

SAS RAID controllers for Linux on the M6-700 or M7-700□□

ESCALA Power7



REFERENCE
86 A1 65FF 01

ESCALA Power7

SAS RAID controllers for Linux on the M6-700 or M7-700□□

This publication concerns the following models:

- Bull Escala E5-700 (Power 750 / 8233-E8B)
- Bull Escala M6-700 (Power 770 / 9117-MMB)
- Bull Escala M7-700 (Power 780 / 9179-MHB)
- Bull Escala E1-700 (Power 710 / 8231-E2B)
- Bull Escala E2-700 / E2-700T (Power 720 / 8202-E4B)
- Bull Escala E3-700 (Power 730 / 8231-E2B)
- Bull Escala E4-700 / E4-700T (Power 740 / 8205-E6B)

References to Power 755 / 8236-E8C models are irrelevant.

Hardware

September 2010

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

REFERENCE
86 A1 65FF 01

The following copyright notice protects this book under Copyright laws which prohibit such actions as, but not limited to, copying, distributing, modifying, and making derivative works.

Copyright © Bull SAS 2007-2010

Printed in France

Suggestions and criticisms concerning the form, content, and presentation of this book are invited. A form is provided at the end of this book for this purpose.

To order additional copies of this book or other Bull Technical Publications, you are invited to use the Ordering Form also provided at the end of this book.

Trademarks and Acknowledgements

We acknowledge the right of proprietors of trademarks mentioned in this book.

The information in this document is subject to change without notice. Bull will not be liable for errors contained herein, or for incidental or consequential damages in connection with the use of this material.

Contents

Safety notices	vii
SAS RAID controllers for Linux	1
What's new in SAS RAID controllers for Linux	1
General information	1
Comparison of general features	2
Comparison of cache features	11
Comparison of HA features	12
SAS overview	13
SAS architecture summary	14
Disk arrays	15
Supported RAID levels	16
Estimating disk array capacities	20
RAID level summary	20
Stripe-unit size	21
Disk array overview	21
Disk array states	23
Physical disk states	24
I/O adapter states	24
Auxiliary write cache adapter	25
RAID controller software	27
Verifying installation of the controller software	27
Linux ipr device driver updates	29
Updating the iprutils package	29
Common IBM SAS RAID controller tasks	31
Starting the iprconfig utility	31
Status of devices, arrays and paths	31
Viewing device status	31
Viewing array status	33
Viewing path status	34
RAID and JBOD formats	35
Formatting to advanced function	35
Formatting to JBOD	36
Creating and deleting disk arrays	36
Creating an IBM SAS RAID disk array	36
Deleting an IBM SAS RAID disk array	37
Adding disks to an existing array	38
Migrating an existing disk array to a new RAID level	38
Hot spare disks	40
Creating hot spare disks	40
Deleting hot spare disks	40
Considerations for solid-state drives (SSD)	41
Multi-initiator and high availability	43
Possible HA configurations	44
Controller functionality	45
Controller functionality attributes	47
Viewing HA controller attributes	48
HA cabling considerations	48
HA performance considerations	49
Configuration and serviceability considerations for HA RAID configurations	49
HA asymmetric access optimization	50

Enabling asymmetric access	55
Asymmetric access status of disk arrays	56
Installing high availability	57
Installing an HA single-system RAID configuration	57
Installing an HA two-system RAID configuration.	58
Functions requiring special attention in an HA two-system RAID configuration	60
Installing an HA two-system JBOD configuration.	61
IBM SAS RAID controller maintenance	63
Usage tips	63
Updating the controller microcode.	63
Rechargeable battery maintenance.	64
Displaying rechargeable battery information	64
Forcing a rechargeable battery error	65
Replacing the rechargeable battery pack	66
Replacing a nonconcurrently maintainable battery pack.	66
Replacing a concurrently maintainable battery pack	68
Replacing the cache directory card.	70
Physical disks	72
Removing a failed disk	72
Installing a new disk	73
Disk failure recovery	75
RAID 0 failure	75
RAID 5 disk recovery	75
Recovering a RAID 5 single-disk failure	75
RAID 5 multiple-disk failure.	75
RAID 6 disk recovery	75
Recovering a RAID 6 single- or dual-disk failure	76
RAID 6 failure of three or more disks.	76
RAID 10 disk recovery	76
Recovering a RAID 10 single-disk failure	76
RAID 10 multiple-disk failure	76
Reclaiming IOA cache storage	77
Problem determination and recovery.	79
Analyzing error logs	79
Basic vi commands	80
Searching logs	81
Sample error logs	81
Generic IOA or device errors	81
Device configuration errors	81
Array errors	82
Cache errors	82
Disk array problem identification	83
Unit reference code tables	83
Maintenance analysis procedures	88
MAP 3310	88
MAP 3311	89
MAP 3312	90
MAP 3313	91
MAP 3321	92
MAP 3330	92
MAP 3333	93
MAP 3334	93
MAP 3335	95
MAP 3337	95
MAP 3342	97
MAP 3343	97
MAP 3344	97
MAP 3345	98

MAP 3346	98
MAP 3348	98
MAP 3349	99
MAP 3350	99
MAP 3351	102
MAP 3352	103
MAP 3353	103
MAP 3390	104
Notices	105
Trademarks	106
Electronic emission notices	106
Class A Notices.	106
Terms and conditions.	110

Safety notices

Safety notices may be printed throughout this guide:

- **DANGER** notices call attention to a situation that is potentially lethal or extremely hazardous to people.
- **CAUTION** notices call attention to a situation that is potentially hazardous to people because of some existing condition.
- **Attention** notices call attention to the possibility of damage to a program, device, system, or data.

World Trade safety information

Several countries require the safety information contained in product publications to be presented in their national languages. If this requirement applies to your country, a safety information booklet is included in the publications package shipped with the product. The booklet contains the safety information in your national language with references to the U.S. English source. Before using a U.S. English publication to install, operate, or service this product, you must first become familiar with the related safety information in the booklet. You should also refer to the booklet any time you do not clearly understand any safety information in the U.S. English publications.

German safety information

Das Produkt ist nicht für den Einsatz an Bildschirmarbeitsplätzen im Sinne § 2 der Bildschirmarbeitsverordnung geeignet.

Laser safety information

IBM® servers can use I/O cards or features that are fiber-optic based and that utilize lasers or LEDs.

Laser compliance

IBM servers may be installed inside or outside of an IT equipment rack.

DANGER

When working on or around the system, observe the following precautions:

Electrical voltage and current from power, telephone, and communication cables are hazardous. To avoid a shock hazard:

- Connect power to this unit only with the IBM provided power cord. Do not use the IBM provided power cord for any other product.
- Do not open or service any power supply assembly.
- Do not connect or disconnect any cables or perform installation, maintenance, or reconfiguration of this product during an electrical storm.
- The product might be equipped with multiple power cords. To remove all hazardous voltages, disconnect all power cords.
- Connect all power cords to a properly wired and grounded electrical outlet. Ensure that the outlet supplies proper voltage and phase rotation according to the system rating plate.
- Connect any equipment that will be attached to this product to properly wired outlets.
- When possible, use one hand only to connect or disconnect signal cables.
- Never turn on any equipment when there is evidence of fire, water, or structural damage.
- Disconnect the attached power cords, telecommunications systems, networks, and modems before you open the device covers, unless instructed otherwise in the installation and configuration procedures.
- Connect and disconnect cables as described in the following procedures when installing, moving, or opening covers on this product or attached devices.

To Disconnect:

1. Turn off everything (unless instructed otherwise).
2. Remove the power cords from the outlets.
3. Remove the signal cables from the connectors.
4. Remove all cables from the devices

To Connect:

1. Turn off everything (unless instructed otherwise).
2. Attach all cables to the devices.
3. Attach the signal cables to the connectors.
4. Attach the power cords to the outlets.
5. Turn on the devices.

(D005)

DANGER

Observe the following precautions when working on or around your IT rack system:

- Heavy equipment—personal injury or equipment damage might result if mishandled.
- Always lower the leveling pads on the rack cabinet.
- Always install stabilizer brackets on the rack cabinet.
- To avoid hazardous conditions due to uneven mechanical loading, always install the heaviest devices in the bottom of the rack cabinet. Always install servers and optional devices starting from the bottom of the rack cabinet.
- Rack-mounted devices are not to be used as shelves or work spaces. Do not place objects on top of rack-mounted devices.



- Each rack cabinet might have more than one power cord. Be sure to disconnect all power cords in the rack cabinet when directed to disconnect power during servicing.
- Connect all devices installed in a rack cabinet to power devices installed in the same rack cabinet. Do not plug a power cord from a device installed in one rack cabinet into a power device installed in a different rack cabinet.
- An electrical outlet that is not correctly wired could place hazardous voltage on the metal parts of the system or the devices that attach to the system. It is the responsibility of the customer to ensure that the outlet is correctly wired and grounded to prevent an electrical shock.

CAUTION

- Do not install a unit in a rack where the internal rack ambient temperatures will exceed the manufacturer's recommended ambient temperature for all your rack-mounted devices.
- Do not install a unit in a rack where the air flow is compromised. Ensure that air flow is not blocked or reduced on any side, front, or back of a unit used for air flow through the unit.
- Consideration should be given to the connection of the equipment to the supply circuit so that overloading of the circuits does not compromise the supply wiring or overcurrent protection. To provide the correct power connection to a rack, refer to the rating labels located on the equipment in the rack to determine the total power requirement of the supply circuit.
- *(For sliding drawers.)* Do not pull out or install any drawer or feature if the rack stabilizer brackets are not attached to the rack. Do not pull out more than one drawer at a time. The rack might become unstable if you pull out more than one drawer at a time.
- *(For fixed drawers.)* This drawer is a fixed drawer and must not be moved for servicing unless specified by the manufacturer. Attempting to move the drawer partially or completely out of the rack might cause the rack to become unstable or cause the drawer to fall out of the rack.

(R001)

CAUTION:

Removing components from the upper positions in the rack cabinet improves rack stability during relocation. Follow these general guidelines whenever you relocate a populated rack cabinet within a room or building:

- Reduce the weight of the rack cabinet by removing equipment starting at the top of the rack cabinet. When possible, restore the rack cabinet to the configuration of the rack cabinet as you received it. If this configuration is not known, you must observe the following precautions:
 - Remove all devices in the 32U position and above.
 - Ensure that the heaviest devices are installed in the bottom of the rack cabinet.
 - Ensure that there are no empty U-levels between devices installed in the rack cabinet below the 32U level.
- If the rack cabinet you are relocating is part of a suite of rack cabinets, detach the rack cabinet from the suite.
- Inspect the route that you plan to take to eliminate potential hazards.
- Verify that the route that you choose can support the weight of the loaded rack cabinet. Refer to the documentation that comes with your rack cabinet for the weight of a loaded rack cabinet.
- Verify that all door openings are at least 760 x 230 mm (30 x 80 in.).
- Ensure that all devices, shelves, drawers, doors, and cables are secure.
- Ensure that the four leveling pads are raised to their highest position.
- Ensure that there is no stabilizer bracket installed on the rack cabinet during movement.
- Do not use a ramp inclined at more than 10 degrees.
- When the rack cabinet is in the new location, complete the following steps:
 - Lower the four leveling pads.
 - Install stabilizer brackets on the rack cabinet.
 - If you removed any devices from the rack cabinet, repopulate the rack cabinet from the lowest position to the highest position.
- If a long-distance relocation is required, restore the rack cabinet to the configuration of the rack cabinet as you received it. Pack the rack cabinet in the original packaging material, or equivalent. Also lower the leveling pads to raise the casters off of the pallet and bolt the rack cabinet to the pallet.

(R002)

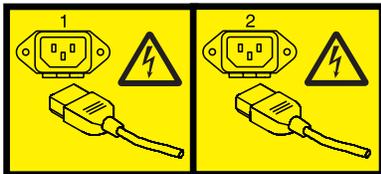
(L001)



(L002)



(L003)



or



All lasers are certified in the U.S. to conform to the requirements of DHHS 21 CFR Subchapter J for class 1 laser products. Outside the U.S., they are certified to be in compliance with IEC 60825 as a class 1 laser product. Consult the label on each part for laser certification numbers and approval information.

CAUTION:

This product might contain one or more of the following devices: CD-ROM drive, DVD-ROM drive, DVD-RAM drive, or laser module, which are Class 1 laser products. Note the following information:

- Do not remove the covers. Removing the covers of the laser product could result in exposure to hazardous laser radiation. There are no serviceable parts inside the device.
- Use of the controls or adjustments or performance of procedures other than those specified herein might result in hazardous radiation exposure.

(C026)

CAUTION:

Data processing environments can contain equipment transmitting on system links with laser modules that operate at greater than Class 1 power levels. For this reason, never look into the end of an optical fiber cable or open receptacle. (C027)

CAUTION:

This product contains a Class 1M laser. Do not view directly with optical instruments. (C028)

CAUTION:

Some laser products contain an embedded Class 3A or Class 3B laser diode. Note the following information: laser radiation when open. Do not stare into the beam, do not view directly with optical instruments, and avoid direct exposure to the beam. (C030)

Power and cabling information for NEBS (Network Equipment-Building System) GR-1089-CORE

The following comments apply to the IBM servers that have been designated as conforming to NEBS (Network Equipment-Building System) GR-1089-CORE:

The equipment is suitable for installation in the following:

- Network telecommunications facilities
- Locations where the NEC (National Electrical Code) applies

The intrabuilding ports of this equipment are suitable for connection to intrabuilding or unexposed wiring or cabling only. The intrabuilding ports of this equipment *must not* be metallically connected to the interfaces that connect to the OSP (outside plant) or its wiring. These interfaces are designed for use as intrabuilding interfaces only (Type 2 or Type 4 ports as described in GR-1089-CORE) and require isolation from the exposed OSP cabling. The addition of primary protectors is not sufficient protection to connect these interfaces metallically to OSP wiring.

Note: All Ethernet cables must be shielded and grounded at both ends.

The ac-powered system does not require the use of an external surge protection device (SPD).

The dc-powered system employs an isolated DC return (DC-I) design. The DC battery return terminal *shall not* be connected to the chassis or frame ground.

SAS RAID controllers for Linux

The PCI-X and PCIe SAS RAID controller is available for various versions of the Linux[®] kernel. Here you will find out how to use and maintain the controller.

What's new in SAS RAID controllers for Linux

Read about new or significantly changed information for SAS RAID controllers for Linux since the previous update of this topic collection.

September 2010

- Updated the “Comparison of general features” on page 2 table to include information about 57CD, 2BE0, 2BE1, and 2BD9 cards.
- Updated the “Comparison of cache features” on page 11 table to include information about 57CD, 2BE0, 2BE1, and 2BD9 cards.
- Updated the “Comparison of HA features” on page 12 table to include information about 57CD, 2BE0, 2BE1, and 2BD9 cards.
- Updated the “Verifying installation of the controller software” on page 27 table to include information about 57CD card.
- Updated the “Unit reference code tables” on page 83 to include information about MAP 3300.

General information

This section provides general information about the IBM SAS RAID controllers for Linux.

The controllers have the following features:

- PCI-X 266 system interface or PCI Express[®] (PCIe) system interface.
- Physical link (phy) speed of 3 Gb per second Serial Attached SCSI(SAS) supporting transfer rates of 300 MB per second.
- Supports SAS devices and non-disk Serial Advanced Technology Attachment (SATA) devices.
- Optimized for SAS disk configurations which utilize dual paths thru dual expanders for redundancy and reliability.
- Controller managed path redundancy and path switching for multiported SAS devices.
- Embedded PowerPC[®] RISC processor, hardware XOR DMA engine, and hardware Finite Field Multiplier (FFM) DMA Engine (for Redundant Array of Independent Disks (RAID) 6).
- Supports nonvolatile write cache for RAID disk arrays on some adapters.
- Support for RAID 0, 5, 6, and 10 disk arrays.
- Supports attachment of other devices such as non-RAID disks, tape, and optical devices.
- RAID disk arrays and non-RAID devices supported as a bootable device.
- Advanced RAID features:
 - Hot spares for RAID 5, 6, and 10 disk arrays
 - Ability to increase the capacity of an existing RAID 5 or 6 disk array by adding disks
 - Background parity checking
 - Background data scrubbing
 - Disks formatted to 528 bytes per sector, providing cyclical redundancy checking (CRC) and logically bad block checking
 - Optimized hardware for RAID 5 and 6 sequential write workloads

- Optimized skip read/write disk support for transaction workloads
- Supports a maximum of 64 advanced function disks with a total device support maximum of 255 (the number of all physical SAS and SATA devices plus the number of logical RAID disk arrays must be less than 255 per controller).

Note: This information refers to various hardware and software features and functions. The realization of these features and functions depends on the limitations of your hardware and software. Linux supports all functions mentioned. If you are using another operating system, consult the appropriate documentation for that operating system regarding support for the mentioned features and functions.

Related reference

“Related information” on page 13

Many other sources of information on Linux, RAID, and other associated topics are available.

“References to Linux” on page 13

Three different versions of Linux are referred to in this set of information.

Comparison of general features

This table presents a comparison of general features of the SAS RAID controller cards for the Linux operating system.

Table 1. Comparison of general features

CCIN (Custom Card Identi- fication Number)	Description	Form factor	Adapter failing function code LED value	Physical links	RAID level supported	Write cache size ⁴	Read cache size
572A	PCI-X 266 Ext Dual-x4 3 Gb SAS adapter	Low profile 64-bit PCI-X	2515	8 (two mini SAS 4x connectors)	0, 5 ² , 6 ² , 10	0	
572B	PCI-X 266 Ext Dual-x4 3 Gb SAS RAID adapter	Long 64-bit PCI-X	2517	8 (two mini SAS 4x connectors)	0, 5, 6, 10	175 MB	
572C	PCI-X 266 Planar 3 Gb SAS adapter	Planar integrated	2502	8 ¹	0	0	
572F	PCI-X 266 Ext Tri-x4 3 Gb SAS RAID adapter	Long 64-bit PCI-X, double wide card set	2519	12 (bottom 3 mini SAS 4x connectors) and 2 (top mini SAS 4x connector for HA only)	0, 5, 6, 10	Up to 1.5 GB (compressed)	Up to 1.5 GB (compressed)
574E	PCIe x8 Ext Dual-x4 3 Gb SAS RAID adapter	PCIe x8	2518	8 (two mini SAS 4x connectors)	0, 5, 6, 10	380 MB	
575C	PCI-X 266 auxiliary cache card	Long 64-bit PCI-X, double wide card set	251D	NA	NA	Up to 1.6 GB (compressed)	Up to 1.6 GB (compressed)

Table 1. Comparison of general features (continued)

CCIN (Custom Card Identi- fication Number)	Description	Form factor	Adapter failing function code LED value	Physical links	RAID level supported	Write cache size ⁴	Read cache size
57B3	PCIe x8 Ext Dual-x4 3 Gb SAS adapter	PCIe x8	2516	8 (two mini SAS 4x connectors)	0, 5 ² , 6 ² , 10	0	
57B7	PCIe x1 auxiliary cache adapter	Planar auxiliary cache	2504	2	NA	175 MB	
57B8	PCI-X 266 Planar 3 Gb SAS RAID adapter	Planar RAID enablement	2505	8 ¹	0, 5, 6, 10	175 MB	
57B9	PCIe x8 Ext Dual-x4 3 Gb SAS adapter and cable card	Combination PCIe x8 and cable card	2D0B	4 (bottom mini SAS 4x connector required to connect via an external AI cable to the top mini SAS 4x cable card connector)	0, 5 ² , 6 ² , 10	0	
57BA	PCIe x8 Ext Dual-x4 3 Gb SAS adapter and cable card	Combination PCIe x8 and cable card	2D0B	4 (bottom mini SAS 4x connector required to connect via an external AI cable to the top mini SAS 4x cable card connector)	0, 5 ² , 6 ² , 10	0	
57C7	PCI Express x8 Planar 3 Gb SAS adapter (disk/media backplane)	Planar	2D14	8	0, 5 ¹ , 6 ¹ , 10		
57CD	PCIe SAS RAID and SSD Adapter 3 Gb x8	Double-wide PCIe x8 with 1-4 integrated solid-state drives (SSD)	2D40	4 (one direct SAS physical link to each integrated SSD)	0 ⁵ , 5, 6 ³		

Table 1. Comparison of general features (continued)

CCIN (Custom Card Identification Number)	Description	Form factor	Adapter failing function code LED value	Physical links	RAID level supported	Write cache size ⁴	Read cache size
57CF	PCI Express x8 Planar 3 Gb SAS RAID adapter (with 175 MB Cache RAID - Dual IOA Enablement Card)	Planar and Enablement Card	2D15	8	0, 5, 6, 10	175 MB	
2BD9	PCI-X266 Planar 3 Gb SAS Adapter (RAID/Cache Storage Controller)	Planar unique 64-bit PCI-X	2D18	6 (two 2x wide ports to shared SAS drives and one 2x wide port to 2BE1 adapter)	0, 5, 6, 10	175 MB	
2BE0	PCI-X266 Planar 3 Gb SAS Adapter (Raid-10 Storage Controller)	Planar unique 64-bit PCI-X	2D16	3 (directly attached SAS drives)	0, 5 ² , 10		
2BE1	PCI-X266 Planar 3 Gb SAS Adapter (Raid/Cache Enablement)	Planar RAID enablement	2D17	8 (two 2x wide ports to shared SAS drives, one 2x wide port to 2BD9 adapter, one physical link to DVD, and optionally one physical link to tape drive)	0, 5, 6, 10	175 MB	

¹ Some systems provide an external mini SAS 4x connector from the integrated backplane controller.

² The write performance of RAID level 5 and RAID level 6 may be poor on adapters which do not provide write cache. Consider using an adapter which provides write cache when using RAID level 5 or RAID level 6 or utilize SSDs where supported to improve write performance.

³ The integrated SSDs are allowed to run in JBOD mode.

⁴ Nonvolatile write cache is supported for RAID disk arrays only.

⁵ RAID 0 requires the use of Software Raid in Linux.

