwd
passwd_file_loc      sp4en0
homedir_server       sp4en0
homedir_path         /home/sp4en0
filecoll_config      true
supman_uid           102
supfilesrv_port      8431
spacct_enable       false
spacct_actnode_thresh  80
spacct_exclude_enable  false
acct_master         0
cw_has_usr_clients   false
code_version         PSSP-3.1
layout_dir           ""
authent_server       spp
backup_cw            ""
ipaddrs_bucw         ""
active_cw            ""
sec_master           ""
ods_server           ""
cell_name            ""
cw_lppsource_name    aix432
cw_dcehostname       ""
#

```

SP accounting is controlled by the following attributes:

- spacct_enable** This attribute indicates whether, by default, accounting is enabled on all nodes that have an accounting enabled attribute of default.

- spacct_actnode_thresh** This attribute indicates the percentage of nodes for which accounting data must be available for merging and reporting of the data for a cycle to take place.

spacct_exexcluse_enable	This attribute indicates whether accounting start and end job records will be generated for jobs having exclusive use of the node.
acct_master	This attribute indicates which node is to act as the accounting master.

By default, SP accounting is not installed by PSSP.

14.5.3 Enabling SP accounting

To enable SP accounting, issue the `smitty site_env_dialog` fast path:

```

Site Environment Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[MORE...15]                                     [Entry Fields]
Home Directory Server Hostname                 [sp3n01]
Home Directory Path                             [/home/filesvr]

File Collection Management                       true +
File Collection daemon uid                       [102] #
File Collection daemon port                     [8431] #

SP Accounting Enabled                           true +
SP Accounting Active Node Threshold             [80] #
SP Exclusive Use Accounting Enabled             false +
Accounting Master                               [0]

Control Workstation LPP Source Name             [aix432]
[BOTTOM]

F1=Help           F2=Refresh           F3=Cancel           F4=List
F5=Reset          F6=Command           F7=Edit            F8=Image
F9=Shell          F10=Exit              Enter=Do

```

Specify `true` in the SP Accounting Enabled field.

Alternatively, issue the `spsitenv` command:

```
# spsitenv sp_acct_enable=true
```

The command issues the `services_config` command on CWS and all nodes. The script starts AIX system accounting on each node. This function is a kernel service; so, there is no particular procedure to start accounting. The accounting information is stored in `/var/adm/pacct` file. Its size is increased every time a process terminates.

Note

If you want to stop AIX system accounting manually, issue the `shutacct` command on the node:

```
# /usr/sbin/acct/shutacct
```

To restart it, issue the `startup` command:

```
# /usr/sbin/acct/startup
```

The `services_config` command also sets the accounting command in root's crontab file on the CWS and all the nodes to create a daily summary. The script also exports the `/var/adm/acct` directory of the node for NFS mount. This directory keeps the accounting information of the nodes. SP accounting master, it is CWS by default, mounts the directory and makes a whole system accounting report.

The `nrunacct` command is issued on nodes from root's crontab from Monday to Friday at 2:00 AM:

```
/usr/lpp/ssp/bin/nrunacct 2>/var/adm/acct/nite/accterr
```

Then, the `crunacct` command is issued on CWS from root's crontab from Monday to Friday at 4:00 AM:

```
/usr/lpp/ssp/bin/crunacct 2>/var/adm/cacct/nite/accterr
```

By these commands, the accounting reports on each node are merged into `/var/adm/cacct/sum/rprtYYYYMMDD` file on CWS.

The accounting data on nodes are erased daily. But, it is accumulated on CWS to generate a monthly summary report.

The `cmonacct` command is issued at the first day of the month, 5:15 AM:

```
/usr/lpp/ssp/bin/cmonacct
```

You can change the execution time of these commands.

To maintain the crontabs file, refer to Chapter 11, "Maintaining the crontabs File" in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

Chapter 15. Online documentations

Due to the vast amount of commands, scripts, and configurations included in IBM Parallel System Support Programs for AIX (PSSP), there are online documentations available to help you out. This is very useful when you can not remember the syntax of a command or need to refer to a section in the IBM publication, and you do not have it in front of you. There are several different ways that you can access the online documentations.

Man pages can be accessed directly at the command line. AIX man pages must first be functioning, and then PSSP just includes its own references for the man pages. Information on how to install and use man pages is covered in this chapter.

Portable Document Format (PDF) files are available along with Adobe Acrobat Reader to view them. PSSP also includes these PDF files of its documentations so that they can be read through Adobe Acrobat Reader. All of these software are included with AIX or PSSP. An explanation is given on how to install and read these software.

HyperText Markup Language (HTML) files and Netscape Navigator are also included with AIX or PSSP. If you wish to browse through the documentations via Netscape Navigator, the instructions are covered in this chapter.

And finally, there is the SP Resource Center that is a Web-based interface that allows you to read the online documentations. Information on how to start up and use the SP Resource Center is covered.

15.1 Man page

Man pages are one of the easiest and quickest way to refresh yourself on the definition and syntax of a command. They provide detailed explanations on all the flags for a specific command and includes several examples of its usage. Another reason why man pages are so convenient is that run directly from the command line and do not require the start up of another application to view the information. This section discusses how to install man pages, how to add the SP man pages, and how to use them.

15.1.1 Installing an AIX man page

In order to have the SP man pages, you must first ensure that your AIX man pages are installed.

Here is a list of the man page file sets located on the AIX 4.3.2 Base Documentation, 5765-C34 CD-ROM:

- bos.html.en_US.nav
- bos.html.en_US.cmds.cmds1
- bos.html.en_US.cmds.cmds2
- bos.html.en_US.cmds.cmds3
- bos.html.en_US.cmds.cmds4
- bos.html.en_US.cmds.cmds5
- bos.html.en_US.cmds.cmds6

Once these file sets are installed, verify that your LANG variable is set correctly.

```
# echo $LANG
en_US
#
```

Now, the AIX man pages should work.

15.1.2 Installing an SP man page

To install the SP man pages, it is simply a matter of installing the file set. The SP man pages are included in the ssp.docs file set. To install the file set, use the `smitty install_latest` fast path. After specifying INPUT device / directory for software field, hit the **F4** key. You will have the following list:

```

Install and Update from LATEST Available Software

Ty+-----+
Pr|                SOFTWARE to install
|
| Move cursor to desired item and press F7. Use arrow keys to scroll.
* |     ONE OR MORE items can be selected.
* |     Press Enter AFTER making all selections.
|
| [MORE...37]
|   @ 3.1.0.0  SP PERL Distribution Package
|   @ 3.1.0.0  SP Problem Management
|   @ 3.1.0.0  SP Supervisor Microcode Package
|   @ 3.1.0.0  SP Sysctl Package
|   @ 3.1.0.0  SP System Monitor Graphical User Interface
|   @ 3.1.0.0  SP System Partitioning Aid
|   @ 3.1.0.0  SP System Support Package
|   + 3.1.0.0  SP man pages and PDF files and HTML files
| [MORE...29]
|
| F1=Help           F2=Refresh           F3=Cancel
F1| F7=Select       F8=Image           F10=Exit
Es| Enter=Do       /=Find             n=Find Next
F9+-----+

```

In this list, select **3.1.0.0 SP man pages and PDF files and HTML files**.

Installing this file set installs PDF files and HTML files also.

15.1.3 Using man pages

When you install a man page, it is placed under the `/usr/lpp/ssp/man` directory. There is a man page for Perl also available, and it is placed under the `/usr/lpp/ssp/perl5` and `/usr/lpp/ssp/perl5/lib` directory.

To use these man pages, their directory must be included in the `MANPATH` environment variable. Add the following entry in your `.profile` file:

```
MANPATH=$MANPATH:/usr/lpp/ssp/man:/usr/lpp/ssp/perl5:/usr/lpp/ssp/perl5/li
b
```

Then, you can use the `man` command. If you want to learn the usage of the `splstdata` command, issue the `man` command as follows:

```
# man splstdata
```

Then, you will see the following output:

splstdata

Purpose

splstdata - Displays configuration data from the System Data Repository (SDR) or system information for each node.

Syntax

```
splstdata  {-A | -n | -s | -b | -a | -u | -v | -h | -i | -d} [-G]
           [{start_frame start_slot {node_count | rest} |
            -N node_group | -l node_list}]
```

OR

```
splstdata  {-e | -f | -p}
```

Flags

One of the following flags must be specified with each invocation of
:

If you are interested in what kind of man pages are available, issue the `ls` command:

```
# ls /usr/lpp/ssp/man/*/*
```

This shows all the man pages.

15.2 PDF

PSSP includes PDF files that can be read online using Adobe Acrobat Reader. The Adobe Acrobat Reader is included with AIX, and the PDF files are included with PSSP. All of the different PSSP documentations are included in PDF format. This means that if you don't have a hard copy of a documentation, you can still access the information on it. This section explains how to install the PDF files, how to install Adobe Acrobat Reader, and how to read the PDF files.

Redbooks Online!

Many of the redbooks are now provided in Adobe PDF format. You can view the complete contents of our redbooks online. To download redbooks, visit our Web site:

<http://www.redbooks.ibm.com/>

Then, click the **Redbooks Online!** button.

15.2.1 Installing PDF files

PDF files are included in ssp.docs file set. To install PDF files, refer to 15.1.2, "Installing an SP man page" on page 430.

The following IBM publications are available in PDF format:

- *RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281
- *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347
- *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348
- *IBM Parallel System Support Programs for AIX: Managing Shared Disks*, SA22-7349
- *IBM Parallel System Support Programs for AIX: Performance Monitoring Guide and Reference*, SA22-7353
- *IBM Parallel System Support Programs for AIX: Diagnosis Guide*, GA22-7350
- *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351
- *IBM Parallel System Support Programs for AIX: Message Reference*, GA22-7352
- *RS/6000 Cluster Technology: Event Management Programming Guide and Reference*, SA22-7354
- *RS/6000 Cluster Technology: Group Service Programming Guide and Reference*, SA22-7355

15.2.2 Installing Adobe Acrobat Reader

There is only one file set needed for Adobe Acrobat Reader and this is included on AIX 4.3.2 Bonus Pack, 5765-C34 CD-ROM. The file set is:

- Adobe.acrobat 3.0.1.0

Install this file set following your normal software installation procedures.

15.2.3 Reading PDF files

When you install ssp.docs file set, the PDF files for PSSP 3.1 will be placed under the /usr/lpp/ssp/docs directory. The following figure shows the contents of this directory:

```
# cd /usr/lpp/ssp/docs
# ls -al
total 22968
drwxr-xr-x  2 bin      bin           512 Mar  1 10:05 .
drwxr-xr-x 43 bin      bin          4096 Mar 11 16:56 ..
-rwxr-xr-x  1 bin      bin        1789823 Oct 19 08:33 pssp_admin_v3r10.pdf
-rwxr-xr-x  1 bin      bin        1036721 Oct 19 08:33 pssp_commands_v1_v3r10.pdf
-rwxr-xr-x  1 bin      bin        1096538 Oct 19 08:33 pssp_commands_v2_v3r10.pdf
-rwxr-xr-x  1 bin      bin         806705 Oct 19 08:33 pssp_diag_v3r10.pdf
-rwxr-xr-x  1 bin      bin         868856 Oct 19 08:34 pssp_event_mgmt_v3r10.pdf
-rwxr-xr-x  1 bin      bin         746670 Oct 19 08:34 pssp_grp_svcs_v3r10.pdf
-rwxr-xr-x  1 bin      bin         906810 Oct 19 08:34 pssp_install_v3r10.pdf
-rwxr-xr-x  1 bin      bin        1237404 Oct 19 08:34 pssp_messages_v3r10.pdf
-rwxr-xr-x  1 bin      bin        1258923 Oct 19 08:34 pssp_perfmon_v3r10.pdf
-rwxr-xr-x  1 bin      bin         779267 Oct 19 08:34 pssp_shared_disks_v3r10.pdf
-rwxr-xr-x  1 bin      bin        1198850 Oct 19 08:34 pssp_sys_plan_v3r10.pdf
#
```

In this directory, there are 11 PDF files available. Each file includes one online documentation of PSSP 3.1. To read them, start the Adobe Acrobat Reader with the `acroread` command:

```
# acroread
```

Usually, Adobe Acrobat Reader is installed in the /usr/lpp/Acrobat3/bin directory. Then, open one of the PDF files. In the case of *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*, you will see the documentation shown in Figure 92 on page 435:

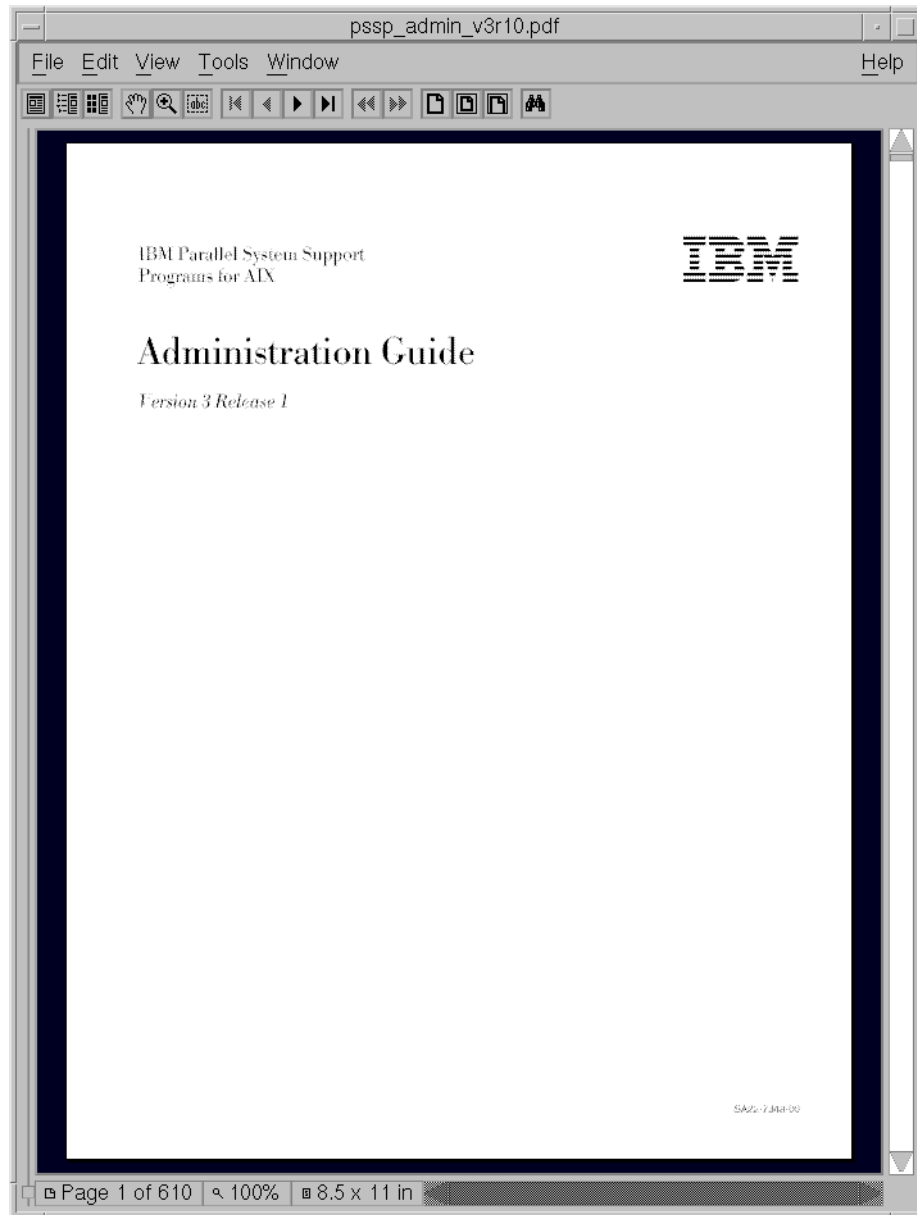


Figure 92. PSSP online documentations in PDF format

15.3 HTML files

PSSP includes HTML files as another way to view the PSSP documentations. AIX includes Netscape Navigator to view the HTML files. This is particularly handy for those users who use a Web browser continually. This section explains how to install the HTML files from the PSSP and how to view these files through Netscape Navigator.

15.3.1 Installing HTML files

HTML files are included in ssp.docs file set. To install HTML files, refer to 15.1.2, "Installing an SP man page" on page 430.

The following IBM publications are available in HTML format:

- *RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281
- *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347
- *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348
- *IBM Parallel System Support Programs for AIX: Managing Shared Disks*, SA22-7349
- *IBM Parallel System Support Programs for AIX: Performance Monitoring Guide and Reference*, SA22-7353
- *IBM Parallel System Support Programs for AIX: Diagnosis Guide*, GA22-7350
- *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351
- *IBM Parallel System Support Programs for AIX: Message Reference*, GA22-7352
- *RS/6000 Cluster Technology: Event Management Programming Guide and Reference*, SA22-7354
- *RS/6000 Cluster Technology: Group Service Programming Guide and Reference*, SA22-7355

15.3.2 Installing Netscape Navigator

Netscape Navigator is included in AIX 4.3.2 Bonus Pack, 5765-C34 CD-ROM. To install Netscape Navigator, you need the following file set installed:

- Netscape.communicator-us.rte 4.0.7.0

Install this file set using your normal software installation procedures.

15.3.3 Reading HTML files

When you install ssp.docs file set, HTML files for PSSP 3.1 will be placed under the /usr/lpp/ssp/html directory. The following figure shows the contents of this directory:

```
# cd /usr/lpp/ssp/html
# ls -al
total 112
drwxr-xr-x 13 bin      bin      512 Mar 01 10:05 .
drwxr-xr-x 43 bin      bin      4096 Mar 11 16:56 ..
drwxr-xr-x  3 bin      bin      1536 Feb 20 21:25 admin
drwxr-xr-x  3 bin      bin      1024 Feb 20 21:26 cmdsv1
drwxr-xr-x  3 bin      bin      1024 Feb 20 21:26 cmdsv2
drwxr-xr-x  3 bin      bin      1536 Feb 20 21:26 diag
drwxr-xr-x  3 bin      bin      512 Feb 20 21:26 evmgt
drwxr-xr-x  3 bin      bin      1024 Feb 20 21:26 grpsvcs
drwxr-xr-x  3 bin      bin      1024 Feb 20 21:26 instmig
drwxr-xr-x  3 bin      bin      1024 Feb 20 21:26 mngdisks
drwxr-xr-x  3 bin      bin      1536 Feb 20 21:26 msgs
drwxr-xr-x  3 bin      bin      1024 Mar 01 10:05 perfmon
drwxr-xr-x  3 bin      bin      1024 Feb 20 21:26 planv2
-rw-r--r--  1 bin      bin      1438 Oct 20 08:59 psspbooks.html
#
```

In this directory, there is a HTML file named psspbooks.html. This file is the main page, and it includes the links for all the PSSP 3.1 online documents. To read them, start the Netscape Navigator by issuing the `netscape` command:

```
# netscape
```

Usually, Netscape Navigator is installed in the /usr/netscape/navigator-us directory. Then, open the psspbooks.html file. You will see the index page shown in Figure 93 on page 438:

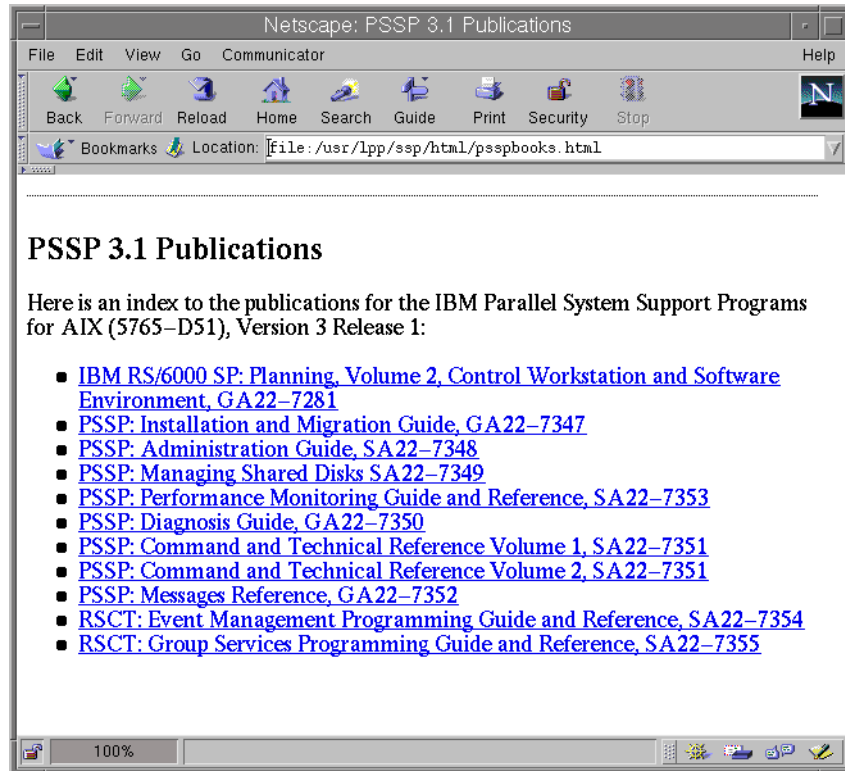


Figure 93. PSSP online documentations in HTML format

15.4 IBM RS/6000 SP Resource Center

RS/6000 SP Resource Center is another online documentation resource. RS/6000 SP Resource Center provides links to your local SP documentations and to other useful Web sites on the Internet. This section describes how you can install the RS/6000 SP Resource Center on your control workstation (CWS) and then provides how you can use the local online documentation.

15.4.1 Installing SP Resource Center

Before you utilize SP Resource Center as an online documents resource, you need to install the `man` command, man pages, Netscape Navigator, and HTML documents. Refer to 15.1, "Man page" on page 429, and 15.3, "HTML files" on page 436 to install them.

To use SP Resource Center, you need to install the `ssp.resctr.rte` file set, which is included with your PSSP.

15.4.2 Starting SP Resource Center

You can start SP Resource Center from SP Perspectives. To start the SP Perspectives, issue the `perspectives` command:

```
# perspectives
```

You will see the SP Perspectives Launch Pad window show in Figure 94.

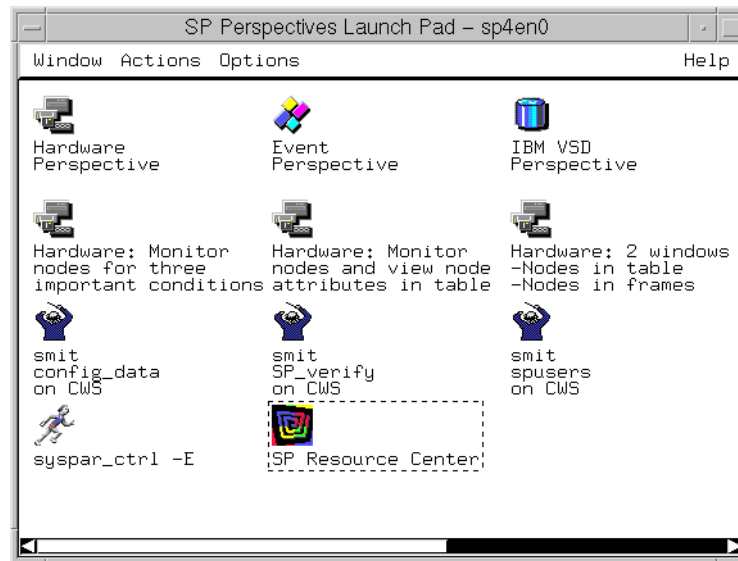


Figure 94. SP Perspectives Launch Pad

There is an icon for SP Resource Center in the SP Perspectives Launch Pad. Double click the **SP Resource Center** icon. You will see the RS/6000 SP Resource Center opening page shown in Figure 95 on page 440:



Figure 95. RS/6000 SP Resource Center

15.4.3 Reading online documentations

To read the PSSP online documentations, click the **Online Books** link in the index frame shown in Figure 95.

Then click the **PSSP** link in the index frame. You will have a list of PSSP online documentations as shown in Figure 96 on page 441.

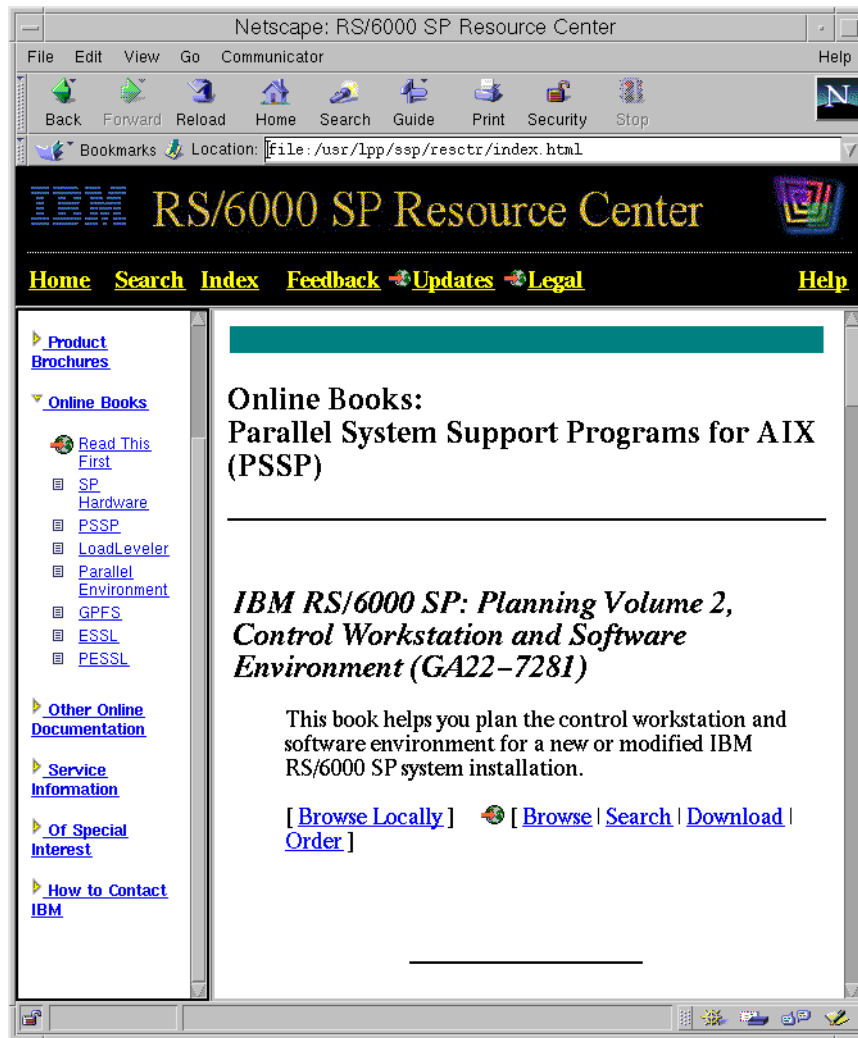


Figure 96. Reading PSSP online document

15.4.4 Using SP man pages

To use the SP man pages, click the **Other Online Documentation** link in the index frame shown in Figure 96. Then click the **Man Pages** link in the index frame. You will have a list of SP man pages as shown in Figure 97 on page 442.

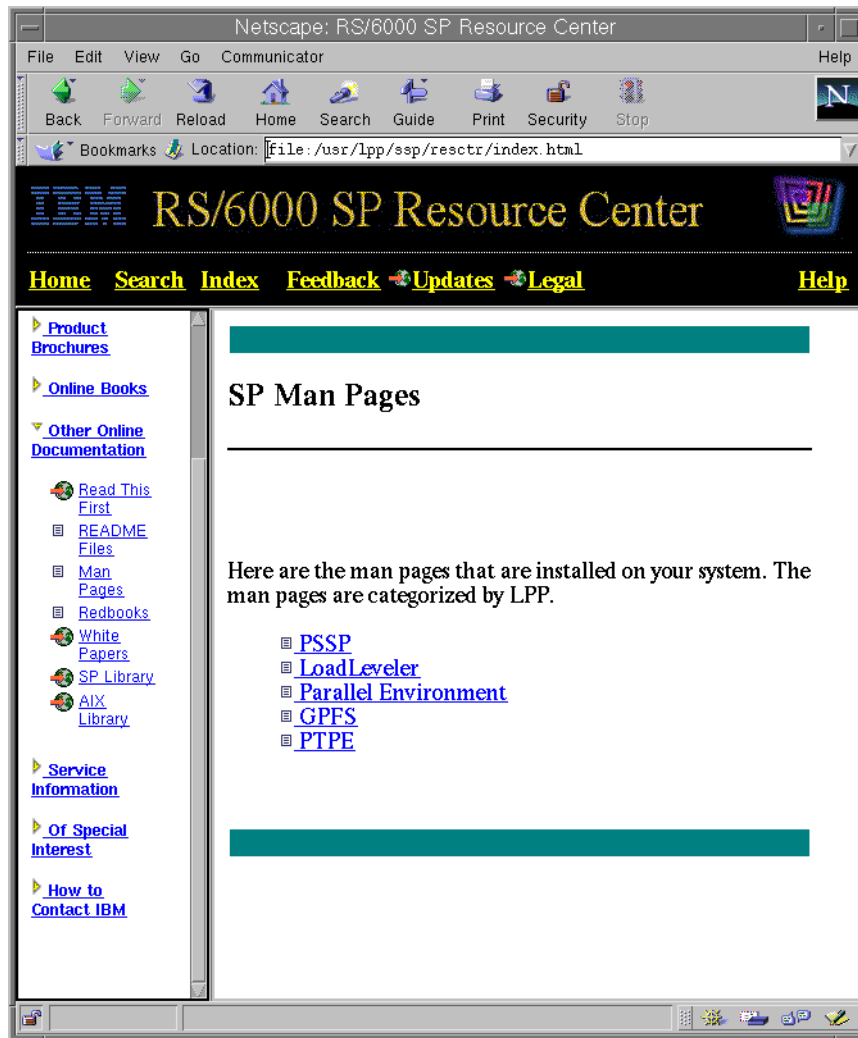


Figure 97. Reading SP man page

Chapter 16. Security

The fact that the SP system is used for scientific and business computing obviously demands a certain security level from the hardware and software components used by it. The SP system uses Kerberos to handle security from its control workstation (CWS) to its nodes. This chapter focuses on the software and provides security-related system management examples.

The different authentication methods are discussed in this chapter, explaining which methods you can use. Also, it explains how you can select and enable the authorization methods of your choice.

For Kerberos to run smoothly on your SP system, there must be several Kerberos daemons running. This chapter describes the responsibilities of these daemons and explains how to manage them.

An authentication database is the core of the Kerberos system. This chapter explains basic operations include initialize, destroy, back up, and restore an authentication database. In addition, it describes how to read the contents of an authentication database and how to change the Kerberos master key to improve the security of your SP system.

Another important aspect to Kerberos administration is the management of principals. This chapter focuses on the maintenance of Kerberos principals including how to check, add, delete, or modify them.

A service key file is used by service principals. If there is a problem with a service key, you can not have service from these principals. If this is the case, you may need to create or check it. To improve the security of your SP system, changing the service key occasionally is a plus.

To be authorized as AIX user and to be authorized as Kerberos principal requires different procedures. This chapter describes how to integrate these two different procedures together.

Finally, the chapter discusses the entire Kerberos system configuration. It covers how to configure, unconfigure, back up, and restore the whole Kerberos system configuration and also provides the unique configuration that integrates the external RS/6000 system to your SP system from the security view point.

For further reading, refer to Chapter 12, "Security Features of the SP System" in *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348.

16.1 Authentication and authorization methods

An *authentication* is the process of validating the identity of either a user of a service or the service itself. An *authorization* is the process of obtaining permission to access resources or perform tasks. In SP security services, authorization is based on the principal identifier.

This section discusses how to configure your chosen authentication methods. You can choose to configure any or all of the following authentication methods:

- Standard AIX
- Kerberos Version 4
- Kerberos Version 5

Kerberos Version 5 is not provided by IBM Parallel System Support Programs for AIX (PSSP) 3.1. Distributed Computing Environment for AIX (DCE) 2.2 (or later) provides a protocol compatible with Kerberos Version 5. You must order, install, and configure DCE if you choose to configure Kerberos Version 5 for authentication method. Standard AIX authentication method comes with AIX 4.3.2.

The rules for configuring the authentication methods presented are:

- All nodes within a single system partition will have the same set of authentication methods enabled.
- The authentication methods enabled on the CWS will be the union of all authentication methods enabled for all system partitions plus any other methods already set.
- For PSSP 3.1, you must install and configure Kerberos Version 4 for all partitions within the SP system.

When specifying several authentication methods within the SP system, they will be resolved in a certain priority. This ranking mirrors the levels of security:

1. Kerberos Version 5
2. Kerberos Version 4
3. Standard AIX

If you are looking for further reading on the available authentication methods and how to configure them, refer to the following IBM publications:

- Chapter 7, “Planning for Security” in *RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281

- Chapter 12, “Security Features of the SP System” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*

16.1.1 Enabling authentication methods

When you configure the SP system, you need to select the active authentication methods for each system partition. To do this, issue the `smitty spauth_methods` fast path:

```

Enable Authentication Methods

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
Enable on Control Workstation Only      no          +
Force change on nodes                   no          +
* System Partition names                 sp3en0     +
* Authentication Methods                 k4 std     +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command      F7=Edit       F8=Image
F9=Shell     F10=Exit        Enter=Do

```

Alternatively, issue the `chauthpar` command:

```
# /usr/lpp/ssp/bin/chauthpar -p sp3en0 k4 std
```

This example enables Kerberos Version 4 and standard AIX authentication methods for `sp3en0` system partition.

The `chauthpar` command issues the AIX `chauthent` command on the CWS. It is necessary to ensure that all methods used by each system partition is also enabled on the CWS. The `chauthpar` command also remotely issues the `chauthent` command on all the nodes that belong to the system partition you specified.

The command writes this information to the `auth_methods` attribute in Syspar SDR class:

```
# SDRGetObjects Syspar
syspar_name ip_address install_image syspar_dir code_version haem_cdb_versi
on auth_install auth_root_rcmd auth_methods
sp5en0 192.168.5.150 default "" PSSP-3.1 920921076,3853
79072,0 k4 k4:std k4:std
#
```

16.1.2 Selecting authorization methods

To select authorization methods for root access to remote commands, issue the `smitty spauth_config` fast path:

```
Select Authorization Methods for Root access to Remote Comands

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                     [Entry Fields]
* System Partition names             sp3en0          +
* Authorization Methods              k4 std         +

F1=Help      F2=Refresh    F3=Cancel    F4=List
F5=Reset     F6=Command    F7=Edit     F8=Image
F9=Shell     F10=Exit     Enter=Do
```

Alternatively, issue the `spsetauth` command:

```
# /usr/lpp/ssp/bin/spsetauth -p sp3en0 -d k4 std
```

This example selects Kerberos Version 4 and standard AIX authorization methods for `sp3en0` system partition for root access to remote commands.

The command writes this information to the `auth_root_rcmd` attribute in Syspar SDR class:

```
# SDRGetObjects Syspar
syspar_name ip_address install_image syspar_dir code_version haem_cdb_versi
on auth_install auth_root_rcmd auth_methods
sp5en0 192.168.5.150 default "" PSSP-3.1 920921076,3853
79072,0 k4 k4:std k4:std
#
```

16.1.3 Listing authentication methods

To list the active authentication methods for a current system partition, issue the `lsauthpar` command:

```
# lsauthpar
Kerberos 4
Standard Aix
#
```

This example shows you are using Kerberos Version 4 and standard AIX authentication methods for a current system partition.

16.2 Daemons

The SP system administrator should know about the daemons used by the authentication servers to be able to monitor and control them:

- kerberos daemon
- kadmind daemon
- kpropd daemon

All the Kerberos daemons reside in the `/usr/lpp/ssp/kerberos/etc` directory and run on the primary or secondary authentication servers. By default, the primary authentication server is assigned to the CWS. By option, you can build the secondary authentication servers.

16.2.1 kerberos daemon

The *kerberos daemon* is the actual authentication server. It is sometimes referred to as the *Key Distribution Center* (KDC) because of its role in distributing the session key. The *session key* is a temporary key used to encrypt parts of authentication protocol messages. Its lifetime is the same as the ticket with which it is created. The KDC actually consists of two separate functions: The *Authentication Service*, which generates *ticket-granting-tickets*

(TGT) based on authentication using the client user's *private key* and the *Ticket-Granting Service (TGS)*, which generates *service tickets*, based on authentication using a TGT.

The kerberos daemon runs on the both primary and secondary authentication server. The daemon listens for requests on the *kerberos4* User Datagram Protocol (UDP) port. If this port is not defined in the */etc/services* file, it uses port 750.

When you start the primary authentication server, you can specify a maximum database age for the authentication database files. This can be used to ensure that you do not start a secondary authentication server with out-of-date information. This could occur in a situation when a secondary authentication server was down, and a authentication database update was scheduled by the primary authentication server.

The daemon is registered as a subsystem within the System Resource Controller (SRC) of AIX.

The daemon keeps its log information in the *kerberos.log* file. It is located in the */var/adm/SPlogs/kerberos* directory on the primary authentication server.

16.2.2 kadmind daemon

The *kadmind daemon* provides network access for authentication database administration via the *kpasswd* or *kadmin* command. This daemon provides the interface for Kerberos Version 4 authentication database administration. It uses the Kerberos master key for authorization.

The *kadmin* daemon runs only on the primary authentication server. The daemon listens to requests on the *kerberos_admin* Transmission Control Protocol (TCP) port. If this port is not defined in the */etc/services* file, it uses port 751.

When a Kerberos client requests a action, the *kadmind* daemon checks access control lists (ACLs) to determine the authorization of the client to perform the requested action. Currently, the following three distinct access types are supported:

Addition This action uses the */var/kerberos/database/admin_acl.add* file. A principal on this list can add new principals to the authentication database.

Retrieval This action uses the */var/kerberos/database/admin_acl.get* file. A principal on this list can retrieve authentication

database entries. Note, that a principal's private key is never returned by the get functions.

Modification This action uses the `/var/kerberos/database/admin_acl.mod` file. A principal on this list can modify entries in the authentication database.

Principals are always granted authorization to change their own password.

The daemon is registered as a subsystem within the SRC of AIX.

The daemon keeps its log information in the `admin_server.syslog` file. It is located in the `/var/adm/SPlogs/kerberos` directory on the primary authentication server.

16.2.3 kpropd daemon

The Kerberos design allows secondary authentication servers to provide authentication database availability in case of primary authentication server failures. Adding principals and changing passwords takes place in the primary authentication database and is propagated to secondary authentication databases through periodic updates only.

Secondary authentication databases are maintained by the *kpropd daemon* that runs only on secondary authentication server and receives the authentication database content in encrypted form from the `kprop` command that runs on the primary authentication server. The daemon listens for requests on the `krb_prop` TCP port. If this port is not defined in the `/etc/services` file, it uses port 754.

The `kpropd` daemon validates the connection by checking the `/etc/krb.conf` file for the local realm for the primary authentication server.

The activity of `kpropd` daemon is logged in the `kpropd.log` file located in the `/var/adm/SPlogs/kerberos` directory.

16.2.4 Kerberos authenticated-applications

There are two service principals used by PSSP components:

hardmon This principal is used by the SP System Monitor daemon, `hardmon`, and SP logging daemon, `splogd`.

rcmd This principal is used by the `Sysctl`.

The `hardmon` daemon runs only on the CWS. The `splogd` daemon can run on other RS/6000 systems. Therefore, for each network interface name, a

service principal uses `hardmon` for its name and the network interface name for its instance.

The remote commands can be issued from or to any RS/6000 system. Therefore, for each network interface name, a service principal uses `rcmd` for its name and the network interface name for its instance.

16.2.5 Starting daemons

To start the Kerberos authentication server, issue the `startsrc` command:

```
# stopsrc -s kerberos
```

To start the authentication database administration server, issue the `startsrc` command:

```
# stopsrc -s kadmind
```

To start the authentication database propagation server, issue the `startsrc` command:

```
# stopsrc -s kpropd
```

16.2.6 Stopping daemons

To stop the Kerberos authentication server, issue the `stopsrc` command:

```
# stopsrc -s kerberos
```

To stop the authentication database administration server, issue the `stopsrc` command:

```
# stopsrc -s kadmind
```

To stop the authentication database propagation server, issue the `stopsrc` command:

```
# stopsrc -s kpropd
```

16.3 Authentication database

An *authentication database* is a set of files containing the names and authentication information of all principals within a realm. A *realm* has one primary authentication database and may have multiple secondary authentication databases. Secondary databases are backup copies of the primary authentication database and may be provided to improve performance or availability.

An authentication database is protected by a *Kerberos master key*. The Kerberos master key is derived from the Kerberos master password supplied initially by the administrator when the primary authentication server is created. This key is saved in the Kerberos master key cache file.

The authentication database is located in the `/var/kerberos/database` directory on the primary authentication server. The ACLs for the authentication database are located there as well. Refer to 16.2.2, “`kadmind` daemon” on page 450, for ACLs. The files containing the authentication database are:

- `principal.dir`
- `principal.ok`
- `principal.pag`

Kerberos has to be treated with the same care as any other important database. Databases used in commercial environments should be backed up regularly. The same applies to the authentication database. Using secondary authentication servers is also a good treatment. But, they only keep read-only copies of the original authentication database. They can not be modified.

16.3.1 Initializing authentication database

To initialize an authentication database, issue the `kdb_init` command. This command initializes the authentication database, creating the necessary initial system principals.

After determining the realm to be created, the command prompts for a master key password. The user must remember this password because it is used for other commands. The master key password is used to encrypt every encryption key stored in the authentication database.

The `kdb_init` command is issued by `setup_authent` command. For a sample operation, refer to Figure 98 on page 485.

16.3.2 Destroying authentication database

To destroy an authentication database, issue the `kdb_destroy` command:

```
# kdb_destroy
You are about to destroy the Kerberos database on this machine.
Are you sure you want to do this (y/n)? y
Database deleted at /var/kerberos/database/principal
#
```

This command removes the authentication database. You first must reply `y` or `Y` to a prompt, to confirm the request, or the command exits without removing the authentication database files. This command can only be issued on the system on which the authentication database resides.

The following files are deleted as an authentication database by this command:

- `principal.dir`
- `principal.pag`

16.3.3 Backing up authentication database

To create an authentication database back up file, issue the `kdb_util` command with the `dump` operand:

```
# kdb_util dump /var/kerberos/database/slavesave
# ls -al /var/kerberos/database/slavesave
total 41
-rw-r--r--  1 root    system    1548 May 23 16:34 slavesave
#
```

The command dumps the authentication database into a text representation in the file specified.

The following files are backed up as an authentication database by this command:

- `principal.dir`
- `principal.ok`
- `principal.pag`

16.3.4 Restoring authentication database

To restore an authentication database using a back up file, issue the `kdb_util` command with the `load` operand:

```

# ls -al /var/kerberos/database/principal*
total 41
-rw----- 1 root    system    4096 May 22 14:03 principal.dir
-rw----- 1 root    system      0 May 22 14:03 principal.ok
-rw----- 1 root    system   14336 May 22 14:04 principal.pag
# kdb_util load /var/kerberos/database/slavesave
# ls -al /var/kerberos/database/principal*
total 41
-rw----- 1 root    system    4096 May 23 16:36 principal.dir
-rw----- 1 root    system      0 May 23 16:36 principal.ok
-rw----- 1 root    system   14336 May 23 16:36 principal.pag
#

```

The command initializes the authentication database with the records described by the text contained in the specified file. Any existing authentication database is overwritten.

The following files are restored as an authentication database by this command:

- principal.dir
- principal.ok
- principal.pag

16.3.5 Reading an authentication database dump

This section shows how an authentication database dump file looks. In daily authentication database administration, this task is normally not needed but to be familiar with it can have certain advantages.

First, you need to create an authentication database dump file. To do this, issue the `kdb_util` command. Then issue the `cat` command to look inside:

```

# kdb_util dump /tmp/authdb
# cat /tmp/authdb | grep -v rand
K M 255 3 3 0 9e876f67 9b8fbed5 203801010459 199905191422 db_creation *
changepw kerberos 255 3 1 0 a035c01a 44830c45 203801010459 199905191422 db_creat
ion *
root admin 255 3 1 0 9e876f67 9b8fbed5 203801010459 199905191423 * *
hardmon sp5en0 255 3 1 0 fc3a34f9 cfe0bb78 203801010459 199905191424 root admin
krbtgt MSC.ITSO.IBM.COM 255 3 1 0 9d301750 10c137e2 203801010459 199905191422 db
_creation *
default * 255 1 1 0 0 0 203801010459 199905191422 db_creation *
hardmon sp5cw0 255 3 1 0 2ecb1bb9 3b39d23d 203801010459 199905191424 root admin
#

```

The output shows all the principals except principals whose name is rcmd.

Take a look at the hardmon.sp5en0 principal for example:

```
hardmon sp5en0 255 3 1 0 fc3a34f9 cfe0bb78 203801010459 199905191424 root
admin
```

The following is an explanation for each attribute from left to right:

- The principal's name is hardmon.
- The principal's instance is sp5en0.
- The maximum ticket lifetime is 255.
- The key version of the Kerberos master key used to encrypt the entry is 3.
- The key version of the principal's secret key is 1.
- The principal's attributes is 0. This attribute is not used currently.
- The two halves of the encrypted private key are fc3a34f9 and cfe0bb78.
- The expiration date (in Universal Coordination Time) is 203801010459. It uses YYYYMMDDHHMM format.
- The date the entry was created or modified is 199905191424. It uses YYYYMMDDHHMM format also.
- The name and instance of the administrator who make the last update is the root.admin principal.

It is interesting to know principals, other than root, hardmon, or rcmd. The following introduces these special principals:

K.M	This principal is defined to hold the master key password as its own. The authentication database is encrypted using that password. The <code>kstash</code> command reads it, and it is also used to secure access to various admin commands.
changepw.kerberos	This principal is the one used by the <code>kadmin</code> daemon. It has a randomly generated password.
krbtgt	This principal name, used with the realm name as the instance, is the service principal used by the Ticket-Granting Service, the <code>kerberos</code> daemon. It has a randomly generated password.
default	This principal (with a null instance) serves as the model or prototype that is cloned when creating a new principal, thus, supplying the default expiration date, maximum ticket lifetime, and attributes. It is assigned

no password, which prevents it from being logged in by the `k4init` command.

16.3.6 Changing the Kerberos master key

This task can be essential to protect your authentication database against intruders. The best precautions one can take is to change the Kerberos master key regularly. This section will show you how to do this.

Step 1: Save the authentication database

Prior to changing the Kerberos master key, make sure to have a backup of the authentication database using the old Kerberos master key in case you forget the new one but remember the old one.

To do this, refer to 16.3.3, “Backing up authentication database” on page 454.

Step 2: Change the Kerberos master key

The `new_master_key` operand of the `kdb_util` command is used to perform this task. Issuing it will dump the authentication database to an ASCII file.

Therefore, the file name must be specified with the command. The system will first prompt you for the old Kerberos master key, and then the new one can be entered.

The following shows what it looks like:

```
# cd /usr/kerberos/etc
# kdb_util new_master_key /var/kerberos/database/newdb

Enter the CURRENT master key.
Enter Kerberos master key:

Now enter the NEW master key. Do not forget it!!
Enter Kerberos master key:
Verifying, please re-enter
Enter Kerberos master key:

Now you must issue `kdb_util load /var/kerberos/database/newdb' to reload the database,
then issue `kstash' to update the master key cache file,
and finally kill the Kerberos daemons `kerberos' and `kadmind' to restart them.
#
```

Step 3: Load the authentication database

Issue the `kdb_util` command for this task. You have to specify the file containing the ASCII authentication database dump as parameter when using the `load` operand:

```
# kdb_util load /var/kerberos/database/newdb
#
```

Step 4: Create the new Kerberos master key cache file

The new Kerberos master key needs to be cached in the `/.k` file on the primary authentication server. To create the Kerberos master key cache file, issue the `kstash` command:

```
# kstash
Enter Kerberos master key:
#
```

The system prompts for the Kerberos master key. Enter it, and you have completed the change of the Kerberos master key.

Step 5: Restart the kerberos daemon

The kerberos daemon still references the old Kerberos master key version. Therefore, you have to stop it and start it by issuing the `stopsrc` and `startsrc` command:

```
# stopsrc -s kerberos
0513-044 The stop of the kerberos Subsystem was completed successfully.
# startsrc -s kerberos
0513-059 The kerberos Subsystem has been started. Subsystem PID is 14972.
#
```

Otherwise, you will get problems accessing the Hardware Monitor, remote commands can not be executed, and the following error message will be presented by the system while trying to issue the `k4init` command for principals:


```
# k4init root.admin
Kerberos Initialization for "root.admin"
kinit: 2504-005 Incorrect Kerberos master key version
#
```

Step 6: Propagate the authentication database

If secondary authentication servers are used in your SP system, the authentication database has to be propagated. This includes the following actions:

1. Issue the `push-kprop` command from the primary authentication server.
2. Copy the `/.k` file to the secondary authentication servers.
3. Stop and start the kerberos daemon on the secondary servers.

16.4 Principal

A *principal* is an entity whose identifier and key are maintained in the authentication database. The principal may be a real AIX user that requires access to a Kerberos authenticated service, or it could be the name of an authenticated service provider.

An *instance* is a qualifier for a principal name. For services, an instance represents a particular occurrence of the server. For users, an instance allows a single user to assume additional (or alternate) roles with different authority.

This section explains how to list and view information on the principals in your authentication database. It also explains how to manage your principals including adding them, deleting them, and changing their attributes.

16.4.1 Listing principal

To list all or some principals in the authentication database, you can issue the `lskp` command on the authentication server. The following example lists all principals who are Kerberos administrators (`.admin`) or a `sp3n01` instance (`.sp3n01`):

```
# /usr/kerberos/etc/lskp .admin .sp3n01
root.admin          tkt-life: 30d      key-vers: 1  expires: 2037-12-31 23:59
rcmd.sp3n01        tkt-life: Unlimited key-vers: 1  expires: 2037-12-31 23:59
#
```

The output shows the name and instance, the maximum ticket lifetime, the key version number, and the expiration date.

If you are not in the authentication server (sp3en0, in this example), you can issue the `lskp` command to the authentication server by issuing the `sysctl` command:

```
# sysctl -h sp3en0 lskp .admin .sp3n01
root.admin          tkt-life: 30d      key-vers: 1  expires: 2037-12-31 23:59
rcmd.sp3n01        tkt-life: Unlimited key-vers: 1  expires: 2037-12-31 23:59
#
```

16.4.2 Getting principal information

To get principal information, there are two methods available:

- Using the `kadmin` command
- Viewing the authentication database dump file

Using the `kadmin` command

To get information on a principal, you can issue the `kadmin` command from any machine.