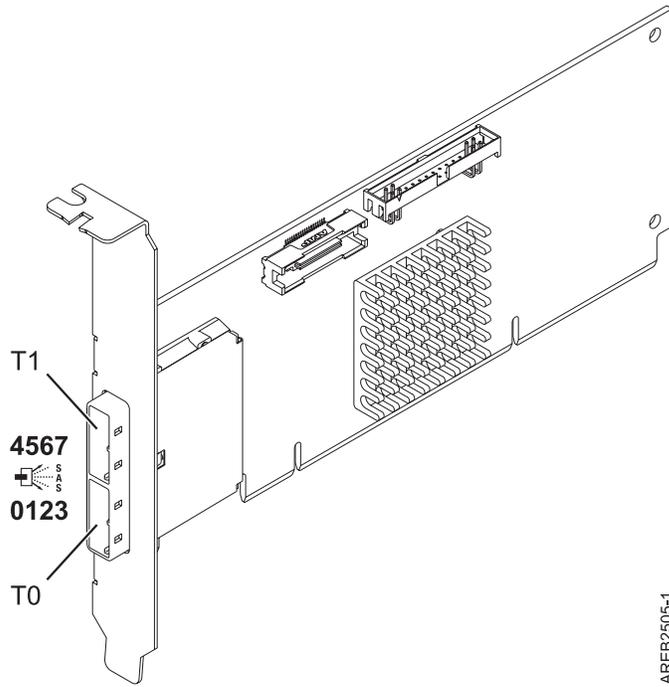


Figure 1. CCIN 572A PCI-X266 External Dual-x4 3 Gb SAS adapter

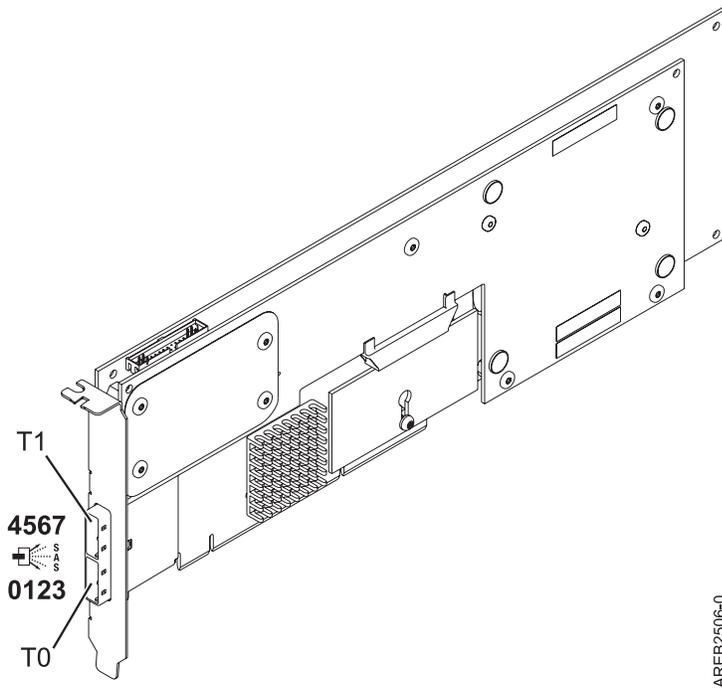
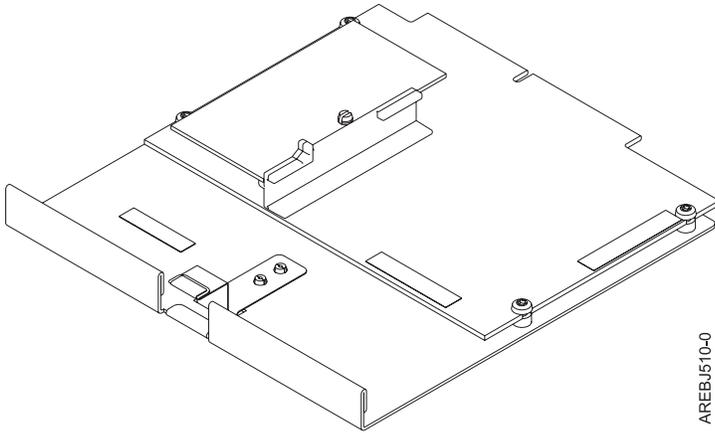
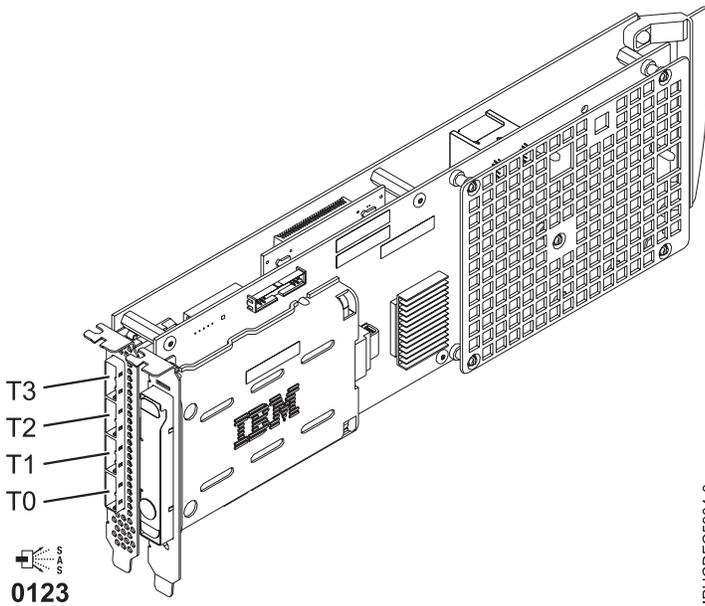


Figure 2. CCIN 572B PCI-X266 Ext Dual-x4 3 Gb SAS RAID adapter



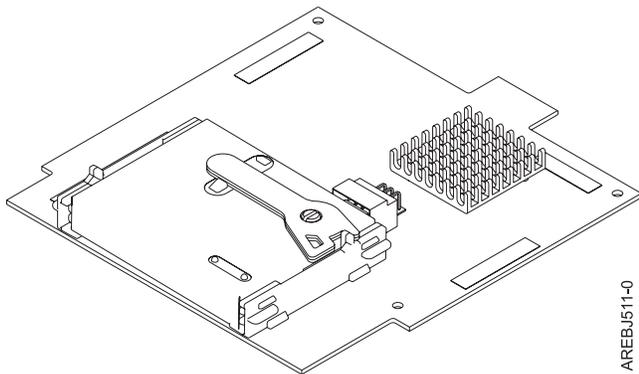
AREBJ510-0

Figure 3. CCIN 57B8 Planar RAID enablement card



IPHCDFC5904-0

Figure 4. CCIN 572F PCI-X266 Ext Tri-x4 3 Gb SAS RAID adapter and CCIN 575C PCI-X266 auxiliary cache adapter



AREBJ511-0

Figure 5. CCIN 57B7 Planar auxiliary cache

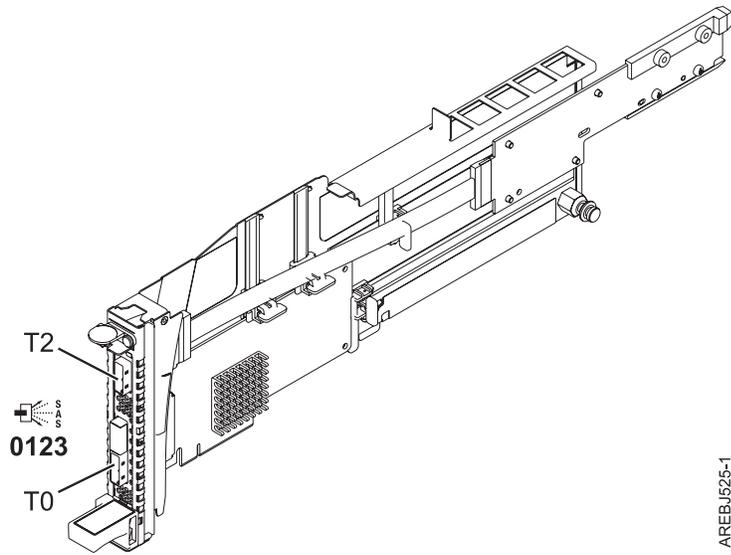


Figure 6. CCIN 57B9 PCI Express x8 Ext Dual-x4 3 Gb SAS adapter and cable card

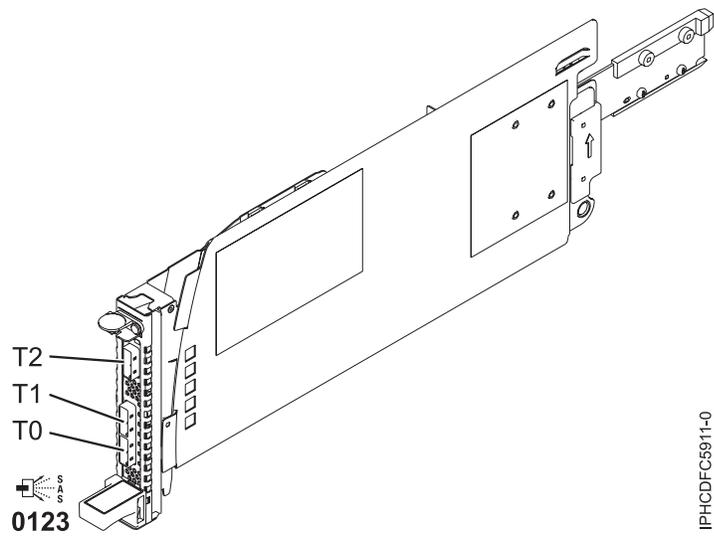


Figure 7. CCIN 57BA PCI Express x8 Ext Dual-x4 3 Gb SAS adapter and cable card

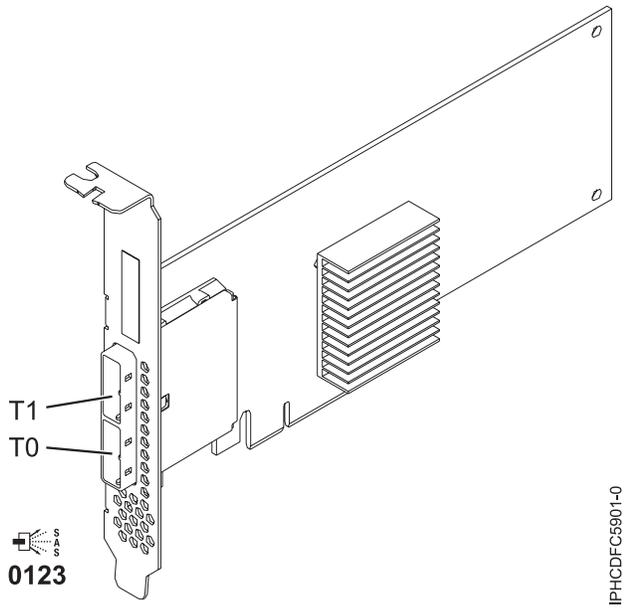


Figure 8. CCIN 57B3 PCI Express x8 Ext Dual-x4 3 Gb SAS adapter

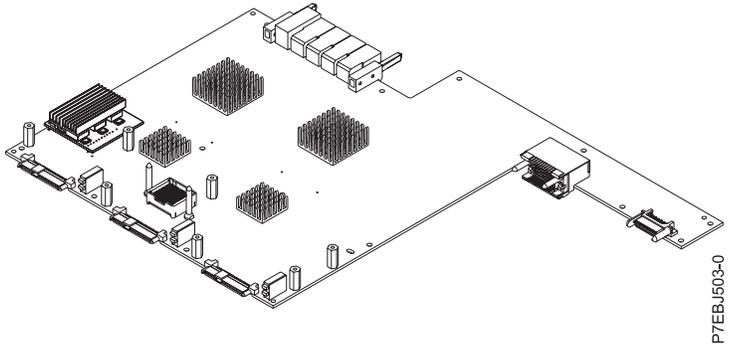


Figure 9. CCIN 57C7 PCI Express x8 Planar 3 Gb SAS Adapter (with 175 MB Cache RAID - Dual IOA Enablement card)

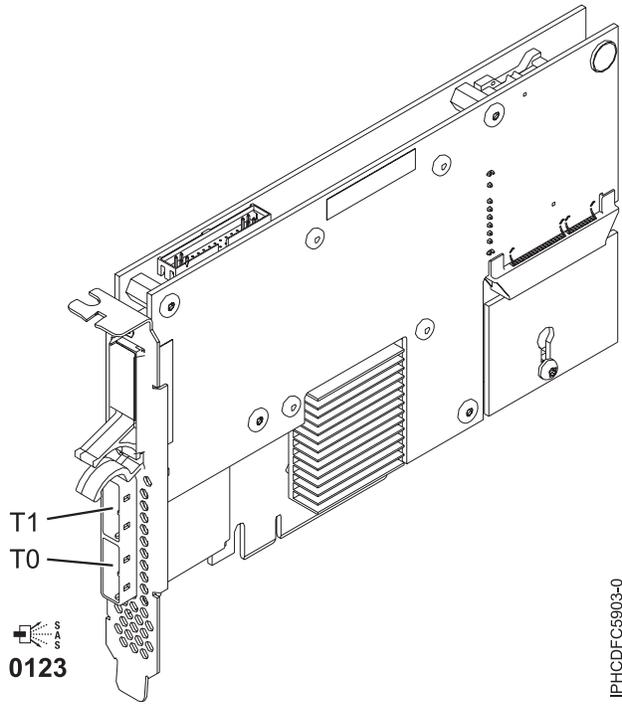


Figure 10. CCIN 574E PCI Express x8 Ext Dual-x4 3 Gb SAS RAID adapter

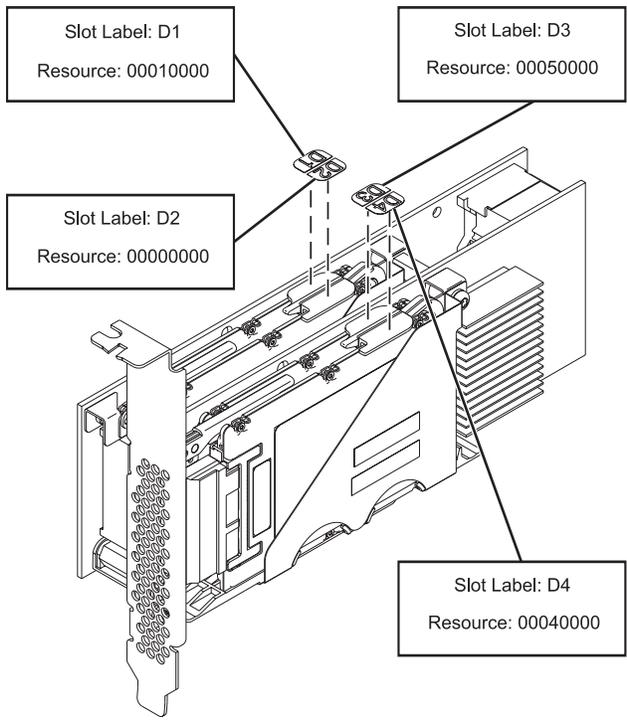


Figure 11. CCIN 57CD PCIe SAS RAID and SSD adapter 3 Gb x8

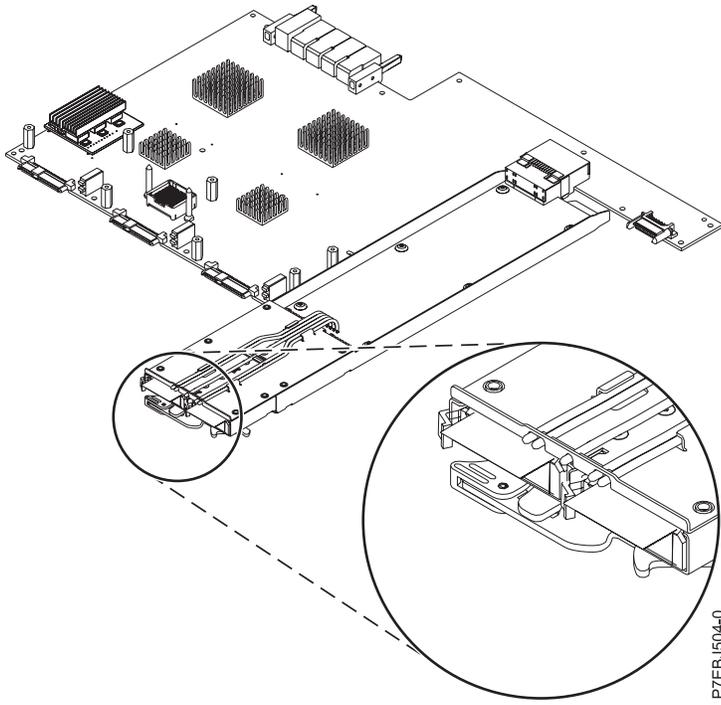


Figure 12. CCIN 57CF PCI Express x8 Planar 3 Gb SAS RAID adapter

RAID/Cache Enablement

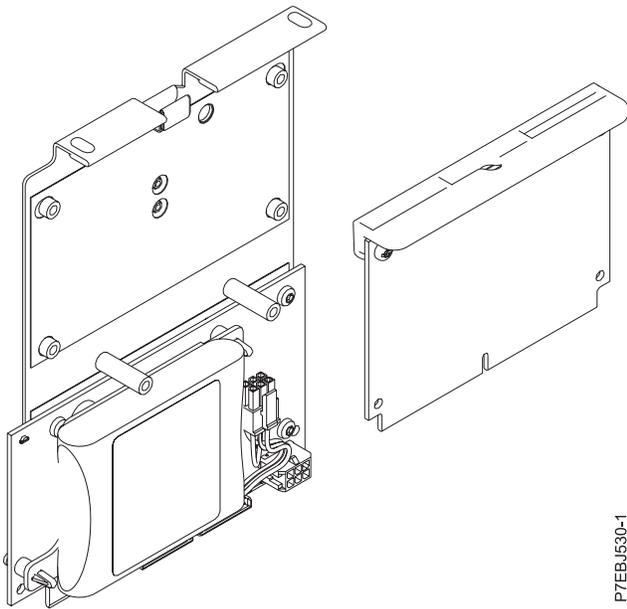
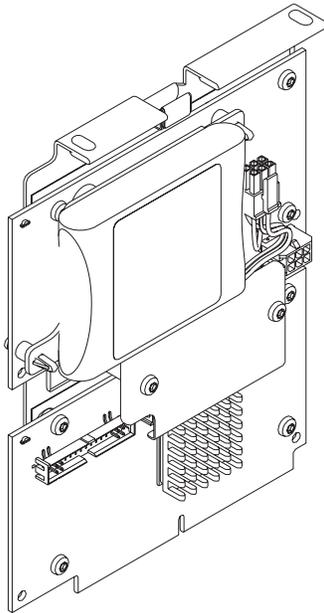


Figure 13. CCIN 2BE0 PCI-X266 Planar 3 Gb SAS adapter

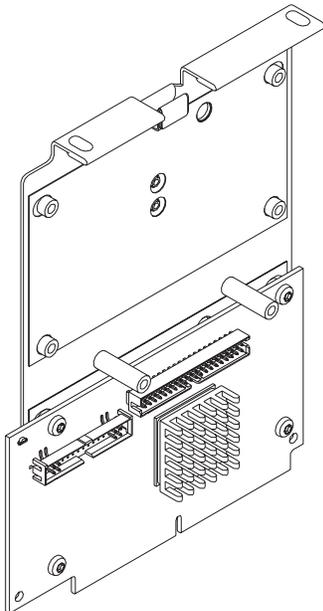
Planar RAID/Cache Storage Controller



P7EBJ531-1

Figure 14. CCIN 2BD9 PCI-X266 Planar 3 Gb SAS RAID adapter

Planar RAID 10 Storage Controller



P7EBJ532-1

Figure 15. CCIN 2BE1 PCI-X266 Planar 3 Gb SAS RAID adapter

Comparison of cache features

This table presents a comparison of cache features of the SAS RAID controller cards for the Linux operating system.

Table 2. Comparison of cache features

CCIN (Custom Card Identification Number)	Cache battery pack technology	Cache battery pack FFC	Cache battery concurrent maintenance	Cache data present LED	Removable cache card	Auxiliary write cache (AWC) support
572A	NA	NA	No	No	No	No
572B	LiIon	2D03	No	No	Yes	No
572C	NA	NA	No	No	No	No
572F	NA	NA	NA	No	No	Yes
574E	LiIon	2D0E	No	Yes	Yes	No
575C	LiIon	2D06 ²	Yes	No	No	Yes
57B3	NA	NA	No	No	No	No
57B7	LiIon	2D05	Yes	Yes	No	Yes
57B8	NA ¹	NA ¹	NA ¹	No	No	Yes
57B9	NA	NA	No	No	No	No
57BA	NA	NA	No	No	No	No
57CD	NA	NA	No	No	No	No
2BD9	Lilon	2D1B	No	Yes	No	Yes
2BE0	NA	NA	No	No	No	No
2BE1	Lilon	2D1B	No	Yes	No	Yes

¹ The controller contains battery-backed cache, but the battery power is supplied by the 57B8 controller through the backplane connections.

² The cache battery pack for both adapters is contained in a single battery FRU which is physically located on the 575C auxiliary cache card.

Comparison of HA features

Use the table here to compare the High Availability features of the SAS RAID controller cards for Linux.

Table 3. Comparison of HA features

CCIN (Custom Card Identification Number)	High Availability (HA) two-system RAID	High Availability (HA) two-system JBOD	High Availability (HA) single-system RAID	Requires HA RAID configuration
572A	Yes ¹	Yes ¹	Yes ¹	No
572B	Yes	No	Yes	Yes
572C	No	No	No	No
57B7	NA	NA	NA	NA
57B8	No	No	No	No
57B9	No	No	No	No
57BA	No	No	No	No
57B3	Yes	Yes	Yes	No
574E	Yes	No	Yes	Yes
572F	Yes	No	Yes	No
575C	NA	NA	NA	NA
57CD	No	No	No	No

Table 3. Comparison of HA features (continued)

CCIN (Custom Card Identification Number)	High Availability (HA) two-system RAID	High Availability (HA) two-system JBOD	High Availability (HA) single-system RAID	Requires HA RAID configuration
2BE0	No	No	No	No
2BE1	Yes	No	Yes	Yes
2BD9	Yes	No	Yes	Yes

¹ Multi-initiator and high availability is supported on the CCIN 572A adapter except for part numbers of either 44V4266 or 44V4404 (feature code 5900).

Note: This information refers to various hardware and software features and functions. The realization of these features and functions depends on the limitations of your hardware and software. Linux supports all functions mentioned. If you are using another operating system, consult the appropriate documentation for that operating system regarding support for the mentioned features and functions.

References to Linux

Three different versions of Linux are referred to in this set of information.

References to the Linux operating system in this topic collection include the versions Linux 2.6, SuSE Linux Enterprise Server 10, RedHat Enterprise Linux 4 and RedHat Enterprise Linux 5. Make sure you are consulting the appropriate section of this set of information for the operating system you are using.

This document might describe hardware features and functions. While the hardware supports them, the realization of these features and functions depends upon support from the operating system. Linux provides this support. If you are using another operating system, consult the appropriate documentation for that operating system regarding support for those features and functions.

Related information

Many other sources of information on Linux, RAID, and other associated topics are available.

The following publications contain related information:

- System unit documentation for information specific to your hardware configuration
- IPR Linux device driver Web site, available at <http://sourceforge.net/projects/iprdd/>
- *RS/6000® eServer™ pSeries® Adapters, Devices, and Cable Information for Multiple Bus Systems*, order number SA38-0516. Also available at <https://techsupport.services.ibm.com/server/library/>
- *Linux Documentation Project* Web site, available at <http://www.tldp.org/>
- *Linux for IBM eServer pSeries* Web site, available at <http://www-03.ibm.com/systems/power/software/linux/index.html>
- *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*, order number SA38-0509. Also available at <https://techsupport.services.ibm.com/server/library>
- *The RAIDbook: A Handbook of Storage Systems Technology*, Edition 6, Editor: Paul Massiglia
- *Penguinppc* Web site, dedicated to Linux on PowerPC architecture, available at <http://penguinppc.org/>

SAS overview

The term *Serial-attached SCSI (SAS)* refers to a set of serial device interconnect and transport protocols.

This set of protocols defines the rules for information exchange between devices. SAS is an evolution of the parallel SCSI device interface into a serial point-to-point interface.

SAS physical links (*phys*) are a set of four wires used as two differential signal pairs. One differential signal transmits in one direction while the other differential signal transmits in the opposite direction. Data can be transmitted in both directions simultaneously.

Phys are contained in *ports*. A port contains one or more phys. A port is a *wide* port if there are multiple phys in the port. A port is a *narrow* port if there is only one phy in the port. A port is identified by a unique SAS worldwide name (also called a SAS address). A SAS controller contains one or more SAS ports.

A *path* is a logical point-to-point link between a SAS initiator port in the controller and a SAS target port in the I/O device (for example, a disk).

A *connection* is a temporary association between a controller and an I/O device through a path. A connection enables communication to a device. The controller can communicate to the I/O device over this connection using either the SCSI command set or the ATA/ATAPI command set, depending on the device type.

An *expander* facilitates connections between a controller port and multiple I/O device ports. An expander routes connections between the expander ports. There is only a single connection through an expander at any given time. Using expanders creates more nodes in the path from the controller to the I/O device.

If an I/O device supports multiple ports, then it is possible to have more than one path to the device when there are expander devices on the path.

A SAS *fabric* refers to the summation of all paths between all controller ports and all I/O device ports in the SAS subsystem.

SAS architecture summary

Elements that interact to enable the structure of the SAS architecture include controllers, ports, and expanders.

The following points are applicable to this description of general SAS architecture:

- A SAS fabric describes all possible paths between all SAS controllers and I/O devices including cables, enclosures and expanders.
- A SAS controller, expander and I/O device contains one or more SAS ports.
- A SAS port contains one or more physical links (*phys*).
- A SAS path is a logical connection between a SAS controller port and I/O device ports.
- SAS devices use the SCSI command set and SATA devices use the ATA/ATAPI command set.