```
# kadmin
Welcome to the Kerberos Administration Program, version 2
Type "help" if you need it.
admin: ?
Available admin requests:

change_password, cpw      Change a user's password
change_admin_password, cap
                           Change your admin password
add_new_key, ank          Add new user to kerberos database
get_entry, get            Get entry from kerberos database
destroy_tickets, dest     Destroy admin tickets
help                      Request help with this program
list_requests, lr, ?     List available requests.
quit, exit, q            Exit program.
admin: get root.admin
Admin password:
Info in Database for root.admin:
    Max Life: 255   Exp Date: Thu Dec 31 23:59:59 2037

    Attribs: 00   key: 0 0
admin: quit
Cleaning up and exiting.
#
```

In this example, you issue the `?` subcommand to list available requests and then issue the `get` subcommand to get `root.admin` entry from authentication database. It shows similar data that you get from the `lskp` command used in 16.4.1, “Listing principal” on page 459.

Using the authentication database dump file

Simply view the contents of the authentication database dump file. The authentication database exists only on the authentication server.

To read the dump file, refer to 16.3.5, “Reading an authentication database dump” on page 455.

16.4.3 Adding a principal

There are four commands available to add a principal to the authentication database:

- `kadmin`
- `add_principal`
- `kdb_edit`
- `mkkp`

The `kdb_edit` and `mkkp` commands are only available on the authentication server.

In this section, you add a principal named `kuser1`, for example.

16.4.3.1 Using the `kadmin` command

To add a principal, `kuser1`, issue the `kadmin` command. At the `admin:` prompt, enter the `ank` subcommand for `kuser1`. Then enter admin password and password for `kuser1`. To check if the principal was created, issue the `get` subcommand. The following is the sample output:

```

# kadmin
Welcome to the Kerberos Administration Program, version 2
Type "help" if you need it.
admin: ank kuser1
Admin password:
Password for kuser1:
Verifying, please re-enter Password for kuser1:
kuser1 added to database.
admin: get kuser1
Admin password:
Info in Database for kuser1.:
    Max Life: 255   Exp Date: Thu Dec 31 23:59:59 2037

    Attribs: 00   key: 0 0
admin: quit
Cleaning up and exiting.
#

```

The `kadmin` command is a good choice to add a principal, but you can not specify the expiration date and the maximum ticket lifetime. It requires admin password for each time you issue a subcommand.

16.4.3.2 Using the `add_principal` command

The big advantage of the `add_principal` command is that it is non-interactive. It allows a large number of principals to be added at one time. The name and initial password of the principals can be described in an ASCII file, and you can specify the file name when you issue the command. This adds the principals to the authentication database and sets the initial passwords at the same time.

The following example uses the `/tmp/addkuser1` file to provide the command with a principal's name and its initial password:

```

# cat /tmp/addkuser1
kuser1 kuser1pw
# add_principal /tmp/addkuser1
# lskp kuser1
kuser1          tkt-life: 30d      key-vers: 1  expires: 2037-12-31 23:59
#

```

The command does not require admin password, but you can not specify the expiration date and the maximum ticket lifetime.

16.4.3.3 Using the `kdb_edit` command

The `kdb_edit` command is the most powerful command and is used by the `setup_authent` script issued during PSSP installation. You can specify all the attributes by this command.

To add a principal `kuser1`, issue the `kdb_edit` command. The system will prompt you as such:

```
Opening database...
Enter Kerberos master key:
```

Enter the Kerberos master key. If you want to avoid input to the Kerberos master key, issue the command with `-n` flag. This indicates that the Kerberos master key is obtained from the `/.k` Kerberos master key cache file.

The command will prompt first for the principal name, which is `kuser1`. In this example, you are not going to add an administrator of the authentication database. So, null instance is enough for the principal. It looks like this:

```
Principal name: kuser1
Instance:
```

The command checks the authentication database, and if the principal does not exist, it asks to create it:

```
<Not found>, Create [y] ?
```

Reply `y`, or just hit the **Enter** key.

The command prompts for the password of the principal. You have to enter it twice because it is verified:

```
New Password:
Verifying, please re-enter
New Password:
```

Now, you are prompted for the expiration date and the maximum ticket lifetime. The attribute named `attributes` is not used currently. Adjust them according to your needs. Accepting the default choices is, in most cases, appropriate:

```
Expiration date (enter yyyy-mm-dd) [ 2037-12-31 ] ?
Max ticket lifetime [ 255 ] ?
Attributes [ 0 ] ?
```

Finally, the command states that everything went fine and prompts you again to enter a principal name.

```
Edit O.K.
```

Principal name:

If you don't have any more principals added, hit the **Enter** key.

The following is the sample output of this operation:

```
# kdb_edit
Opening database...

Enter Kerberos master key:

Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: kuser1
Instance:

<Not found>, Create [y] ?

Principal: kuser1, Instance: , kdc_key_ver: 1
New Password:
Verifying, please re-enter
New Password:

Principal's new key version = 1
Expiration date (enter yyyy-mm-dd) [ 2037-12-31 ] ?
Max ticket lifetime [ 255 ] ?
Attributes [ 0 ] ?
Edit O.K.
Principal name:
#
```

16.4.3.4 Using the mkkp command

To add a principal, you use the `mkkp` command. The command allows you to specify the expiration date and the maximum ticket lifetime but not an initial password. After adding a principal, you need to initialize the password:

```
# mkkp -e 2037-12-31 -l 255 kuser1
# lskp kuser1
kuser1          tkt-life: 30d      key-vers: 1  expires: 2037-12-31 23:59
#
```

To initialize the password, refer to 16.4.6, “Changing password” on page 468.

16.4.4 Giving access authorization

When you get a principal information, add a principal, or modify principal information, you request these actions from the `kadmind` daemon. The daemon checks the following ACL files, respectively, if you have access:

- `/var/kerberos/database/admin_acl.get`
- `/var/kerberos/database/admin_acl.add`
- `/var/kerberos/database/admin_acl.mod`

For more details about `kadmin` daemons, refer to 16.2.2, “`kadmind` daemon” on page 450.

If you want to give accesses to an AIX user, for example, `kuser1`, you need to do the following:

1. Add a Kerberos principal whose name is `kuser1` and instance that is `admin`. Refer to 16.4.3, “Adding a principal” on page 461, to add a principal.

2. If you give an addition access to the `kuser1.admin` principal:

```
# print "kuser1.admin" >> /var/kerberos/database/admin_acl.add
```

3. If you give a retrieval access to the `kuser1.admin` principal:

```
# print "kuser1.admin" >> /var/kerberos/database/admin_acl.get
```

4. If you give a modification access to the `kuser1.admin` principal:

```
# print "kuser1.admin" >> /var/kerberos/database/admin_acl.mod
```

After having added the `kuser1.admin` principal, check if it works. Log in to AIX as a `kuser1` user, then issue the `k4init` command to be authenticated as a `kuser1.admin` principal. Then issue the `lsksp` command to list the `kuser1.admin` principal:

```

AIX Version 4
(C) Copyrights by IBM and by others 1982, 1996.
login: kuser1
kuser1's Password:
*****
*
*
* Welcome to AIX Version 4.3!
*
*
* Please see the README file in /usr/lpp/bos for information pertinent to
* this release of the AIX Operating System.
*
*
*****

$ k4init kuser1.admin
Kerberos Initialization for "kuser1.admin"
Password:
$ /usr/lpp/ssp/kerberos/etc/lskp kuser1.admin
ksh: /usr/lpp/ssp/kerberos/etc/lskp: 0403-006 Execute permission denied.
$

```

The `lskp` command fails, even though the `kuser1` is authenticated as the `kuser1.admin` principal because the `kuser1` does not have AIX root authority. To perform the `lskp` command, issue the command with the `sysctl` command:

```

$ sysctl -h sp3en0 lskp kuser1.admin
kuser1.admin      tkt-life: 30d      key-vers: 1  expires: 2037-12-31 23:59
$

```

This time, it should work.

16.4.5 Delete principal

To delete the principal from the authentication database, use the following steps. In this section, you delete a principal named `kuser1.admin`, for example.

Step 1: Delete a principal

To delete the principal from the authentication database, issue the `rmlkp` command:


```
# rmkp -v -n kuser1.admin
kuser1.admin was removed.
Removed entries were saved in /var/kerberos/database/rmkp.save.4184
#
```

The `-n` flag bypasses prompting for confirmation, and the `-v` flag displays informational messages.

As you may have noticed from the output, the `rmkp` command saves a deleted principal in the authentication database directory using its process ID (PID) as a file extension.

If you do not have an AIX root authority, but are a Kerberos administrator, you can issue the `rmkp` command with the `sysctl` command:

```
$ sysctl -h sp3en0 rmkp -v -n kuser1.admin
kuser1.admin was removed.
Removed entries were saved in /var/kerberos/database/rmkp.save.4184
$
```

In this example, `sp3en0` is the name of the primary authentication server.

Step 2: Cleaning up kadmind ACLs

If you registered the `kuser1.admin` principal as an authentication database administrator, you should delete the entry of the `kuser1.admin` principal from the following files:

- `admin_acl.get`
- `admin_acl.add`
- `admin_acl.mod`

The files are located in the `/var/kerberos/database` directory on the authentication server.

Step 3: Cleaning up \$HOME/.klogin ACLs

You may registered the `kuser1.admin` principal to `$HOME/.klogin` file on the CWS or nodes. If this is the case, you should delete the entry of the `kuser1.admin` principal from this file.

Step 4: Cleaning up SP System Monitor ACL

You may registered the kuser1.admin principal to use the SP System Monitor, in other words, use the `spmon`, `hmcmds`, or `hmon` command or SP Perspectives. If this is the case, you should delete the entry of the kuser1.admin principal from the following file:

- `/spdata/sys1/spmon/hmacls`

This file is located on the CWS.

You need to issue the `hmadm` command to let SP System Monitor daemon know the change:

```
# hmadm setacls
```

Step 5: Cleaning up Sysctl ACLs

You may registered the kuser1.admin principal to perform special tasks using the `sysctl` command. If this is the case, you should delete the entry of the kuser1.admin principal from the following files:

- `/etc/sysctl.acl`
- `/etc/sysctl.pman.acl`
- `/etc/sysctl.rootcmds.acl`
- `/etc/sysctl.vsd.acl` (for IBM Virtual Shared Disk)
- `/etc/sysctl.mmcmd.acl` (for IBM General Parallel File System for AIX)

They are located in the CWS and the nodes.

16.4.6 Changing password

There are three commands available to change the password for a principal:

- `kpasswd`
- `kadmin`
- `kdb_edit`

The `kdb_edit` command is only available on the authentication server.

In this section, you change a password for the kuser1 principal.

16.4.6.1 Using the kpasswd command

The `kpasswd` command is fairly straightforward, and it prompts your for the new password:

```
# kpasswd -n kuser1
Old password for kuser1:
New Password for kuser1:
Verifying, please re-enter New Password for kuser1:
Password changed.
#
```

To issue this command, you need to know the old password for the principal. This means the command is for end-users to change their password.

16.4.6.2 Using the `kadmin` command

The alternative method is to issue the `kadmin` command and its `cpw` subcommand:

```
# kadmin
Welcome to the Kerberos Administration Program, version 2
Type "help" if you need it.
admin: cpw kuser1
Admin password:
New password for kuser1:
Verifying, please re-enter New password for kuser1:
Password changed for kuser1.
admin: quit
Cleaning up and exiting.
#
```

Unlike the `kpasswd` command, this command does not ask you the old password. Instead, you need to know the password for the admin principal. This means the command is for the system administrator.

16.4.6.3 Using the `kdb_edit` command

The `kdb_edit` command allows you to change all the attributes including the password. Reply `y` to the Change password [n] ? prompt to change the password:

```

# kdb_edit -n
Opening database...
Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: kuser1
Instance:

Principal: kuser1, Instance: , kdc_key_ver: 1
Change password [n] ? y

New Password:
Verifying, please re-enter
New Password:

Principal's new key version = 2
Expiration date (enter yyyy-mm-dd) [ 2037-12-31 ] ?
Max ticket lifetime [ 255 ] ?
Attributes [ 0 ] ?
Edit O.K.
Principal name:
#

```

If you issue the `kdb_edit` command with the `-n` flag, the command does not ask you for the Kerberos master key.

16.4.7 Changing expiration date

There are two commands available to change the expiration date for a principal:

- `chkp`
- `kdb_edit`

The `chkp` and `kdb_edit` commands are only available on the authentication server.

In this section, you change an expiration date for the `kuser1` principal.

16.4.7.1 Using the `chkp` command

To change the expiration date for the `kuser1` principal, issue the `chkp` command with the `-e` flag:

```

# lskp kuser1
kuser1          tkt-life: 30d      key-vers: 2  expires: 2037-12-31 23:59
# chkp -e 1999-12-30 kuser1
# lskp kuser1
kuser1          tkt-life: 30d      key-vers: 2  expires: 1999-12-30 23:59
#

```

Use the format *yyyy-mm-dd* to specify the expiration date. In this example, you changed expiration date from 2037-12-31 to 1999-12-30.

16.4.7.2 Using the `kdb_edit` command

To change the expiration date for the `kuser1` principal, issue the `kdb_edit` command:

```

# lskp kuser1
kuser1          tkt-life: 30d      key-vers: 2  expires: 1999-12-30 23:59
# kdb_edit -n
Opening database...
Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: kuser1
Instance:

Principal: kuser1, Instance: , kdc_key_ver: 1
Change password [n] ?
Expiration date (enter yyyy-mm-dd) [ 1999-12-30 ] ? 2001-09-28
Max ticket lifetime [ 255 ] ?
Attributes [ 0 ] ?
Edit O.K.
Principal name:
# lskp kuser1
kuser1          tkt-life: 30d      key-vers: 2  expires: 2001-09-28 23:59
#

```

In this example, you changed the expiration date from 1999-12-30 to 2001-09-28.

16.4.8 Changing maximum ticket lifetime

There are two commands available to change the maximum ticket lifetime for a principal:

- `chkp`
- `kdb_edit`

The `chkp` and `kdb_edit` commands are only available on the authentication server.

Whichever command you use, Table 13 shows a representative sample of relationship between the lifetime operand and approximate duration:

Table 13. Lifetime operand and approximate duration

Lifetime Operand	Approximate Duration
141	1 day
151	2 days
170	1 week
180	2 weeks
191	1 month

For a complete list of the possible ticket lifetime values, refer to Chapter 12, “Security Features of the SP System” in *IBM Parallel System Support Programs for AIX: Administration Guide, SA22-7348*.

In this section, you change the maximum ticket lifetime for the `kuser1` principal.

16.4.8.1 Using the `chkp` command

To change the maximum ticket lifetime for the `kuser1` principal, issue the `chkp` command with the `-l` flag:

```
# lskp kuser1
kuser1          tkt-life: 30d          key-vers: 2  expires: 2001-09-28 23:59
# chkp -l 141 kuser1
# lskp kuser1
kuser1          tkt-life: 1d+01:26 key-vers: 2  expires: 2001-09-28 23:59
#
```

This example changed the maximum ticket lifetime from 30 days to one day plus one hour and 26 minutes.

16.4.8.2 Using the `kdb_edit` command

This procedure is similar to the procedure used for changing an expiration date. Issue the `kdb_edit` command. Instead of replying to the Expiration date prompt, reply to the Max ticket lifetime prompt this time:

```

# lskp kuser1
kuser1          tkt-life: 1d+01:26 key-vers: 2  expires: 2001-09-28 23:59
# kdb_edit -n
Opening database...
Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: kuser1
Instance:

Principal: kuser1, Instance: , kdc_key_ver: 1
Change password [n] ?
Expiration date (enter yyyy-mm-dd) [ 2001-09-29 ] ?
Max ticket lifetime [ 141 ] ? 170
Attributes [ 0 ] ?
Edit O.K.
Principal name:
# lskp kuser1
kuser1          tkt-life: 7d+08:50 key-vers: 2  expires: 2001-09-28 23:59
#

```

This example changed the maximum ticket lifetime from one day plus one hour and 26 minutes to seven weeks plus eight hours and 50 minutes.

You can change the expiration date and the maximum ticket lifetime at the same time.

16.4.9 Changing default values

When you add a principal, Kerberos provides you with default values. For example, it provides 2037-12-31 for the expiration date and 255 for the maximum ticket lifetime. If you want to change these default values, there is a way.

The principal named default is created when you initialized the authentication database. Its expiration date and maximum ticket lifetime are used for default values. To change the default values, change the default principals of these values. To do this, issue the `kdb_edit` command:

```

# kdb_edit -n
Opening database...
Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: default
Instance:

Principal: default, Instance: , kdc_key_ver: 1
Change password [n] ?
Expiration date (enter yyyy-mm-dd) [ 2037-12-31 ] ? 1999-12-31
Max ticket lifetime [ 255 ] ? 170
Attributes [ 0 ] ?
Edit O.K.
Principal name:
#

```

In this example, you changed the expiration date to 1999-12-31 and the maximum ticket lifetime to 170 (about one week).

To activate these default values, you need to restart the Kerberos daemons, `kerberos` and `kadmind`. Issue the `stopsrc` and `startsrc` commands:

```

# stopsrc -s kadmind
0513-044 The stop of the kadmind Subsystem was completed successfully.
# stopsrc -s kerberos
0513-044 The stop of the kerberos Subsystem was completed successfully.
# startsrc -s kerberos
0513-059 The kerberos Subsystem has been started. Subsystem PID is 14972.
# startsrc -s kadmind
0513-059 The kadmind Subsystem has been started. Subsystem PID is 19750.
#

```

Check the default values by issuing the `kdb_edit` command:


```

# kdb_edit -n
Opening database...
Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: new
Instance: user

<Not found>, Create [y] ? y

Principal: new, Instance: user, kdc_key_ver: 1
New Password:
Verifying, please re-enter
New Password:

Principal's new key version = 1
Expiration date (enter yyyy-mm-dd) [ 1999-12-31 ] ?
Max ticket lifetime [ 170 ] ?
Attributes [ 0 ] ?
Edit O.K.
Principal name:
#

```

The command uses the default values you specified. For the default principal, refer to 16.3.5, “Reading an authentication database dump” on page 455.

16.5 Service key file

A *service key file* (also referred to as a *srvtab file*) is a file containing the names and private keys of the local instances of services. It is accessible only to processes that run under the UID of the user owning the server daemons. On an SP system, all services run as root.

The service key file resides in the SP system as the `/etc/krb-srvtab` file.

16.5.1 Creating a service key file

To create a service key file, perform the following steps. In this section, you are going to create a service key file for node 1 (sp3n01):

Step 1: Set the node to customize mode

Set the node to customize mode from disk mode. To do this, issue the `smitty server_dialog fast` path:

```

                                Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                [Entry Fields]
Start Frame                      [1]                      #
Start Slot                       [1]                      #
Node Count                       [1]                      #

OR

Node List                        []

Response from Server to bootp Request customize          +
Volume Group Name                []
Run setup_server?                no                       +

F1=Help      F2=Refresh      F3=Cancel      F4=List
F5=Reset     F6=Command     F7=Edit      F8=Image
F9=Shell     F10=Exit       Enter=Do

```

Alternatively, issue the `spbootins` command:

```
# /usr/lpp/ssp/bin/spbootins -r customize -s no 1 1 1
```

Make sure that you do not need to run the `setup_server` command.

Step 2: Issue the `create_krb_files` script

Issue the `create_krb_files` command on CWS. If your SP system uses a boot/install server (BIS) for a target node, issue the script on the BIS instead:

```
# create_krb_files
create_krb_files: tftpaccess.ctl file and client srvtab files created/updated
on server node 0.
#
```

This script generates a new service key file in the `/tftpboot` directory. The file name is `initial_hostname-new-srvtab`, where `initial_hostname` is the SP Ethernet interface name of the node.

The new service key file should look like the following:

```
-r----- 1 nobody system 118 Jan 18 14:26 sp3n01-new-srvtab
```

Step 3: Copy the new service key file

Copy the new service key file from CWS or BIS to the target node. When you copy the file, change its name to `/etc/krb-srvtab` and change its attributes as follows:

```
-rw----- 1 root      system      118 Jan 18 14:26 /etc/krb-srvtab
```

Step 4: Set the node to disk mode

Do not forget to set back the node to disk mode. To do this, issue the `smitty server_dialog fast` path:

Boot/Install Server Information

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]	
Start Frame	[1]	#
Start Slot	[1]	#
Node Count	[1]	#
OR		
Node List	[]	
Response from Server to bootp Request	disk	+
Volume Group Name	[]	
Run setup_server?	no	+

F1=Help	F2=Refresh	F3=Cancel	F4=List
F5=Reset	F6=Command	F7=Edit	F8=Image
F9=Shell	F10=Exit	Enter=Do	

Alternatively, issue the `spbootins` command:

```
# /usr/lpp/ssp/bin/spbootins -r disk -s no 1 1 1
```

Make sure that you do not need to run the `setup_server` command.

16.5.2 Listing service key information

To get the contents of a service key file, there are two commands available:

- `ksrvutil`
- `k4list`

Using the *ksrvutil* command

The *ksrvutil* command shows you the key version, the key itself, and the principal in the service key file. To do this, issue the command with the *-k* flag and the *list* operand:

```
# ksrvutil -k list
Version      Key          Principal
  1          e6b62fe3 8f4046f1  hardmon.sp3en0@MSC.ITSO.IBM.COM
  1          5e38435b 6bd5617a  rcmd.sp3en0@MSC.ITSO.IBM.COM
  1          8fdfbf70 9e4007ad  rcmd.sp3cw0@MSC.ITSO.IBM.COM
  1          85e62f62 c23194bc  hardmon.sp3cw0@MSC.ITSO.IBM.COM
#
```

When you issue the command without the *-k* flag, it does not display the key information.

Using the *k4list* command

The *k4list* command shows the service, instance, realm, and key version in the service key file. To do this, issue the command with the *-srvtab* flag:

```
# k4list -srvtab
Server key file: /etc/krb-srvtab
Service      Instance    Realm      Key Version
-----
hardmon      sp3en0     MSC.ITSO.IBM.COM 1
rcmd         sp3en0     MSC.ITSO.IBM.COM 1
rcmd         sp3cw0     MSC.ITSO.IBM.COM 1
hardmon      sp3cw0     MSC.ITSO.IBM.COM 1
#
```

When you issue the command without the *-srvtab* flag, it displays the contents of the ticket cache file instead.

16.5.3 Changing service keys

There are several reasons when you need to change the service keys.

- When you believe system security has been compromised.
- If your system policy requires changing the service keys in every certain period.
- You need to ensure that the service keys contained in the authentication database and in the server key files on all workstations and SP nodes always match.

To change service keys, issue the `ksrvutil` command with `change` operand on each node or CWS where the service keys need to be changed. The command uses a randomly generated password to derive the new service key. The service keys are updated in the `/etc/krb-srvtab` file on the local machine and in the authentication database on the CWS.

The following shows a sample operation:

```
# ksrvutil -k list
Version      Key          Principal
  1      3b94cd73 a29b9462  hardmon.sp5en0@MSC.ITSO.IBM.COM
  1      8529dc2f d6d675e5  rcmd.sp5en0@MSC.ITSO.IBM.COM
# kdb_util dump /tmp/kdbdump
# cat /tmp/kdbdump | grep sp5en0
hardmon sp5en0 255 3 1 0 fc3a34f9 cfe0bb78 203801010459 199905191424 root admin
rcmd sp5en0 255 3 1 0 352b77a2 5ba8ab22 203801010459 199905191424 root admin
# ksrvutil -k change

Principal: hardmon.sp5en0@MSC.ITSO.IBM.COM; version 1
Changing to version 2.
Old key: 3b94cd73 a29b9462; new key: 34da0440 f4947929
Key changed.

Principal: rcmd.sp5en0@MSC.ITSO.IBM.COM; version 1
Changing to version 2.
Old key: 8529dc2f d6d675e5; new key: 8cd6b379 4ce02031
Key changed.
Old keyfile in /etc/krb-srvtab.old.
# ksrvutil -k list
Version      Key          Principal
  2      34da0440 f4947929  hardmon.sp5en0@MSC.ITSO.IBM.COM
  2      8cd6b379 4ce02031  rcmd.sp5en0@MSC.ITSO.IBM.COM
# kdb_util dump /tmp/kdbdump
# cat /tmp/kdbdump | grep sp5en0
hardmon sp5en0 255 3 2 0 6359ebd ff420 203801010459 199905232043 hardmon sp5en0
rcmd sp5en0 255 3 2 0 4f7d24e4 5c3e2704 203801010459 199905232043 rcmd sp5en0
#
```

The sample operation checks the service key version, both in the `/etc/krb-srvtab` file and in the authentication database, before and after changing the service keys. Before changing the service keys, the service key version was 1; after changing the service keys, it becomes 2. You also know each principal, `hardmon.sp5en0` and `rcmd.sp5en0`, changed its service key by itself.

If you use the `-i` flag for the `ksrvutil` command, it prompts for yes or no before changing each service key.

16.6 Ticket cache file

A *ticket cache file* stores the ticket-granting ticket (TGT) and ticket for service instance during the registration process. The file is pointed to by the KRBTKFILE environment variable. If this variable is not defined, the Kerberos tickets used by a principal are located in /tmp/tktuid, where uid specifies the AIX user identification number. For example, the ticket cache file for the root user is /tmp/tkt0.

These tickets may require administration, and this section explains how to get TGT automatically and how to destroy tickets.

16.6.1 Getting a ticket-granting ticket automatically

After having authenticated as an AIX user, the normal way of getting TGT is to issue the `k4init` command. The following example shows the case of root AIX user and root.admin Kerberos principal:

```
AIX Version 4
(C) Copyrights by IBM and by others 1982, 1996.
login: root
root's Password:
*****
*
*
* Welcome to AIX Version 4.3!
*
*
* Please see the README file in /usr/lpp/bos for information pertinent to
* this release of the AIX Operating System.
*
*
*****

# k4init root.admin
Kerberos Initialization for "root.admin"
Password:
#
```

Another automated way of getting TGT is to integrate a Kerberos authentication process to the regular AIX log in process. To achieve this, perform the following steps:

Step 1: Register a Kerberos authentication program

You need to add a Kerberos authentication program to an AIX log in configuration file. In this case, the program is /usr/kerberos/bin/kinit, also

known as the `k4init` command. Modify the `login.cfg` file located in the `/etc/security` directory. The following is an excerpt from this file:

```
*****
*
* Authentication methods:
*
* auth_method:
*     program = /any/program
*
* auth_method corresponds to a custom authentication method specified in an
* AUTH1 or AUTH2 attribute in /etc/security/user, and /any/program is the
* program to run in order to do the authentication.
*
*****

*auth_method:
*     program =
KRB:
    program = /usr/kerberos/bin/kinit
```

A stanza called `KRB` is added to the file. The `program` attribute specifies the program to be executed including the full path name. You can do this using a text editor.

Step 2: Integrate root user log in process

After having registered the Kerberos authentication program, you need to integrate the program to the root user's log in process so that it is executed every time the root user logs in to the AIX. The principal used by the `k4init` command is `root.admin`. To achieve this, you need to issue the `chsec` command.

```
# chsec -f /etc/security/user -s root -a auth1="SYSTEM,KRB;root.admin"
#
```

You can verify the changes made by looking at the user file located in the `/etc/security` directory. The following is an excerpt from it:

```
root:
    admin = true
    SYSTEM = "compat"
    loginretries = 0
    account_locked = false
    auth1 = SYSTEM,KRB;root.admin
```

The following is an explanation about auth1 and auth2 parameters excerpted from the /etc/security/user file:

```
* auth1      Defines primary authentication methods for a user. This
*            attribute describes Version 3 style authentication methods.
*            Commands login, telnet, rlogin, and su support these
*            authentication methods.
*
*            Possible values: SYSTEM,NONE,token;username.
*
*            SYSTEM : Describes normal password authentication in
*                    Version 3. Version 4 has extended this
*                    definition to include loadable modules and
*                    an authentication grammar. See SYSTEM
*                    attribute description below.
*
*            NONE   : No authentication.
*
*            token;username : A generic name for a custom
*                    authentication method defined in
*                    /etc/security/login.cfg.
*
*            Example:
*            If auth1 is:
*                auth1 = SYSTEM,mylogin;mary
*
*            And the stanza in /etc/security/login.cfg is:
*                mylogin:
*                    program = /etc/myprogram
*
*            This will do password authentication, and then
*            invoke the program /etc/myprogram with "mary"
*            as the first parameter.
*
* auth2      Defines the secondary authentication methods for a user.
*            It is not a requirement to pass this method to login.
*            See auth1 description above for examples.
```

Step 3: Verify the log in process

Verify the log in process. Now, it asks you for a password for an AIX user and then a Kerberos principal:


```

AIX Version 4
(C) Copyrights by IBM and by others 1982, 1996.
login: root
root's Password:
Kerberos Initialization for "root.admin"
Password:
*****
*
*
* Welcome to AIX Version 4.3!
*
*
* Please see the README file in /usr/lpp/bos for information pertinent to
* this release of the AIX Operating System.
*
*
*****
#

```

The ticket lifetime is set to 30 days by default. So, once you are authenticated as a Kerberos principal, you normally do not need to issue the `k4init` command for 30 days. Kerberos uses the ticket cache file, `/etc/tkt0` (for `root.admin`), instead of asking you for a password. Issuing the `k4init` command is for getting a new TGT, at the same time it destroys all the old tickets. So, even you have still valid tickets, issuing the `k4init` command every time you log in improves the security of your SP system.

16.6.2 Destroying tickets

Once you issue the `k4init` command, Kerberos creates a ticket cache file. By default setting, the ticket lifetime is set to 30 days. It means even you log off your AIX user, valid tickets are still on the system. It might degrade your SP system security.

There are two ways to destroy current tickets:

- Issue the `k4destroy` command. This destroys Kerberos tickets in the current ticket cache file.
- Issue the `k4init` command. This will override existing TGT with new TGT.

16.7 Configuration

If you have damage on your authentication database, you can initialize it or load a authentication database dump file. These topics are discussed in 16.3, "Authentication database" on page 452, but the authentication database is

only a part of Kerberos system of you SP system. If you have damage on areas other than the authentication database itself and you can not find out what is the problem, you may need to reconfigure the whole Kerberos system. To prepare for unexpected disaster, you may need to back up the entire Kerberos system. This section describes how to manage Kerberos system configuration. It covers configuring, unconfiguring, backing up, and restoring the Kerberos system.

This section also shows you unique Kerberos configuration. You can integrate external RS/6000 system to the Kerberos system of your SP system. This means you can use Kerberized remote command to/from external RS/6000 systems. This gives you an advanced system management method.

16.7.1 Configuring the Kerberos system

This task is normally part of the PSSP installation process. However there can be some reasons for performing it again. An example is a complete authentication database crash on the CWS, being the only authentication server for your SP system. Another scenario would be the decision to use the CWS as a primary authentication server currently using other machines for this.

This section assumes that you are going to configure the Kerberos system using the CWS as the primary authentication server and Kerberos Version 4 provided with PSSP.

Step 1: Unconfiguring the Kerberos system

If your SP system has a certain history of being an authentication server, first of all, unconfigure the Kerberos system. To learn how to do this, refer to 16.7.2, “Unconfiguring the Kerberos system” on page 490.

Step 2: Setup the authentication server

To set up the authentication server, you usually need to issue several commands. To make this operation simple, PSSP provides you with the `setup_authent` script as a wrapper. The reason for this is that the script enables you to configure a more complex Kerberos system, for example, using the Andrew File System (AFS), if required.

The `setup_authent` script uses three steps to perform its task:

1. Create the authentication database and set up Kerberos daemons.

In this step, the `setup_authent` script issues the `kdb_init` and `kstash` commands.

2. Add an administrative principal (`root.admin`).

In this step, the `setup_authent` script issues the `kdb_edit` command.

3. Log in as `root.admin` principal and configure authentication server.

In this step, the `setup_authent` script issues the `k4init` command and the `setup_server` command.

Issue the `setup_authent` command. Figure 98 is a sample output for the first step:

```
# setup_authent
*****
                Creating the Kerberos Database
*****

Invoking the kdb_init and kstash utilities to create the database.

You must decide on a master password for the database. You will be
prompted to enter it twice. Save this password in a very secure
place, since it is used to encrypt all keys in the database and you
will need it for other administrative tasks.

After you complete this task, the Kerberos daemons will be started:
kerberos for ticket-granting services, kadmind for administration.

For more information see the kdb_init and kstash man pages.
*****
You will be prompted for the database Master Password.
It is important that you NOT FORGET this password.

Enter Kerberos master key:

Enter Kerberos master key:

0513-004 The Subsystem or Group, kerberos, is currently inoperative.
0513-083 Subsystem has been Deleted.
0513-071 The kerberos Subsystem has been added.
0513-059 The kerberos Subsystem has been started. Subsystem PID is 21066.
0513-004 The Subsystem or Group, kadmind, is currently inoperative.
0513-083 Subsystem has been Deleted.
0513-071 The kadmind Subsystem has been added.
0513-059 The kadmind Subsystem has been started. Subsystem PID is 17246.
```

Figure 98. The `setup_authent` script (1 of 3)

The `setup_authent` script prompts you for the Kerberos master key. To be precise, you enter the Kerberos master password. The script derives the Kerberos master key from the Kerberos master password. Enter the Kerberos master password twice as requested by the script. Do not forget this password, because the authentication database information will be encrypted with it.

The Kerberos daemons, `kerberos` and `kadmind`, are added to the SRC and started.

Figure 99 is a sample output for the second step:

```
*****
          Defining an Administrative Principal to Kerberos

The kdb_edit utility is used to define the initial Kerberos users. You
must define a user whose UID is 0 as a Kerberos database administrator.
This user will have to login to Kerberos with this name prior to
performing installation tasks that result in execution of the
setup_server command, during installation or whenever network interfaces
have been added or renamed in the SP system configuration.

kdb_edit prompts you separately for the name and the instance. First
enter the user name, specifying the login name of the user who will be
the primary Kerberos administrator for the local realm. When you are
prompted for the instance, you must enter admin. You must assign a
Kerberos password for this user and enter it twice (you may use the AIX
login password). To take default values on other options, hit <Enter>.

You may create any number of other Kerberos principals at this time.
To exit kdb_edit, hit <Enter> when prompted for another principal name.
For more information see the kdb_edit man page.
*****
Opening database...
Previous or default values are in [brackets] ,
enter return to leave the same, or new value.

Principal name: root
Instance: admin

<Not found>, Create [y] ?

Principal: root, Instance: admin, kdc_key_ver: 1
New Password:
Verifying, please re-enter
New Password:

Principal's new key version = 1
Expiration date (enter yyyy-mm-dd) [ 2037-12-31 ] ?
Max ticket lifetime [ 255 ] ?
Attributes [ 0 ] ?
Edit O.K.
Principal name:
```

Figure 99. The `setup_authent` script (2 of 3)

The `setup_authent` script prompts you to add an administrative principal. Its name must be `root`, and its instance must be `admin`. Both are concatenated by a dot (`.`) in the Kerberos nomenclature. This principal is associated with AIX root user.

Because the authentication database is completely new, in other words the root.admin principal has not yet added, reply `y` or just hit the **Enter** key for the `<Not found>`, `Create [y] ?` prompt.

You will be prompted for the password for the root.admin principal. Enter it twice for verification. Do not confuse it with the Kerberos master password. This password is used to be authenticated as root.admin. The Kerberos master password is used to create the Kerberos master key.

The script asks you to enter the expiration date, the maximum ticket lifetime, and the attributes. You may enter your favorite value or use the default value. You can change these values later. To do this, refer to 16.4.7, “Changing expiration date” on page 470 or 16.4.8, “Changing maximum ticket lifetime” on page 471.

The Principal name: prompt asks you to enter another principal name. If you do not have any more principals to be added at this point, press the **Enter** key. This terminates the second step.

Figure 100 on page 488 is a sample output for the third step.

```

*****
Logging into Kerberos as an admin user

You must assume the role of a Kerberos administrator <user>.admin to
complete the initialization of Kerberos on the local system. The kinit
command is invoked and will prompt you for the password. If you are
setting up your primary server here, you have just defined it. If you
have defined multiple administrative principals, or if your primary
authentication server is on another system, you must first enter the
name of an administrative principal who has root privilege (UID 0). You
need to be authenticated as this administrator so that this program
can create the principals and service key files for the authenticated
services that run on the SP system.

For more information, see the kinit man page.
*****
Kerberos Initialization for "root.admin"
Password:
setup_server: Running services_config script to configure SSP services.This may
take a few minutes...
rc.ntp: Starting ntp daemon(xntpd)
0513-029 The supfilesrv Subsystem is already active.
Multiple instances are not supported.
setup_CWS: Control Workstation setup complete.
mknimmast: Node 0 (sp5en0) already configured as a NIM master.
create_krb_files: tftpaccess.ctl file and client srvtab files created/updated
on server node 0.
mknimres: Copying /usr/lpp/ssp/install/bin/pssp_script to /spdata/sys1/install/p
ssp/pssp_script.
mknimres: Copying /usr/lpp/ssp/install/config/bosinst_data.template to /spdata/s
ys1/install/pssp/bosinst_data.
mknimres: Copying /usr/lpp/ssp/install/config/bosinst_data_prompt.template to /s
pdata/sys1/install/pssp/bosinst_data_prompt.
mknimres: Copying /usr/lpp/ssp/install/config/bosinst_data_migrate.template to /
spdata/sys1/install/pssp/bosinst_data_migrate.
mknimclient: 0016-242: Client node 1 (sp5n01.msc.itso.ibm.com) already defined o
n server node 0 (sp5en0).
mknimclient: 0016-242: Client node 5 (sp5n05.msc.itso.ibm.com) already defined o
n server node 0 (sp5en0).
mknimclient: 0016-242: Client node 9 (sp5n09.msc.itso.ibm.com) already defined o
n server node 0 (sp5en0).
mknimclient: 0016-242: Client node 13 (sp5n13.msc.itso.ibm.com) already defined
on server node 0 (sp5en0).
export_clients: File systems exported to clients from server node 0.
allnimres: Node 1 (sp5n01.msc.itso.ibm.com) prepared for operation: customize.
allnimres: Node 5 (sp5n05.msc.itso.ibm.com) prepared for operation: customize.
allnimres: Node 9 (sp5n09.msc.itso.ibm.com) prepared for operation: customize.
allnimres: Node 13 (sp5n13.msc.itso.ibm.com) prepared for operation: customize.
setup_server: Processing complete (rc= 0).
setup_authent: All previously installed nodes have been set up to be customized
- ready for re-boot.
#

```

Figure 100. The setup_authent script (3 of 3)

In the third step, the `setup_authent` command will configure the authentication server using the data in the System Data Repository (SDR).

To be authenticated as an administrative principal, enter the password for the `root.admin` principal you entered in the second step.

Then the script issues the `setup_server` command, and finally, you get the shell prompt.

At this point, the following configuration is completed:

- The ticket cache file for `root.admin` is created. Check this by the `k4list` command:

```
# k4list
Ticket file:  /tmp/tkt0
Principal:    root.admin@MSC.ITSO.IBM.COM

   Issued           Expires           Principal
May 19 18:09:30  Dec 26 18:09:30  krbtgt.MSC.ITSO.IBM.COM@MSC.ITSO.IBM.COM
#
```

- The following Kerberos files are created:
 - `/.k`
 - `/.klogin`
 - `/etc/krb.conf`
 - `/etc/krb.realms`
 - `/etc/krb-srvtab`
 - `/var/kerberos/database/admin_acl.add`
 - `/var/kerberos/database/admin_acl.get`
 - `/var/kerberos/database/admin_acl.mod`
 - `/var/kerberos/database/principal.dir`
 - `/var/kerberos/database/principal.ok`
 - `/var/kerberos/database/principal.pag`
- All the nodes are set to customize mode. Check this by issuing the `splstdata` command.

```

# splstdata -b
List Node Boot/Install Information

node#      hostname  hdw_enet_addr  srvr  response  install_disk
last_install_image  last_install_time  next_install_image  lppsource_name
pssp_ver          selected_vg
-----
  1 sp5n01      02608CF57A7C  0  customize  hdisk0
    default Thu_May_13_16:07:49  default  aix432
    PSSP-3.1          rootvg
  5 sp5n05      02608CE880F1  0  customize  hdisk0
    default Thu_May_13_16:04:14  default  aix432
    PSSP-3.1          rootvg
  9 sp5n09      02608C2D08D7  0  customize  hdisk0
    default Thu_May_13_16:06:48  default  aix432
    PSSP-3.1          rootvg
 13 sp5n13      02608CE8FCB3  0  customize  hdisk0
    default Thu_May_13_16:06:19  default  aix432
    PSSP-3.1          rootvg
#

```

Step 3: Customize the nodes

The nodes are not available to use the Kerberos system yet. To make them ready, you need to customize the nodes. To do this, network boot the nodes or issue the `pssp_script` script on each node. Refer to 5.1.3, “Customizing a node with or without rebooting” on page 164, for this operation.

At this moment, you can not use any of the remote commands if you are not using the `/.rhosts` file.

Step 4: Distribute the `/.klogin` file

Distribute the `/.klogin` file on the authentication server (CWS) to all the nodes so that the `root.admin` principal and the other available principals can be authenticated.

At this moment, you still can not use any of the remote commands if you are not using the `/.rhosts` file.

16.7.2 Unconfiguring the Kerberos system

There may be some reasons for you to unconfigure the Kerberos system in your SP system. For example, your Kerberos system has catastrophic damage, and you need to recreate it. Or, you decided to assign primary authentication server to the other machine.

This section provides you with the information on how to unconfigure the Kerberos system in your SP system.

Step 1: Back up the Kerberos system

Before you unconfigure the Kerberos system, make backup files. It saves you time if you have to face a fall back. Refer to 16.7.3, “Backing up the Kerberos system automatically” on page 492, to learn how to do this.

Step 2: Delete the Kerberos files on nodes

Delete the following Kerberos files on the nodes using the `rm` command:

- `/.klogin`
- `/etc/krb-srvtab`
- `/etc/krb.conf`
- `/etc/krb.realms`
- `/tmp/tkt*`

Step 3: Stop the Kerberos daemons

Stop the Kerberos daemons, `kadmind` and `kerberos`, by issuing the `stopsrc` command:

```
# stopsrc -s kadmind
0513-044 The stop of the kadmind Subsystem was completed successfully.
# stopsrc -s kerberos
0513-044 The stop of the kerberos Subsystem was completed successfully.
#
```

Step 4: Delete your current tickets

Delete your current tickets by issuing the `k4destroy` command:

```
# k4destroy
Tickets destroyed.
#
```

Step 5: Delete the Kerberos files on the authentication server

Delete the following Kerberos files on the authentication server by issuing the `rm` command:

```
# rm /.k
# rm /.klogin
# rm /etc/krb.conf
# rm /etc/krb.realms
# rm /etc/krb-srvtab
# rm /tmp/tkt*
#
```

The authentication database is located in the `/var/kerberos/database` directory. Delete all the files in this directory by issuing the `rm` command:

```
# rm /var/kerberos/database/*
#
```

16.7.3 Backing up the Kerberos system automatically

When your Kerberos system has catastrophic damage, you may need to configure the Kerberos system from scratch as described in 16.7.1, “Configuring the Kerberos system” on page 484. After configuration, you need to customize the Kerberos system to make it as it was. To avoid this operation, it is a better idea to create back up files occasionally.

This section provides you with a sample shell script to perform this operation instead of step by step instruction. Using the shell script is of a great value, because it can be added to the crontab of the primary authentication server and be executed in regular intervals. This automates the backup of the Kerberos system.

Step 1: Create a directory for backups

You need a place to create backup files. This section uses the subdirectory named `my_backup` in the authentication database directory. To create this subdirectory, issue the `mkdir` command:

```
# mkdir -p /var/kerberos/database/my_backup
```

Step 2: Back up all necessary files

The following shell script can be used for the backup. The script makes a copy of all necessary files for the Kerberos system. This provides the easiest way to restore the Kerberos system at a working level:

```

#!/usr/bin/ksh

# Set constant values

CRED=$RANDOM
BACKUPDIR=/var/kerberos/database/my_backup/$CRED
KDBDIR=/var/kerberos/database
LOGFILE=$BACKUPDIR/backup.log.$CRED
ERROR=0

# Create back up directory if not existent

if [[ ! -d $BACKUPDIR ]]
then
    mkdir -p $BACKUPDIR
fi

# Use time stamp for log

print "\n\n" > $LOGFILE
date >> $LOGFILE

# Copy all Kerberos files to /var/kerberos/database/my_backup/$CRED

for file in /.k /.klogin /etc/krb-srvtab /etc/krb.realms /etc/krb.conf \
    $KDBDIR/admin_acl.add $KDBDIR/admin_acl.mod $KDBDIR/admin_acl.get \
    $KDBDIR/principal.dir $KDBDIR/principal.pag $KDBDIR/principal.ok
do
    cp $file $BACKUPDIR
    if [[ $? != 0 ]]
    then
        ERROR=1
        print "ERROR: Copying $file to $BACKUPDIR\n" >> $LOGFILE
    fi
done

# Dump an ASCII version of the authentication database

/usr/lpp/ssp/kerberos/etc/kdb_util dump $BACKUPDIR/my_dump

# Complete the log file and return with OK

if [[ $ERROR -eq 0 ]]
then
    print "INFO: Kerberos system backup completed without errors\n" >> $LOGFILE
    return 0
fi

# Return with error

return 1

```

The script creates the `my_backup/credential` directory in the authentication database directory of the authentication server, where *credential* is a random number.

A log file records the script's activity using a time stamp. The name of the file is `backup.log.credential`.

The script copies the following files to the backup directory:

- `/.k`
- `/.klogin`
- `/etc/krb-srvtab`
- `/etc/krb.realms`
- `/etc/krb.conf`
- `/var/kerberos/database/admin_acl.add`
- `/var/kerberos/database/admin_acl.mod`
- `/var/kerberos/database/admin_acl.get`
- `/var/kerberos/database/principal.dir`
- `/var/kerberos/database/principal.pag`
- `/var/kerberos/database/principal.ok`

The script dumps an ASCII version of the authentication database to the backup directory also. The script uses `my_dump` for the name of the dump file.

Finally, a message is written to the log file if all went fine. If the script has problems copying files, this will be written to the log file as well.

Step 3: Add the script to the crontab

Add the script to the crontab and choose a reasonable interval.

16.7.4 Restoring the Kerberos system

There may be hard times when the Kerberos system is messed up completely. There are two ways for recovery. One way is to configure the Kerberos system from scratch as described in 16.7.1, "Configuring the Kerberos system" on page 484. The other way is to restore the Kerberos system using the back up files described in 16.7.3, "Backing up the Kerberos system automatically" on page 492, if you created.

This section shows the second way, restoring the Kerberos system using the back up files created in the `/var/kerberos/database/backup/42` directory, for example.

Step 1: Stop the Kerberos daemons

Before you start restoring the Kerberos system, make sure you stop the Kerberos daemons, `kerberos` and `kadmind`, by issuing the `stopsrc` command:

```
# stopsrc -s kadmind
0513-044 The stop of the kadmind Subsystem was completed successfully.
# stopsrc -s kerberos
0513-044 The stop of the kerberos Subsystem was completed successfully.
#
```

Step 2: Restore all necessary files

Restore all necessary files from your back up directory. Issue the `cp` command to perform this operation:

```
# cp /var/kerberos/database/my_backup/42/.k /
# cp /var/kerberos/database/my_backup/42/.klogin /
# cp /var/kerberos/database/my_backup/42/krb* /etc
# cp /var/kerberos/database/my_backup/42/admin_acl.* /var/kerberos/database
# cp /var/kerberos/database/my_backup/42/principal.* /var/kerberos/database
#
```

You may load the ASCII version of authentication database dump file, `my_dump`, by issuing the `kdb_util` command instead of copying the `principal.dir` and `prindipal.pag` files.

Step 3: Start the Kerberos daemons

After restoring all necessary files, start the Kerberos daemons, `kerberos` and `kadmind`, by issuing the `startsrc` command:

```
# startsrc -s kerberos
0513-059 The kerberos Subsystem has been started. Subsystem PID is 14972.
# startsrc -s kadmind
0513-059 The kadmind Subsystem has been started. Subsystem PID is 19750.
#
```

16.7.5 Adding an external RS/6000 system to Kerberos realm

If you can issue a Kerberized remote command from the CWS or nodes to external RS/6000 systems, it might be useful. If you can issue a Kerberized remote command to the CWS or nodes from external RS/6000 systems, it might be useful also.

If you create a principal for your external RS/6000 systems, you can implement this environment. The principal must be added to the authentication database, must have the service key file, and must be associated to the Kerberos realm. A *domain* is a collection of systems over

which an administrator exercises control. A *realm* is a domain that shares an authentication database and servers. There is a single name-space for principal name/instance parts within a realm. A realm is also a logical collection of clients and servers registered in the authentication database.

This section assumes that your SP system has one CWS and four nodes. They are connected by single SP Ethernet segment. The CWS has a Token-Ring adapter also. The CWS and your external RS/6000 system are connected by a Token-Ring network. You are going to integrate the external RS/6000 system to your Kerberos realm and make it possible to issue the `dash` command on all your system.

Table 14 shows the adapter type, host name, IP address, IP domain name, and Kerberos realm name for each network adapter:

Table 14. Network adapter attributes

Adapter Type	Host Name	IP Address	IP Domain	Kerberos Realm
SP Ethernet	sp5n01	192.168.5.1	msc.itso.ibm.com	MSC.ITSO.IBM.COM
SP Ethernet	sp5n05	192.168.5.5	msc.itso.ibm.com	MSC.ITSO.IBM.COM
SP Ethernet	sp5n09	192.168.5.9	msc.itso.ibm.com	MSC.ITSO.IBM.COM
SP Ethernet	sp5n13	192.168.5.13	msc.itso.ibm.com	MSC.ITSO.IBM.COM
SP Ethernet	sp5en0	192.168.5.0	msc.itso.ibm.com	MSC.ITSO.IBM.COM
Token-Ring	sp5cw0	9.12.0.5	itso.ibm.com	MSC.ITSO.IBM.COM
Token-Ring	rs6k	9.12.0.66	itso.ibm.com	Will be assigned as MSC.ITSO.IBM.COM

Figure 101 on page 497 shows the physical location and connection of the network adapters.

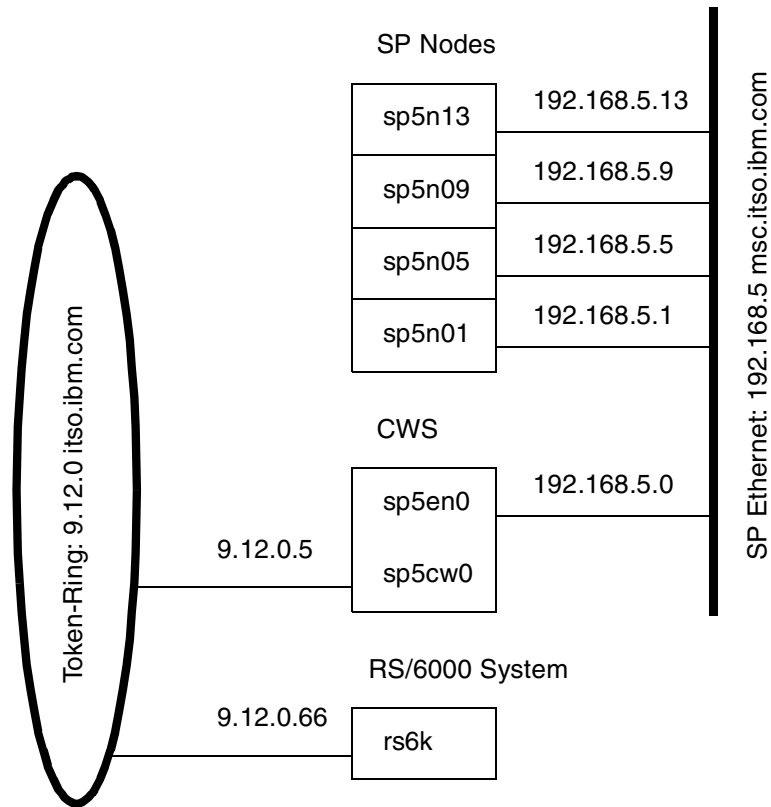


Figure 101. Sample configuration

The entire steps in this section can be categorized into three major parts:

1. Preparing an external RS/6000 system
2. Enabling the remote command from SP system to the external RS/6000 system
3. Enabling the remote command from the external RS/6000 system to the SP system

16.7.5.1 Preparing an external RS/6000 system

All the operations need to be performed on the external RS/6000 system.

Step 1: Instal the required file set

To make your external RS/6000 system ready for the Kerberos environment, you need to install the following file sets:

- ssp.clients
- ssp.perlpkg

They are included in PSSP 3.1. Make sure your AIX version is 4.3.2 or greater. The ssp.clients file set includes SP Authenticated Client Commands, and the ssp.perlpkg file set includes SP PERL Distribution Package. The ssp.perlpkg is required because the `dsh` command is written by Perl.

Step 2: Enable authentication methods

You need to enable Kerberos Version 4 as an authentication method. Check the current methods issuing the `lsauthent` command, then issue the `chauthent` command to enable Kerberos Version 4:

```
# lsauthent
Standard Aix
# chauthent -k4 -std
# lsauthent
Kerberos 4
Standard Aix
#
```

Step 3: Check the routing table

Normally, SP nodes use the CWS as default gateway; so, it may not have IP forwarding problem from SP nodes to the external RS/6000 system. On the other hand, you may need to add routing information to the routing table of the external RS/6000 system. Check the current routing information by issuing the `netstat` command, then add the routing information by issuing the `route` command.


```

# netstat -rn
Routing tables
Destination      Gateway          Flags  Refs      Use  If  PMTU  Exp  Groups

Route Tree for Protocol Family 2 (Internet):
default          9.12.0.1        UG     1         3295  tr0  -    -
9.12/24          9.12.0.66       U      22        39272  tr0  -    -
127/8            127.0.0.1       U       3          559   lo0  -    -

Route Tree for Protocol Family 24 (Internet v6):
::1              ::1             UH     0          0    lo0 16896  -
# route add -net 192.168.5.0 9.12.0.5
9.12.0.5 net 192.168.5.0: gateway 9.12.0.5
# netstat -rn
Routing tables
Destination      Gateway          Flags  Refs      Use  If  PMTU  Exp  Groups

Route Tree for Protocol Family 2 (Internet):
default          9.12.0.1        UG     1         3418  tr0  -    -
9.12/24          9.12.0.66       U      23        39272  tr0  -    -
127/8            127.0.0.1       U       3          559   lo0  -    -
192.168.5/24    9.12.0.5        UG     0           7    tr0  -    -

Route Tree for Protocol Family 24 (Internet v6):
::1              ::1             UH     0          0    lo0 16896  -
#

```

16.7.5.2 Enabling the remote command from SP system

All the operations need to be performed on the authentication server, that is, the CWS.

Step 4: Create a principal for the rcmd service

You want to issue the remote command on all of your system. To make this possible, you need to add a principal. Its name is rcmd, and its instance is rs6k. The rcmd.rs6k principal is a service of the rcmd and an instance of the Token-Ring adapter in the external RS/6000 system. The rcmd service principal is one of the two default service principals added to the authentication database during PSSP installation. For more detail, refer to 16.2.4, “Kerberos authenticated-applications” on page 451.

To add the rcmd.rs6k principal to the authentication database, issue the `kadmin` command and verify it by issuing the `lskp` command.

```

# kadmin
Welcome to the Kerberos Administration Program, version 2
Type "help" if you need it.
admin: ank rcmd.rs6k
Admin password:
Password for rcmd.rs6k:
Verifying, please re-enter Password for rcmd.rs6k:
rcmd.rs6k added to database.
admin: quit
Cleaning up and exiting.
# lskp rcmd.rs6k
rcmd.rs6k          tkt-life: Unlimited key-vers: 1  expires: 2037-12-31 23:59
#

```

For the Password for rcmd.rs6k: prompt, you can enter any word. The password is used to derive the service key, and you can randomly change it by using the `ksrvutil` command later. For changing service keys, refer to 16.5.3, “Changing service keys” on page 478.

For adding a principal, refer to 16.4.3, “Adding a principal” on page 461. For listing a principal, refer to 16.4.1, “Listing principal” on page 459.

Step 5: Extract the service keys from the authentication database

The principal is added to the authentication database. Now the service keys for the external RS/6000 system have to be extracted from the authentication database.

The purpose of this step is the same as described in 16.5.1, “Creating a service key file” on page 475. But, you can not use this method in this case because the method refers to the SDR. Your external RS/6000 system is not part of the SP system; therefore, it is not part of the SDR.

Instead, the `ext_srvtab` command extracts the service keys to a file from the authentication database. Move to any directory you want to save the service key file. This example uses the `/tftpboot` directory. Then issue the `ext_srvtab` command with the `-n` flag. The flag specifies the instance you want to extract. The following is an example operation:

```

# cd /tftpboot
# ext_srvtab -n rs6k
Generating 'rs6k-new-srvtab' ....
# ksrvutil -k list -f rs6k-new-srvtab
Version      Key              Principal
-----
1           49fe0475 9262ab76  rcmd.rs6k@MSC.ITSO.IBM.COM
#

```

The command creates the `rs6k-new-srvtab` service key file in the current directory. This example checks the contents of the service key file by issuing the `ksrvutil` command with the `-f` flag. If you do not use this flag, the command refers to `/etc/krb-srvtab` file instead of `./rs6k-new-srvtab`.

Step 6: Modify the Kerberos files on the authentication server

The authentication database has information about the `rcmd.rs6k` principal. There are two more files to be associated with the principal:

- `/.klogin`
- `/etc/krb.realms`

The `/.klogin` file specifies remote principals that can use a local user account. Add the `rcmd.rs6k` principal, then check it as follows:

```
# print "rcmd.rs6k@MSC.ITSO.IBM.COM" >> /.klogin
# cat /.klogin
root.admin@MSC.ITSO.IBM.COM
rcmd.sp5en0@MSC.ITSO.IBM.COM
rcmd.sp5n01@MSC.ITSO.IBM.COM
rcmd.sp5n05@MSC.ITSO.IBM.COM
rcmd.sp5n09@MSC.ITSO.IBM.COM
rcmd.sp5n13@MSC.ITSO.IBM.COM
rcmd.rs6k@MSC.ITSO.IBM.COM
#
```

The word `MSC.ITSO.IBM.COM` is a name of a realm. Do not confuse it with a domain name for an IP address.

The `/etc/krb.realms` file specifies the translations from host names to authentication realms. Add the `rs6k.itso.ibm.com` host name, then check it as follows:

```
# print "rs6k.itso.ibm.com MSC.ITSO.IBM.COM" >> /etc/krb.realms
# cat /etc/krb.realms
sp5cw0.itso.ibm.com MSC.ITSO.IBM.COM
rs6k.itso.ibm.com MSC.ITSO.IBM.COM
#
```

Step 7: Distribute the Kerberos files to the nodes

To propagate the change to the nodes, distribute the modified Kerberos files:

```
# hostlist -av | pcp -w - /.klogin
# hostlist -av | pcp -w - /etc/krb.realms
#
```

Step 8: Distribute the Kerberos files to the external RS/6000

To propagate the change to the external RS/6000 system, distribute the essential Kerberos files:

- /.klogin
- /etc/krb.conf
- /etc/krb.realms
- /tftpboot/rs6k-new-srvtab

Make sure they reside in the same directories on the external RS/6000 system, except for the /tftpboot/rs6k-new-srvtab file. The /tftpboot/rs6k-new-srvtab service key file must be renamed to /etc/krb-srvtab when distributed. The following is a sample operation issuing the `ftp` command:

```

# ftp rs6k
Connected to rs6k.itso.ibm.com.
220 rs6k FTP server (Version 4.1 Tue Sep 8 17:35:59 CDT 1998) ready.
Name (rs6k:root):
331 Password required for root.
Password:
230 User root logged in.
ftp> prompt
Interactive mode off.
ftp> asc
200 Type set to A; form set to N.
ftp> put /.klogin
200 PORT command successful.
150 Opening data connection for /.klogin.
226 Transfer complete.
208 bytes sent in 0.000965 seconds (210.5 Kbytes/s)
local: /.klogin remote: /.klogin
ftp> mput /etc/krb.*
200 PORT command successful.
150 Opening data connection for /etc/krb.conf.
226 Transfer complete.
73 bytes sent in 0.001565 seconds (45.55 Kbytes/s)
local: /etc/krb.conf remote: /etc/krb.conf
200 PORT command successful.
150 Opening data connection for /etc/krb.realms.
226 Transfer complete.
75 bytes sent in 0.000525 seconds (139.5 Kbytes/s)
local: /etc/krb.realms remote: /etc/krb.realms
ftp> bin
200 Type set to I.
ftp> put /tftpboot/rs6k-new-srvtab /etc/krb-srvtab
200 PORT command successful.
150 Opening data connection for /etc/krb-srvtab.
226 Transfer complete.
37 bytes sent in 0.000519 seconds (69.62 Kbytes/s)
local: /tftpboot/rs6k-new-srvtab remote: /etc/krb-srvtab
ftp> quit
221 Goodbye.
#

```

16.7.5.3 Enabling the remote command from the RS/6000 system

All the operations need to be performed on the external RS/6000 system.

Step 9: Modify Kerberos files

To issue the `dsh` command from the external RS/6000 system, you need to modify the following Kerberos files:

- `/.klogin`
- `/etc/krb.conf`

The external RS/6000 system needs to recognize the `rcmd.sp5cw0` principal. Add the principal, then check it as follows:

```
# print "rcmd.sp5cw0@MSC.ITSO.IBM.COM" >> /.klogin
# cat /.klogin
root.admin@MSC.ITSO.IBM.COM
rcmd.sp5en0@MSC.ITSO.IBM.COM
rcmd.sp5n01@MSC.ITSO.IBM.COM
rcmd.sp5n05@MSC.ITSO.IBM.COM
rcmd.sp5n09@MSC.ITSO.IBM.COM
rcmd.sp5n13@MSC.ITSO.IBM.COM
rcmd.rs6k@MSC.ITSO.IBM.COM
rcmd.sp5cw0@MSC.ITSO.IBM.COM
#
```

The `/etc/krb.conf` file contains the SP authentication configuration. The first line contains the name of the local authentication realm. The second line contains the location of an authentication server for a realm. The following is the contents of the `/etc/krb.conf` file on the authentication server:

```
MSC.ITSO.IBM.COM
MSC.ITSO.IBM.COM sp5en0.msc.itso.ibm.com admin server
```

The authentication server (CWS) uses host name `sp5en0.msc.itso.ibm.com` for the SP Ethernet network. For the Token-Ring network, however, it uses `sp5cw0.itso.ibm.com`. Change the host name of the authentication server as follows:

```
MSC.ITSO.IBM.COM
MSC.ITSO.IBM.COM sp5cw0.itso.ibm.com admin server
```

16.7.5.4 Check if it works

To verify if it works fine, issue the `dsh` command to the external RS/6000 system from the CWS or node:

```
# k4destroy
Tickets destroyed.
# k4init root.admin
Kerberos Initialization for "root.admin"
Password:
# dsh -w rs6k date
rs6k: Sat May 22 14:08:19 EDT 1999
#
```

You need to destroy tickets first.

Then issue the `dsh` command to the CWS and nodes from the external RS/6000 system:

```
# k4init root.admin
Kerberos Initialization for "root.admin"
Password:
# /usr/lpp/ssp/bin/dsh -w sp5cw0,sp5n01,sp5n05,sp5n09,sp5n13 date
sp5cw0: Sat May 22 14:10:05 EDT 1999
sp5n01: Sat May 22 14:10:15 EDT 1999
sp5n05: Sat May 22 14:10:15 EDT 1999
sp5n09: Sat May 22 14:10:15 EDT 1999
sp5n13: Sat May 22 14:10:15 EDT 1999
#
```

The alternative to the way presented in this section would have been to issue the `setup_authent` command to integrate an external RS/6000 system into the Kerberos realm. For a description of this method, refer to Chapter 2, “Installing and Configuring a New RS/6000 SP System” in *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347.

This section has intentionally chosen the method presented here, because it gives you the opportunity to look under the covers. If you are planning to enter additional adapters to your SP system and make their network interfaces under the control of Kerberos, this method can help you.

If you are going to use Kerberos for adapters used by the IBM High Availability Cluster Multi-Processing for AIX Enhanced Scalability (HACMP/ES), refer to Chapter 6, “Configuration Examples” in *HACMP Enhanced Scalability Handbook*, SG24-5328.

Part 8. Managing shared disks

Chapter 17. IBM Recoverable Virtual Shared Disk

IBM Virtual Shared Disk is a subsystem that lets application programs executing on different nodes of a system partition access a raw logical volume as if it were local at each of the nodes. If you use *IBM Recoverable Virtual Shared Disk* component and twin-tailed disks or disk arrays in addition, you can allow a secondary node to take over the server function from the primary node when certain types of failure occur.

This chapter explains how to install IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk on the control workstation (CWS) and on the nodes. It explains how you can configure IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk by using the IBM Virtual Shared Disk Perspective.

Information is also provided on how to monitor IBM Recoverable Virtual Shared Disk on your SP system. It covers IBM Recoverable Virtual Shared Disk daemons, resource variables, and log files.

17.1 Configuration

To configure IBM Recoverable Virtual Shared Disk, at least one twin-tailed disk or disk array is required. A *twin-tailed disk* is a disk or group of disks that is attached to two nodes of an SP system. For recoverability purposes, only one of these nodes provides IBM Virtual Shared Disk server service at any given time. The secondary node takes over IBM Virtual Shared Disk server service if the primary node fails, is powered off, or if you need to change the server node temporarily for administrative reasons. Both nodes must be in the same system partition.

To learn more about configuration of IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk and IBM Virtual Shared Disk Perspective refer to the following IBM publications:

- *IBM Parallel System Support Programs for AIX: Managing Shared Disks*, SA22-7349
- Chapter 6, “The IBM Virtual Shared Disk Perspective” in *SP Perspectives: A New View of Your SP System*, SG24-5180

17.1.1 Installing IBM Virtual Shared Disk

The file sets required for IBM Virtual Shared Disk are not installed automatically. File sets listed in Table 15 should be installed on both CWS and nodes.

Table 15. File sets for IBM Virtual Shared Disk (CWS and nodes)

File Set	Description
vsd.vsd	IBM Virtual Shared Disk Device Driver
vsd.sysctl	IBM Virtual Shared Disk sysctl commands
vsd.cmi	IBM Virtual Shared Disk centralized Management Interface (SMIT)
vsd.hsd	IBM Virtual Shared Disk Hashed Shared Disk

File sets listed in Table 16 should be installed only on CWS.

Table 16. File sets for IBM Virtual Shared Disk (CWS only)

File Set	Description
ssp.vsdgui	IBM Virtual Shared Disk Graphical User Interface (IBM Virtual Shared Disk Perspective)
ssp.vsdgui.loc.<locale_name>	IBM Virtual Shared Disk GUI Locale
ssp.vsdgui.msg.<locale_name>	IBM Virtual Shared Disk GUI Messages

17.1.2 Installing IBM Recoverable Virtual Shared Disk

The file sets required to IBM Recoverable Virtual Shared Disk are not installed automatically. File sets listed in Table 17 should be installed on both CWS and nodes.

Table 17. File sets for IBM Recoverable Virtual Shared Disk (CWS and nodes)

File Set	Description
vsd.rvsd.rvsdd	IBM Recoverable Virtual Shared Disk Daemon
vsd.rvsd.hc	IBM Recoverable Virtual Shared Disk Connection Manager
vsd.rvsd.scripts	IBM Recoverable Virtual Shared Disk Recovery Scripts

17.1.3 Getting authorization

You can use IBM Virtual Shared Disk Perspective to configure IBM Virtual Shared Disk. The operations related to IBM Virtual Shared Disk configuration require that you are authenticated as a Kerberos principal and that you have an authorization to use the `sysctl` command. This means you can issue the

k4init command successfully, and your Kerberos principal is listed in the /etc/sysctl.vsd.acl Access Control List (ACL) file on the CWS and all the nodes. If you are going to use the root.admin principal, the /etc/sysctl.vsd.acl file should appear like the following:

```
# cat /etc/sysctl.vsd.acl
#acl#

# These are the users that can issue sysctl_vsdXXX command on this node
# Name must have a Kerberos name format which defines user@realm
# Please check your security administrator to fill in correct realm name
# you may find realm name from /etc/krb.conf

# _PRINCIPAL root@PPD.POK.IBM.COM
# _PRINCIPAL root.admin@PPD.POK.IBM.COM
# _PRINCIPAL rcmd@PPD.POK.IBM.COM
# _PRINCIPAL userid@PPD.POK.IBM.COM
# _PRINCIPAL root.admin@SP4ENO
#
```

You can use the `pcp` command to distribute the ACL file to all the nodes:

```
# pcp -a /etc/sysctl.vsd.acl
```

17.1.4 Starting IBM Virtual Shared Disk Perspective

Start the IBM Virtual Shared Disk Perspective by issuing the `spvsd` command:

```
# spvsd
```

Then, you will see the IBM Virtual Shared Disk Perspective window as shown in Figure 102 on page 512.

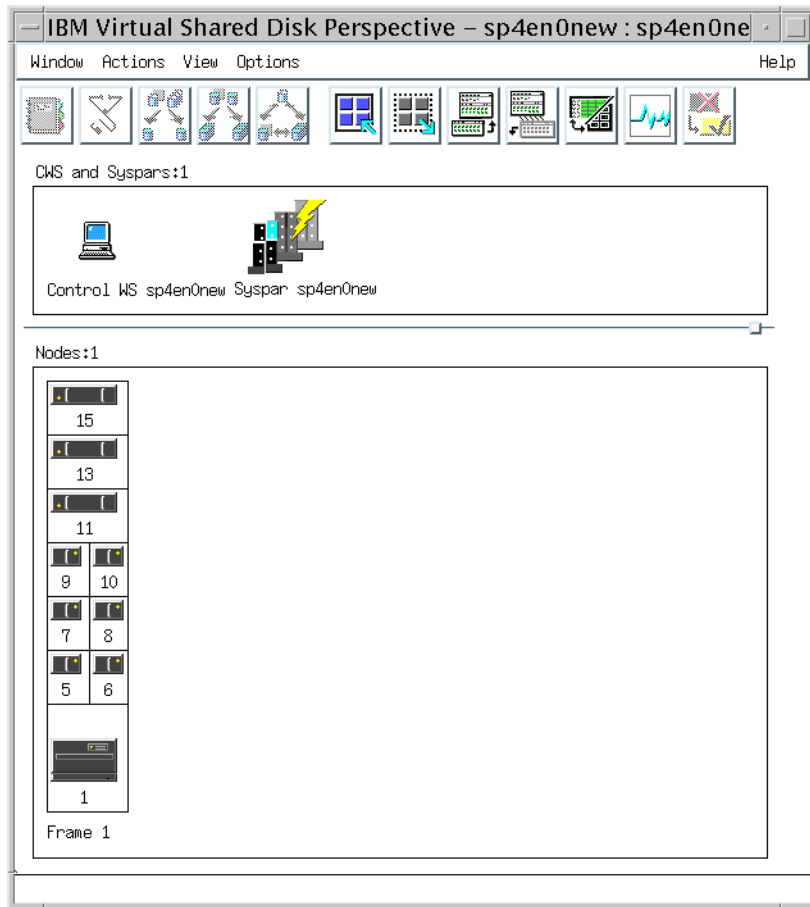












Figure 102. IBM Virtual Shared Disk Perspective

Table 18 shows the tool bar icons and their names.

Table 18. IBM Virtual Shared Disk Perspective tool bar icons

Icon	Name of Icon
	Properties
	Run diagnostics

Icon	Name of Icon
	Filter to show related nodes in a new pane
	Filter to show related IBM VSDs in a new pane
	Filter to show related IBM HSDs in a new pane
	Bring up the Filter to Show Related Objects dialog
	Select all objects
	Deselect all selected objects
	Add a pane
	Delete the current pane from this window

17.1.5 Designating a node as an IBM Virtual Shared Disk node

If you are going to use a node for IBM Recoverable Virtual Shared Disk, it is required to be designated as IBM Virtual Shared Disk node. This section uses all the nodes for IBM Recoverable Virtual Shared Disk.

To select all nodes, click on the Nodes pane and click on the **Select all objects** icon:



Then click **Actions** in the menu bar and select **Designate as an IBM VSD Node...** from the list. The Designate as an IBM VSD Node dialog box shown in Figure 103 will be displayed:

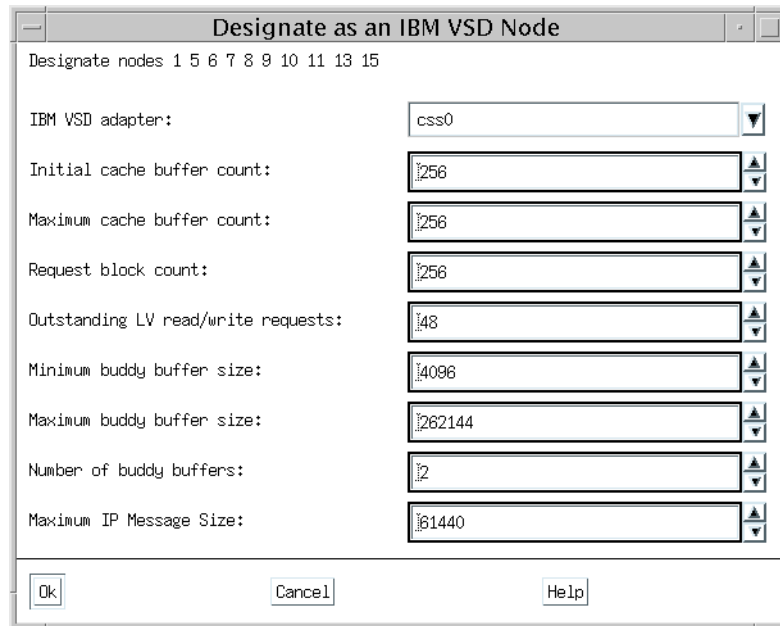


Figure 103. Designate as an IBM VSD node dialog box

When you click the **OK** button, a disk icon is added on each node icon in the Nodes pane as shown in Figure 104 on page 515:

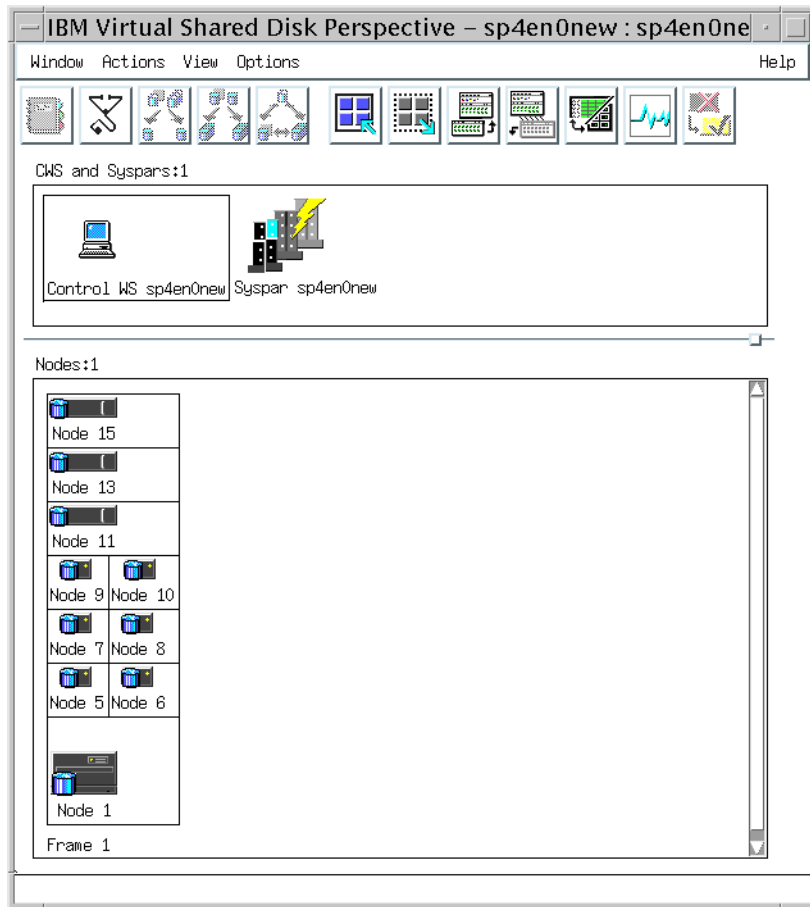


Figure 104. Designated nodes as IBM VSD nodes

17.1.6 Creating IBM Virtual Shared Disk

To create IBM Virtual Shared Disk, you need an IBM VSDs pane. If you do not have this pane, click the **Add a pane** icon:



The Add Pane dialog box shown in Figure 105 on page 516 will be displayed:

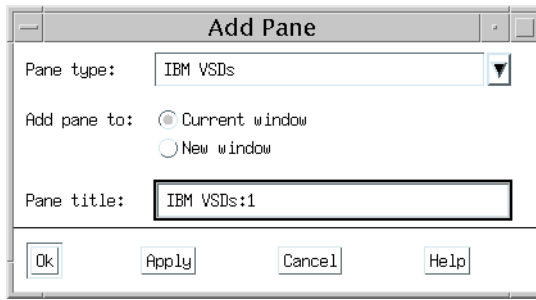


Figure 105. Add Pane dialog box

Select **IBM VSDs** in the Pane type: field. Then click **OK**. The IBM VSDs pane shown in Figure 106 on page 517 will be added to the current window.

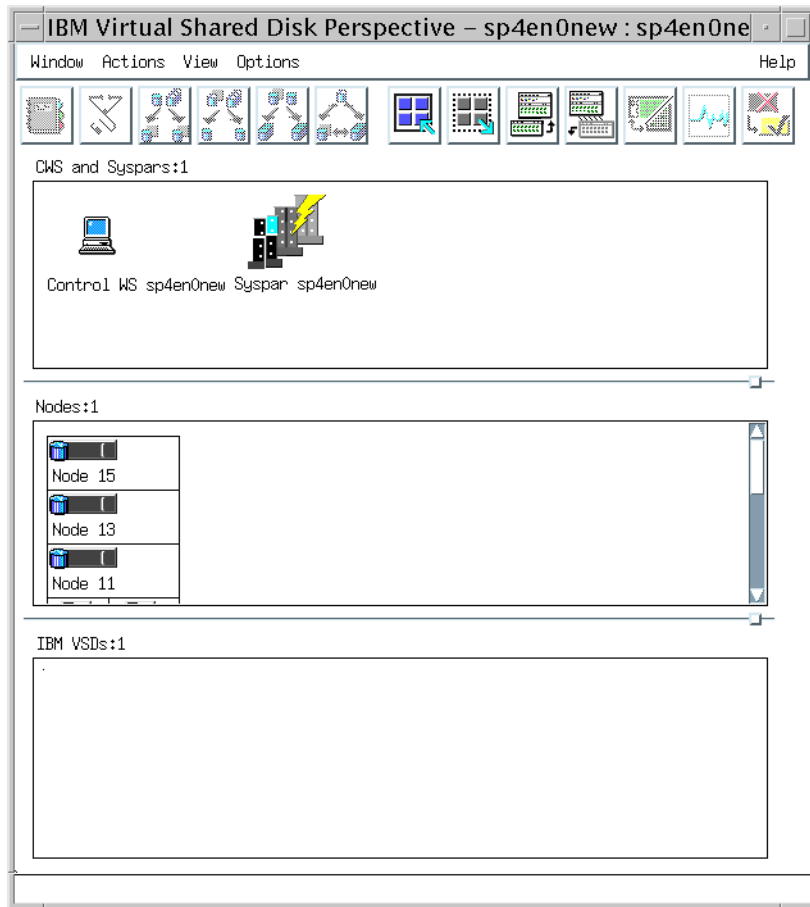


Figure 106. IBM Virtual Shared Disk Perspective with IBM VSDs pane

To create IBM Virtual Shared Disk, click on the **IBM VSDs** pane to give it a focus. Then click **Actions** in menu bar and select **Create...** from the list. The Create IBM VSDs dialog box shown in Figure 107 on page 518 will be displayed.

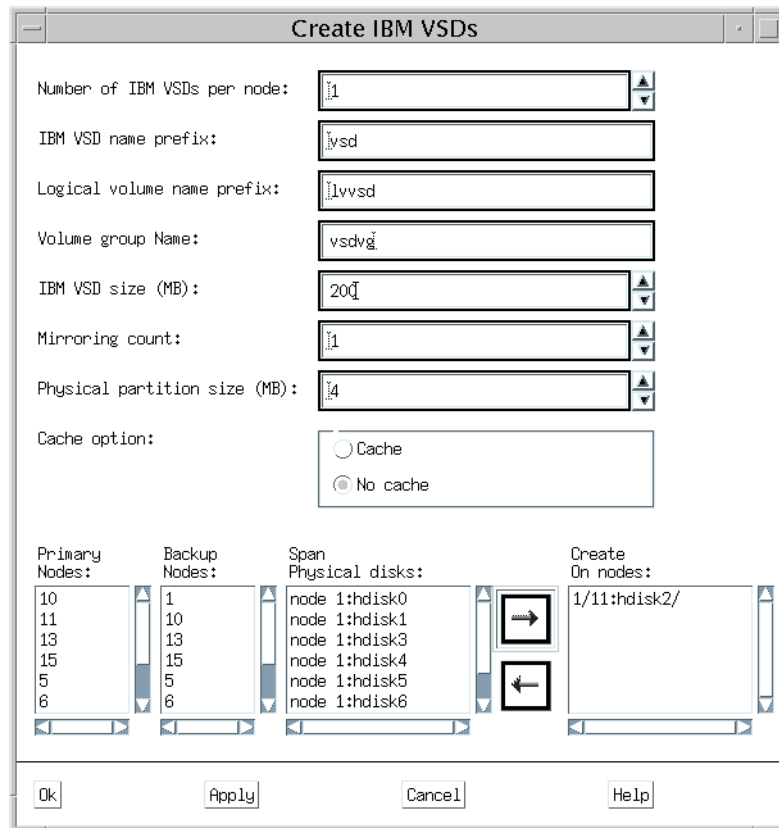


Figure 107. Create IBM VSDs dialog box

There are at least two fields you need to modify the default value and one field you need to specify the value.

By default, rootvg is specified in the Volume group Name: field. You can not put the rootvg or any other volume group that contains bootable logical volumes on a twin-tailed disk. You need to specify a volume group other than rootvg.

You need to change the size of IBM Virtual Shared Disk by modifying the default value in the IBM VSD size (MB): field.

To create IBM Virtual Shared Disk, select the primary node from the Primary Nodes: list box and select the secondary node from the Backup Nodes: list box. Then select the physical disk from the Span Physical disks: list box. In this example, hdisk2 on the node 1 is a twin-tailed disk, and it is shared by

node 1 and node 11. Finally, click the **right arrow** icon. It generates the parameters passed to the `createvsd` command. This parameter is displayed in the Create On nodes: list box.

When you click the **OK** button, the IBM Virtual Shared Disk is created on node 1. In this example, IBM Virtual Shared Disk named `vsd1n1` is created, and it appears in IBM VSDs pane in the window as shown in Figure 108.

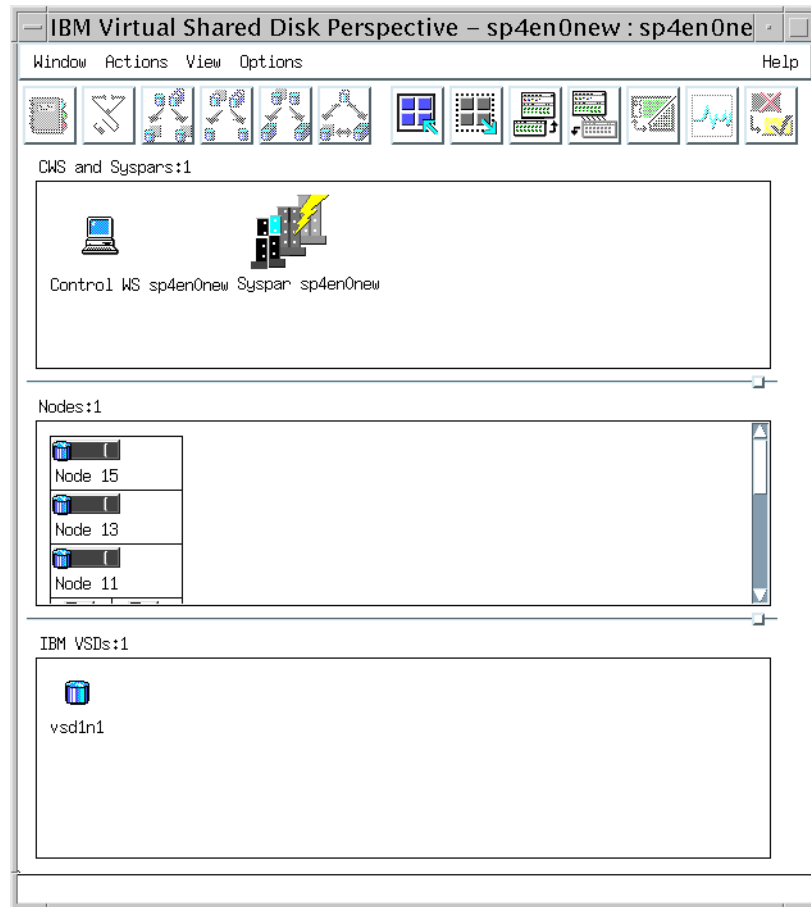


Figure 108. IBM Virtual Shared Disk `vsd1n1` created

17.1.7 Configuring and activating IBM VSD and IBM RVSD

If you have IBM Recoverable Virtual Shared Disk subsystem running on each IBM Virtual Shared Disk node, you can configure and activate both IBM Virtual Shared Disk and IBM Recoverable Virtual Shared Disk at the same

time. You do not need to additional steps to configure IBM Virtual Shared Disk.

To activate the IBM Virtual Shared Disk on all nodes, click the Nodes pane to give it a focus, then click the **Select all objects** icon:



Then click **Actions** in menu bar and select **Control IBM RVSD Subsystem...** from the list. The Control IBM RVSD Subsystem dialog box show in Figure 109 will be displayed.

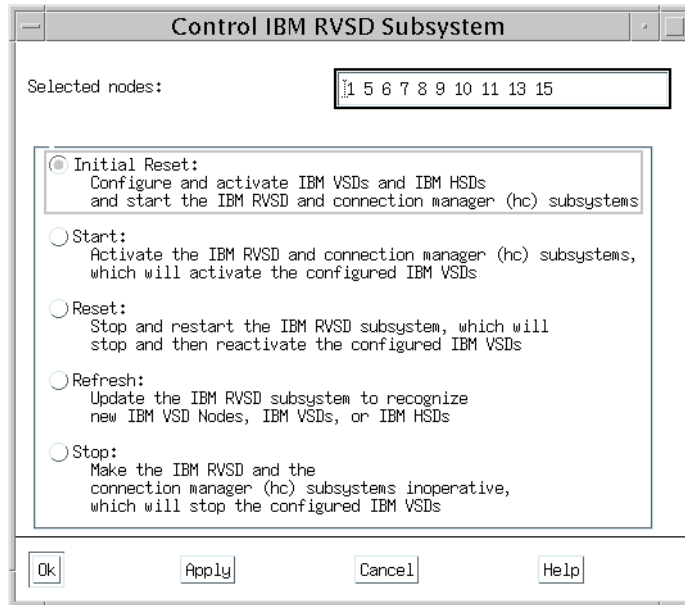


Figure 109. Control IBM RVSD subsystem

Select the **Initial Reset** and click **OK**. It takes a few minutes to complete the task.

17.1.8 Verify IBM Virtual Shared Disk

You can check the status of IBM Virtual Shared Disk by using the `lsvsd` command located in the `/usr/lpp/csd/bin` directory. It can be executed from IBM Virtual Shared Disk Perspective.

To check the status of the IBM Virtual Shared Disk on all nodes, click the Nodes pane to give it a focus, then click the **Select all objects** icon:



Then click **Actions** in menu bar and select the **Run Command...** from the list. You will have the Run Command on Nodes dialog box shown in Figure 110.

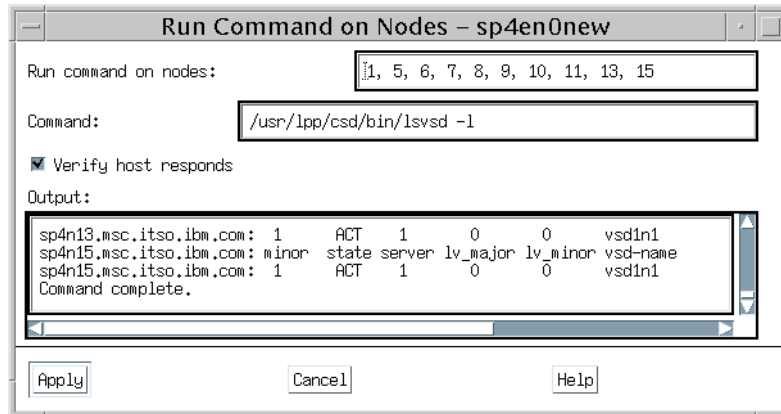


Figure 110. Run Command on Nodes dialog box

Fill in the Command: field as shown in Figure 110, then click the **Apply** button. The output will be displayed in the Output: field. The state value should be ACT for all the nodes.

IBM Virtual Shared Disk also provides the `vsdvtls` command to verify if you can write to the IBM Virtual Shared Disk. The command is located in the `/usr/lpp/csd/bin` directory.

Caution

The `vsdvtls` command writes data to the IBM Virtual Shared Disk. Do not perform this task after you have put real data on IBM Virtual Shared Disk. This is only appropriate after a software installs a creation of a new IBM Virtual Shared Disk.

The command has to be issued with the name of IBM Virtual Shared Disk and requires your reply if you want to continue or not.

```
# /usr/lpp/csd/bin/vsdvts vsdln1
NOTE: This command will change the content of vsdln1 !!!
Do you want to continue [N] ?
```

Therefore, if you execute this command by issuing the `dsh` command or use the Run Command on Nodes dialog box shown in Figure 110 on page 521, you need to use the `yes` command to pass the character `yes` to the command:

```
# dsh -w sp4n01 "yes | /usr/lpp/csd/bin/vsdvts vsdln1"
sp4n01: NOTE: This command will change the content of vsdln1 !!!
sp4n01: Do you want to continue [N] ?
sp4n01: vsdvts: Step 1: Writing file /unix to VSD vsdln1.
sp4n01: dd if=/unix of=/dev/rvsdln1 count=256 bs=4096 seek=1
sp4n01: vsdvts: Step 1 Successful!
sp4n01: vsdvts: Step 2: Reading data back from the VSD.
sp4n01: dd of=/tmp/vsdvts.12498 if=/dev/rvsdln1 count=256 bs=4096 skip=1
sp4n01: vsdvts: Step 2 Successful!
sp4n01: vsdvts: Step 3: Verifying data read from the VSD.
sp4n01: dd if=/unix count=256 bs=4096 | cmp -s - /tmp/vsdvts.12498
sp4n01: vsdvts: Step 3 Successful!
sp4n01:
sp4n01:
sp4n01: VSD Verification Test Suite Successful!
sp4n01: 256+0 records in.
sp4n01: 256+0 records out.
sp4n01: 256+0 records in.
sp4n01: 256+0 records out.
sp4n01: 256+0 records in.
sp4n01: 256+0 records out.
#
```

17.2 Monitoring IBM Recoverable Virtual Shared Disk

It is important to monitor IBM Recoverable Virtual Shared Disk so that you can tell if a fail over has occurred, and then if it was not planned, you need to figure out why it failed over. IBM Recoverable Virtual Shared Disk can be monitored in several manners.

First, you may need to monitor IBM Recoverable Virtual Shared Disk daemon itself if it is running and if it has a membership of Group Services (GS) subsystem. Event Management (EM) subsystem provides many resource variables for IBM Recoverable Virtual Shared Disk. You can use them to monitor IBM Recoverable Virtual Shared Disk from various points of view. This section also covers information written in the log file.

17.2.1 The recovery subsystem

The IBM Recoverable Virtual Shared Disk recovery subsystem that operates as the `rvsdd` daemon uses the utilities of the Group Services subsystem. The `rvsdd` daemon joins the Group Services and creates the group named `ha.vsd`. You can see the current status of Group Services by issuing the `lssrc` command:

```
# dsh -w sp4n01 lssrc -ls hags
sp4n01: Subsystem      Group      PID      Status
sp4n01: hags          hags      4618    active
sp4n01: 4 locally-connected clients. Their PIDs:
sp4n01: 18022 16704 8170 6012
sp4n01: HA Group Services domain information:
sp4n01: Domain established by node 1.
sp4n01: Number of groups known locally: 3
sp4n01:
sp4n01:      Number of      Number of local
sp4n01: Group name      providers      providers/subscribers
sp4n01: cssMembership      10             1             1
sp4n01: ha_em_peers      11             1             0
sp4n01: ha.vsd          10             1             0
#
```

When a node goes down, or a disk adapter or cable fails, the `rvsdd` daemon notifies all surviving processes in the remaining IBM Virtual Shared Disk server nodes through the Group Services so that they can begin recovery action.

Monitoring the `rvsdd` daemon is important for a successful recovery action. To monitor the `rvsdd` daemon, you can use a variety of methods. Refer to Chapter 10, “Monitoring hardware” on page 249, Chapter 11, “Managing Events” on page 279.

17.2.2 Using resource variables

Event management (EM), one of the RS/6000 Cluster Technology (RSCT) components, provides several resource variables related to IBM Virtual Shared Disk.

To list up all the IBM Virtual Shared Disk-related resource variable names and their short description, issue the `haemqvar` command.

```

# export PATH=$PATH:/usr/sbin/rsct/bin
# haemqvar -d "IBM.PSSP.VSD" "" "*"
IBM.PSSP.VSDdrv.timeout_error    Timeouts.
IBM.PSSP.VSDdrv.request_rework   Requests rework.
IBM.PSSP.VSDdrv.request_block_shortage  Requests queued waiting for a request block.
IBM.PSSP.VSDdrv.rejected_responds  Rejected responses.
IBM.PSSP.VSDdrv.rejected_requests  Rejected requests.
IBM.PSSP.VSDdrv.rejected_no_buddy_buffer  Rejected no buddy buffer.
IBM.PSSP.VSDdrv.pbuf_shortage     Requests queued waiting for a pbuf.
IBM.PSSP.VSDdrv.num_suspended     The number of VSDs that are suspended on this node.
IBM.PSSP.VSDdrv.num_stopped       The number of VSDs that are stopped on this node.
IBM.PSSP.VSDdrv.num_not_active     The number of VSDs that are not active on this node.
IBM.PSSP.VSDdrv.num_active        The number of VSDs that are active on this node.
IBM.PSSP.VSDdrv.indirect_io       I/O is not performed directly from mbuf.
IBM.PSSP.VSDdrv.comm_buf_shortage  Shortage on the Communication Buf pool.
IBM.PSSP.VSDdrv.cache_shortage     Requests queued waiting for a cache block.
IBM.PSSP.VSDdrv.buddy_buffer_shortage  Requests queued waiting for a buddy buffer.
IBM.PSSP.VSDdrv.avg_buddy_wait     Average buddy buffer wait_queue size.
IBM.PSSP.VSDdrv.RVSD_status       The state of the RVSD on this node.
IBM.PSSP.VSDdrv.9_retry_count      Retries 9.
IBM.PSSP.VSDdrv.8_retry_count      Retries 8.
IBM.PSSP.VSDdrv.7_retry_count      Retries 7.
IBM.PSSP.VSDdrv.6_retry_count      Retries 6.
IBM.PSSP.VSDdrv.5_retry_count      Retries 5.
IBM.PSSP.VSDdrv.4_retry_count      Retries 4.
IBM.PSSP.VSDdrv.3_retry_count      Retries 3.
IBM.PSSP.VSDdrv.2_retry_count      Retries 2.
IBM.PSSP.VSDdrv.1_retry_count      Retries 1.
IBM.PSSP.VSD.state                VSD state: STOPPED/ACTIVE/SUSPENDED.
IBM.PSSP.VSD.server                Current VSD server.
IBM.PSSP.VSD.remote_req_write      Remote write.
IBM.PSSP.VSD.remote_req_read       Remote read.
IBM.PSSP.VSD.physical_req_write     Physical write.
IBM.PSSP.VSD.physical_req_read     Physical read.
IBM.PSSP.VSD.local_req_write       Local write.
IBM.PSSP.VSD.local_req_read        Local read.
IBM.PSSP.VSD.client_req_write      Client write.
IBM.PSSP.VSD.client_req_read       Client read.
IBM.PSSP.VSD.cache_hits            Cache hits.
IBM.PSSP.VSD.bytes_write           Total bytes (write).
IBM.PSSP.VSD.bytes_read            Total bytes (read).
IBM.PSSP.VSD.blocks_rw            The total number of data block read or written.
#

```

If you are interested in the particular resource variable, issue the `haemqvar` command as follows:

```

# haemqvar "" IBM.PSSP.VSDdrv.RVSD_status "*"
Variable Name:  IBM.PSSP.VSDdrv.RVSD_status
Value Type:    Quantity
Data Type:     long
Initial Value: 0
Class:         IBM.PSSP.VSD
Locator:       NodeNum
Variable Description:
    The state of the RVSD on this node.