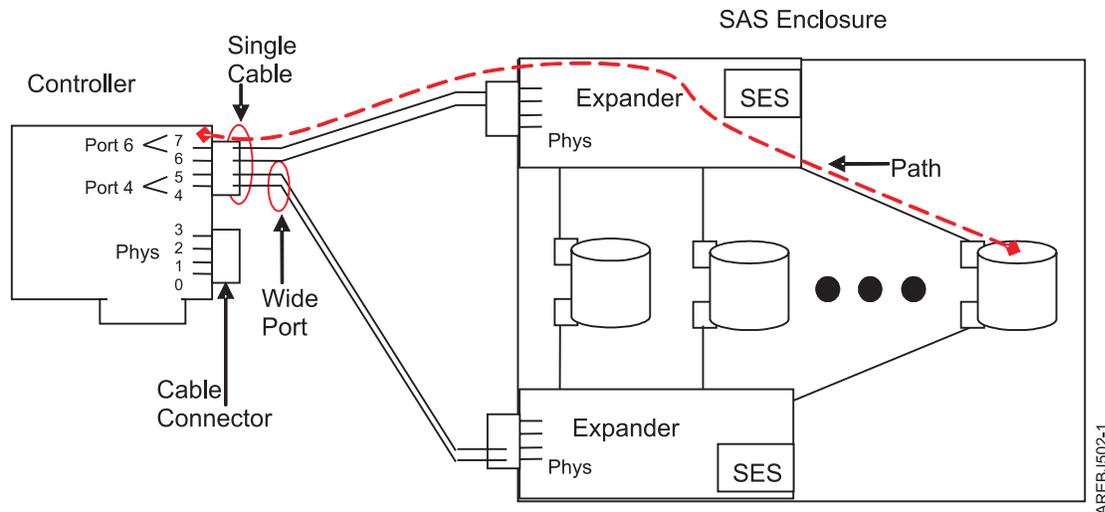


Figure 16. Example of the SAS subsystem

The example of a SAS subsystem in the preceding figure illustrates some general concepts.

This controller has eight SAS phys connections. Four of those phys are connected into two different wide ports. (One connector contains four phys grouped into two ports; the connectors signify a physical wire connection.) The four-phys connector can contain between one and four ports depending on the type of cabling used.

The uppermost port in the figure shows a controller wide port number 6 consisting of phy numbers 6 and 7. Port 6 connects to an expander which attaches to one of the I/O devices dual ports.

The dashed red line indicates a path between the controller and an I/O device. There is another path from the controller's port number 4 to the other port of the I/O device. These two paths provide two different possible connections for increased reliability by using redundant controller ports, expanders and I/O device ports. The SCSI Enclosure Services (SES) is a component of each expander.

Disk arrays

RAID technology is used to store data across a group of disks known as a disk array.

Depending on the RAID level selected, the technique of storing data across a group of disks provides the data redundancy required to keep data secure and the system operational. If a disk failure occurs, the disk can usually be replaced without interrupting normal system operation. Disk arrays also have the potential to provide higher data transfer and input and output (I/O) rates than those provided by single large disks.

Each disk array can be used by Linux in the same way as it would a single SCSI disk. For example, after creating a disk array, you can use Linux commands to make the disk array available to the system by partitioning and creating file systems on it.

The SAS controller and I/O devices are managed by the `iprconfig` utility. The `iprconfig` utility is the interface to the RAID configuration, monitoring, and recovery features of the controller and I/O devices.

If a disk array is to be used as the boot device, it might be necessary to prepare the disks by booting into Rescue mode and creating the disk array before installing Linux. You might want to perform this procedure when the original boot drive is to be used as part of a disk array.

For a RAID 0 array of three disks, data would be written in the following pattern.

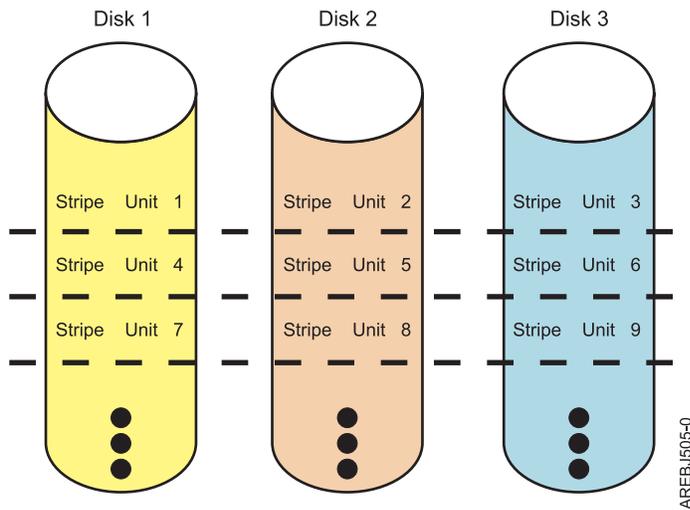


Figure 18. RAID 0

RAID 0 offers a high potential I/O rate, but it is a nonredundant configuration. As a result, there is no data redundancy available for the purpose of reconstructing data in the event of a disk failure. There is no error recovery beyond what is normally provided on a single disk. Unlike other RAID levels, the array controller never marks a RAID 0 array as degraded as the result of a disk failure. If a physical disk fails in a RAID 0 disk array, the disk array is marked as failed. All data in the array must be backed up regularly to protect against data loss.

Related concepts

“Supported RAID levels” on page 16

The level of a disk array refers to the manner in which data is stored on the disk and the level of protection that is provided.

RAID 5

RAID 5 stripes data across all disks in the array.

In addition to data, RAID 5 also writes array parity data. The parity data is spread across all the disks. For a RAID 5 array of three disks, array data and parity information are written in the following pattern:

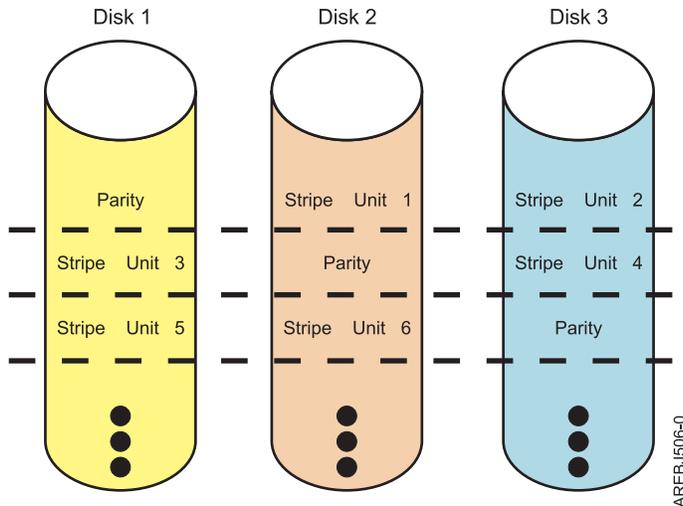


Figure 19. RAID 5

If a disk fails in a RAID 5 array, you can continue to use the array normally. A RAID 5 array operating with a single failed disk is said to be operating in degraded mode. Whenever data is read from a degraded disk array, the array controller recalculates the data on the failed disk by using data and parity blocks on the operational disks. If a second disk fails, the array will be placed in the failed state and will not be accessible.

Related concepts

“Supported RAID levels” on page 16

The level of a disk array refers to the manner in which data is stored on the disk and the level of protection that is provided.

RAID 6

RAID 6 stripes data across all disks in the array.

In addition to data, RAID 6 also writes array P and Q parity data. The P and Q parity data, which is based on Reed Solomon algorithms, is spread across all the disks. For a RAID 6 array of four disks, array data and parity information are written in the following pattern:

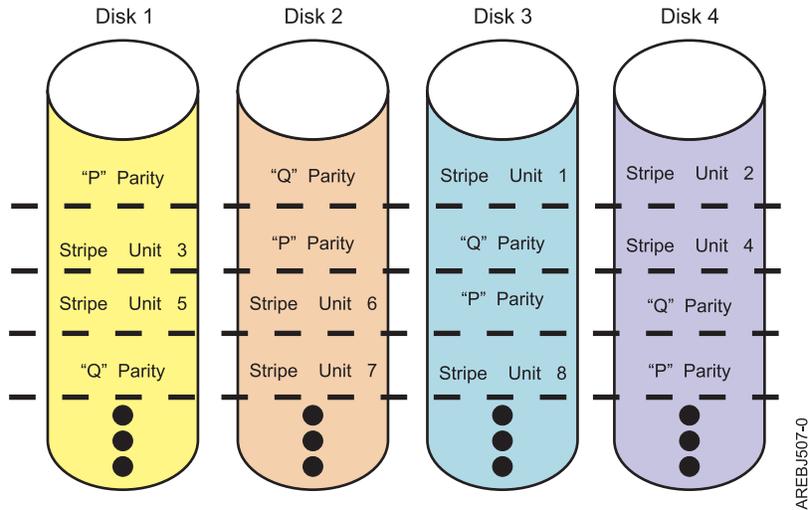


Figure 20. RAID 6

If one or two disks fail in a RAID 6 array, you can continue to use the array normally. A RAID 6 array operating with one or two failed disks is said to be operating in degraded mode. Whenever data is read from a degraded disk array, the array controller recalculates the data on the failed disks by using data and parity blocks on the operational disks. A RAID 6 array with a single failed disk has similar protection to that of a RAID 5 array with no disk failures. If a third disk fails, the array will be placed in the failed state and will not be accessible.

Related concepts

“Supported RAID levels” on page 16

The level of a disk array refers to the manner in which data is stored on the disk and the level of protection that is provided.

RAID 10

RAID 10 uses mirrored pairs to redundantly store data.

The array must contain an even number of disks. Two is the minimum number of disks that you need to create a RAID 10 array. A two-disk RAID 10 array is equal to a RAID 1 array. The data is striped across the mirrored pairs. For example, a RAID 10 array of four disks would have data written to it in the following pattern:

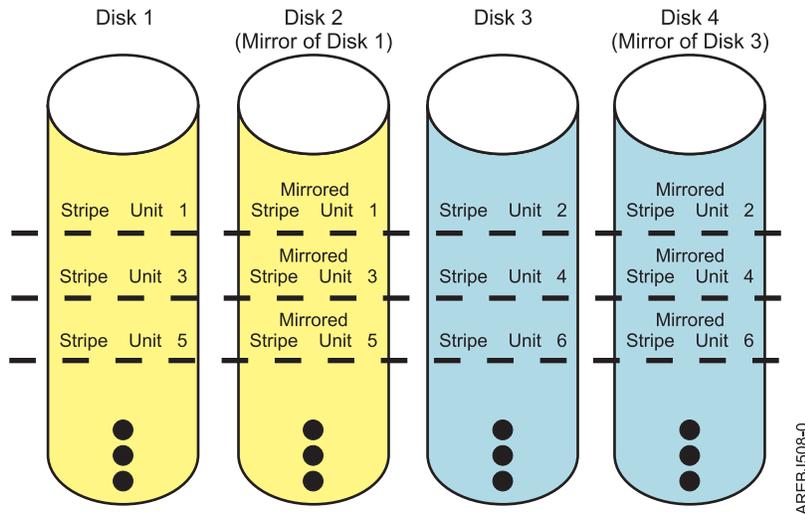


Figure 21. RAID 10

RAID 10 can tolerate multiple disk failures. If one disk in each mirrored pair fails, the array is still functional, operating in degraded mode. You can continue to use the array normally because for each failed disk, the data is stored redundantly on its mirrored pair. However, if both members of a mirrored pair fail, the array will be placed in the failed state and will not be accessible.

Related concepts

“Supported RAID levels” on page 16

The level of a disk array refers to the manner in which data is stored on the disk and the level of protection that is provided.

Estimating disk array capacities

The capacity of a disk array depends on the capacity of the advanced function disks used and the RAID level of the array.

In order to estimate the capacity of a disk array, you must know the capacity of the advanced function disks and the RAID level of the array.

1. For RAID 0, multiply the number of disks by the disk capacity.
2. For RAID 5, multiply one fewer than the number of disks by the disk capacity.
3. For RAID 6, multiply two fewer than the number of disks by the disk capacity.
4. For RAID 10, multiply the number of disks by the disk capacity and divide by 2.

Note:

- If disks of different capacities are used in the same array, all disks are treated as if they have the capacity of the smallest disk.
- SAS RAID controllers support up to 18 member disks in each RAID array.

Related concepts

“RAID level summary”

Compare RAID levels according to their capabilities.

RAID level summary

Compare RAID levels according to their capabilities.

The following information provides data redundancy, usable disk capacity, read performance, and write performance for each RAID level.

Table 4. RAID level summary

RAID level	Data redundancy	Usable disk capacity	Read performance	Write performance	Min/Max devices per array
RAID 0	None	100%	Very good	Excellent	1/18
RAID 5	Very good	67% to 94%	Very good	Good	3/18
RAID 6	Excellent	50% to 89%	Very good	Fair to good	4/18
RAID 10	Excellent	50%	Excellent	Very good	2/18 (even numbers only)

RAID 0

Does not support data redundancy, but provides a potentially higher I/O rate.

RAID 5

Creates array parity information so that the data can be reconstructed if a disk in the array fails. Provides better capacity than RAID level 10 but possibly lower performance.

RAID 6

Creates array "P" and "Q" parity information so that the data can be reconstructed if one or two disks in the array fail. Provides better data redundancy than RAID 5 but with slightly lower capacity and possibly lower performance. Provides better capacity than RAID level 10 but possibly lower performance.

RAID 10

Stores data redundantly on mirrored pairs to provide maximum protection against disk failures. Provides generally better performance than RAID 5 or 6, but has lower capacity.

Note: A two-drive RAID level 10 array is equivalent to RAID level 1.

Related tasks

"Estimating disk array capacities" on page 20

The capacity of a disk array depends on the capacity of the advanced function disks used and the RAID level of the array.

Stripe-unit size

With RAID technology, data is striped across an array of physical disks.

Striping data across an array of physical disks complements the way the operating system requests data. The granularity at which data is stored on one disk of the array before subsequent data is stored on the next disk of the array is called the stripe-unit size. The collection of stripe units, from the first disk of the array to the last disk of the array, is called a stripe.

You can set the stripe-unit size of a disk array to 16 KB, 64 KB, 256 KB or 512 KB. You might be able to maximize the performance of your disk array by setting the stripe-unit size to a value that is slightly larger than the size of the average system I/O request. For large system I/O requests, use a stripe-unit size of 256 KB or 512 KB. The recommended stripe size for most applications is 256 KB.

Disk array overview

Disk arrays are groups of disks that work together with a specialized array controller to potentially achieve higher data transfer and input and output (I/O) rates than those provided by single large disks.

Disk arrays are groups of disks that work together with a specialized array controller to potentially achieve higher data transfer and input and output (I/O) rates than those provided by single large disks. The array controller keeps track of how the data is distributed across the disks. RAID 5, 6, and 10 disk arrays also provide data redundancy so that no data is lost if a single disk in the array fails.

Note: This topic and the `iprconfig` utility use common terminology for disk formats:

- **JBOD**

A *JBOD* disk is formatted to 512 bytes/sector. JBOD stands for “Just a Bunch Of Disks”. A JBOD disk is assigned a `/dev/sdX` name and can be used by the Linux operating system.

- **Advanced function**

An *advanced function* disk is formatted to 528 bytes/sector. This format allows disks to be used in disk arrays. An advanced function disk cannot be used by the Linux operating system directly. The Linux operating system can use an advanced function disk only if it is configured into a disk array.

Disk arrays are accessed in Linux as standard SCSI disk devices. These devices are automatically created when a disk array is created, and deleted whenever a disk array is deleted. The individual physical disks that comprise disk arrays (or are candidates to be used in disk arrays), which are formatted for advanced function, are hidden from Linux and are accessible only through the `iprconfig` utility. Linux sees all JBOD disks. These disks must be formatted for advanced function before they can be used in disk arrays. For information on formatting JBOD disks to make them available for use in disk arrays, see “Formatting to JBOD” on page 36.

An advanced function disk can be used in one of three roles:

- **Array Member:** An advanced function disk which is configured as a member of a disk array.
- **Hot Spare:** An advanced function disk assigned to a controller waiting to be used by the controller to automatically replace a Failed disk in a Degraded disk array.
- **Array Candidate:** An advanced function disk which has not been configured as a member of a disk array or a hot spare.

The **Display Hardware Status** option in the `iprconfig` utility can be used to display these disks and their associated resource names. For details regarding how to view the disk information, see “Viewing device status” on page 31. The following sample output is displayed when the **Display Hardware Status** option is invoked:

Table 5. Disk array states (continued)

State	Description
Missing	The disk array was not detected by the host operating system.
Offline	The disk array has been placed offline due to unrecoverable errors.
Failed	The disk array is no longer accessible because of disk failures or configuration problems.

Related concepts

“I/O adapter states”

There are three possible states for the I/O adapter.

“Physical disk states”

There are six possible states for the physical disks.

Physical disk states

There are six possible states for the physical disks.

The six possible states for physical disks are described in the following table:

Table 6. Physical disk states

State	Description
Active	The disk is functioning properly.
Failed	The IOA cannot communicate with the disk or the disk is the cause of the disk array being in the degraded state.
Offline	The disk array has been placed offline due to unrecoverable errors.
Missing	The disk was not detected by the host operating system.
R/W Protected	The device cannot process a read nor write operation. A disk may be in this state because of a cache, device configuration, or any other problem that could cause a data integrity exposure.
Format Required	The disk unit must be formatted to become usable on this IOA.

Related concepts

“I/O adapter states”

There are three possible states for the I/O adapter.

“Disk array states” on page 23

A disk array can be in one of seven states.

I/O adapter states

There are three possible states for the I/O adapter.

The three possible states for I/O adapters are described in the following table:

Table 7. I/O adapter states

State	Description
Operational	The IOA is functional.
Not Operational	The device driver cannot successfully communicate with this IOA.
Not Ready	The IOA requires a microcode download.

Related concepts

“Disk array states” on page 23

A disk array can be in one of seven states.

“Physical disk states” on page 24

There are six possible states for the physical disks.

Auxiliary write cache adapter

The Auxiliary Write Cache (AWC) adapter provides a duplicate, nonvolatile copy of write cache data of the RAID controller to which it is connected.

Protection of data is enhanced by having two battery-backed (nonvolatile) copies of write cache, each stored on separate adapters. If a failure occurs to the write cache portion of the RAID controller, or if the RAID controller itself fails in such a way that the write cache data is not recoverable, the AWC adapter provides a backup copy of the write cache data to prevent data loss during the recovery of the failed RAID controller. The cache data is recovered to the new replacement RAID controller and then written out to disk before resuming normal operations.

The AWC adapter is not a failover device that can keep the system operational by continuing disk operations when the attached RAID controller fails. The system cannot use the auxiliary copy of the cache for runtime operations even if only the cache on the RAID controller fails. The AWC adapter does not support any other device attachment and performs no other tasks than communicating with the attached RAID controller to receive backup write cache data. The purpose of the AWC adapter is to minimize the length of an unplanned outage, due to a failure of a RAID controller, by preventing loss of critical data that might have otherwise required a system reload.

It is important to understand the difference between multi-initiator connections and AWC connections. Connecting controllers in a multi-initiator environment refers to multiple RAID controllers connected to a common set of disk enclosures and disks. The AWC controller is not connected to the disks, and it does not perform device media accesses.

Should a failure of either the RAID controller or the Auxiliary Cache occur, it is extremely important that the Maintenance Analysis Procedures (MAPs) for the errors in the Linux error log be followed precisely. Required service information can be found at “Problem determination and recovery” on page 79.

The RAID controller and the AWC adapter each require a PCI-X or PCIe bus connection. They are required to be in the same partition. The two adapters are connected by an internal SAS connection. For the Planar RAID Enablement and Planar Auxiliary Cache features, the dedicated SAS connection is integrated into the system planar.

If the AWC adapter itself fails or the SAS link between the two adapters fails, the RAID controller will stop caching operations, destage existing write cache data to disk, and run in a performance-degraded mode. After the AWC adapter is replaced or the link is reestablished, the RAID controller automatically recognizes the AWC, synchronizes the cache area, resumes normal caching function, and resumes writing the duplicate cache data to the AWC.

The AWC adapter is typically used in conjunction with RAID protection. RAID functions are not affected by the attachment of an AWC. Because the AWC does not control other devices over the bus and communicates directly with its attached RAID controller over a dedicated SAS bus, it has little, if any, performance impact on the system.

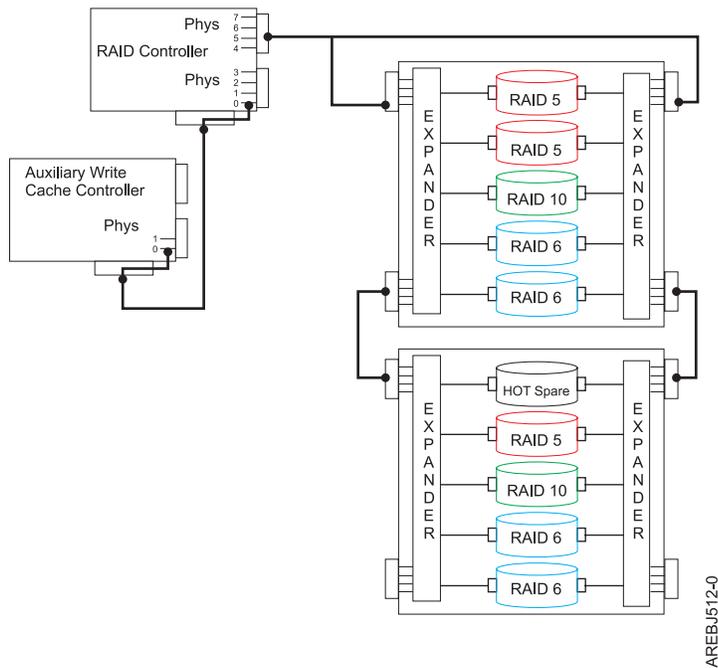


Figure 22. Example RAID and AWC controller configuration

Related information

Many other sources of information on Linux, RAID, and other associated topics are available.

The following publications contain related information:

- System unit documentation for information specific to your hardware configuration
- IPR Linux device driver Web site, available at <http://sourceforge.net/projects/iprdd/>
- *RS/6000 eServer pSeries Adapters, Devices, and Cable Information for Multiple Bus Systems*, order number SA38-0516. Also available at <https://techsupport.services.ibm.com/server/library/>
- *Linux Documentation Project* Web site, available at <http://www.tldp.org/>
- *Linux for IBM eServer pSeries* Web site, available at <http://www-03.ibm.com/systems/power/software/linux/index.html>
- *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*, order number SA38-0509. Also available at <https://techsupport.services.ibm.com/server/library>
- *The RAIDbook: A Handbook of Storage Systems Technology*, Edition 6, Editor: Paul Massiglia
- *Penguinppc* Web site, dedicated to Linux on PowerPC architecture, available at <http://penguinppc.org/>

RAID controller software

A device driver and a set of utilities must be installed so that the controller can be identified and configured by Linux.

Note: References to the Linux operating system in this topic collection include the versions Linux 2.6, SuSE Linux Enterprise Server 10 and SuSE Linux Enterprise Server 11, RedHat Enterprise Linux 4, RedHat Enterprise Linux 5, and RedHat Enterprise Linux 6. Make sure you are consulting the appropriate section of this set of information for the operating system you are using.

This document might describe hardware features and functions. While the hardware supports them, the realization of these features and functions depends upon support from the operating system. Linux provides this support. If you are using another operating system, consult the appropriate documentation for that operating system regarding support for those features and functions.

For the controller to be identified and configured by Linux, the requisite device support software must be installed. Software for the controller consists of a device driver and a set of utilities.

The device driver is usually compiled as a kernel module named `ipr.ko`. The user utilities are usually packaged in a RedHat Package Manager (RPM) called `iprutils`. The requisite software for the controller is often preinstalled as part of the normal Linux installation.

If the software package is not installed, software verification will fail. The missing packages can be installed from your Linux operating system CD-ROM. If you are missing components or need newer versions, obtain them from your Linux distributor or online at SourceForge.net.

The controller executes onboard microcode. The `iprconfig` utility in `iprutils` RPM can be used to update the microcode being used by the controller. For more information regarding the `iprconfig` utility, see "Updating the controller microcode" on page 63.

Verifying installation of the controller software

Verify that the `ipr` device driver for the controller is installed.

Consult the following table for the minimum `ipr` device driver version required for each supported adapter:

Table 8. Minimum ipr device driver support

Custom card identification number (CCIN)	Minimum supported main line Linux version		Minimum supported Red Hat Enterprise Linux version		Minimum supported SUSE Enterprise Linux version	
	Device driver ipr version	Kernel version	Device driver ipr version	RHEL Version	Device driver ipr version	SLES Version
572A No dual adapter	2.1.2	2.6.16	2.0.11.5 / 2.2.0.1	RHEL4 U6 / RHEL5 U1	2.2.0.1	SLES10 SP 1
572A With dual adapter	2.4.1	2.6.22	2.0.11.6 /2.2.0.2	RHEL4 U7 /RHEL5 U2	2.2.0.2	SLES10 SP2
572B	2.4.1	2.6.22	2.0.11.6 /2.2.0.2	RHEL4 U7 /RHEL5 U2	2.2.0.2	SLES10 SP2
572C	2.1.2	2.6.16	2.0.11.4 /2.2.0	RHEL4 U5 /RHEL5	2.2.0.1	SLES10 SP1

Table 8. Minimum ipr device driver support (continued)

Custom card identification number (CCIN)	Minimum supported main line Linux version		Minimum supported Red Hat Enterprise Linux version		Minimum supported SUSE Enterprise Linux version	
	Device driver ipr version	Kernel version	Device driver ipr version	RHEL Version	Device driver ipr version	SLES Version
572F/575C	2.4.1	2.6.22	2.0.11.6 / 2.2.0.2	RHEL4 U7 / RHEL5	2.2.0.2	SLES10 SP2
574E	2.4.1	2.6.22	2.0.11.6 /2.2.0.2	RHEL4 U7 /RHEL5 U2	2.2.0.2	SLES10 SP2
57B3	2.4.1	2.6.22	2.0.11.6 /2.2.0.2	RHEL4 U7 /RHEL5 U2	2.2.0.2	SLES10 SP2
57B7	2.3.0	2.6.20	2.0.11.5 /2.2.0.1	RHEL4 U6 /RHEL5 U1	2.2.0.1	SLES10 SP 1
57B8	2.3.0	2.6.20	2.0.11.5 /2.2.0.1	RHEL4 U6 /RHEL5 U1	2.2.0.1	SLES10 SP 1
57B9	2.1.2	2.6.16	2.0.11.5 /2.2.0.1	RHEL4 U6 /RHEL5 U1	2.2.0.1	SLES10 SP 1
57BA	2.1.2	2.6.16	2.0.11.5 /2.2.0.1	RHEL4 U6 /RHEL5 U1	2.2.0.1	SLES10 SP 1
57C7	2.4.1	2.6.22	2.0.11.6 /2.2.0.2	RHEL4 U7 /RHEL5 U2	2.2.0.2	SLES10 SP 2
57CF	2.4.1	2.6.22	2.0.11.6 /2.2.0.2	RHEL4 U7 /RHEL5 U2	2.2.0.2	SLES10 SP 2
57CD	2.4.1	2.6.22	2.2.0.5 /2.4.3	RHEL5 U6 /RHEL6	2.2.0.3 /2.4.3	SLES10 SP 3/SLES11 SP 1

Consult the following table for the minimum version of the iprconfig utility required for each supported function:

Table 9. Versions of iprconfig for SAS functions

SAS function	Version of iprconfig
Base SAS support	2.0.15.8
Displaying SAS drive path status	2.2.5
High availability dual adapter support	2.2.8
Migrating RAID array between RAID level 0 and 10	2.2.10
Active-Active (tentative name)	2.2.11

You can verify several parameters of the device driver software.

- To verify that the ipr device driver for the controller is installed, type `modinfo ipr | grep version`. Data is displayed in this form:

```
version:      2.4.1
srcversion:   D720FC6935DA2B2638B177A
```
- To verify the ipr device driver version, type `# modinfo -F version ipr`. You will see data in this form:

```
2.4.1
```

3. To verify that the iprconfig utility is installed, type `iprconfig -version`. Output from this command will indicate if the utility is installed, and if so, version information will be displayed. You will see data in this form:

```
iprconfig: 2.2.13 (November 17, 2008)
```

Linux ipr device driver updates

The controller device driver for Linux, ipr, is part of the Linux kernel. Linux distributors periodically release updated device drivers for their Linux kernels.

Updated device drivers are usually packaged in an RPM. The updated device drivers provided by the Linux distributors are well-tested with other parts of the Linux distribution. Contact your Linux distributor for detailed information about the procedure for obtaining and updating the ipr device driver.

The Linux device drivers provided by your Linux distribution might be a few versions behind due to their test and release cycles. The most up-to-date ipr device driver can be found in the main line Linux kernel tree. Main line Linux kernel trees can be downloaded from Linux Kernel Archives.

The ipr device driver is located in the `drivers/scsi/` directory. Users can obtain the ipr device driver from a Linux kernel source tree and follow the instructions in `Documentation/kbuild/modules.txt` in the Linux kernel source tree to recompile the ipr device driver.

Attention:

- The ipr device driver depends on other parts of Linux kernel to work properly. Updating ipr device driver without updating other parts of Linux kernel might cause compiling errors or runtime errors and result in system crash and data loss. Only experienced users should update the ipr device driver from source code.
- A device driver is considered part of the Linux kernel. Updating a device driver may void a Linux service contract provided by your Linux distributor or service provider. Check your Linux distributor or service provider for their device driver update policy.

Updating the iprutils package

You can update the iprutils package.

You can download the user space package iprutils source code from SourceForge.net by clicking on the **Download IPR Linux Device Driver** button. The iprutils source code is packaged in RPM and tar.gz formats. Users can download either package.

You must install certain development headers and libraries before compiling iprutils:

Table 10. Required development headers and libraries

RHEL4 and RHEL5	SLES10
<ul style="list-style-type: none">• glibc-headers• kernel-headers• ncurses-devel• pciutils-devel• libsysfs-devel	<ul style="list-style-type: none">• glibc-devel• ncurses-devel• pci-utils-devel• sysfsutils

When building iprutils in an RPM package, use the correct `<RPM_DIR>` for your system. If your system is running RHEL4 or RHEL5, use `/usr/src/redhat`. If your system is running SLES10, use `/usr/src/packages`.

1. Login to the Linux system as root.

2. Enter `rpmbuild -rebuild -target=ppc iprutils-x.y.z-1.src.rpm` where *x.y.z* is the iprutils version to build the iprutils package.
3. Enter `cd <RPM_DIR>/RPMS/ppc`.
4. Enter `rpm -U iprutils-x.y.z-1.ppc.rpm` to update iprutils.

- This task requires rewriting data to the whole disk. The required time varies depending on the disk capacities. For large disks, it might take hours to complete.

Formatting to JBOD

Before a disk is recognized as a standalone disk, it must be formatted to JBOD.

The steps in this procedure result in a disk formatted for JBOD.

1. Run the `iprconfig` utility by typing `iprconfig`.
2. Select **Work with disk arrays**.
3. Select **Format device for JBOD function (512)**.
4. From the list of eligible disk units, choose the disks you want to format to JBOD and press Enter.
Attention: Continuing with this option will format the disks. All data on the disks will be lost.
5. To proceed with the format, type `c` to confirm. To return to the previous menu without formatting the disks, type `q`.

Notes:

- After the formatting is complete, the disks are ready for use in disk arrays. Resource names (such as `/dev/sdb`) might change when the system is rebooted. This could affect kernel command line entries and `fstab` entries. You might want to reboot the server now.
- This task requires rewriting data to the whole disk. The required time varies depending on the disk capacities. For large disks, it might take hours to complete.

Note: Resource names (such as `/dev/sdb`) might change as a result of reformatting. This could affect kernel command line entries and `fstab` entries. You might want to reboot the server now.

Creating and deleting disk arrays

Create and delete IBM SAS RAID disk arrays by following the procedures in this information set.

Follow the instructions provided in this section to create and delete IBM SAS RAID disk arrays.

Creating an IBM SAS RAID disk array

A disk array is created using a set of disks that are formatted for advanced function.

Disk arrays with data redundancy (RAID 5, RAID 6 and RAID 10) are unprotected until parity is finished being built. For disk arrays with no data redundancy (RAID 0), no parity is calculated and the array is built immediately.

Complete the following steps to create an IBM SAS RAID disk array:

1. Run the `iprconfig` utility by typing `iprconfig`.
2. Select **Work with disk arrays**.
3. Select **Create a disk array**.
4. Select the controller under which you would like to create a disk array, and then press Enter.
5. Select the disk units to be included in the disk array, and then press Enter. The information you see will be similar to the following:

```

+-----+
|                                     |
|                               Select Protection Level and Stripe Size         |
|                                     |
| Default array configurations are shown. To change                          |
| setting hit "c" for options menu. Highlight desired                       |
| option then hit Enter                                                     |
|                                     |
| c=Change Setting                                                           |
|                                     |
| Protection Level . . . . . : RAID 5                                       |
| Stripe Size . . . . . : 256 k                                           |
| Queue Depth (default = 12). . . . . : 12                                |
|                                     |
|                                     |
| Press Enter to Continue                                                  |
|                                     |
| e=Exit  q=Cancel                                                         |
|                                     |
+-----+

```

6. Select the protection level that you want for the array. For more information about selecting an appropriate RAID level, see “Supported RAID levels” on page 16.
7. Select the stripe size that you want, in kilobytes, for the array. For more information about selecting an appropriate stripe size, see “Stripe-unit size” on page 21.
8. Press Enter to continue.

Attention: All data on the selected drives will be lost when the disk array is created. If you are sure you want to create the disk array, press Enter. If you do not want to create the disk array, type q to cancel.

If you choose to create the disk array, a status screen is displayed until the operation is completed. If you want to create additional disk arrays, you can exit the status screen by typing e.

When a disk array has been built, it is available to Linux like any SCSI disk. To find the resource name that has been assigned to the disk array, see the **Display Disk Array Status** screen.

Note: Resource names (such as /dev/sdb) might change as a result of creating a new disk array. This could affect kernel command line entries and fstab entries. You might want to reboot the server now.

Deleting an IBM SAS RAID disk array

Delete a disk array carefully to avoid losing data.

To preserve the data on the disk array, you must first back up all data that you want to save.

Attention: After a disk array is deleted, it cannot be accessed. All data will be lost and cannot be recovered.

A disk array cannot be deleted if it is currently being rebuilt or synched. However, it is possible to delete a disk array that is currently in use. Make sure the disk array is not being used before deleting it.

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk arrays**.
3. Select **Delete a disk array**.
4. From the listed disk arrays, select the one you want to delete, and then press Enter.

Attention: All data on the selected drives will be lost when the disk array is deleted. If you are sure you want to delete the disk array, press Enter. If you do not want to delete the disk array, type q to cancel.

Remember: Resource names (such as /dev/sdb) might change when the system is rebooted. This could affect kernel command line entries and fstab entries.

Adding disks to an existing array

The IBM SAS RAID controller supports adding disks to existing RAID level 5 or 6 disk arrays. Adding disks to an existing RAID 5 or 6 array can dynamically increase the capacity of a disk array while preserving existing data in the disk array. Extensive use of this feature, however, will result in a performance penalty because the data will not be restriped.

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk arrays**.
3. Select **Add a device to a disk array**.
4. Select the disk array to which you want to add the disks.
5. Select the eligible disks to be included into the disk array, and then press Enter.

Attention: All data on the selected drives will be lost when the disks are added to the existing array. If a particular disk is not included in the list, it may not be a candidate that can be added to the array because of one or more of the following reasons:

- The disk's capacity is less than that of the smallest disk already in the array.
- The disk has not been formatted for Advanced Function.
- The disk belongs to another disk array or is configured as a Hot Spare.

6. To include the disks, press Enter. To exclude the disks from the disk array, type q to cancel.

Migrating an existing disk array to a new RAID level

The SAS RAID controller supports migrating an existing RAID 0 or 10 disk array to RAID 10 or 0, respectively. This allows you to dynamically change the level of protection of a disk array while preserving its existing data.

When migrating RAID 0 to RAID 10, additional disks must be included into the RAID 10 disk array in order to provide the additional level of protection. The number of additional disks will be equal to the number of disks in the original RAID 0 disk array. The capacity of the disk array will remain unchanged and the disk array remains accessible during the migration. The disk array is not protected by RAID 10 until the migration completes.

When migrating RAID 10 to RAID 0, no additional disks are included into the RAID 0 disk array. The number of disks in the resulting RAID 0 disk array will be reduced to half the number of disks in the original RAID 10 disk array. The capacity of the disk array will remain unchanged and the disk array remains accessible during the migration.

To migrate an existing array to a new level, do the following:

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk arrays**.
3. Select **Migrate disk array protection**.
4. Select the disk array to which you want to migrate the disk array protection level.
5. Press c and use up arrow and down arrow keys to change the protection level, and press Enter. A screen will display similar to the following:

7. Optional: If fewer member disks are needed, for example when migrating from RAID 10 to RAID 0, the adapter will select the extra disks and take them out of the disk array automatically. No user intervention is needed.
8. Press **Enter** to confirm the RAID migration.
The migration between different RAID levels takes anywhere from several seconds to several hours. A progress screen displays the migration progress. Migration progress can also be found in the **Display Disk Array Status** screen and **Display Hardware Status** screen.

Hot spare disks

Use hot spare disks to automatically replace failed disks in a RAID environment. .

Upon a disk failure in a RAID array is detected, the RAID adapter will look into all disks it controls for a hot spare disk. If one is available, the RAID adapter will replace the role of the failed disk in the RAID array with the hot spare disk. The data and parity are rebuilt from the survival member disks automatically. The RAID array is changed back to optimized state when the rebuild is complete.

Hot spare disks are useful only if their capacity is greater than or equal to that of the smallest-capacity disk in an array that becomes Degraded. In order to assign a disk for use as a hot spare, it must be formatted for advanced function.

Creating hot spare disks

Create a hot spare disk to automatically replace a failed disk in a RAID environment.

Format a disk for advanced function if you want to use it as a hot spare disk.

Complete the following steps to create hot spare disks:

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk arrays**.
3. Select **Create a hot spare**.
4. Select the adapter for which you want to create hot spares and press Enter.
5. Select the disks that you want to designate as hot spares and press Enter.
Attention: All data on the selected drives will be lost when the disks are configured as hot spares.
6. To configure the disks as hot spares, press Enter. If you do not want to configure the disks as hot spares, type q to cancel.

Related tasks

“Formatting to advanced function” on page 35

Before a disk can be used in a disk array or as a hot spare disk, it must be formatted for advanced function.

Deleting hot spare disks

Delete a hot spare disk to free it for use in other disk arrays.

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk arrays**.
3. Select **Delete a hot spare device**.
4. Select the hot spare disks to be deleted and press Enter.
5. To reconfigure the disks, press Enter. If you do not want to reconfigure the disks as hot spares, type q to cancel.

Considerations for solid-state drives (SSD)

The purpose of this section is to cover important controller functionality when using solid-state drives (SSD).

Hard disk drives (HDD) use a spinning magnetic platter to store data in magnetic fields in nonvolatile storage. SSDs are storage devices that use nonvolatile solid-state memory (usually flash memory) to emulate HDDs. HDDs have an inherent latency and access time caused by mechanical delays in the spinning of the platter and movement of the head. SSDs greatly reduce the latency and time to access the stored data. The nature of solid-state memory is such that read operations can be performed faster than write operations and that write cycles are limited. Using techniques such as wear leveling and overprovisioning, enterprise class SSDs are designed to withstand many years of continuous use.

SSD usage specifications

While using SSDs consider the following specifications:

- Intermixing of SSDs and HDDs within the same disk array is not allowed. A disk array must contain all SSDs or all HDDs.
- It is important to properly plan for hot-spare devices when using arrays of SSDs. An SSD hot-spare device is used to replace a failed device in an SSD disk array and an HDD hot spare is used for an HDD disk array
- Although SSDs can be used in a RAID 0 disk array, it is preferred that SSDs to be protected by RAID levels 5, 6, or 10.
- See Installing and configuring solid-state drives to identify specific configuration and placement requirements related to the SSD devices.

RAID 0 automatically created disk array

SSD devices are included in 528 bytes per sector array-candidate format. During the controller boot process, any SSD array candidate that is not already part of a disk array is automatically created as a single-drive RAID 0 disk array. Following are the two options for changing the RAID 0 disk array to a protected RAID level (5, 6, or 10):

- The RAID 0 disk array can be migrated to a RAID 10 disk array by using the technique described in Migrating an existing disk array to a new RAID level.
- The automatically created RAID 0 disk array can be deleted (see Deleting a disk array) and a new SSD disk array can be created with a different level of RAID protection (see Creating a disk array).

Adapter cache control

Adapter caching improves overall performance with disk drives. In some configurations, adapter caching might not improve performance when you use SSD disk arrays. In these situations, adapter caching can be disabled by using the Change Configuration of Adapter window.

To disable adapter caching, complete the following steps:

1. Run the `iprconfig` utility by typing `iprconfig`.
2. Select **Work with adapter configuration**.
3. Select the desired adapter by typing 1, and press Enter.
4. Select **IOA Caching Mode**.
5. Type `c`.
6. Select **Disabled** and press Enter. A display similar to the following example appears in the window.

```
+-----+
|                                     |
|               Change configuration of Adapter               |
|                                     |
|Current Adapter configurations are shown. To change setting hit 'c' for options
|menu. Highlight desired option then hit Enter.              |
| c=Change Setting                                           |
|                                     |
| Adapter:  0001:00:01.0/33  IBM      574E001SISIOA          |
|                                     |
| Preferred Dual Adapter State . . . . . : Primary          |
| Active/Active Mode . . . . . : Enabled                    |
| IOA Caching Mode . . . . . : Default                      |
|                                     |
|                                     |
|                                     |
|Or leave blank and press Enter to cancel                    |
|e=exit  q=cancel                                           |
|                                     |
+-----+
```

Multi-initiator and high availability

You can increase availability using multi-initiator and high availability to connect multiple controllers to a common set of disk expansion drawers.

The terms *multi-initiator* and *high availability* (HA) in the SAS arena refer to connecting multiple controllers (typically two controllers) to a common set of disk expansion drawers for the purpose of increasing availability. IBM SAS RAID controllers support high availability with up to two controllers in one or two systems.

IBM SAS RAID controllers support the following configurations:

HA two-system configuration

An HA two-system configuration provides a high-availability environment for system storage by enabling two systems or partitions to have access to the same set of disks and disk arrays. This feature is typically used with high availability cluster applications such as the IBM High-Availability Cluster Multiprocessing application (HACMP™). The IBM HACMP software provides a commercial computing environment that ensures that mission-critical applications can recover quickly from hardware and software failures.

The HA two-system configuration is intended for using disk arrays. The disks must be formatted to 528 bytes per sector. Any RAID level, or combination of RAID levels, can be used.

Use of disks without RAID (referred to as JBOD) is also possible. The disks must be formatted to 512 bytes per sector. This JBOD alternative is supported only on particular controllers and requires unique setup and cabling.

HA single-system configuration

An HA single-system configuration provides for redundant controllers from a single system to the same set of disks and disk arrays. This feature is typically used with Multi-Path I/O (MPIO). MPIO support in Linux is provided by Device-Mapper Multipath and can be used to provide a redundant IBM SAS RAID controller configuration with RAID-protected disks.

When using an HA single-system configuration, the disks must be formatted to 528 bytes per sector and used in one or more disk array. Any RAID level, or combination of RAID levels, may be used. Disks formatted to 512 bytes per sector are not supported in an HA single-system configuration.

Disks formatted to 512 bytes per sector are allowed in an HA single-system configuration. These disks are detected through both SAS controllers with two different paths and assigned with two different `/dev/sdX` names, for example, `/dev/sde` and `/dev/sdf`. The disks can be accessed using either path. However, there is no additional coordination between the two SAS controllers to automatically switch between the two paths. Multipath IO software such as Device-Mapper Multipath needs to be configured to monitor and switch paths. Using Device-Mapper Multipath to provide HA function is a software-only solution that is independent of IBM SAS RAID controllers. Its use is beyond the scope of this document.

Not all controllers support all configurations. Refer to the comparison table for high availability features and look for controllers that have the entry **Yes** for **HA two-system RAID**, **HA two-system JBOD**, or **HA single-system RAID**.

Possible HA configurations

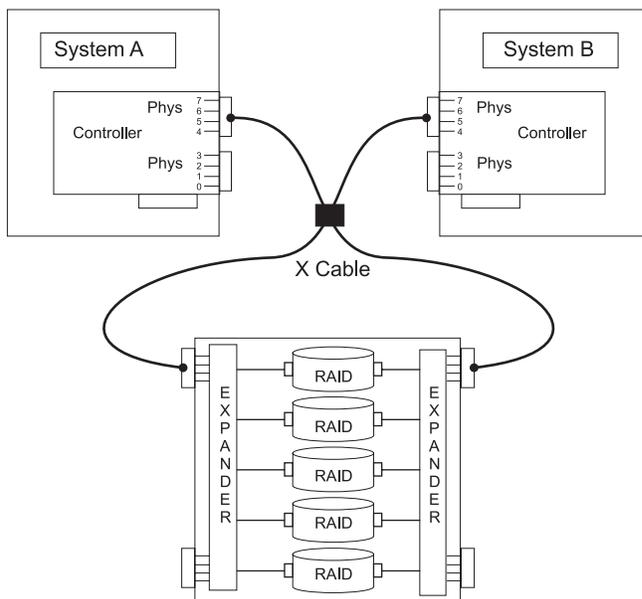
Compare RAID and JBOD features as used with single- and two-system HA configurations.

Table 11. SAS RAID and JBOD HA configurations

Multi-initiator configuration	HA two system (for example PowerHA™ for AIX®)	HA single system (for example MPIO)
RAID (disks formatted 528 bytes per sector)	<ul style="list-style-type: none"> • Maximum of two controllers • Both controllers must have same write cache capability and write cache sizes • Both controllers must support “HA two-system RAID” • Controllers are in different systems or partitions 	<ul style="list-style-type: none"> • Maximum of two controllers • Both controllers must have same write cache capability and write cache sizes • Both controllers must support “HA single-system RAID” • Controllers are in the same system or partition
JBOD (disks formatted 512 bytes per sector)	<ul style="list-style-type: none"> • Maximum of two controllers • Both controllers must support HA two-system JBOD • Controllers are in different systems or partitions • Requires unique setup and cabling 	<ul style="list-style-type: none"> • Not supported

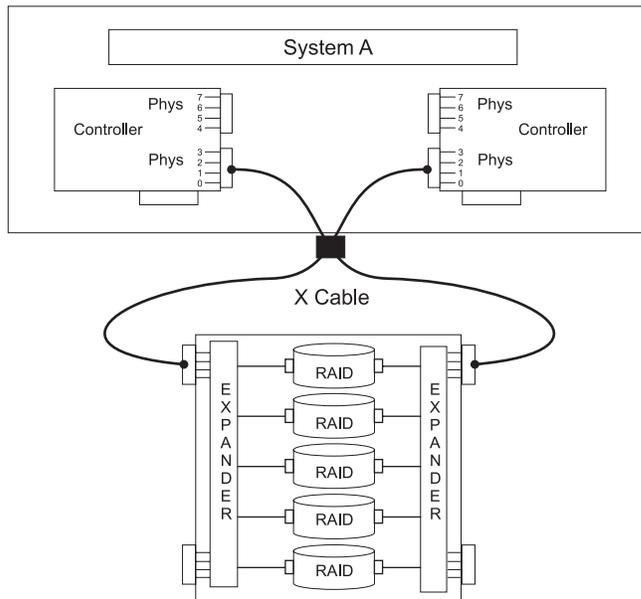
The following figures illustrate an example of each configuration.

HA Two System RAID



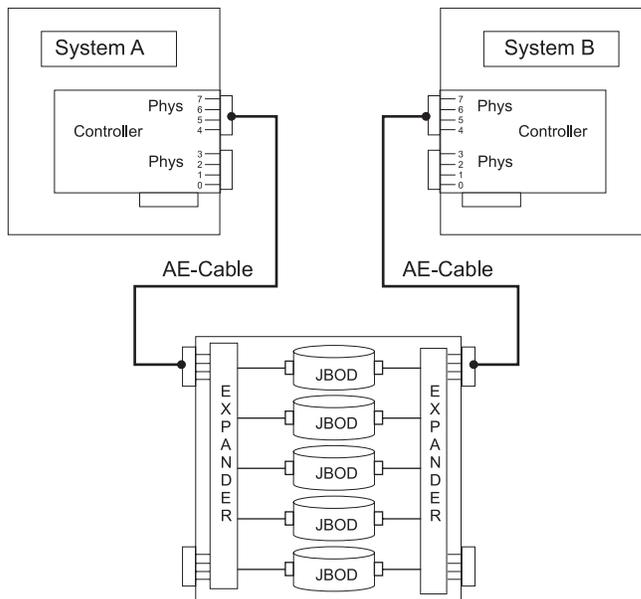
AREBJ514-0

HA Single System RAID



AREBU515-0

HA Two System JBOD



AREBU516-0

Controller functionality

Consider these factors when using multi-initiator and HA functionality.