    The VSD subsystem can be queried for the state of RVSD on the node.

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

    The possible states are:
    (-2) RVSD_Can_Not_Tell
    (-1) RVSD_NotAvail
    ( 0) RVSD_Idle
    ( 1) RVSD_Activating
    ( 2) RVSD_Node_Join
    ( 3) RVSD_Node_Recovery
    ( 4) RVSD_Fence
    ( 5) RVSD_Quorum
    ( 6) RVSD_Tail_Recoveryn
    ( 7) RVSD_Wait_GS
    ( 8) RVSD_Refresh

Example: To be informed whenever node 5 is changing states, one could
register for the event like the following:

Resource Variable Name: IBM.PSSP.VSDdrv.RVSD_status
Resource ID:           NodeNum=5
Expression:            X != X@P

Resource ID wildcarding:

The "NodeNum" resource ID element may be wildcarded in order to
apply a query or event registration to all nodes within the domain.
Resource ID:  NodeNum=int
NodeNum: The number of the node for which the information applies.
#

```

This example displays the detailed description of the IBM.PSSP.VSDdrv.RVSD_status resource variable.

If you want to know the current status of the resource variable, issue the haemqvar command as follows:

```

# haemqvar -c "" IBM.PSSP.VSDdrv.RVSD_status "NodeNum=9"
9 IBM.PSSP.VSDdrv.RVSD_status NodeNum=9 0
#

```

This example displays the status of the IBM.PSSP.VSDdrv.RVSD_status resource variable on node 9. You know the state of IBM Recoverable Virtual Shared Disk on node 9 is idle.

The `haemqvar` command is not the only method to monitor resource variables. To learn more about resource variables, refer to Chapter 10, “Monitoring hardware” on page 249 Chapter 11, “Managing Events” on page 279.

17.2.3 Using log files

The log file can be an important information resource. IBM Recoverable Virtual Shared Disk uses the `vsd.debuglog` file located in the `/var/adm/csd` directory as its log file.

To monitor the activity of IBM Recoverable Virtual Shared Disk, you can issue the `tail` command. By issuing the command to the log file on the primary and secondary IBM Virtual Shared Disk server nodes, you can monitor what kind of scripts are executed when fail over occurs.

The following is the `vsd.debuglog` log file on the secondary node. At 11:34:54, the primary node is halted by the `halt` command to simulate the node down:

```
# tail -f /var/adm/csd/vsd.debuglog
11/01/98 11:34:54: ./vsd.DOWN1 membership=1 5 6 7 8 9 10 11 13 15 local-node=11
input=1
11/01/98 11:34:54: self=false Memb_Before=1 5 6 7 8 9 10 11 13 15 Memb_After= 5
6 7 8 9 10 11 13 15
11/01/98 11:34:54: ctlvsd -k 1
11/01/98 11:34:54: suspendvsd reported 1 total_suspends 0 failures
11/01/98 11:34:54: stopvsd reported 0 total_stops 0 failures
11/01/98 11:34:54: ./vsd.DOWN1 script done
11/01/98 11:34:55: ./vsd.DOWN2 membership=1 5 6 7 8 9 10 11 13 15 local-node=11
input=1
11/01/98 11:34:56: self=false Memb_Before=1 5 6 7 8 9 10 11 13 15 Memb_After= 5
6 7 8 9 10 11 13 15
11/01/98 11:34:56: reserve on volume group vsdvg called
11/01/98 11:34:57: varyonvg -bn vsdvg complete; RC = 0; output =
11/01/98 11:34:58: reserve on volume group vsdvg complete
11/01/98 11:34:58: resumevsd reported 1 total_resumes 0 failures
11/01/98 11:34:58: ./vsd.DOWN2 script done
```

You may notice that a couple of scripts executed to fail over the service. In this example, the `vsd.DOWN1` and `vsd.DOWN2` scripts are executed.

The `rvsdd` daemon executes scripts when Group Services notifies the event to the daemon. According to the event, the daemon executes appropriate scripts. Scripts have a prefix `vsd.` and are located in the `/usr/lpp/csd/bin`

directory. The vsd.debuglog file is used as output of the scripts. The following is the scripts that are used by the rvsdd daemon:

```
# cd /usr/lpp/csd/bin
# ls -l vsd.[A-Z]*
-rwxr-x--- 2 bin bin 37535 Oct 20 09:08 vsd.CSER1
-rwxr-x--- 2 bin bin 37535 Oct 20 09:08 vsd.CSER2
-rwxr-x--- 4 bin bin 51591 Oct 20 09:08 vsd.DOWN1
-rwxr-x--- 4 bin bin 51591 Oct 20 09:08 vsd.DOWN2
-rwxr-x--- 4 bin bin 8690 Oct 20 09:08 vsd.FENCE0
-rwxr-x--- 4 bin bin 8690 Oct 20 09:08 vsd.FENCE1
-rwxr-x--- 4 bin bin 8690 Oct 20 09:08 vsd.FENCE2
-rwxr-x--- 4 bin bin 8690 Oct 20 09:08 vsd.FENCE3
-rwxr-x--- 1 bin bin 40100 Oct 20 09:08 vsd.REFRESH
-rwxr-x--- 4 bin bin 51591 Oct 20 09:08 vsd.UP1
-rwxr-x--- 4 bin bin 51591 Oct 20 09:08 vsd.UP2
#
```

As you can see in the number of links column (the second from left), some of them are the same script. It branches the flow by using its name as parameter:

vsd.UP1, vsd.UP2, vsd.DOWN1, and vsd.DOWN2

They correspond to the node up event and node down event.

vsd.CSER1 and vsd.CSER2

CSER stands for Change Server. They correspond to the event that the node is active but can not write to IBM Virtual Shared Disk by I/O error (EIO), for example.

vsd.FENCE0, vsd.FENCE1, vsd.FENCE2, vsd.FENCE3

To preserve data integrity during application recovery, IBM Virtual Shared Disk can be fenced. They correspond to this event.

vsd.REFRESH

Issue appropriate commands to make the IBM VSD device driver aware that the configuration of IBM Virtual Shared Disk has changed.

All of them are a shell scripts. You can read them directly to learn their detailed function.

Appendix A. Special notices

This publication is intended to help IBM customers, IBM business partners, and IBM I/T specialists concerned with RS/6000 SP system management. The information in this publication is not intended as the specification of any programming interfaces that are provided by IBM Parallel System Support Programs for AIX. See the PUBLICATIONS section of the IBM Programming Announcement for IBM Parallel System Support Programs for AIX for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

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, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

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

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee

that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

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

AIX	ESCON
IBM ®	LoadLeveler
Micro Channel	RS/6000
SP	

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

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

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States and/or other countries.

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Appendix B. Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

B.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see “How to get ITSO redbooks” on page 535.

- *PSSP 3.1 Announcement*, SG24-5332
- *SP Perspectives: A New View of Your SP System*, SG24-5180
- *HACMP Enhanced Scalability Handbook*, SG24-5328
- *Understanding and Using the SP Switch*, SG24-5161
- *The RS/6000 SP Inside Out*, SG24-5374
- *RS/6000 SP Software Maintenance*, SG24-5160
- *AIX Version 4.3 Differences Guide*, SG24-2014
- *Using ADSM to Back Up Database*, SG24-4335

B.2 Redbooks on CD-ROMs

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at <http://www.redbooks.ibm.com/> for information about all the CD-ROMs offered, updates, and formats.

CD-ROM Title	Collection Kit Number
System/390 Redbooks Collection	SK2T-2177
Networking and Systems Management Redbooks Collection	SK2T-6022
Transaction Processing and Data Management Redbooks Collection	SK2T-8038
Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
AS/400 Redbooks Collection	SK2T-2849
Netfinity Hardware and Software Redbooks Collection	SK2T-8046
RS/6000 Redbooks Collection (BkMgr)	SK2T-8040
RS/6000 Redbooks Collection (PDF Format)	SK2T-8043
Application Development Redbooks Collection	SK2T-8037
IBM Enterprise Storage and Systems Management Solutions	SK3T-3694

B.3 Other publications

These publications are also relevant as further information sources:

- *RS/6000 SP: Planning, Volume 1, Hardware and Physical Environment*, GA22-7280
- *RS/6000 SP: Planning, Volume 2, Control Workstation and Software Environment*, GA22-7281
- *RS/6000 SP: Maintenance Information, Volume 1, Installation and Relocation*, GA22-7375
- *RS/6000 SP: Maintenance Information, Volume 2, Maintenance Analysis Procedures*, GA22-7376
- *RS/6000 SP: Maintenance Information, Volume 3, Locations and Service Procedures*, GA22-7377
- *RS/6000 SP: Maintenance Information, Volume 4, Parts Catalog*, GA22-7378
- *IBM Parallel System Support Programs for AIX: Installation and Migration Guide*, GA22-7347
- *IBM Parallel System Support Programs for AIX: Administration Guide*, SA22-7348
- *IBM Parallel System Support Programs for AIX: Managing Shared Disks*, SA22-7349
- *IBM Parallel System Support Programs for AIX: Performance Monitoring Guide and Reference*, SA22-7353
- *IBM Parallel System Support Programs for AIX: Diagnosis Guide*, GA22-7350
- *IBM Parallel System Support Programs for AIX: Command and Technical Reference*, SA22-7351
- *IBM Parallel System Support Programs for AIX: Message Reference*, GA22-7352
- *RS/6000 Cluster Technology: Event Management Programming Guide and Reference*, SA22-7354
- *RS/6000 Cluster Technology: Group Service Programming Guide and Reference*, SA22-7355
- *Diagnostic Information for Multiple Bus Systems*, SA38-0509
- *OEM- Diagnostic Information for Micro Channel Bus System*, SA23-2765

- *AIX Version 4.3 System Management Guide: Operating System and Devices*, SC23-4126
- *AIX Version 4.3 Network Installation Management Guide and Reference*, SC23-4113

The following Web sites are also relevant as further information sources:

- <http://www.redbooks.ibm.com>
- <http://service.boulder.ibm.com/rs6k/sp/status>
- <http://w3.viewblue.ibm.com>
- <http://service.software.ibm.com/support/rs6000>
- <http://www.rs6000.ibm.com/support/micro>
- <http://www.neosoft.com/tcl>
- <http://www.scriptics.com>
- <http://expect.nist.gov>
- <http://www.hursley.ibm.com/~ssa/rs6k>
- <http://www.cise.ufl.edu/ftp/perl/CPAN/CPAN.html>
- <http://www.cs.cmu.edu/afs/cs.cmu.edu/project/mach/public>
- <http://www.software.ibm.com/year2000/papers/aixy2k.html>
- <http://www.elink.ibm.link.ibm.com/pbl/pbl>
- <http://w3.itso.ibm.com>
- <http://w3.ibm.com>

How to get ITSO redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** <http://www.redbooks.ibm.com/>

Search for, view, download, or order hardcopy/CD-ROM redbooks from the redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this redbooks site.

Redpieces are redbooks in progress; not all redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

Send orders by e-mail including information from the redbooks fax order form to:

	e-mail address
In United States	usib6fpl@ibmmail.com
Outside North America	Contact information is in the "How to Order" section at this site: http://www.elink.ibm.com/pbl/pbl/

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the "How to Order" section at this site: http://www.elink.ibm.com/pbl/pbl/

- **Fax Orders**

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	Fax phone number is in the "How to Order" section at this site: http://www.elink.ibm.com/pbl/pbl/

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the redbooks Web site.

IBM Intranet for Employees

IBM employees may register for information on workshops, residencies, and redbooks by accessing the IBM Intranet Web site at <http://w3.itso.ibm.com/> and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access MyNews at <http://w3.ibm.com/> for redbook, residency, and workshop announcements.

IBM Redbook fax order form

Please send me the following:

Title	Order Number	Quantity
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

First name _____ Last name _____

Company _____

Address _____

City _____ Postal code _____ Country _____

Telephone number _____ Telefax number _____ VAT number _____

Invoice to customer number _____

Credit card number _____

Credit card expiration date _____ Card issued to _____ Signature _____

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

List of abbreviations

ACL	Access Control List	HTML	HyperText Markup Language
ADSM	Adstar Distributed Storage Management	IBM	International Business Machines Corporation
AFS	Andrew File System	ITSO	International Technical Support Organization
APAR	Authorized Program Analysis Report	LAN	Local Area Network
BFF	AIX Backup File Format	LCD	Liquid Crystal Display
BIS	Boot Install Server	LED	Light Emitting Diode
BOS	Base Operating System	LPP	Licensed Program Product
BSD	Berkeley Software Distribution	MAC	Medium Access Control
CE	IBM Customer Engineer	MACN	Monitor and Control Node (control workstation)
CHRP	IBM PowerPC Common Hardware Reference Platform	MCA	Microchannel Bus Architecture Model
CuAT	Customized Attribute	MIT	Massachusetts Institute of Technology
CWS	Control Workstation	NFS	Network File System
DES	Data Encryption Standard	NIM	Network Installation Management
DNS	Domain Name System	NIS	Network Information Services
EM	Event Management	NTP	Network Time Protocol
EMCDB	Event Management Configuration Database	ODM	Object Data Manager
FDDI	Fiber Distributed Data Interface	PCI	Peripheral Component Interface Bus Architecture Model
FRU	Field Replacement Unit	PDF	Portable Document Format
FTP	File Transfer Protocol	PDS	Public Domain Software
GUI	Graphical User Interface	PERL	Practical Extraction and Report Language
GS	Group Services	POST	Power On Self Test
HACMP/ES	IBM High Availability Cluster Multi-Processing for AIX Enhanced Scalability	PP	Physical Partition

PSSP	IBM Parallel System Support Programs for AIX	TFTP	Trivial File Transfer Protocol
PTF	Program Temporary Fix	TK	Tcl-based Tool Kit for X-windows
RS6k	RS/6000 Architecture	TS	Topology Services
RSCT	RS/6000 Cluster Technology	URL	Universal Resource Locator
RSPC	IBM PowerPC	VFOP	Virtual Front Operator Panel
RVSD	Recoverable Virtual Shared Disk	VPD	Vital Product Data
SAMI	Service and Manufacturing Interface	WWW	World Wide Web
SBS	Structured Byte String		
SCSI	Small Computer System Interface		
SDR	System Data Repository		
SMIT	System Management Interface Tool		
SMUX	SNMP Multiplexing Protocol		
SNMP	Simple Network Management Protocol		
SPOC	Single Point of Control		
SPOT	Shared Product Object Tree		
SRC	System Resource Controller		
SSA	Serial Storage Architecture		
SUP	Software Update Protocol		
SWVPD	Software Vital Product Data		
TCL	Tool Command Language		
TCLX	Tool Command Language Extended		

Index

Numerics

7027-HSD SCSI-2 Fast/Wide disk 186
7133 IBM SSA Disk Subsystem 188

A

add a node 18
add an SP frame 11
add an SP-attached server 38
adding physical disk 178
Adobe Acrobat Reader 432
 install 434
ADSM 121
Adstar Distributed Storage Management 121
AIX Error Log 327, 328
AIX man page 429
 install 429
alias IP address 356
alternative boot system image
 define 183
 install 185
 switch 186
annotate 29, 50, 63, 68, 73
APAR 102
archive the SDR 11
attach an SP-attached server 76
authentication 446
authentication database 452
 back up 454
 destroy 453
 initialize 453
 read 455
 restore 454
authentication method
 enable 447
 list 449
Authentication Service 449
authorization 446
authorization method
 select 448
Authorized Program Analysis Report 102
autojoin 346
automounter 418
 information 419
 stop using 423
 use SP switch 422

B

backup
 /spdata on CWS 119
 rootvg on CWS 118
 rootvg on SP node 122
Backup File Format 138
basecode version 245
BFF 138
boot configuration 183
boot from external disk
 SCSI 186
 SSA 188
BOOTP 154
Bootstrap Protocol 154
BSD syslog 327, 330

C

CDE 401
CHRP 223, 226
Commands
 acroread 434
 add_principal 462
 aixterm 299, 300
 bootlist
 -m -o 185
 cfgmgr 155
 chauthent 447
 chauthpar
 -p 361, 447
 chgcsc 347
 chkp
 -e 470
 -l 472
 chps
 -s 325
 chsec
 -f -s -a 481
 chuser 397
 chypdom 406
 cmonacct 428
 compress 127
 create_krb_files 476
 crontab
 -e 387
 -l 387
 crunacct 428
 cshUTDOWN

ALL 370
 -F ALL 358
 -F -N 197
 -X 371
 cstartup
 ALL 361, 369
 -X 370
 date 417
 diag 215
 dsh
 -N 373
 dterror.ds 299
 Eannotator 30, 51, 63, 68, 73
 Eclock 30
 -d 351
 -r 351
 Efence 346, 373
 -autojoin 346
 Eprimary 344
 -backup 345
 Equiesce 346
 errclear 313, 328
 errpt 328
 -a -i -y 159
 -a -l 353
 Estart 36, 56, 345
 Euncence 346
 Eunfence 36, 56
 Eunpartition 358
 ext_srvtab 500
 extendvg 178
 -f 326
 find 312, 314
 ftp 103
 bin 107
 mget 107, 108
 prompt 107
 haemqvar 302
 -c 304
 -d 302
 hmadm
 setacls 233, 468
 hmcmds
 -G 101
 -G -v -u 101
 -v -G basecode 246
 -v -G boot_supervisor 245
 -v -G exec_supervisor 246
 -v -G runpost 244
 -v -G setid 244
 -v -u -G microcode 247
 hmmon
 -G -Q 265, 268, 270, 272
 -G -q 268, 269, 272, 275
 -G -Q -s 267, 270, 273
 -G -Q -s -v 268, 269, 271, 275
 hmreinit 22
 hostlist 374
 ifconfig 356, 362
 delete 362
 install_cw 119, 350, 355
 installp 80
 inutoc 108, 139
 k4destroy 491
 k4init 417, 480
 k4list
 -srvtab 478
 kadmin 460, 461, 469
 kdb_destroy 453
 kdb_edit 463, 469, 471, 472
 kdb_init 453
 kdb_util
 dump 454, 455
 load 454, 458
 new_master_key 457
 kpasswd
 -n 468
 kprop 451
 ksrvutil
 -k change 479
 -k list 478
 kstash 458
 ln 424
 lsattr
 -E -l 323, 346
 lsauthpar 449
 lscfg 214
 -pv 88, 89
 -vl 76, 89, 90
 -vl rmt0 93
 -vl ssa0 94
 lsdev
 -Cc 187, 188, 323
 lskp 459
 lspp 78
 lsnim 113, 134
 -l 137, 138, 140, 141
 lsps

- a 320, 323, 325
- s 320
- lssrc 523
- lsvg 178, 324, 326
 - l 180, 181
- make 410
- man 431
- migratepv
 - l 182
- mkclient 409
- mkdev 12, 39
- mkdir
 - p 117
- mkkp
 - e -l 464
- mkmaster 407
- mknfsexp 421
 - d -t -N 111
- mkps 326
- mkslave 408
- mksysb
 - i 118
- mkuser 393
- mkvg
 - t 178
- mount 112
- mv 424
- netscape 437
- netstat
 - rn 498
- ngaddto 372
- ngclean 372
- ngcreate 372
- ngdelete 372
- ngdelfrom 372
- ngfind 372
- nglist 372
- ngnew 373
- ngresolve 373
- nim
 - o 113
 - o -a 116
- no 423
- nodecond 35, 55, 124
- nrunacct 428
- odmget 150
 - CuVPD 217
- passwd 404
- passwd.aix 404
- pcp 356, 511
- pdf
 - a 315
- perspectives 439
- pmandef
 - C 297
 - c 297
 - e 297
 - event command 299
 - event expression 298
 - expression 297
 - handle name 297
 - r 297
 - rearm command 299
 - rearm expression 298
 - resource ID 297, 298
 - resource variable 297, 298
 - s 297
 - u 301
- pmanquery 300
- psyclr
 - a -y 331
- psyslprt
 - w 330
- push-kprop 204, 206
- reducevg 182
- restore 318
- restvg
 - q -f 121
- rmdev 205
- rmkp 466
- rmlvcopy 181
- route 423
 - add -net 498
- ruser
 - a 399
 - d 400
- s1term 153
- savevg
 - i -f 120
 - X -i -f 181
- SDR_config
 - v -d 17, 23, 45
- SDRArchive 11
- SDRGetObjects 125
 - Frame 219, 221
 - Node 222, 226
 - NodeControl 224
 - Switch 228

- Switch_partition 351
- services_config 427
- setclock 411, 417
- setup_authent 485
- showled 170
- shutacct 314, 428
- shutdown
 - F -r ALL 371
 - Fr 81, 207
- skulker 314
- sort 312, 314
- sp_passwd 404
- spacs_cntrl 402
 - block 402
 - unblock 402
- spadaptrs 28, 49, 202
- spapply_config 359
 - v 359
- spbootins
 - r customize 77, 164, 204
 - r disk 113, 186
 - r install 35, 55, 124, 185
- spbootlist
 - l 186
- spchvgobj 32, 52, 123, 191, 195
- spcustomize_syspar 358
- spdeladap 205
- spdelfram 60, 71
- spdelnode 66
- spdisplay_config 356
 - R -n 356
- spethernt 25, 46
- spevent 286
- spframe 13, 40
- sphardware 234
- sphostnam 34, 53, 208
- sphrdwrad 26, 47, 77, 198
- splm 332
 - a -k -t -d -y -l -r 336
 - a -t -d -y 334
- splstdata
 - a 25, 28, 46, 49
 - b 27, 48, 124
 - e 380, 390, 412, 419, 425
 - f 14, 42, 219, 221
 - n 20, 34, 54, 224, 227
 - s 30, 51, 64, 69, 73, 229
 - v 32, 52
 - v -l 192
- spmirrorvg
 - l 193
- spmuser 392
- spmkgobj 190
 - r -h -n -i -v -P 184
- spmon
 - G -d 14, 263
 - G -Key 242
 - G -key 242
 - G -p 243
 - G -r 243
- spmon_ctest 13, 41
- spruser 394
- spsetauth
 - d -p 360, 448
- spsitenv 405, 414, 421, 423, 427
- spsvrmgr
 - G -r 23, 100
 - G -r status 17
 - G -u 17, 24, 101
- spsyspar 362
- spunmirrorvg
 - l 196
- spvsd 511
- startup 428
- su 314
- supper
 - files 383
 - log 388
 - rlog 388
 - serve 383
 - status (BIS) 382
 - status (CWS) 382
 - status (node) 382
 - update 386
 - when 384
 - where 384
- sysctl 468
 - h 460, 466
- syspar_ctrl
 - G -r 62, 68, 72
- tail 526
- umount 80, 113
- vdid3
 - i 347
- vi 314
- vsdvts 521
- who 311
- xntpdc 415

- c peers 415
- p 415
- yes 522
- yppasswd 409
- zcat 127
- Common Desktop Environment 401
- Common Hardware Reference Platform 223, 226
- configure SP Ethernet 77
- control_flow 157
- CuAt 119
- Customized Attribute 119

D

Daemons

- /usr/lpp/ssp/bin/xntpd 410
- /usr/sbin/xntpd 410
- cssadm 348
- fault_service_Worm_RTG_SP 169, 351
- hardmon 451
- kadmind 450
- kerberos 449
- kpropd 451
- qdaemon 313
- rvsdd 523
- splogd 451
- supfilesrv 389
- syslogd 330
- ypbind 407, 408
- ypserv 407, 408

DCE 446

- DEFAULT_SPOT_OPTIONS 149
- DEFAULT_SPOT_OPTIONS_41 149
- DEFAULT_SPOT_OPTIONS_42 149
- delete a frame 59
- delete a node 65
- delete an SP-attached server 70
- deleting physical disk 180
- detach an SP-attached server 82
- device driver 140
- Device Specific.(L1) 94
- Device Specific.(Z1) 94
- Device Specific.(Z2) 95

Directories

- /dev 312
- /etc/amd 84
- /etc/auto 84
- /etc/ha 84
- /etc/objrepos 149, 217

- /etc/SP 84, 352
- /etc/ssp 84
- /home/\$homedir_server 420
- /spdata 84
- /spdata/sys1/ha/css 352
- /spdata/sys1/install/name/lppsource 108, 138, 175
- /spdata/sys1/install/name/spot/spot_name 141, 176
- /spdata/sys1/install/pssplpp 79, 165
- /spdata/sys1/logtables 332
- /spdata/sys1/sdr/archives 11
- /spdata/sys1/spmon 232
- /spdata/sys1/ucode 97
- /ftfboot 85
- /tmp/YMMDD/arch_sysman.tab/sysman 334
- /u 421
- /u/username 393
- /usr/lpp/Acrobat3/bin 434
- /usr/lpp/csd/bin 521, 526
- /usr/lpp/ssp/css 347
- /usr/lpp/ssp/docs 434
- /usr/lpp/ssp/html 437
- /usr/lpp/ssp/install/bin 170
- /usr/lpp/ssp/perl5 431
- /usr/lpl/ssp/perl5/lib 431
- /usr/netscape/navigator-us 437
- /usr/sbin/rsct 84
- /var/adm/acct 314, 428
- /var/adm/csd 526
- /var/adm/ras 158, 313
- /var/adm/SPlogs 84, 338
- /var/adm/SPlogs/kerberos 450
- /var/adm/SPlogs/SPconfig 214
- /var/adm/SPlogs/sysman 168, 334
- /var/ha 84
- /var/ha/log 338
- /var/preserve 314
- /var/spool 313
- /var/sysman 84
- /var/sysman/sup 387
- /var/sysman/sup/lists 385
- /var/tmp 313
- disk configuration 177
- Distributed Computing Environment 446
- DNS 208
- domain 495
- Domain Name System 208
- DSHPATH 375

E

- EM 522
- enable SP accounting 427
- errdemon daemon 328
- escape sequence 409
- ESCON-adapter 205
- event 279
 - event definition
 - check 292
 - register 292
 - unregister 295
 - event expression 283
- Event Management 522
- Event Management subsystem 279
- Event Perspective 285
 - Condition notebook 289
 - Event expression 290
 - Rearm expression (optional) 290
 - Event Definition notebook 288
 - Description 289
 - Event definition name 289
 - Name 289
 - Remaining resource ID elements specified 289
 - Event Notification Log window 292, 294
 - Show Resource Variable Details window 291
 - Resource variable description 291
 - Resource variable name 291
 - View Event Notification (Rearm) window 294
 - View Event Notification window 293
- expect 126
- expression 282

F

- factor 178
- file collection technology 379
- File sets
 - Adobe.acrobat.3.0.1.0 434
 - bos 139
 - bos.64bit 139
 - bos.acct 425
 - bos.diag 139
 - bos.html.en_US.cmds.cmds1 430
 - bos.html.en_US.cmds.cmds2 430
 - bos.html.en_US.cmds.cmds3 430
 - bos.html.en_US.cmds.cmds4 430
 - bos.html.en_US.cmds.cmds5 430
 - bos.html.en_US.cmds.cmds6 430

- bos.html.en_US.nav 430
- bos.mp 139
- bos.net 139
- bos.net.nis.client 406
- bos.net.nis.server 406
- bos.sysmgt 139
- bos.sysmgt.nim.client 149
- bos.sysmgt.nim.spot 149
- bos.terminfo 139
- bos.terminfo.all.data 139
- bos.up 139
- devices.base.all 139
- devices.buc.all 139
- devices.common.all 139
- devices.graphics.all 139
- devices.mca.all 139
- devices.rs6ksmp.base 139
- devices.scsi.all 139
- devices.sio.all 139
- devices.sys.all 139
- devices.tty.all 139
- Netscape.communicator-us.rte.4.0.7.0 436
- perfagent.tools 78, 83
- perfagent.tools.2.2.32.x 165
- rsct.basic.hacmp 80, 83
- rsct.basic.rte 80, 83
- rsct.basic.sp 80, 83
- rsct.clients.hacmp 80, 83
- rsct.clients.rte 80, 83
- rsct.clients.sp 80, 83
- ssp.basic 79, 80, 83, 165
- ssp.clients 80, 83, 498
- ssp.css 83, 350
- ssp.docs 430
- ssp.ha_topsvcs.compat 83
- ssp.perlpkg 80, 498
- ssp.perlpkgx 83
- ssp.pman 83
- ssp.resctr.rte 438
- ssp.st 83
- ssp.sysctl 83
- ssp.sysman 83
- ssp.unicode 97
- ssp.vsdgui 510
- ssp.vsdgui.loc 510
- ssp.vsdgui.msg 510
- vsd.cmi 510
- vsd.hsd 510
- vsd.rvsd.hc 510

vsd.rvsd.rvsdd 510
 vsd.rvsd.scripts 510
 vsd.sysctl 510
 vsd.vsd 510
 xIC.rte 139
 File systems
 /dev/install_images 310
 /dev/install_pssplpp 310
 /dev/spot_name 310
 /spdata 119
 /usr 113, 142
 Files
 .k 463, 489, 494
 .klogin 85, 489, 491, 494, 501
 .resident 403
 .rhosts 85, 113
 .toc 139
 /etc/group 403
 /etc/passwd 403
 /etc/security/group 403
 /etc/security/passwd 403
 /usr/lpp/ssp/samples/script.cust 320
 admin_acl.add 450, 465, 489, 494
 admin_acl.get 450, 465, 489, 494
 admin_acl.mod 451, 465, 489, 494
 admin_server.syslog 451
 auto.log 425
 auto.u 393, 421
 bootlog 155
 bootptab.info 62, 161
 bosinst.data 157
 bosinstlog 158
 c_sh_lib 146
 configfb.log 166
 cshutSeq 370
 cssadm.debug 349, 350
 cssadm.stderr 350
 cssadm.stdout 350
 cstartSeq 369
 cw_allowed 401
 devinst.log 158
 errlog 159, 328
 errtmpl 159
 failedlogin 311
 ftpusers 398
 hmacs 232, 468
 hostname.YYMMDDhmmss.css.snap.tar.Z
 352
 hosts 24, 209
 image.data 159, 318
 LOGICAL_VOLUME 319
 PP 319
 PP_SIZE 319
 TYPE 319
 initial_hostname-new-srvtab 476
 inittab 82, 166, 167
 kerberos.log 450
 kpropd.log 451
 krb.conf 85, 451, 489, 491, 494, 504
 krb.realms 85, 489, 491, 494, 501
 krb-srvtab 85, 475, 489, 491, 494
 login.cfg 481
 logmgmt.acls 328
 netsvc.conf 209
 newpass.log 393
 nim.installp 144
 nim_attr 149
 nim_attr.vc 150
 nim_object 150
 nim_object.vc 150
 nim_pdatr 150
 nim_pdatr.vc 150
 niminfo 145
 nimlog 144
 node_hostname.config_info 175
 node_hostname.install_info 175
 node_number.lscfg 214
 node-name.config.log.PID 166, 175
 node-name.configfb.log.PID 175
 nologin 396
 ntp.conf 411
 pacct 314, 427
 passwd 409
 passwd.id.idx 403
 passwd.idx 403
 passwd.nm.idx 403
 pmandefaults 296
 principal.dir 489, 494
 principal.ok 453, 489, 494
 principal.pag 453, 489, 494
 profile 401
 psspbooks.html 437
 rc.net 356
 rc.nfs 406
 rc.switch 169
 resolv.conf 209
 script.cust 320
 SDR_dest_info 78, 165, 169

- server_name 389
- services 450, 451
 - port 750 450
 - port 751 450
 - port 754 451
- smit.log 309
- smit.script 111, 309
- SPdaemon.log 330
- spmuser.default 393
- spmon_ctest.log 13, 41
- spot.out.PID 144
- spot.out.process_ID 142
- ssp.public.README 126
- sulog 314
- summlog 352
 - index 353
 - label 353
 - node name 352
 - partition 353
 - snap 352
 - time stamp 352
- summlog.old 352
- switch_node_number 167
- sysctl.acl 468
- sysctl.mmcmd.acl 468
- sysctl.pman.acl 285, 468
- sysctl.rootcmds.acl 468
- sysctl.vsd.acl 468, 511
- syslog.conf 330
- sysman.tab 332
- tftpaccessctl 161
- toc 139
- trcfile 313
- user 396, 482
- user.admin 385, 421
 - execute exec-command (filename) 386
 - symlinkall 385
 - upgrade filename 386
- vsd.debuglog 526
- wtmp 313
- FixDist 103
- frame information 12, 60
 - delete 71
- frame number 4
- frame supervisor microcode 98
- frame supervisor microcode level 17
- frameControllerNotResponding 252
- framePowerOff 252
- FRU 90

G

- Group Services 522
- GS 522

H

- ha.vsd 523
- HACMP/ES 505
- hardware monitor 249
 - Hardware Perspective 250
- hardware control 231
 - Hardware Perspective 233
 - fence 239
 - key switch position 239
 - network boot 241
 - open TTY 240
 - power off 239
- hmcmds
 - boot supervisor card 245
 - command 243
 - download supervisor microcode 247
 - execute basecode version 246
 - POST 244
 - SP frame ID 244
 - switch basecode version 245
- spmon
 - command 241
 - key switch position 242
 - power 243
 - reset 243
- hardware Ethernet address 26, 47
 - delete 62, 67, 72
- hardware information 213
 - control workstation 218
 - SDR 218
 - SP frame 219
 - SP node 213, 225
 - SP switche 228
 - SP-attached server 220
 - VPD 213
- hardware monitor
 - Hardware Perspective 250
 - original condition 263
 - real-time monitor 251, 255, 259
 - snapshot 253, 257, 261
 - SP frames or SP-attached servers 251
 - SP node or SP-attached server 254
 - SP Switch board 258
- hmmon

- command 265
- SP frame 265
- SP node 270
- SP Switch board 272
- SP-attached server 268
- spmon
 - command 263
- Hardware Perspective 235
 - Add Pane dialog box 235
 - Fence or Unfence Nodes dialog box 240
 - frame properties 220, 222
 - Frames and Switches pane 237, 251
 - Network Boot Nodes dialog box 241
 - node properties 225, 228
 - Power Off, Reset, Shutdown or Fence Nodes dialog box 239
 - Set Monitoring for Frames and Switches notebook 251, 259
 - Set Monitoring for Nodes notebook 255
 - SP node notebook 238
 - SP switch board properties 230
 - View or Modify Properties Frame notebook 253
 - View or Modify Properties Node notebook 256
 - View or Modify Properties SwitchBoard notebook 260
- hardware variable 267, 271, 274
- host name resolution 24
- hostResponds 35, 55, 255
- HTML 429, 436
 - install 436
- HyperText Markup Language 429

I

- IBM General Parallel File System for AIX 468
- IBM Recoverable Virtual Shared Disk 509
 - configure 509
 - configure and activate 519
 - install 510
 - log file 526
 - monitor 522
 - recovery subsystem 523
- IBM Virtual Shared Disk 509
 - configure and activate 519
 - create 515
 - install 510
 - resource variable 523
 - verify 520
- IBM Virtual Shared Disk node
 - designate 513
- IBM Virtual Shared Disk Perspective 510
 - Add Pane dialog box 515
 - Control IBM RVSD Subsystem dialog box 520
 - Create IBM VSDs dialog box 517
 - Designate as an IBM VSD Node dialog box 514
 - Run Command on Nodes dialog box 521
 - start 511
- IBM.PSSP.pm.Errlog 340
- IBM.PSSP.VSDdrv.RVSD_status 525
- initial host name 33, 53, 207
- instance 459
- ipforwarding 423

K

- KDC 449
- Kerberos 126, 445
- Kerberos daemons
 - start 452
 - stop 452
- Kerberos master key 453
 - change 457
- Kerberos master key cache file 453
- Kerberos system 484
 - back up 492
 - configure 484
 - restore 494
 - unconfigure 490
- Kerberos Version 4 446
- Kerberos Version 5 446
- kerberos_admin 450
- kerberos4 450
- Key Distribution Center 449
- krb_prop 451

L

- licensed program product 102
- locale 157
- location code 184, 187
- logical device name 184
- LPP 102
- lpp_source object 138
 - check 140
 - lppsource_name 138
 - option 147
 - requirement 146
 - spot_name 141
 - update 140

lpp_source type 135
lppsource_name 31, 51

M

Machine Type and Model 94
machines class 134
man page 429
 use 431
managing events 279
 Event Perspective 285
 haemqvar
 explanation 302
 list 302
 value 304
 haemqvar command 301
 pmandef
 command 296
 subscribe 297
 pmanquery
 list 300
managing SP switch 343
MANPATH 431
microcode for devices 87, 93
 check level 93
 SSA adapter and disk 94
 tape drive 93
 upgrade 95
 download 96
 Web 95
mirroring root volume group 190
mkysb 113

N

Netscape Navigator 436, 437
 install 436
network adapter
 add 199
 define 205
network boot image 141, 154
 network boot adapter 141
 platform 141
 processor architecture 141
network configuration 197
Network Information Service 405
Network Installation Management 113, 133
Network Time Protocol 410
networks class 134
NIM 113, 133

 configuration file 145
 NIM client 145
 NIM master 145
 log file 144
NIM client 134
 create 137
 delete 136
NIM information in ODM 149
NIM master 134
NIS 405
node 0 5
node customization 163
 changes
 CWS 166
 node 167
 isolate problem 169
 pssp_script 170
 psspsb_script 172
 with reboot 164
 without reboot 164
node group 368
 shut down 370
 start 369
node information 19, 24, 46
 delete 66
node installation 153
 isolate problem 160
 monitoring offline 155
 bootlog file 155
 bosinst.data file 157
 bosinstlog file 158
 devinst.log file 158
 errlog file 159
 image.data file 159
 monitoring online 153
 s1term 153
 three digit code 155
node number 5
node placement scheme 6
node supervisor microcode 98
node supervisor microcode level 23
nodeEnvProblem 256
nodePowerDown 256
nodePowerLED 256
NTP 126, 410
numbering scheme 4
 frame 4
 node 5
 slot 4

switch port 6

O
Object Data Manager 119, 217
ODM 119, 217
 switch_node_number 167
online documentation 429

P
pageSpaceLow 256
Part Number 94
PDF 429, 432
 install 433
PDS 126
Perl 126, 305
physical partition 178
physical volume 179
pman 296
Portable Document Format 429
POST 244
Power-On Self Tests 244
PP 178
pre-defined event definition 287
primary backup node 343
primary node 343
principal 459
 access authorization 465
 add 461
 change password 468
 default values 473
 delete 466
 expiration date 470
 information 460
 list 459
 maximum ticket lifetime 471
printer spool 313
private key 450
Problem Management subsystem 296
Program Temporary Fix 102, 140
PTF 102, 140
 apply to AIX 108
 CWS and nodes 109
 SPOT 113
 apply to LPP 116
 CWS and node 117
 get 102
 anonymous FTP servers 105
 Web 103
 PTF set 103
 public domain software 126
 PV 179

R
real time monitor 249
realm 452, 496
rearm event notification
 check 293
rearm expression 283
receive pool size 346
remote commands 452
REQUIRED_SIMAGES 147
REQUIRED_SIMAGES_41 147
REQUIRED_SIMAGES_42 147
resource ID 280
 element name 280
 element value 280
resource variable 279
 resource 280
 resource attribute 280
resources class 134
restore
 /spdata on CWS 120
 rootvg mirroring node 125
 rootvg on CWS 118
 rootvg on SP node 122
root volume group mirroring
 configure 190
 discontinue 195
 initiate 192
ROS Level and ID 94
rpoolsize 347
RS/6000 Cluster Technology 279
RS/6000 SP Resource Center 438
RS-232 control line 12
RSCT 279

S
s1term connection 39
SAMI connection 39
SBS 282, 340
Scripts
 cde_cw_restrict_login 401
 cleanup.logs.nodes 337
 cleanup.logs.ws 337
 css.snap 352
 cw_restrict_login 400

- install_swlog 355
- ntp_config 411
- pssp_script 81, 165
 - log file 171
 - three digit code 170
- psspfb_script
 - log file 173
 - three digit code 172
- rc.sp 168
- rc.switch 167, 169
- st_set_switch_number 167
- tuning.cust 168
- vsd.CSER1 527
- vsd.CSER2 527
- vsd.DOWN1 526, 527
- vsd.DOWN2 526, 527
- vsd.FENCE0 527
- vsd.FENCE1 527
- vsd.FENCE2 527
- vsd.FENCE3 527
- vsd.REFRESH 527
- vsd.UP1 527
- vsd.UP2 527
- SDR 11, 218
 - Frame
 - frame_type 219, 221
 - hardware_protocol 219, 221
 - s1_tty 219, 221
 - slots 219, 221
 - Node 163
 - description 223, 226
 - hardware_control_type 224, 227
 - platform 223, 226
 - processor_type 223, 226
 - SP 380, 390, 412, 419, 425
 - acct_master 427
 - amd_comfig 420
 - amd_config 392
 - change home directory server 420
 - filecoll_config 381, 391
 - homedir_path 420
 - homedir_server 420
 - ntp_config 413
 - ntp_server 414
 - ntp_versoin 414
 - spacct_actnode_thresh 426
 - spacct_enable 426
 - spacct_exclude_enable 427
 - supfilesrv_port 381
 - supman_uid 381
 - usermgmt_config 391
 - Switch
 - switch_level 229
 - switch_name 229
 - switch_type 229
 - switch_respond 346
 - Syspar 447, 448
 - auth_methods 447
 - auth_root_rcmd 448
 - Volume_Group 125
 - security 445
 - send pool size 346
 - Service and Manufacturing Interface (SAMI) 39
 - service key file 475
 - change 478
 - create 475
 - information 477
 - service principals 451
 - hardmon 451
 - rcmd 451
 - service tickets 450
 - services_config 405
 - session key 449
 - Shared Product Object Tree 113
 - short host name 53
 - SIMAGES_OPTIONS 148
 - SIMAGES_OPTIONS_41 148
 - SIMAGES_OPTIONS_42 148
 - slave server 407
 - slot 0 5, 233
 - slot 17 5, 233
 - slot number 4
 - smitty
 - add_adapt_dialog 27, 48, 201
 - annotator 29, 50, 63, 68, 73
 - bffcreate 200
 - bootlist_dialog 186
 - changevg_dialog 31, 51, 122, 190, 195
 - chuser 397
 - chypdom 406
 - createvg_dialog 183, 189
 - delete_frame_dialog 60, 71
 - delete_node_dialog 66
 - hostname_dialog 33, 53, 207
 - hrdwrad_dialog 26, 47, 198
 - install_latest 79, 430
 - installp 109
 - maktty 12, 39

- mkclient 408
- mkftusersfast 398
- mkmaps 409
- mkmaster 406
- mknfsmnt 424
- mkslave 407
- mksysb 118
- mktcpip 77
- nim_res_op 114
- nosp_frame_dialog 40
- perrpt 328
- png_create 371
- primary_node_dialog 344
- reducevg 179
- remove 83
- restvg 120, 179
- rmftusers 399
- rmlvcopy 181
- savevg 119, 179
- server_dialog 34, 54, 77, 123, 185, 203
- site_env_dialog 405, 414, 420, 422, 427
- sp_eth_dialog 24, 46
- sp_frame_dialog 12
- spauth_config 448
- spauth_methods 360, 447
- spauth_rcmd 360
- spscreate_archive 333
- sperrlog 329
- spgather_archive 336
- spmuser 392
- sprmuser 394
- spsyslog 331
- start_mirroring 193
- stop_mirroring 196
- supervisor 98, 100
- syspar_apply 359
- syspar_cust 357
- snapshot 249
- software resources management 309
 - Event Perspective 315
 - /tmp 316
 - /var 316
 - file system 315
 - file system 309
 - / 309, 311
 - /home 310
 - /spdata 310
 - /tmp 310
 - /usr 309
 - /var 310, 313
 - log file 327
 - Event Perspective 338
 - pmandef command 339
 - paging space 317
 - adjusting 320, 322
 - Event Perspective 327
 - getting information 318
 - sizing 318
 - pmandef
 - /tmp 316
 - /var 317
 - pmandef command 316
 - Software Update Protocol 379
 - SP accounting 425
 - information 425
 - SP Ethernet 197
 - replace 197
 - SP Ethernet address 25, 46
 - SP man page 430
 - install 430
 - SP Perspectives 439
 - SP Perspectives Launch Pad 439
 - SP Resource Center
 - install 438
 - SP Switch adapter 27, 48
 - SP Switch address 28
 - SP Switch information 30, 64
 - delete 69, 73
 - SP Switch topology file 29, 50, 63, 68, 73, 351
 - SP System Monitor 232, 250
 - SP System Monitor Access Control List 232, 250
 - SP User Management 390
 - add user 392
 - control log in 395
 - delete user 394
 - stop using 404
 - using NIS 405
 - SP-attached server information 40
 - spoolsize 347
 - SPOT 113
 - log file 142
 - spot object 141
 - check 141
 - option 149
 - update 144
 - spot type 135
 - srvtab file 475
 - SSA drive serial number 188

- standalone type 135
- start the SP Switch 36
- Structured Byte String 282, 340
- SUP 126, 380
- supervisor microcode 97
 - check level 98
 - downgrade 101
 - get 97
 - upgrade 100
- switch port number (SP Switch) 6
- switch port number (SP Switch-8) 9
- switch port number (SP-attached server) 9
- switch supervisor microcode 98
- switchResponds 37, 56, 256
- system backup 117
- System Data Repository 11, 218
- system firmware 87, 88
 - check level 88
 - 200 MHz SMP high node 89
 - 332 MHz SMP wide/thin node 89
 - SP-attached server 88
 - download 91
 - upgrade 92
 - Web 90
- system partition 355
 - apply 355
 - delete 361
 - original partition configuration/layout 362
 - system partition configuration 356
 - system partition layout 356
 - System Partitioning Aid Perspective 362
 - Define System Partition dialog box 364
 - Display Existing Configurations window 367
 - System partitions loaded from SDR pane 365

T

- target_disk_data 157
- Tcl 126
- TclX 126
- TCP 450
- TFTP 141, 154
- TGS 450
- three digit code 173
- ticket
 - destroy 483
- ticket cache file 480
- Ticket-Granting Service 450
- ticket-granting-ticket 449

- time synchronization 410
 - change NTP time server 414
 - change system time 416
 - information 412
 - monitor NTP status 415
 - rule 412
- Tk 126
- tmpFull 256
- Token-Ring adapter 199
- trace log 313
- Transmission Control Protocol 450
- Trivial File Transfer Protocol 141, 154
- twin-tailed disk 509

U

- UDP 450
- user account 314
- User Datagram Protocol 450

V

- varFull 256
- VFOP 233
- Virtual Front Operator Panel 233
- Vital Product Data 213
- VPD 213

W

- WCOLL 373
- working collective 373
- wrapper 136
 - allnimres 136
 - delnimclient 136
 - l 137, 198
 - delnimmast 136
 - mknimclient 136
 - l 137, 199
 - mknimint 136
 - mknimmast 136
 - mknimres 136
 - unallnimres 136

ITSO redbook evaluation

RS/6000 SP System Management: Power Recipes for PSSP 3.1
SG24-5628-00

Your feedback is very important to help us maintain the quality of ITSO redbooks. **Please complete this questionnaire and return it using one of the following methods:**

- Use the online evaluation form found at <http://www.redbooks.ibm.com/>
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Which of the following best describes you?

Customer **Business Partner** **Solution Developer** **IBM employee**
 None of the above

Please rate your overall satisfaction with this book using the scale:
(1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor)

Overall Satisfaction _____

Please answer the following questions:

Was this redbook published in time for your needs? Yes___ No___

If no, please explain:

What other redbooks would you like to see published?

Comments/Suggestions: (THANK YOU FOR YOUR FEEDBACK!)

SG24-5628-00
Printed in the U.S.A.

RS/6000 SP System Management: Power Recipes for PSSP 3.1

SG24-5628-00