Use of the multi-initiator and HA functionality requires controller and Linux software support. Controller support is shown in the Comparison of HA features table. Look for controllers that have HA two-system RAID, HA two-system JBOD, or HA single-system RAID marked as Yes for the configuration that you want. The Linux software levels required for multi-initiator support are identified in Table 9 on page 28.

Specific controllers are intended only to be used in either an HA two-system RAID or HA single-system RAID configuration. Use the Comparison of HA features table to look for controllers that have Requires HA RAID configuration marked as Yes. This type of controller may not be used in an HA two-system JBOD or a stand-alone configuration.

Controllers connected in a RAID configuration must have the same write cache size (given they support write cache). A configuration error will be logged if the controllers' write caches are not the same size.

When reconfiguring a controller previously configured in a different HA configuration, it is recommended to configure the High-Availability Mode of the controllers to RAID or JBOD before attaching the SAS cables.

For all HA RAID configurations, one controller functions as the primary controller and performs management of the physical devices, such as creating a disk array or downloading disk microcode. The other controller functions as the secondary controller and is not capable of physical device management.

If the secondary controller detects the primary controller going offline, it will switch roles to become the primary controller. When the original primary controller comes back online, it will become the secondary controller. The exception to this case is if the original primary controller was previously designated as the "preferred" primary controller.

Both controllers are capable of performing direct I/O accesses to the disk arrays for purposes of read and write operations, but at any given time only one controller in the pair is "optimized" for the disk array. The controller that is optimized for a disk array is the one that directly accesses the physical devices for I/O operations. The controller that is non-optimized for a disk array will forward read and write requests through the SAS fabric to the optimized controller. See "HA asymmetric access optimization" on page 50 for more information on setting and viewing disk array optimization.

The primary controller logs most errors related to problems with a disk array. Some disk array errors may also be logged on the secondary if a disk array is optimized on the secondary, at the time the error occurred.

Typical reasons for the primary and secondary controllers to switch roles from what was expected or preferred are as follows:

- Controllers will switch roles for asymmetric reasons. For example, one controller detects more disk drives than the other. If the secondary controller is able to find devices that are not found by the primary controller, an automatic transition (failover) occurs. The controllers will communicate with each other, compare device information, and switch roles.
- Powering off the primary controller or the system that contains the primary controller causes an automatic transition (failover) to occur.
- Failure of primary controller or the system that contains the primary controller causes an automatic transition (failover) to occur.
- If the preferred primary controller is delayed in becoming active, the other controller assumes the role of primary controller. After the preferred primary controller becomes active, an automatic transition (failover) occurs.
- If the primary controller loses contact with the disks that are also accessible by the secondary controller, an automatic transition (failover) occurs.

- Downloading controller microcode might cause an automatic transition (failover) to occur. Such a transition is because the controller would reset itself to activate the new microcode. The controller will be temporarily offline until the reset is complete. Failover to another controller can prevent disk access disruption.

For all JBOD configurations, both controllers function only as stand-alone controllers and do not communicate directly with each other.

Users and their applications are responsible to ensure orderly read and write operations to the shared disks or disk arrays, for example, by using device reservation commands (persistent reservation is not supported).

Controller functionality attributes

Compare important attributes of controller functionality.

Table 12. SAS controller functionality

Controller functionality	HA two-system RAID configuration	HA two-system JBOD configuration	HA single-system RAID configuration
512 bytes per sector disks supported	No ¹	Yes	No ¹
528 bytes per sector disks supported	Yes	No	Yes
Mirrored write cache between controllers (for controllers that have write cache)	Yes		Yes
Mirrored RAID parity footprints between controllers	Yes		Yes
Dual paths to disks	Yes	No	Yes
Target mode initiator device support	Yes	No	No
Only IBM qualified disk drives supported	Yes	Yes	Yes
Only IBM qualified disk expansion drawer supported	Yes	Yes	Yes
Tape or optical devices supported	No	No	No
Boot support	No	No	Yes
Operating mode ²	Primary Adapter or Secondary Adapter ³	StandaloneAdapter ³	Primary Adapter or Secondary Adapter ³
Preferred Dual Initiator Operating mode ²	None (no preference) or Primary ³	None (no preference) or Primary ³	None (no preference) or Primary ³
Dual Initiator Configuration ²	Default ³	JBOD HA Single Path ³	Default ³
High-Availability Mode	RAID ³	JBOD ³	RAID ³
Manage HA Access Characteristics ⁴	Yes	No	Yes

1. 512 bytes-per-sector disks are supported by software.
2. Can be viewed by using the Change Adapter Configuration screen.
3. This option can be set by using the Change Adapter Configuration screen.
4. For information on managing the HA asymmetric access state for a disk array, see "HA asymmetric access optimization" on page 50.

Viewing HA controller attributes

You can view HA configuration-related information from the **Change Adapter Configuration** menu located under the `iprconfig` menu option titled **Work with adapter configuration**.

1. Run `iprconfig`.
2. Type 7 to select **Work with adapter configuration**.
3. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then, press Enter.
4. In the menu option titled **Change Configuration of Adapter**:
 - a. Configure **Preferred Dual Adapter State** to **None** or **Primary**.
 - b. Configure **High-Availability Mode** to **RAID** or **JBOD**.

The screen displayed will look similar to the following example.

```
+-----+
|                                     |
|               Change Configuration of Adapter               |
|                                     |
| Current Disk configurations are shown. To change setting hit 'c' for options |
| menu. Highlight desired option then hit Enter.              |
|   c=Change Setting                                          |
|                                     |
| Adapter: 0000:c8:01.0/0   IBM       572A001SISIOA          |
|                                     |
| Preferred Dual Adapter State . . . . . : None              |
| High-Availability Mode . . . . . : Normal                  |
| Active/Active Mode . . . . . : Disabled                   |
|                                     |
|                                     |
| Or leave blank and press Enter to cancel                    |
| e=Exit   q=Cancel                                          |
|-----+

```

Note: For additional details on how to set up a configuration, see “Installing an HA single-system RAID configuration” on page 57, “Installing an HA two-system RAID configuration” on page 58 or “Installing an HA two-system JBOD configuration” on page 61.

HA cabling considerations

There are different types of cables to consider with high availability.

Correct cabling is one of the most important aspects of planning for a multi-initiator and HA configuration. For RAID configurations with disk expansion drawers, using X cables provides the redundancy of two wide SAS ports between each controller and the disk expansion drawer, along with two narrow SAS ports for each disk drive. For RAID configurations with internal SAS disk slots, using YR cables provides the redundancy of two narrow SAS ports between each controller and the internal disk enclosure, along with two narrow SAS ports for each disk drive. For JBOD configurations, using AE

cables provides one wide SAS port between each controller and the disk expansion drawer, along with one narrow SAS port for each disk drive. Thus there is better SAS fabric redundancy for RAID configurations versus JBOD configurations.

To see examples of how to cable HA configurations, refer to the section entitled Serial attached SCSI cable planning. Look for the following sections for examples of how to cable HA configurations:

- Two SAS adapters to disk expansion drawer – RAID
- Two SAS adapters to disk expansion drawer – JBOD

Attention: Some systems have SAS RAID adapters integrated on the system boards and use a Cache RAID - dual IOA Enablement Card (for example, FC5662) to enable storage adapter write cache and dual storage input/output adapter (IOA) (HA RAID mode). For these configurations, installation of the Cache RAID - Dual IOA Enablement Card places the two integrated adapters into an HA RAID configuration. There are no separate SAS cables required to connect the two integrated SAS RAID adapters to each other.

HA performance considerations

Controller failures can impact performance.

The controller is designed to minimize performance impacts when running in an HA configuration. When using RAID 5, 6, and 10, parity footprints are mirrored between the controller's nonvolatile memory, which causes only a slight impact to performance. For controllers with write cache, all cache data is mirrored between the controller's nonvolatile memories, which also causes only a slight impact to performance.

If one controller fails in an HA configuration, the remaining controller will disable write caching and begin to keep an additional copy of parity footprints on disk. This can significantly impact performance, particularly when using RAID 5 and RAID 6.

Configuration and serviceability considerations for HA RAID configurations

There are configuration and serviceability differences between the primary and secondary controllers.

There are configuration and serviceability differences between the primary controller (which performs direct management of the physical devices) and the secondary controller (which runs as a client of the primary controller). This functional difference requires many of the configuration and serviceability functions to be performed on the primary controller because it is the only controller with that can perform the commands. Attempting these commands on a secondary controller is not recommended and might return unexpected results.

The following are common System Management Interface Tool (SMIT) tasks within the IBM SAS Disk Array Manager that must be performed from the primary controller:

- Under the SMIT menu option titled **IBM SAS Disk Array Manager**:
 - **Create an Array Candidate pdisk and Format to 528 Byte Sectors**
 - **Create a SAS Disk Array**
 - **Delete a SAS Disk Array**
 - **Add Disks to an Existing SAS Disk Array**
 - **Reconstruct a SAS Disk Array**
- Under the SMIT menu option titled **Change/Show SAS pdisk Status**:
 - **Create a Hot Spare**
 - **Delete a Hot Spare**

- Create an Array Candidate pdisk and Format to 528 Byte Sectors
- Delete an Array Candidate pdisk and Format to 512 Byte Sectors
- Under the SMIT menu option titled **Diagnostics and Recovery Options**:
 - Certify Physical Disk Media
 - Download Microcode to a Physical Disk
 - Format Physical Disk Media (pdisk)
 - SCSI and SCSI RAID Hot Plug Manager
 - Reclaim Controller Cache Storage
- Under the SMIT menu option titled **Show SAS Controller Physical Resources**:
 - Show Fabric Path Graphical View
 - Show Fabric Path Data View

Perform other SMIT functions not listed in the preceding lists (for example, Controller Rechargeable Battery Maintenance) on the appropriate controller.

HA asymmetric access optimization

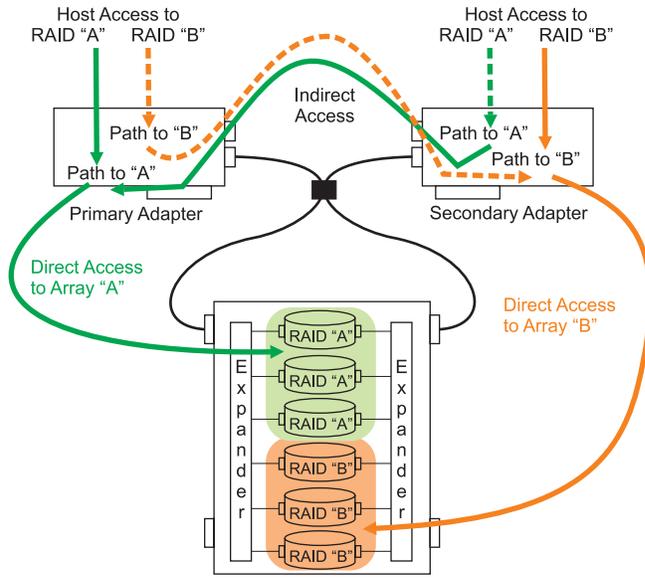
HA access characteristics can balance the controller workload.

Important: Update each controller to the latest SAS controller microcode from the code download Web site. For instructions, see “Updating the controller microcode” on page 63. This update must be performed to obtain the latest, critical fixes to ensure proper functioning.

By default, the primary adapter is optimized for all RAID arrays. The secondary adapter is non-optimized for all RAID arrays. The workload is not balanced between the two controllers. With either of the HA RAID configurations, maximum performance might be achieved by defining HA asymmetric access state for each disk array such that the workload is balanced between the two controllers. Setting the HA asymmetric access state for a disk array specifies which controller is preferred to be optimized for the disk array and perform direct reads and writes to the physical devices.

HA Access Optimization

The example assumes this setup:
 RAID Array "A" Optimized on Primary
 RAID Array "B" Optimized on Secondary



AREBU526-0

Figure 23. HA access optimization

Once the HA asymmetric access is enabled, HA asymmetric access states are displayed on the **Array Asymmetric Access** screen under the menu option titled **Work with Disk Arrays**, similar to the following screen; note that the status column shows the primary adapter view of the HA asymmetric access states of RAID arrays in HA single system RAID configuration:


```

+-----+
|           Change Asymmetric Access Configuration of Array           |
|                                                                     |
|Current array asymmetric access configuration is shown. To change setting hit |
| 'c' for options menu. Highlight desired option then hit Enter.      |
|  c=Change Setting                                                  |
|                                                                     |
|Array: /dev/sdh                                                     |
|Current asymmetric access state: Optimized                          |
|Saved asymmetric access state: Not Set                              |
|                                                                     |
|Preferred Asymmetric Access State. . . . : Not Set                  |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|                                                                     |
|Or leave blank and press Enter to cancel                            |
|                                                                     |
|e=Exit  q=Cancel                                                    |
+-----+

```

This display shows the HA asymmetric access state for the disk arrays selected. For each disk array listed, the current and preferred HA asymmetric access states are indicated. The *current* value shows how the disk array is currently accessed from the controller that was selected. The *preferred* value is the desired access state that is saved with the disk array configuration. Selecting the remote controller would show the opposite settings for the current and preferred access states.

There are three possible access state settings used for the HA asymmetric access state:

Optimized

The selected controller performs direct access for this disk array. This gives I/O operations performed on selected controller optimized performance compared to the remote controller. The selected disk array will be set to Non-Optimized on the remote adapter.

Non-Optimized

The selected controller performs indirect access for this disk array. This gives I/O operations performed on selected controller non-optimized performance compared to the remote controller. The selected disk array will be set to Optimized on the remote adapter.

Not Set

Neither an Optimized nor Non-Optimized access state has been set for this disk array. By default the disk array will be optimized on the primary controller.

The HA asymmetric access state can only be displayed and modified from the primary controller. The **Preferred Asymmetric Access** state can be modified when the disk array is selected from the primary controller. Changing the **Preferred Asymmetric Access** state from the primary controller stores the settings in the disk array and will automatically set the opposite settings on the secondary controller.

The controller will always try to switch the **Current Asymmetric Access** state of the disk array to match the **Preferred Asymmetric Access** state. This switching is done in the background by the controller; therefore, there may be delays between setting the **Preferred Asymmetric Access** state and seeing **Current Asymmetric Access** state switch. There are also situations where the controller will not switch HA asymmetric access state due to scenarios involving configuration errors, failed components, and certain RAID configuration activities.

By default all disk arrays are created with a **Preferred Asymmetric Access** state of Not Set. In order to maximize performance it is recommended that, when appropriate, multiple disk arrays be created and optimized equally between the controller pair. This is accomplished by setting the **Preferred Asymmetric Access** to Optimized for half of the disk arrays and Non-Optimized to the other half.

```

+-----+
|           Change/Show HA Access Characteristics of a SAS Disk Array           |
|                                                                              |
|Type or select values in entry fields.                                     |
|Press Enter AFTER making all desired changes.                             |
|                                                                              |
|                                                                              |
|                                                                              |
|                                                                              |
|          [Entry Fields]          |
| Disk Array          hdisk4          |
| Controller          sissas0         |
| Current Access      Optimized        |
| Preferred Access    Cleared          + |
|                                                                              |
|                                                                              |
|                                                                              |
|                                                                              |
|                                                                              |
|                                                                              |
|                                                                              |
|F1=Help      F2=Refresh      F3=Cancel      F4=List      |
|F5=Reset     F6=Command     F7=Edit       F8=Image     |
|F9=Shell     F10=Exit       Enter=Do      |
+-----+

```

The controller will always try to switch the **Current Access** state of the disk array to match the **Preferred Access** state. This switching is done in the background by the controller; therefore, there may be delays between setting the **Preferred Access** state and seeing **Current Access** state switch. There are also situations where the controller will not switch HA access characteristics due to scenarios involving configuration errors, failed components, and certain RAID configuration activities.

By default all disk arrays are created with a **Preferred Access** state of **Cleared**. In order to maximize performance it is recommended that, when appropriate, multiple disk arrays be created and optimized equally between the controller pair. This is accomplished by setting the **Preferred Access** to **Optimized** for half of the disk arrays and **Non Optimized** to the other half.

The HA access characteristics can also be altered by the **Preferred HA Access State Setting in the Change/Show SAS RAID Controller** screen (see HA configuration-related information above.) This field gives the option to either **Preserve** or **Clear** the **Preferred Access** settings for all disk arrays on the controller pair.

Enabling asymmetric access

HA asymmetric access, an advanced function, is not enabled by default. Use the procedure here to enable it.

Perform the following steps on the primary controller to enable HA asymmetric access:

Installing high availability

Use the procedures in this section when performing HA installations.

Installation procedures are described for an HA two-system RAID configuration, an HA single-system RAID configuration, and an HA two-system JBOD configuration.

Installing an HA single-system RAID configuration

Use this procedure to help you to install an HA single-system RAID configuration.

To avoid problems during installation, follow the steps exactly as written.

Attention: Disk arrays can be created either before or after the HA RAID configuration is set up. See “Configuration and serviceability considerations for HA RAID configurations” on page 49 for important considerations.

1. Install and update the `iprutils` package on each system or partition. See “Updating the `iprutils` package” on page 29 for more information.
2. Power off the system or partition and install the SAS controllers into the system or partition. If you do not want to power off the machine or partition, Linux hot plug can be used to install the SAS controllers without powering off the system or partition.
Attention: Do not attach any cables to the SAS controllers at this time.
3. Update each controller to the latest SAS controller microcode from the code download Web site. See “Updating the controller microcode” on page 63.
4. To prevent errors while connecting the cables, configure the SAS controllers to the HA single-system RAID configuration before attaching any cables to the SAS controllers:
 - a. Run `iprconfig`
 - b. Type 7 to select Work with adapter configuration.
 - c. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then press Enter to select it.
 - d. In the menu option titled **Change Configuration of Adapter**, configure **Preferred Dual Adapter State** to None and **High-Availability Mode** to RAID.
5. Attach an X cable from the shared disk expansion drawer to the same SAS connector on each controller. To see examples of how to cable the HA configurations, see Serial attached SCSI cable planning. For examples of how to cable the HA configurations, see “Two SAS adapters to disk expansion drawer – RAID.”
6. Power on the system or partition if it was powered off. Verify that the cabling and functioning of the controllers are correct by using the **Display Hardware Status** screen. Each SAS controller in dual-initiator configuration should show an optional block of Remote Adapter information for the other SAS controller.
 - a. Run `iprconfig`
 - b. Type 1 to select **Display Hardware Status**.
 - c. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then press Enter to select it. An optional block of Remote Adapter information for the other SAS controller in the remote system or partition is shown in the bottom of the **IOA Hardware Resource Information Details** screen, similar to the following:

Attention: Disk arrays can be created either before or after the HA RAID configuration is set up. See “Configuration and serviceability considerations for HA RAID configurations” on page 49 and “Functions requiring special attention in an HA two-system RAID configuration” on page 60 for important considerations.

1. Install and update the `iprutils` package on each system or partition. See “Updating the `iprutils` package” on page 29 for more information.
2. Power off the system or partition and install the SAS controllers into the system or partition. If you do not want to power off the machine or partition, Linux hot plug can be used to install the SAS controllers without powering off the system or partition.

Attention: Do not attach any cables to the SAS controllers at this time.

3. Update each controller to the latest SAS controller microcode from the code download Web site. See “Updating the controller microcode” on page 63.
4. To prevent errors while connecting the cables, configure the SAS controllers to the HA two-system RAID configuration before attaching any cables to the SAS controllers:
 - a. Run `iprconfig`
 - b. Type 7 to select Work with adapter configuration.
 - c. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then press Enter to select it.
 - d. In the menu option titled **Change Configuration of Adapter**, configure **Preferred Dual Adapter State** to None and **High-Availability Mode** to RAID.
5. Attach an X cable from the shared disk expansion drawer to the same SAS connector on each controller. To see examples of how to cable the HA configurations, see Serial attached SCSI cable planning. For examples of how to cable the HA configurations, see [Two SAS adapters to disk expansion drawer – RAID](#).
6. Power on the system or partition if it was powered off. Verify that the cabling and functioning of the controllers are correct by using the **Display Hardware Status** screen. Each SAS controller in dual-initiator configuration should show an optional block of Remote Adapter information for the other SAS controller.
 - a. Run `iprconfig`
 - b. Type 1 to select **Display Hardware Status**.
 - c. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then press Enter to select it. An optional block of Remote Adapter information for the other SAS controller in the remote system or partition is shown in the bottom of the **IOA Hardware Resource Information Details** screen, similar to the following:

primary controller might also require some manual intervention on the system or partition containing secondary controller to get visibility to the new configuration.

The following table lists some of the common functions and the required steps to perform on the secondary controller:

Table 13. Configuration steps for the secondary controller

Function performed on primary controller	Required configuration on secondary controller
Create pdisk (Format to 528 Byte Sectors)	If device was previously a JBOD hdisk: <code>rmdev -dl hdiskX</code> Then configure new pdisk device: <code>cfgmgr -l sissasX</code>
Delete pdisk (Format to 512 Byte Sectors)	Remove pdisk device: <code>rmdev -dl pdiskX</code> Then configure new hdisk device: <code>cfgmgr -l sissasX</code>
Create Disk Array	Configure new hdisk device: <code>cfgmgr -l sissasX</code>
Delete Disk Array	Remove array hdisk device: <code>rmdev -dl hdiskX</code>
Add disks to Disk Array	No configuration steps needed
Reconstruct Disk Array	No configuration steps needed
Create/Delete Hot Spare disk	No configuration steps needed
Add disk (Hot Plug Manager)	Configure new disk device: <code>cfgmgr -l sissasX</code>
Remove disk (Hot Plug Manager)	Remove disk device: <code>rmdev -dl pdiskX</code>
Reclaim Controller Cache Storage	No configuration steps needed

Installing an HA two-system JBOD configuration

Use this procedure to help you to install an HA configuration for two systems that both have JBOD-formatted disks.

To avoid problems during installation, follow the steps exactly as written.

Attention: Consider using an HA two-system RAID configuration rather than an HA two-system JBOD configuration due to the increased redundancy, performance, and reliability provided by the RAID subsystem.

Attention: Both controllers must have the High-Availability Mode option set to JBOD before being connected to the disk drives, all of the disk drives must be formatted to JBOD (512 bytes per sector) format, and the correct cabling must be used. Remove any devices from the disk expansion drawer that are not set to JBOD or reformat them to JBOD 512 bytes per sector.

1. Install and update the `iprutils` package on each system or partition. See “Updating the `iprutils` package” on page 29 for more information.
2. Power off the system or partition and install the SAS controllers into the system or partition. If you do not want to power off the machine or partition, Linux hot plug can be used to install the SAS controllers without powering off the system or partition.

Attention: Do not attach any cables to the SAS controllers at this time.
3. Update each controller to the latest SAS controller microcode from the code download Web site. See “Updating the controller microcode” on page 63.
4. To prevent errors while connecting the cables, configure the SAS controllers to the HA two-system RAID configuration before attaching any cables to the SAS controllers:
 - a. Run `iprconfig`
 - b. Type 7 to select Work with adapter configuration.
 - c. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then press Enter to select it.

- d. In the menu option titled **Change Configuration of Adapter**, configure **Preferred Dual Adapter State** to None and **High-Availability Mode** to JBOD.
5. Attach an AE cable from the shared disk expansion drawer to the same SAS connector on each controller. To see examples of how to cable the HA configurations, see Serial attached SCSI cable planning.
6. Power on the system or partition if it was powered off. Verify that the cabling and functioning of the controllers are correct by using the **Display Hardware Status** screen. Each SAS controller in dual-initiator configuration should show an optional block of Remote Adapter information for the other SAS controller.
 - a. Run iprconfig
 - b. Type 1 to select **Display Hardware Status**.
 - c. Use the up/down arrow keys to move the cursor to the SAS controller and type 1; then press Enter to select it. An optional block of Remote Adapter information for the other SAS controller in the remote system or partition is shown in the bottom of the **IOA Hardware Resource Information Details** screen, similar to the following:

```

+-----+
|                                     |
|                               IOA Hardware Resource Information Details |
|                                     |
| Manufacturer . . . . . : IBM |
| Machine Type and Model . . . . . : 572B001SISIOA |
| Firmware Version . . . . . : 03200046 |
| Serial Number. . . . . : 07125793 |
| Part Number. . . . . : 0000042R4591 |
| Plant of Manufacturer. . . . . : 0022 |
| Cache Size . . . . . : 175 MB |
| DRAM Size. . . . . : 0EE MB |
| Resource Name. . . . . : /dev/sg20 |
|                                     |
| Physical location |
| PCI Address. . . . . : 0002:00:01.0 |
| SCSI Host Number . . . . . : 1 |
|                                     |
| Current Dual Adapter State . . . . . : Primary |
| Preferred Dual Adapter State . . . . . : No Preference |
| Remote Adapter Manufacturer. . . . . : IBM |
| Remote Adapter Machine Type And Model. . . . . : 572B001SISIOA |
| Remote Adapter Serial Number . . . . . : 07125687 |
|                                     |
| Current Asymmetric Access State. . . . . : Disabled |
|                                     |
| Press Enter to Continue |
|                                     |
| e=Exit  q=Cancel |
|                                     |
+-----+

```

IBM SAS RAID controller maintenance

Maintenance procedures include microcode updates, failure recovery procedures, and working with the battery pack.

Follow the recommended maintenance procedures in this section to help maintain your RAID controller for Linux.

Usage tips

Follow recommended usage tips to help avoid controller and disk array problems.

The following usage tips will help you maintain your controller.

- Always perform a normal system shutdown before physically replacing or moving the RAID adapter or members of disk arrays. A normal shutdown of the system will flush the adapter's write cache and remove dependencies between the adapter and the physical disks. Using the **modprobe -r ipr** command will have the same effect as a system shutdown. For systems that support PCI hot plug, PCI hot plug will also have the same effect as a system shutdown.

Note: A disk that is a failed member of a Degraded disk array can be replaced and the disk array rebuilt while the system continues to run.

- You can physically move disks from one adapter to another. However, if the physical disks are members of a disk array, be sure to move all the disks as a group. Prior to attempting a disk movement, ensure that the disk array is not in a Degraded state because of a disk failure.
- When physically removing disks that are members of a disk array and there is no need to preserve data and no intent to use the disk array again, delete the disk array before removing the disks. This action avoids disk array-related problems the next time these disks are used.
- Always use the "Concurrent device remove" screen to remove and replace a physical disk. For instructions on how to remove and replace a disk, see "Removing a failed disk" on page 72.
- If a disk array is being used as a boot device and the system fails to boot because of a suspected disk array problem, boot into Rescue mode. Linux error logs, the iprconfig utility, and other tools are available to help determine and resolve the problem with the disk array.
- Do not attempt to correct problems by swapping adapters and disks unless you are directed to do so by the service procedures. This is likely to make the problems worse. For additional information regarding problem determination, see "Problem determination and recovery" on page 79.
- Do not confuse the cache directory card, which is a small rectangular card with a round button-shaped battery, for a cache card. The nonvolatile write cache memory is integrated into the main adapter card or integrated into the removable cache card on the newest adapters. The memory itself is not removable, and is battery-backed by the large, rechargeable cache battery pack. The cache directory card contains only a secondary copy of the write cache directory, no cache data, and should never be removed except under very specific cases as described in the MAPs.
- Do not unplug or exchange a cache battery pack without following the procedures as outlined in this section or in the MAPs.
- If multiple errors occur at approximately the same time, look at them as a whole to determine if there might be a common cause.

Updating the controller microcode

Disk drives attached to this RAID adapter must contain the latest microcode level. Microcode can be updated using the iprconfig utility.

Download the latest microcode level for your drive model and controller from IBM Systems Support. If you download new microcode from the Web site, follow the instructions in the readme file or Desc link on the Web site to install it. Use the installation procedure below only if there are no instructions available.

The iprconfig utility allows for multiple levels of adapter and device microcode to exist on the system at the same time. After the adapter or device to update microcode is selected, all available microcode levels are presented to choose to download.

Note: For either an auxiliary cache adapter or a storage I/O adapter which attaches to an auxiliary cache adapter, it is recommended that both the storage I/O adapter and auxiliary cache adapter be updated.

1. Install the package by typing `rpm -ivh pci.101402BD.20-01200041-1.Linux.noarch.rpm`.
2. Update the microcode by typing `iprconfig`.
3. Select **Download microcode** and press Enter.
4. Select the device or adapter you wish to update by typing a 1, and then press Enter. Multiple devices can be selected.
5. Select the microcode image to be downloaded to the device or adapter by typing a 1 and then press Enter.
6. Press Enter again to confirm the download.
7. If multiple devices were selected, repeat from Step 5 for the next device.

Rechargeable battery maintenance

Proper battery maintenance procedures can help to prevent data loss.

Perform battery maintenance tasks only when directed to do so by an isolation procedure or a maintenance analysis procedure (MAP).

Displaying rechargeable battery information

You can view information about the rechargeable battery by using the iprconfig utility.

1. Run the iprconfig utility by typing `iprconfig`.
2. Select **Work with disk unit recovery**.
3. Select **Work with resources containing cache battery packs**.
4. Select your adapter and type 1.
5. Press Enter to display battery information. The Battery Information screen displayed will look similar to the following:

Replacing the rechargeable battery pack

You can replace battery packs that are maintainable both concurrently and nonconcurrently.

The battery pack replacement procedure you follow depends on whether your battery pack is concurrently maintainable or nonconcurrently maintainable.

Note: When replacing the cache battery pack, the battery must be disconnected for at least 60 seconds before connecting the new battery. This is the minimum amount of time needed for the card to recognize that the battery has been replaced.

CAUTION:

The battery is a lithium ion battery. To avoid possible explosion, do not burn. Exchange only with the IBM-approved part. Recycle or discard the battery as instructed by local regulations. In the United States, IBM has a process for the collection of this battery. For information, call 1-800-426-4333. Have the IBM part number for the battery unit available when you call.

Attention: Data loss can occur; if the cache battery pack is not already in the error state, follow the steps described in “Forcing a rechargeable battery error” on page 65 before proceeding. After the battery is in the error state, replace it with a new battery by completing the appropriate replacement procedure.

Attention:

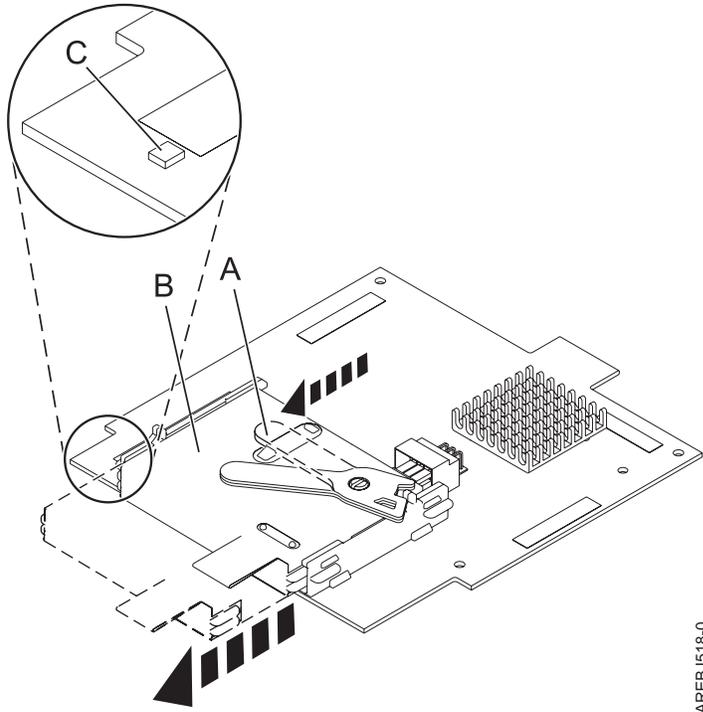
- Static electricity can damage this device and your system unit. To avoid damage, keep this device in its anti-static protective bag until you are ready to install it. To reduce the possibility of electrostatic discharge, follow the precautions listed below:
 - Limit your movement. Movement can cause static electricity to build up around you.
 - Handle the device carefully, holding it by its edges or its frame.
 - Do not touch solder joints, pins, or exposed printed circuitry.
 - Do not leave the device where others can handle and possibly damage the device.
 - While the device is still in its anti-static package, touch it to an unpainted metal part of the system unit for at least two seconds. (This drains static electricity from the package and from your body.)
 - Remove the device from its package and install it directly into your system unit without setting it down. If it is necessary to set the device down, place it on its static-protective package. (If your device is a controller, place it component-side up.) Do not place the device on your system unit cover or on a metal table.
 - Take additional care when handling devices during cold weather, as heating reduces indoor humidity and increases static electricity.

Replacing a nonconcurrently maintainable battery pack

Use this procedure to replace a battery pack that is nonconcurrently maintainable or to replace a battery pack that is not concurrently maintainable or is concurrently maintainable but not accessible due to the structure of the system enclosure.

Determine whether it is safe to replace the cache battery pack by referring to “Displaying rechargeable battery information” on page 64. It is safe to replace the cache battery pack when the Battery Information screen displays Yes at the entry “Battery pack can be safely replaced.”

1. Remove the controller from the system. See your system documentation for instructions.
2. Place the controller on an ESD protected surface.



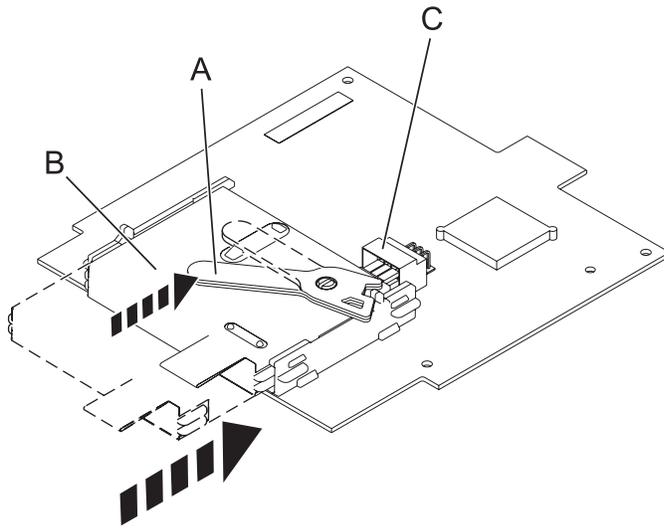
AREBJ518-0

Figure 24. Removing the cache battery

- A** Cache battery lever
- B** Cache battery pack
- C** Cache present LED

3. Move the cache battery lever **A** away from the connector to disengage the battery from the connector.
4. Continue to slide the cache battery pack out of the mounting guides and remove it from the controller.

Note: Ensure that the cache battery pack is disconnected for at least 60 seconds before connecting the new battery. This is the minimum amount of time needed for the card to recognize that the battery has been replaced.



AREBJ519-0

Figure 25. Replacing the cache battery

- A** Cache battery lever
- B** Cache battery pack
- C** Cache battery connector

5. Before installing the new cache battery pack, move the lever to the unlatched position (away from the connector).
6. Slide the new cache battery pack into the mounting guides on the controller until it is seated in the battery connector.
7. Once the battery is seated in the connector, move the lever to the latched position to fully seat the battery into the connector.
8. Install the controller into the system. See your system documentation for instructions.

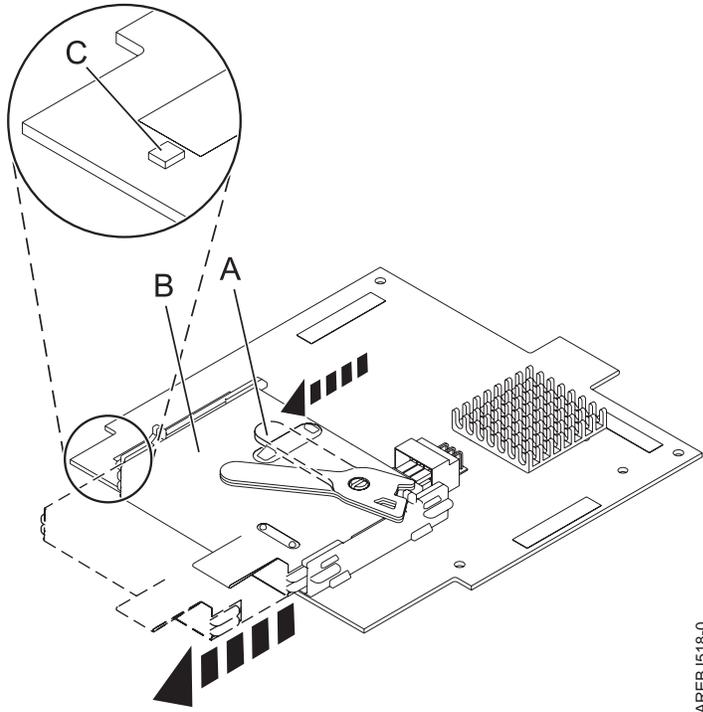
Replacing a concurrently maintainable battery pack

Use this procedure to replace a battery pack that is concurrently maintainable.

Note: If your system enclosure does not allow concurrent access to the battery, refer to “Replacing a nonconcurrently maintainable battery pack” on page 66 to replace the battery pack.

Determine whether it is safe to replace the cache battery pack by referring to “Displaying rechargeable battery information” on page 64. It is safe to replace the cache battery pack when the Battery Information screen displays Yes at the entry “Battery pack can be safely replaced.”

1. Remove the controller from the system. See your system documentation for instructions.
2. Place the controller on an ESD protected surface.



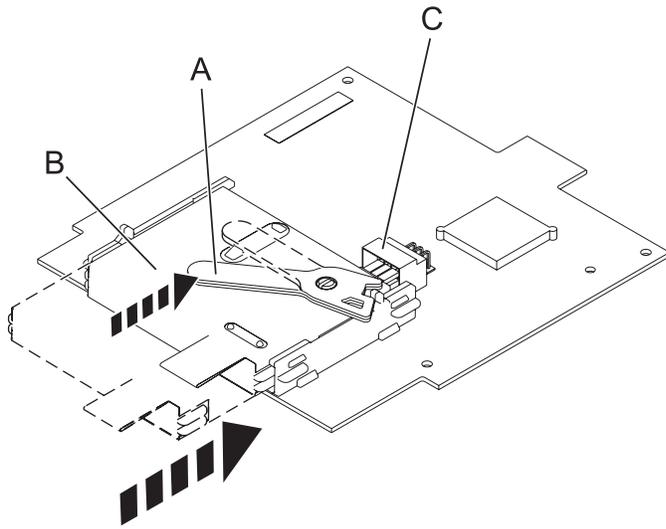
AREBJ518-0

Figure 26. Removing the cache battery

- A** Cache battery lever
- B** Cache battery pack
- C** Cache present LED

3. Move the cache battery lever **A** away from the connector to disengage the battery from the connector.
4. Continue to slide the cache battery pack out of the mounting guides and remove it from the controller.

Note: Ensure that the cache battery pack is disconnected for at least 60 seconds before connecting the new battery. This is the minimum amount of time needed for the card to recognize that the battery has been replaced.



AREBJ519-0

Figure 27. Replacing the cache battery

- A** Cache battery lever
- B** Cache battery pack
- C** Cache battery connector

5. Before installing the new cache battery pack, move the lever to the unlatched position (away from the connector).
6. Slide the new cache battery pack into the mounting guides on the controller until it is seated in the battery connector.
7. Once the battery is seated in the connector, move the lever to the latched position to fully seat the battery into the connector.
8. Restart the adapter's write cache by completing the following steps:
 - a. Run the `iprconfig` utility by typing `iprconfig`.
 - b. Select **Work with disk unit recovery**.
 - c. Select **Work with resources containing cache battery packs**.
 - d. Start the IOA cache on the adapter whose battery you just replaced by typing `3`.
 - e. Press Enter.

Replacing the cache directory card

Replace the cache directory card only if you are directed to do so by an isolation procedure or a maintenance analysis procedure (MAP).

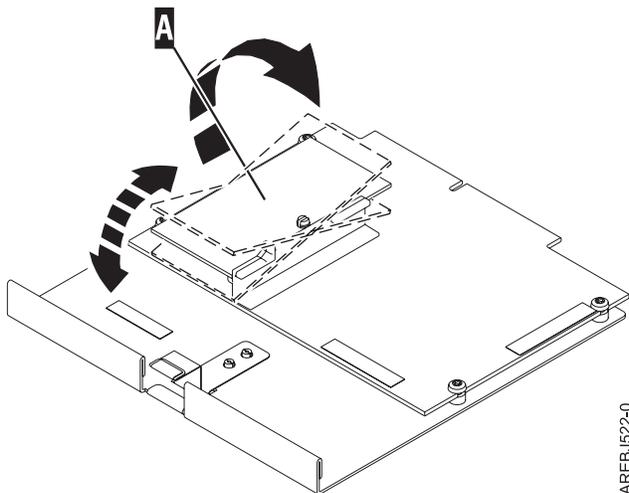
Attention: This procedure should only be performed if directed from an isolation procedure or a maintenance analysis procedure (MAP).

Attention:

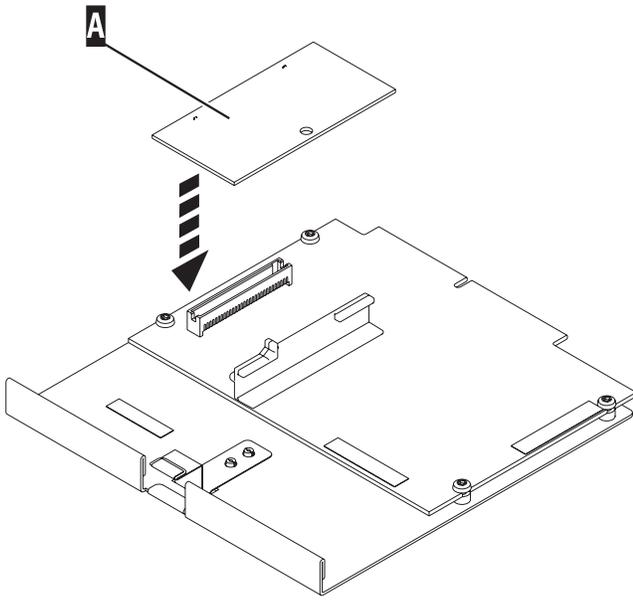
- Static electricity can damage this device and your system unit. To avoid damage, keep this device in its anti-static protective bag until you are ready to install it. To reduce the possibility of electrostatic discharge, follow the precautions listed below:
 - Limit your movement. Movement can cause static electricity to build up around you.
 - Handle the device carefully, holding it by its edges or its frame. Do not touch solder joints, pins, or exposed printed circuitry.
 - Do not leave the device where others can handle and possibly damage the device.
 - While the device is still in its anti-static package, touch it to an unpainted metal part of the system unit for at least two seconds. (This drains static electricity from the package and from your body.)
 - Remove the device from its package and install it directly into your system unit without setting it down. If it is necessary to set the device down, place it on its static-protective package. (If your device is a controller, place it component-side up.) Do not place the device on your system unit cover or on a metal table.
 - Take additional care when handling devices during cold weather, as heating reduces indoor humidity and increases static electricity.

To replace the cache directory card, perform the following steps:

1. Remove the controller using the remove procedures for the model or expansion unit on which you are working.
2. Locate the cache directory card **A**.



3. Unseat the connector on the cache directory card by wiggling the two corners over the connector using a rocking motion. Then, lift the cache directory card off the connector and out of the guides on the cache battery rail.
4. Install the replacement card **A** by inserting it into the guides on the cache battery rail and then seating it on the connector.



AREB.1523-0

5. Install the controller using the installation procedures for the model or expansion unit on which you are working.

Physical disks

Replace failed disks as soon as possible, even if a rebuild was initiated on a hot spare by the controller.

Use the **Concurrent remove device** option in the iprconfig utility to replace physical disks attached to the controller.

Removing a failed disk

Use the iprconfig utility to remove a failed disk.

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk unit recovery**.
3. Select **Concurrent Device Remove**. The resulting screen is similar to the following:

Concurrent Device Remove			
Choose a single location for remove operations			
1=Select			
OPT Name	PCI/SCSI Location	Description	Status
sda	0000:00:01.0/0:6:2:0	Physical Disk	Active
	0000:00:01.0/0:6:3:		Empty
	0000:00:01.0/0:6:4:		Empty
sdb	0000:00:01.0/0:6:5:0	Physical Disk	Active
	0000:00:01.0/0:6:6:		Empty
	0000:00:01.0/0:6:7:		Empty
	0000:00:01.0/0:3:0:		Empty
sdc	0002:00:01.0/1:2:1:0	Physical Disk	Active
sdd	0002:00:01.0/1:2:2:0	Physical Disk	Active
	0002:00:01.0/1:0:3:		Empty
	0002:00:01.0/1:2:4:0	Advanced Function Disk	Active
	0002:00:01.0/1:2:5:0	Advanced Function Disk	Active
	0002:00:01.0/1:2:6:0	Advanced Function Disk	Active
	0002:00:01.0/1:2:7:0	Hot Spare	Active
	0002:00:01.0/1:2:8:0	RAID 6 Array Member	Active
	0002:00:01.0/1:2:9:0	RAID 6 Array Member	Active
	0002:00:01.0/1:2:10:0	RAID 6 Array Member	Active
	0002:00:01.0/1:2:11:0	RAID 6 Array Member	Active

e=Exit q=Cancel t=Toggle

4. Select the device to be removed by typing 1; then, press Enter.
5. Verify that the selected device is the device to be removed; the Identify indicator should now be set. Press Enter.

Attention: iprconfig will allow removal of devices that are currently being used by the system. Ensure that the device is not in use by the system prior to removal to avoid data loss.

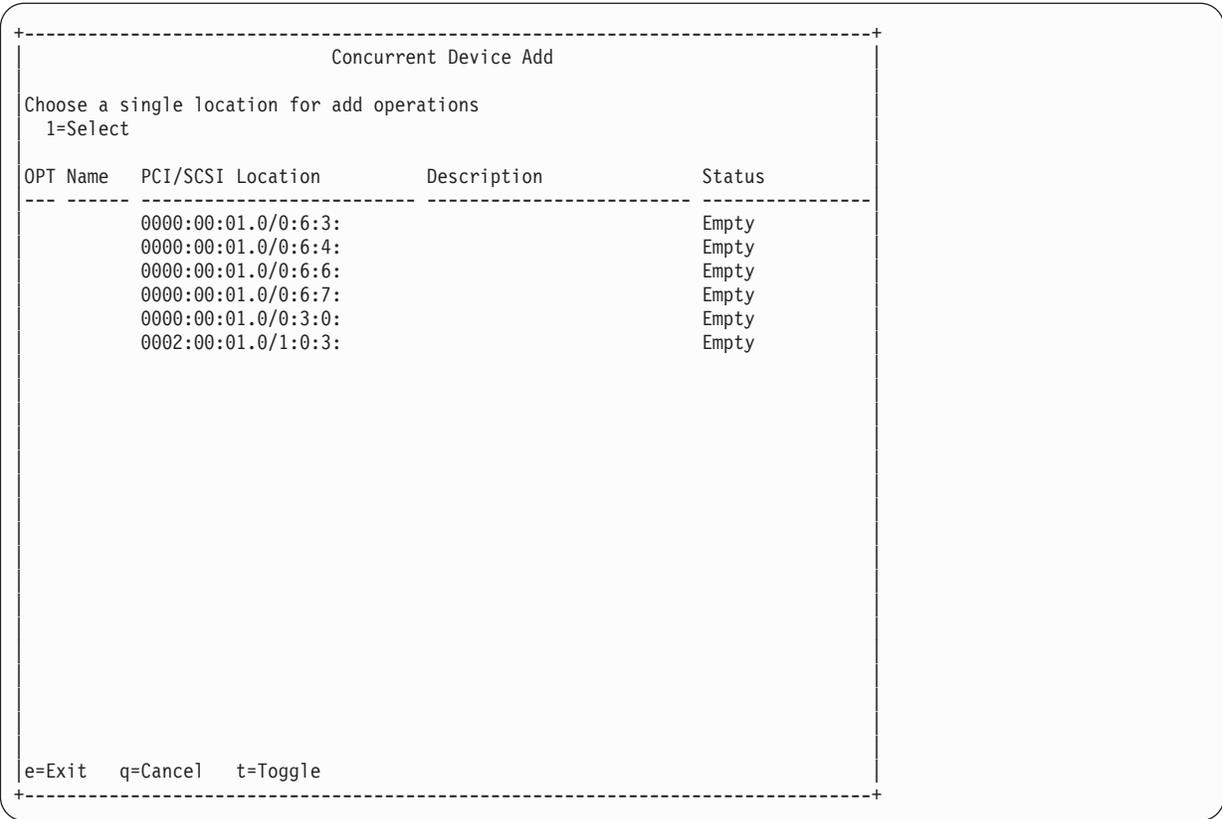
6. Remove the physical disk from the system.

Attention: Do not install the replacement disk at this time.

Installing a new disk

Use the iprconfig utility to install a new disk after the failed disk has been removed.

1. Run the iprconfig utility by typing iprconfig.
2. Select **Concurrent Device Add**. The resulting screen is similar to the following:



3. Select the location to install the device by typing 1; then, press Enter. If you do not see the location to install, press t to toggle the dual path locations.
4. Verify the location of the device to install. The identify indicator should now be set. Press Enter.
5. Install the physical disk into the system.
6. Press Enter.

If a hot spare disk was available and kept your disk array protected throughout the failure, no additional steps are required because the hot spare has replaced the failed disk unit in the disk array, and the newly installed disk is available to be configured as a hot spare.

If a hot spare rebuild was not initiated by the controller, a rebuild needs to be initiated on the newly installed disk.

Related tasks

“Initiating a rebuild on a newly-installed disk”

If a hot spare rebuild is not initiated by the controller when you have installed a new disk, you must initiate a rebuild before the disk is recognized by the system.

Initiating a rebuild on a newly-installed disk

If a hot spare rebuild is not initiated by the controller when you have installed a new disk, you must initiate a rebuild before the disk is recognized by the system.

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk unit recovery**.
3. Select the disk you want to rebuild; then, press Enter.

Attention: Data currently on the disk will be overwritten.

4. To rebuild data, press Enter. If you do not want to rebuild data, type q to cancel.

Disk failure recovery

The IBM SAS RAID controller handles disk failures differently, depending on the RAID level involved.

The recovery procedure you use will depend on the level of the RAID adapter.

For a description of disk array and physical disk states, see “Disk array overview” on page 21. For a description of RAID levels, see “Supported RAID levels” on page 16.

RAID 0 failure

RAID 0 does not provide data protection. A single disk failure causes a RAID 0 array to transition to the Failed state.

If a single disk failure causes a RAID 0 array to transition to the Failed state, you must delete the disk array, replace the disk that is Failed, and recreate the disk array. You must then recreate the file systems on the disk array, and copy data back to the restored disk array from your backup media.

RAID 5 disk recovery

RAID 5 protects data with parity information distributed on all member disks.

The data can be recovered from the remaining disks if one disk fails.

Recovering a RAID 5 single-disk failure

RAID 5 protects the data with parity information distributed on all member disks. The data can be recovered from the remaining disks if one disk fails.

When a single disk in a RAID 5 disk array fails, the disk array status changes to Degraded. The disk array remains functional because the data on the failed disk can be rebuilt using parity and data on the remaining disks. If a hot spare disk is available, the controller can rebuild the data on the disk automatically. If a hot spare disk is not available, you must replace the failed disk and then initiate a rebuild.

Complete the following steps to initiate a rebuild:

1. Run the `iprconfig` utility by typing `iprconfig`.
2. Select **Work with disk unit recovery**.
3. Select **Rebuild disk unit data**.
4. Select the disks you want to rebuild, then press Enter.
5. To rebuild data, press Enter. If you do not want to rebuild data, type `q` to cancel.

Attention: Data currently on the disk will be overwritten.

RAID 5 multiple-disk failure

If a second disk in a RAID 5 disk array fails, the array also fails and its data is not accessible.

If a second disk in a RAID level 5 disk array fails, you must replace the failed disks, then delete and recreate the disk array. You must then recreate the file systems on the disk array and copy data to the restored disk array from your backup media.

RAID 6 disk recovery

RAID 6 protects data with dual parity information distributed on all member disks.

The data is recoverable from the remaining disks if one or two disks fail. If a third disk in a RAID 6 disk array fails, the data is not recoverable.

Recovering a RAID 6 single- or dual-disk failure

Whenever one or two disks in a RAID 6 disk array transitions to the Failed state, the disk array status changes to Degraded. The disk array remains functional because the data on the failed disk can be rebuilt using P and Q parity and data on the remaining disks.

Whenever a disk is marked as Failed, replace it as soon as possible. If a hot spare disk is available, the controller can rebuild the data on the disk automatically. If a hot spare disk is not available, you will need to replace the failed disk and then initiate a rebuild.

Perform the following steps to initiate a rebuild:

1. Run the iprconfig utility by typing `iprconfig`.
2. Select **Work with disk unit recovery**.
3. Select **Rebuild disk unit data**.
4. Select the disks you want to rebuild, then press Enter.
5. To rebuild data, press Enter. If you do not want to rebuild data, type `q` to cancel.

Attention: Data currently on the disk will be overwritten.

RAID 6 failure of three or more disks

If a third disk in a RAID 6 disk array fails, the array is marked as Failed and its data is not accessible.

If a third disk in a RAID 6 disk array fails, you must replace the failed disks, then delete and recreate the disk array. You must then recreate the file systems on the disk array and copy data to the restored disk array from your backup media.

RAID 10 disk recovery

RAID 10 protects the data on mirrored disk pairs.

Data can be recovered unless both disks of a mirrored disk pair fails.

Recovering a RAID 10 single-disk failure

When a single disk in a RAID 10 disk array fails, the disk array status changes to Degraded. The disk array remains functional because the data on the Failed disk is also stored on the other member of its mirrored pair.

When ever a disk fails, replace it as soon as possible. If a hot spare disk is available, the controller can rebuild the data on the disk automatically. If a hot spare disk is not available, you will need to replace the failed disk and then initiate a rebuild.

Complete the following steps to initiate a rebuild:

1. Run the iprconfig utility by typing `iprconfig`.
2. Select **Work with disk unit recovery**.
3. Select **Rebuild disk unit data**.
4. Select the disks you want to rebuild, then press Enter.
5. To rebuild data, press Enter. If you do not want to rebuild data, type `q` to cancel.

Attention: Data currently on the disk will be overwritten.

RAID 10 multiple-disk failure

The procedure for handing a RAID 10 multiple-disk failure depends on whether the failed disk is on the same mirrored pair or on a different mirrored pair.

If each failed disk belongs to a different mirrored pair, then the array is in the Degraded state because the data on each failed disk is still available from its mirrored pair. The recovery procedure for this case is the same as that for a single-disk failure case where you initiate the rebuild operation after you replace the failed disks.

When both members of a mirrored pair fail, the disk array status changes to Failed. You must delete the disk array, replace the failed disks, and then recreate the disk array. You must then recreate the file systems on the disk array and copy data to the restored disk array from your backup media.

Reclaiming IOA cache storage

A reclamation procedure for IOA cache storage is sometimes used to resolve cache problems. Perform this procedure only if directed to do so by a Maintenance Analysis Procedure (MAP).

Attention: Data may be lost. When an auxiliary cache adapter connected to the RAID controller logs a 9055 URC in the error log, the Reclaim procedure does not result in lost sectors. Otherwise, the Reclaim procedure will result in lost sectors.

Complete the following steps to reclaim IOA cache storage:

1. Run the iprconfig utility by typing iprconfig.
2. Select **Work with disk unit recovery**.
3. Select **Reclaim IOA cache storage**.
4. Select the desired adapter, then press Enter. A screen display will then show the devices that might be affected.
5. To reclaim IOA cache storage, type c to confirm. If you do not want to reclaim IOA cache storage, type q to cancel. A screen display similar to one of the following will then be displayed.

```
+-----+
|                                     |
|               Reclaim IOA Cache Storage Results               |
|                                     |
| IOA cache storage reclamation has completed. The number of   |
| lost sectors could not be determined.                          |
|                                     |
| Press Enter to continue.                                       |
|                                     |
| e=Exit  q=Cancel                                              |
|-----+-----+
|               Reclaim IOA Cache Storage Results               |
|                                     |
| IOA cache storage reclamation has completed. Use the number  |
| of lost sectors to decide whether to restore data from the   |
| most recent save media or to continue with possible data loss. |
|                                     |
| Number of lost sectors . . . . . :      1572608              |
|                                     |
| Press Enter to continue.                                       |
|                                     |
| e=Exit  q=Cancel                                              |
|-----+-----+
```

Problem determination and recovery

Several tools are available that the trained service representative can use to identify and recover from problems.

Trained service representatives use error logs, Unit Reference Code tables, and Maintenance Analysis Procedures (MAPs) in problem determination and recovery procedures.

Note: The procedures contained in this section are intended for service representatives specifically trained on the system unit and subsystem that is being serviced. Additionally, some of the service actions in this topic might require involvement of the system administrator. For additional sources of related service procedures required to service this subsystem, see “Related information” on page 13.

If a problem arises related to disk arrays and associated physical disks, use the following to identify the problem:

- **ipr** error log entries associated with the problem, viewed using the **Analyze Log** option in `iprconfig` or directly from `/var/log/messages`
- Disk array and physical disk status, viewed using the `iprconfig` utility

Entries in the error log contain a Unit Reference Code (URC) and a description of the problem. Each URC corresponds to an action that should be taken. It is sometimes recommended that you perform a Maintenance Analysis Procedure (MAP) to further determine what actions should be taken to resolve the problem.

This topic collection provides many of these MAPs. The MAPs are intended to address only problems directly related to disk arrays and SAS cabling problem isolation. MAPs related to other device or adapter problems, when applicable, are located in other Linux or system documentation.

Analyzing error logs

The experienced user can search through this file with commands, such as `grep` and `sed`, but the `iprconfig` utility also offers a convenient menu for analyzing the error log.

Errors logged by the adapters and devices are stored in the `/var/log/messages` file. Options on the error log screen that are not mentioned here are used to gather certain information from the error log and present it in the default editor. This is used mostly during Maintenance Analysis Procedures to search for Unit Reference Codes.

1. Select **Analyze Log** from the main menu in `iprconfig`. A screen will display information similar to the following:

Searching logs

Use commands similar to these to perform a search on an error log.

The steps in this procedure are presented as a sample search for the URC "3400." Replace the variable "3400" with your own search term.

1. From the Kernel Messages Log screen, select the **Use vi to view most recent ipr error messages** option.
2. Type **G** to move to the last line of the file.
3. Type **?3400** and press **Enter** to search backward for the pattern "3400."
4. If the first instance found is not part of an error log entry header, repeat the command by typing **n** as necessary.

Sample error logs

When viewed through iprconfig, ipr error messages are displayed in various formats.

The first portion of the log entry is the header, which contains important information such as the Unit Reference Code, the error class, the IOA under which the error occurred, the error description text, and the hardware location. The last block of data in the entry is the IOA error data, which is data recorded by the IOA to assist with recovery, if the recommended service action did not solve the problem.

Generic IOA or device errors

Sample generic IOA or device error log

The generic IOA or device error log provides information in the following format:

```
+-----+
2:0:9:0: FFFE: Soft device bus error recovered by the IOA
00000000: 01080000 00000900 FFFFFFFF 1104E092
00000010: 00000000 00000002 00000000 00000000
00000020: 00000000 00000000 00000000 00000000
00000030: 00000000 00000000 0034D780 00000900
00000040: 00000311 041F1CC7 0000E092 0000E092
00000050: 0000E092 0000E092 0000E092 0000E092
00000060: 28282828 28282828 D000100F 00052800
00000070: 041F1CC7 00000100 00000000 90280000
00000080: 00000500 00000000 0C804000 00CCDDDD
00000090: 00000000 00000000 49000000 00315208
000000A0: 83000000 40010000 80000000 28E30000
000000B0: 000001F4 00000000 00000000 00000000
000000C0: 00000000 00000000 00000000 00000000
000000D0: 00000000 00000000 00000000 00000000
000000E0: 00000000 00000000 00000000 00000000
000000F0: 00000000 00000000 00000000 0034D780
00000100: 00000000 FFFFFFFF 00000000 00000000
00000110: 00000000 53544154 E00000A2 0034D780
+-----+
```

Device configuration errors

Sample device configuration error log

The device configuration error log provides information in the following format:

```

+-----+
0001:61:01.0: 9002: IOA reserved area LRC error
Device Errors Detected/Logged: 1/1
-----
Device 1: 2:0:9:0
Vendor/Product ID: IBM      ST336605LC
  Serial Number: 000227CE
    WWN: 5000CCA00336F4AC
-----New Device Information-----
Vendor/Product ID:
  Serial Number: 00000000
    WWN: 0000000000000000
Cache Directory Card Information:
Vendor/Product ID:
  Serial Number: 00000000
    WWN: 0000000000000000
Adapter Card Information:
Vendor/Product ID:
  Serial Number: 00000000
    WWN: 0000000000000000
Additional IOA Data: 00000000 00000000 00000000 00000000 00000000
+-----+

```

Array errors

Sample array error log

The array error log provides information in the following format:

```

+-----+
0001:61:01.0: 9030: Array no longer protected due to missing or failed disk unit
-----
RAID 10 Array Configuration: 2:255:0:0
-----
Exposed Array Member 0:
Vendor/Product ID: IBM      IC35L036UCDY10-0
  Serial Number: E3V1WEAB
    WWN: 5000CCA00336F4AC
Current Location: 2:0:4:0
Expected Location: 2:0:4:0
-----
Array Member 1:
Vendor/Product ID: IBM      IC35L036UCDY10-0
  Serial Number: E3V0J55B
    WWN: 5005076C0400C703
Current Location: 2:0:5:0
Expected Location: 2:0:5:0
+-----+

```

Cache errors

Sample cache error log

The cache error log provides information in the following format:

```

+-----+
0001:61:01.0: 9010: Cache data associated with attached devices cannot be found
-----Current Configuration-----
Cache Directory Card Information:
Vendor/Product ID: IBM    5703001
Serial Number: 03060038
                WWN: 5005076C03023F00
Adapter Card Information:
Vendor/Product ID: IBM    5703001
Serial Number: 03060038
                WWN: 5005076C03023F00
-----Expected Configuration-----
Cache Directory Card Information:
Vendor/Product ID:
Serial Number:
                WWN: 0000000000000000
Adapter Card Information:
Vendor/Product ID: IBM    5703001
Serial Number: 03060038
                WWN: 5005076C03023F00
Additional IOA Data: 00000000 00000000 152101D0
+-----+

```

Disk array problem identification

A disk array problem is uniquely identified by a Unit Reference Code (URC).

A Unit Reference Code (URC) is used to indicate the specific problem that has occurred. This code is required in order to determine which MAP to use.

A URC is provided in the ipr error log with each error. This URC describes the exact problem that has been detected and is considered the primary means of identifying a problem. However, the **Display Hardware Status** screen within the iprconfig utility is also a useful tool in identifying a problem or confirming a problem described by the error log. For additional information about this status screen, see “Viewing device status” on page 31.

When you have obtained the URC, you can then determine which Maintenance Analysis Procedure (MAP) to use.

Unit reference code tables

Use the information in this section, along with the unit reference code (URC) obtained from the ipr error log, to determine which MAP to use.

Table 15. URCs.

Note: This table includes only URCs that are associated with Maintenance Analysis Procedures contained in this information set.

URC	Description text	Service action	Failing items
102E	Out of alternate sectors for disk storage	Exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform “MAP 3351” on page 102.	Disk drive
3002	Addressed device failed to respond to selection	Exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform “MAP 3351” on page 102.	Disk drive

Table 15. URCs (continued).

Note: This table includes only URCs that are associated with Maintenance Analysis Procedures contained in this information set.

URC	Description text	Service action	Failing items
3010	Disk device returned wrong response to IOA	Exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform "MAP 3351" on page 102.	Disk drive
3020	Storage subsystem configuration error IOA detected a SAS fabric configuration error	Perform "MAP 3350" on page 99.	
3029	A device replacement has occurred	No action required.	
3100	Device bus error	A SAS fabric error occurred. Perform "MAP 3350" on page 99.	
3109	IOA timed out a device command	Perform "MAP 3350" on page 99.	
3400	Device, I/O Adapter, Any device on I/O bus, Signal cables	Exchange the failing items in the Failing Items list one at a time.	Device, I/O Adapter, Any device on I/O signal cables
4010	Incorrect connection between cascaded expanders	Perform "MAP 3342" on page 97.	
4020	Connections exceed IOA design limits	Perform "MAP 3343" on page 97.	
4030	Incorrect multipath connection	Perform "MAP 3344" on page 97.	
4040	Incomplete multipath connection between IOA and enclosure	Perform "MAP 3344" on page 97.	
4041	Incomplete multipath connection between enclosure and device	Perform "MAP 3346" on page 98.	
4050	Enclosure does not support a required multipath function	Perform "MAP 3348" on page 98.	
4060	Multipath redundancy level got worse	Perform "MAP 3353" on page 103.	
4061	Multipath redundancy level got better	No action required	
4100	Hard device bus fabric error	Perform "MAP 3352" on page 103.	
4101	Soft device bus fabric error	Perform "MAP 3352" on page 103.	
4110	Unsupported enclosure function	Perform "MAP 3345" on page 98.	

Table 15. URCs (continued).

Note: This table includes only URCs that are associated with Maintenance Analysis Procedures contained in this information set.

URC	Description text	Service action	Failing items
7001	IOA sector reassignment successful	If three 7001 messages have occurred for the same disk drive location, then exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform "MAP 3351" on page 102.	Disk drive
8008	A permanent cache battery pack failure occurred Impending cache battery pack failure	Perform MAP 3300	Cache battery pack
8009	A permanent cache battery pack failure occurred Impending cache battery pack failure	Perform MAP 3300	Cache battery pack
8150	Permanent IOA failure	Exchange the failing items in the Failing Items list one at a time.	I/O adapter
	PCI bus error	If two errors have occurred for the same I/O adapter in 24 hours, exchange the failing items in the Failing Items list one at a time.	Any device on I/O signal cables
8151	IOA microcode error	Update adapter microcode. See "Updating the controller microcode" on page 63. If the problem is not resolved, exchange the failing items in the Failing Items list one at a time.	I/O adapter
8157	IOA error requiring IOA reset to recover	If two 8157 messages have occurred for the same I/O adapter location, exchange the failing items in the Failing Items list one at a time.	I/O adapter
9000	IOA reserved area data check	Perform "MAP 3390" on page 104.	
9001	IOA reserved area invalid data pattern	Perform "MAP 3337" on page 95.	
9002	IOA reserved area LRC error	Perform "MAP 3390" on page 104.	
9008	IOA does not support functions expected by devices	Perform "MAP 3330" on page 92.	
9020	Array missing two or more devices with only one device present	Perform "MAP 3311" on page 89.	

Table 15. URCs (continued).

Note: This table includes only URCs that are associated with Maintenance Analysis Procedures contained in this information set.

URC	Description text	Service action	Failing items
9021	Array missing two or more devices with two or more devices present	Perform "MAP 3311" on page 89.	
9022	Exposed array is missing a required device	Perform "MAP 3311" on page 89.	
9023	Array members not at required physical locations	Perform "MAP 3312" on page 90.	
9024	Array not functional due to present hardware configuration	Perform "MAP 3390" on page 104.	
9025	Disk unit is not supported at its physical location	Perform "MAP 3310" on page 88.	
9026	Array not functional due to present hardware configuration	Perform "MAP 3390" on page 104.	
9027	Array is missing a device and parity is out of sync	Perform "MAP 3313" on page 91.	
9028	Maximum number of arrays already exist	Perform "MAP 3390" on page 104.	
9029	Incorrect hardware configuration change has been detected	Perform "MAP 3390" on page 104.	
9030	Array no longer protected due to missing or failed disk unit	Perform "MAP 3310" on page 88.	
9031	Array protection temporarily suspended, protection resuming	Perform "MAP 3310" on page 88.	
9032	Array exposed but still protected	Perform "MAP 3310" on page 88.	
9040	Array protection temporarily suspended, protection resuming	No action required. The array is synching. Wait until the synch is complete.	
9041	Array protection temporarily suspended	Background array parity checking detected and errors corrected. Perform "MAP 3390" on page 104.	
9042	Corrupt array parity detected on specified device	Perform "MAP 3390" on page 104.	
9054	IOA resources not available due to previous problems	Perform "MAP 3321" on page 92.	
9060	One or more disk pairs are missing from an array	Perform "MAP 3311" on page 89.	
9061	One or more disks are missing from an array	Perform "MAP 3311" on page 89.	

Table 15. URCs (continued).

Note: This table includes only URCs that are associated with Maintenance Analysis Procedures contained in this information set.

URC	Description text	Service action	Failing items
9062	One or more disks are missing from an array	Perform "MAP 3311" on page 89.	
9063	Maximum number of functional arrays has been exceeded	Perform "MAP 3390" on page 104.	
9075	Incomplete multipath connection between IOA and remote IOA	Perform "MAP 3349" on page 99	
9081	IOA detected device error	Exchange the failing items in the Failing Items list one at a time.	Disk drive, I/O adapter
9082	IOA detected device error	Exchange the failing items in the Failing Items list one at a time.	Disk drive I/O adapter signal cables backplane any device on I/O signal cables
9090	Disk unit has been modified after the last known status	Perform "MAP 3333" on page 93.	
9091	Incorrect hardware configuration change has been detected	Perform "MAP 3333" on page 93.	
9092	Disk unit requires initialization before use	Perform "MAP 3334" on page 93.	
FF3D	Soft IOA error recovered by the IOA Soft PCI bus error recovered by the IOA	If 10 FF3D messages have occurred for the same I/O Adapter physical location within a week, then exchange the failing items in the Failing Items list one at a time.	I/O adapter
FFF3	Disk media format bad	Perform "MAP 3335" on page 95.	
FFF4	Disk device problem Command to logical unit failed Device microcode is corrupt Data transfer overlength error Data transfer underlength error	Exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform "MAP 3351" on page 102.	Disk drive

Table 15. URCs (continued).

Note: This table includes only URCs that are associated with Maintenance Analysis Procedures contained in this information set.

URC	Description text	Service action	Failing items
FFF6	Failure prediction threshold exceeded	Exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform "MAP 3351" on page 102.	Disk drive
	Device hardware error recovered by the device Device hardware error recovered by the IOA	No action required.	
FFF7	Media error recovered by IOA rewrite procedures	No action required.	
	Media error recovered by device rewrite procedures		
FFF9	Soft media error. Sector reassignment recommended	No action required.	
	Device sector reassign successful		
FFFA	Undefined device response recovered by the IOA	If 10 FFFA messages have occurred for the same disk drive location in a one-week time period, then exchange the failing items in the Failing Items list one at a time. If the problem is not resolved, perform "MAP 3351" on page 102.	Disk drive
FFFB	SCSI bus was reset	No action required.	
FFFE	Soft device bus error recovered by the IOA	If 10 statistical FFFE messages have occurred for the same disk drive location in a one-week time period, perform "MAP 3350" on page 99.	

Maintenance analysis procedures

Use these procedures to resolve adapter, cache, or disk array problems associated with your controller.

For help in determining which MAP to use, see "Unit reference code tables" on page 83.

MAP 3310

Use this MAP to resolve the following problems:

- Disk unit is not supported at its physical location (URC 9025)
- Array no longer protected due to missing or failed disk unit (URC - 9030)
- Array protection temporarily suspended (URC - 9031)
- Disk Array is degraded due to missing or failed disk (URC - 9032)

Step 3310–1

Identify the disk array by examining the error log and the Display Disk Array Status screen. For information regarding accessing the Display Disk Array Status screen, see “Viewing array status” on page 33.

Go to “Step 3310–2”

Step 3310–2

Does a disk array have a state of *Degraded*?

NO Go to “Step 3310–3.”

YES Go to “Step 3310–4.”

Step 3310–3

The affected disk array should have a status of either *Rebuilt* or *Active* due to the use of a *Hot Spare* disk. Create a new *Hot Spare* disk for the disk array by doing the following:

1. Identify the failed disk using the **Display Hardware Status** screen. For further details, see “Viewing array status” on page 33. The failed disk should have a status of *Failed*.
2. Remove the failed disk and replace it with a new disk that will be used as the *Hot Spare*. For further details, see “Physical disks” on page 72.
3. The new disk must be formatted for *Advanced Function* in order to be used as a hot spare. If the new disk needs to be formatted, see “RAID and JBOD formats” on page 35.
4. Assign the new disk as a hot spare for the disk. For further details, see “Hot spare disks” on page 40.

When the problem is resolved, go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

Step 3310–4

The *Failed* disk should be replaced by using the following procedure:

1. Identify the failed disk by using the Display Hardware Status screen. For further details, see “Viewing array status” on page 33. The failed disk should have a status of *Failed*.
2. Remove the failed disk, replace it with a new disk, and rebuild the disk unit data. For further details, see “Physical disks” on page 72.

Note: The replacement disk should have a capacity that is greater than or equal to that of the smallest capacity disk in the *Degraded* disk array.

When the problem is resolved, go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3311

Use this MAP to resolve the following problems:

- Array not functional due to present hardware configuration (URC 9020 / 9021 / 9022)
- One or more disk pairs are missing from an array (URC 9060)
- One or more disks are missing from an array (URC 9061 / 9062)

Step 3311–1

Identify the disks missing from the disk array by examining the error log and the **Display Hardware Status** screen. For further details, see “Viewing array status” on page 33.

Go to “Step 3311–2.”

Step 3311–2

There are three possible ways to correct the problem. Perform only one of the following three options, listed in the order of preference:

- Locate the missing disks and install them in the correct physical locations in the system. Then, IPL the system or logical partition (it may be necessary to boot into Rescue mode), or reset the adapter by performing the following:
 1. Find the scsi host number associated with the adapter by using the **Display Hardware Status** screen. For more details see Viewing the Status of Disks and Disk Arrays.
 2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.
- Delete the disk array. For further details, see “Deleting an IBM SAS RAID disk array” on page 37.
Attention: All data on the disk array will be lost.
- Format the remaining members of the disk array, as follows:
Attention: All data on the disk array will be lost.
 1. the **iprconfig** utility by typing `iprconfig`.
 2. Select **Work with disk unit recovery**.
 3. Select **Initialize and format disk unit**.
 4. Select the drives you wish to format then press **Enter**.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3312

Use this MAP to resolve the following problem:

- Array members not at required resource address (URC 9023)

Step 3312–1

Identify the disks that are not at their required physical locations by using the error log entry and the **Display Hardware Status** screen. For further details, see “Viewing device status” on page 31.

Go to “Step 3312–2.”

Step 3312–2

There are three possible ways to correct the problem. Perform only one of the following three options, listed in the order of preference:

- Locate the missing disks and install them in the correct physical locations in the system. Then, IPL the system or logical partition (it may be necessary to boot into *Rescue* mode), or reset the adapter by performing the following:
 1. Find the scsi host number associated with the adapter by using the **Display Hardware Status** screen. For more details see Viewing the Status of Disks and Disk Arrays.

- 2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.
- Delete the disk array. For further details, see “Deleting an IBM SAS RAID disk array” on page 37.
Attention: All data on the disk array will be lost.
- Format the remaining members of the disk array, as follows:
Attention: All data on the disk array will be lost.
 1. Run the **iprconfig** utility by typing `iprconfig`.
 2. Select **Work with disk unit recovery**.
 3. Select **Initialize and format disk unit**.
 4. Select the drives you wish to format then press **Enter**.

When the problem is resolved, go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3313

Use this MAP to resolve the following problem:

- Array not functional due to present hardware configuration (URC 9027)

Step 3313–1

Identify the adapter and disks related to the error by examining the error log and the **Display Hardware Status** screen. For further details, see “Viewing device status” on page 31.

Go to “Step 3313–2.”

Step 3313–2

Have the disks or adapter card been physically moved recently?

NO Contact your service support organization.

YES Go to “Step 3313–3.”

Step 3313–3

There are three possible ways to correct the problem. Perform only *one* of the following three options, listed in the order of preference:

- Restore the adapter and disks back to their original configuration. Then, IPL the system or logical partition (it may be necessary to boot into *Rescue* mode), or reset the adapter by performing the following:
 1. Find the scsi host number associated with the adapter by using the **Display Hardware Status** screen. For more details see “Viewing device status” on page 31.
 2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.
- Delete the disk array. For further details, see “Deleting an IBM SAS RAID disk array” on page 37.
Attention: All data on the disk array will be lost.
- Format the remaining members of the disk array, as follows:
Attention: All data on the disk array will be lost.
 1. Select **Work with disk unit recovery**.
 2. Select **Initialize and format disk unit**.

3. Select the drives you wish to format then press **Enter**.
4. Run the **iprconfig** utility by typing `iprconfig`.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3321

Use this MAP to resolve the following problems:

- IOA resources not available due to previous problems (URC 9054)

Step 3321–1

Perform the following:

1. Remove any new or replacement disks which have been attached to the adapter.
2. Take action on the other errors which have occurred at the same time as this error.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3330

Use this MAP to resolve the following problem:

- IOA does not support functions expected by devices (URC 9008)

Step 3330–1

Identify the affected disks by examining the error log and the **Display Hardware Status** screen. For further details, see “Viewing device status” on page 31.

Go to “Step 3330–2.”

Step 3330–2

Have the adapter card or disks been physically moved recently?

NO Contact your service support organization.

YES Go to “Step 3330–3.”

Step 3330–3

There are two possible ways to correct the problem. Perform only *one* of the following two options, listed in the order of preference:

- Restore the adapter and disks back to their original configuration; then, either IPL the system or logical partition (it may be necessary to boot into *Rescue* mode) or reset the adapter by performing the following:
 1. Find the SCSI host number associated with the adapter by using the **Display Hardware Status** screen. For more details see Viewing the Status of Disks and Disk Arrays.
 2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.
- Format the remaining members of the disk array, as follows:

Attention: Attention: All data on the disk array will be lost.

1. Run the **iprconfig** utility by typing `iprconfig`.
2. Select **Work with disk unit recovery**.
3. Select **Initialize and format disk unit**.
4. Select the drives you wish to format then press **Enter**.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3333

Use this MAP to resolve the following problem:

- Disk unit has been modified after the last known status (URC 9090) or Incorrect hardware configuration change has been detected (URC 9091)

Step 3333–1

IPL the system or logical partition (it may be necessary to boot into *Rescue* mode), or reset the adapter by performing the following:

1. Find the scsi host number associated with the adapter by using the **Display Hardware Status** screen. For more details see Viewing the Status of Disks and Disk Arrays.
2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.

Take action on any new errors which surface.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3334

Use this MAP to resolve the following problem:

- Disk unit requires initialization before use (URC 9092)

The possible causes are:

- Disk is a previously failed disk from a disk array and was automatically replaced by a *Hot Spare* disk.
- Disk is a previously failed disk from a disk array and was removed and later reinstalled on a different adapter or different location on this adapter.
- Appropriate service procedures were not followed when replacing disks or reconfiguring the adapter, such as not using the **Device Concurrent Maintenance** screen in **iprconfig** when concurrently removing and installing disks (see “Physical disks” on page 72) or not performing a normal power down of the system prior to reconfiguring disks and adapters.
- Disk is member of a disk array, but was detected subsequent to the adapter being configured.
- Disk has multiple or complex configuration problems.

Step 3334–1

Identify the affected disks by examining the error log and the **Display Hardware Status** screen. For further details, see “Viewing device status” on page 31.

Go to “Step 3334–2” on page 94.

Step 3334–2

Are there other disk or adapter errors that have occurred at about the same time as this error?

NO Go to “Step 3334–4.”

YES Go to “Step 3334–3.”

Step 3334–3

Take action on the other errors that have occurred at the same time as this error.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

Step 3334–4

Have the disks or adapter card been physically moved recently?

NO Go to “Step 3334–5.”

YES Go to “Step 3334–6.”

Step 3334–5

Is the data on the disks needed for this or any other system?

NO Go to “Step 3334–7” on page 95.

YES Go to “Step 3334–6.”

Step 3334–6

There are three possible ways to correct this problem. Perform only *one* of the following three options:

- IPL the system or logical partition (it may be necessary to boot into *Rescue* mode), or reset the adapter by performing the following:
 1. Find the SCSI host number associated with the adapter by using the **Display Hardware Status** screen. For more details see Viewing the Status of Disks and Disk Arrays.
 2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.
- Restore the adapter and disks back to their original configuration. Then, IPL the system or logical partition (it may be necessary to boot into *Rescue* mode), or reset the adapter by performing the following:
 1. Find the scsi host number associated with the adapter by using the Display Hardware Status screen. For more details see “Viewing device status” on page 31.
 2. Use the `echo 1 > /sys/class/scsi_host/hostX/reset_host` command to reset the adapter, where X is the scsi host number from the previous step.
- Remove the disks from this adapter

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

Step 3334–7

There are two possible ways to correct the problem. Perform only *one* of these options.

- Format the disks, as follows:
 - Attention:** All data on the disk array will be lost.
 - 1. Run the **iprconfig** utility by typing `iprconfig`.
 - 2. Select **Work with disk unit recovery**.
 - 3. Select **Initialize and format disk unit**.
 - 4. Select the drives you wish to format then press **Enter**.
- If the disks are members of a disk array, delete the disk array. See “Deleting an IBM SAS RAID disk array” on page 37.

Note: In some rare scenarios, deleting the disk array will not have no effect on a disk and the disk must be formatted instead.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3335

Use this MAP to resolve the following problem:

- Disk media format bad (URC FFF3)

The possible causes are:

- Disk was being formatted and was powered off during this process.
- Disk was being formatted and was reset during this process.

Step 3335–1

Identify the affected disk by examining the error log and the **Display Hardware Status** screen. For further details, see “Viewing device status” on page 31.

Go to “Step 3335–2.”

Step 3335–2

Format the disks, as follows:

- Attention:** All data on the disks will be lost.
- 1. Run the **iprconfig** utility by typing `iprconfig`.
- 2. Select **Work with disk unit recovery**.
- 3. Select **Initialize and format disk unit**.
- 4. Select the drives you wish to format then press **Enter**.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3337

Use this MAP to resolve the following problem:

- IOA detected device error (URC 9001)

Step 3337-1

A device configuration error has been detected. The configuration sectors on the device may be incompatible with the current I/O adapter.

Has the I/O adapter been replaced with a different type of I/O adapter, or have the devices been moved from a different type of I/O adapter to this one recently?

NO Go to "Step 3337-2."

YES Go to "Step 3337-3."

Step 3337-2

Contact your service support organization.

Exit this procedure.

Step 3337-3

Identify the affected disks by examining the error log and the **Display Hardware Status** screen. For further details, see "Viewing array status" on page 33.

Go to "Step 3337-4."

Step 3337-4

If the I/O Adapter has been replaced by a different type of I/O Adapter, reinstall the original adapter.

Go to "Step 3337-5."

Step 3337-5

If the disks involved have been moved from a different type of I/O Adapter to this one, return them to their original I/O Adapter.

Go to "Step 3337-6."

Step 3337-6

Attention: The next step will cause data loss on the disks involved. If the data on these disks is important, perform back up procedures now.

Delete the disk arrays. For further details, see "Deleting an IBM SAS RAID disk array" on page 37.

Go to "Step 3337-7."

Step 3337-7

Restore the initial hardware configuration that caused the error.

Go to "Step 3337-8" on page 97.

Step 3337–8

Create new disk arrays, then restore the saved data. For further details, see “Creating an IBM SAS RAID disk array” on page 36.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 eServer pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3342

Use this MAP to resolve the following problem:

- Configuration error; incorrect connection between cascaded enclosure

Step 3342–1

Incorrect SAS fabric connections. Contact your service provider.

MAP 3343

Use this MAP to resolve the following problem:

- Configuration error; connections exceed controller design limits

Step 3343–1

Incorrect SAS fabric connections. Contact your service provider.

MAP 3344

Use this MAP to resolve the following problems:

- Configuration error, incorrect multipath connection
- Configuration error, incorrect multipath connection between controller and enclosure detected

The possible causes are:

- Incorrect cabling to device enclosure.
- A failed connection caused by a failing component in the SAS fabric between, and including, the adapter and device enclosure.

Considerations:

- Remove power from the system before connecting and disconnecting cables or devices, as appropriate, to prevent hardware damage or erroneous diagnostic results.
- Some systems have the disk enclosure or removable media enclosure integrated in the system with no cables. For these configurations the SAS connections are integrated onto the system boards and a failed connection can be the result of a failed system board or integrated device enclosure.

Attention: Replacing RAID adapters is not recommended without assistance from your hardware service support organization when SAS fabric problems exist. The adapter might contain nonvolatile write cache data and configuration data for the attached disk arrays. Additional problems can be created by replacing an adapter when SAS fabric problems exist.

Step 3344–1

Incorrect SAS fabric connections. Contact your service provider.

MAP 3345

Use this MAP to resolve the following problem:

- Unsupported enclosure function detected

The possible causes are:

- Device enclosure or adapter microcode levels are not up to date
- Unsupported type of device enclosure or device

Step 3345–1

Ensure device enclosure or adapter microcode levels are up to date. If this does not correct the problem, contact your service provider.

MAP 3346

Use this MAP to resolve the following problem:

- Configuration error; incomplete multipath connection between enclosures and device detected

The possible causes are:

- A failed connection caused by a failing component within the device enclosure, including the device itself.

Note: The adapter is not a likely cause of this problem.

Considerations:

- Remove power from the system before connecting and disconnecting cables or devices, as appropriate, to prevent hardware damage or erroneous diagnostic results.
- Some systems have the disk enclosure or removable media enclosure integrated in the system with no cables. For these configurations the SAS connections are integrated onto the system boards and a failed connection can be the result of a failed system board or integrated device enclosure.

Attention: Removing functioning disks in a disk array is not recommended without assistance from your Hardware Service Support organization. A disk array may become degraded or failed if functioning disks are removed and additional problems may be created.

Step 3346–1

Ensure device enclosure and SAS disks are properly plugged in their slots. If this does not correct the problem, contact your service provider.

MAP 3348

Use this MAP to resolve the following problem:

- Attached enclosure does not support required multipath function

Step 3348–1

Contact your service provider.

MAP 3349

Use this map to resolve the following problem:

- Incomplete multipath connection between controller and remote controller

Step 3349–1

Contact your service provider.

MAP 3350

Use the following to perform SAS fabric problem isolation.

Considerations:

- Remove power from the system before connecting and disconnecting cables or devices, as appropriate, to prevent hardware damage or erroneous diagnostic results.
- Some systems have SAS and PCI-X bus interface logic integrated onto the system boards and use a pluggable RAID Enablement Card (a non-PCI form factor card) for these SAS/PCI-X buses. For these configurations, replacement of the RAID Enablement Card is unlikely to solve a SAS fabric-related problem because the SAS interface logic is on the system board.

Attention: Replacing RAID adapters is not recommended without assistance from your service support organization when SAS fabric problems exist. Because the adapter may contain non-volatile write cache data and configuration data for the attached disk arrays, additional problems can be created by replacing an adapter when SAS fabric problems exist.

Attention: Removing functioning disks in a disk array is not recommended without assistance from your service support organization. A disk array may become *Degraded* or *Failed* if functioning disks are removed. Also, additional problems may be created.

Identify the SAS fabric on which the problem is occurring by examining the error log entry.

Go to “Step 3350–2.”

Step 3350–2

Have changes been made to the SAS configuration recently?

NO Go to “Step 3350–5” on page 100.

YES Go to “Step 3350–3.”

Step 3350–3

Check for the following problems:

- Cabling problems such as configurations that exceed the maximum cable lengths
- Ensure the SAS configuration does not have multi-initiators (for example, set up for a high-availability configuration)

Note: Multi-initiator support is not provided at this time.

For more details about supported SAS cabling, refer to *RS/6000 pSeries Adapters, Devices, and Cable Information for Multiple Bus Systems*.

Did you find a problem?

NO Go to “Step 3350–5.”

YES Go to “Step 3350–4.”

Step 3350–4

1. Power off the system or logical partition.
2. Correct the problem.
3. Power on the system or logical partition. If you cannot power on normally, boot to *Rescue* mode. Examine the error log.

Did a SAS fabric-related failure occur?

NO Go to “Step 3350–14” on page 101.

YES Go to “Step 3350–5.”

Step 3350–5

Determine if any of the disk arrays on the adapter are in an *Degraded* state. For further details, see “Viewing array status” on page 33.

Does any disk array have a state of *Degraded* ?

NO Go to “Step 3350–7.”

YES Go to “Step 3350–6.”

Step 3350–6

1. Identify the failed disks by first finding disk arrays with a state of *Degraded* , then finding disks on those arrays with a state of *Failed*.
2. Remove the failed disks from each *Degraded* disk array. For further details, see “Physical disks” on page 72.
3. Reboot the system or logical partition. If you cannot power on normally, boot to *Rescue* mode. Examine the error log.

Did a SAS fabric related failure occur?

NO Go to “Step 3350–14” on page 101.

YES Go to “Step 3350–7.”

Step 3350–7

Are there any non-essential removable media devices (such as Tape, CDROM, and DVDROM) on the SAS fabric?

NO Go to “Step 3350–10” on page 101.

YES Go to “Step 3350–8” on page 101.

Step 3350–8

1. Power off the system or logical partition.
2. Remove one of the non-essential removable media devices.
3. Power on the system or logical partition. If you cannot power on normally, boot to *Rescue* mode. Examine the error log.

Did a SAS fabric related failure occur?

NO Go to “Step 3350–9.”

YES Go to “Step 3350–7” on page 100.

Step 3350–9

The last removable media device removed from the SAS fabric may be the cause of the SAS fabric problems. Follow the repair procedures for that device.

Go to “Step 3350–14.”

Step 3350–10

Are there any non-essential disks that are not disk array members (such as 512 byte/sector *JBOD* disks, *Hot Spare* disks, or *Advanced Function* disks) on the SAS fabric?

NO Go to “Step 3350–13.”

YES Go to “Step 3350–11.”

Step 3350–11

1. Remove one of the non-essential disk devices. For further details, see “Physical disks” on page 72.
2. Reboot the system or logical partition. If you cannot power on normally, boot to *Rescue* mode. Examine the error log.

Did a SAS fabric related failure occur?

NO Go to “Step 3350–12.”

YES Go to “Step 3350–10.”

Step 3350–12

The last disk removed from the SAS fabric may be the cause of the SAS fabric problems. Follow the repair procedures for that device.

Go to “Step 3350–14.”

Step 3350–13

Contact your service support organization.

Exit this procedure

Step 3350–14

1. Reinstall any good devices that were removed during the course of this MAP.

2. Reboot the system or logical partition. If you cannot power on normally, boot to *Rescue* mode. Examine the error log.
3. Take action on any other non-SAS fabric related errors if they exist.

MAP 3351

Use the following to determine what other Field Replaceable Units (FRUs) besides a disk, may need to be replaced in order to solve a problem.

You are sent here when an *Advanced Function* disk, a physical disk in 528 bytes/sector format, was identified as the primary FRU to replace in order to resolve a problem. However, if replacing the disk did not resolve the problem then other FRUs may need to be replaced.

Considerations:

- Remove power from the system before connecting and disconnecting cables or devices, as appropriate, to prevent hardware damage or erroneous diagnostic results.
- Some systems have SAS and PCI-X bus interface logic integrated onto the system boards and use a pluggable RAID Enablement Card (a non-PCI form factor card) for these SAS/PCI-X busses. For these configurations, replacement of the RAID Enablement Card is unlikely to solve a SAS fabric related problem since the SAS interface logic is on the system board.

Attention: Replacing RAID adapters is not recommended without assistance from your service support organization when SAS fabric problems exist. Because the adapter may contain non-volatile write cache data and configuration data for the attached disk arrays, additional problems can be created by replacing a adapter when SAS fabric problems exist.

Attention: Removing functioning disks in a disk array is not recommended without assistance from your service support organization. A disk array may become *Degraded* or *Failed* if functioning disks are removed. Also, additional problems may be created.

Step 3351-1

Identify the SAS fabric on which the problem is occurring by examining the error log entry.

Go to "Step 3351-2."

Step 3351-2

While the error persists, replace the components of the failing SAS fabric in the following order.

1. Cable (if present)
2. Adapter (if SAS interface logic is on the adapter) or system board (if SAS interface logic is on the system board)
3. DASD backplane (if present)

To replace a component and see if the problem was corrected, do the following:

1. Power off the system or logical partition
2. Replace a component listed above
3. Power on the system or logical partition.
4. If you cannot power on normally, boot to *Rescue* mode. Examine the error log.

When the problem is resolved then go to MAP 0410: Repair Checkout, in *RS/6000 pSeries Diagnostic Information for Multiple Bus Systems*.

MAP 3352

Use this MAP to resolve the following problems:

- Device bus fabric error
- Temporary device bus fabric error

The possible causes are:

- A failed connection caused by a failing component in the SAS fabric between, and including, the adapter and device enclosure.
- A failed connection caused by a failing component within the device enclosure, including the device itself.

Considerations:

- Remove power from the system before connecting and disconnecting cables or devices, as appropriate, to prevent hardware damage or erroneous diagnostic results.
- Some systems have SAS and PCI-X/PCIe bus interface logic integrated onto the system boards and use a pluggable RAID Enablement Card (a non-PCI form factor card) for these SAS/PCI-X/PCIe buses. For these configurations, replacement of the RAID Enablement Card is unlikely to solve a SAS related problem because the SAS interface logic is on the system board.
- Some systems have the disk enclosure or removable media enclosure integrated in the system with no cables. For these configurations the SAS connections are integrated onto the system boards and a failed connection can be the result of a failed system board or integrated device enclosure.

Attention: Replacing RAID adapters is not recommended without assistance from your service provider when SAS fabric problems exist. Because the adapter may contain non-volatile write cache data and configuration data for the attached disk arrays, additional problems can be created by replacing a adapter when SAS fabric problems exist.

Attention: Removing functioning disks in a disk array is not recommended without assistance from your service provider. A disk array may become *Degraded* or *Failed* if functioning disks are removed and additional problems may be created.

Step 3352–1

Contact your service provider.

MAP 3353

Use this map to resolve the following problem:

- Multipath redundancy level got worse

The possible causes are:

- A failed connection caused by a failing component in the SAS fabric between, and including, the adapter and device enclosure.
- A failed connection caused by a failing component within the device enclosure, including the device itself.

Note: The failed connection was previously working, and may have already recovered.

Considerations:

- Remove power from the system before connecting and disconnecting cables or devices, as appropriate, to prevent hardware damage or erroneous diagnostic results.

- Some systems have SAS and PCI-X/PCIe bus interface logic integrated onto the system boards and use a pluggable RAID Enablement Card (a non-PCI form factor card) for these SAS/PCI-X/PCIe buses. For these configurations, replacement of the RAID Enablement Card is unlikely to solve a SAS related problem because the SAS interface logic is on the system board.
- Some systems have the disk enclosure or removable media enclosure integrated in the system with no cables. For these configurations the SAS connections are integrated onto the system boards and a failed connection can be the result of a failed system board or integrated device enclosure.

Attention: Replacing RAID adapters is not recommended without assistance from your service provider when SAS fabric problems exist. Because the adapter may contain non-volatile write cache data and configuration data for the attached disk arrays, additional problems can be created by replacing a adapter when SAS fabric problems exist.

Attention: Removing functioning disks in a disk array is not recommended without assistance from your service provider. A disk array may become *Degraded* or *Failed* if functioning disks are removed and additional problems may be created.

Step 3353–1

Contact your service provider.

MAP 3390

The problem that occurred is uncommon or complex to resolve. Information should be gathered and assistance obtained from your service support organization.

Step 3390–1

Make a copy of the entire `/var/log/messages` and `/var/log/boot.msg` files.

Go to “Step 3390–2.”

Step 3390–2

Collect the current disk array configuration. For further details, see “Viewing device status” on page 31.

Go to “Step 3390–3.”

Step 3390–3

Collect any `ipr` dump files that may be applicable to the problem. They will be located in `/var/log/`.

Go to “Step 3390–4”

Step 3390–4

Contact your service support organization.

Exit this procedure.

Notices

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

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

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

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: THIS INFORMATION IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

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

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

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning products not produced by this manufacturer was obtained from the suppliers of those products, their published announcements or other publicly available sources. This manufacturer has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to products not produced by this manufacturer. Questions on the capabilities of products not produced by this manufacturer should be addressed to the suppliers of those products.

All statements regarding the manufacturer's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

The manufacturer's prices shown are the manufacturer's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

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

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

The drawings and specifications contained herein shall not be reproduced in whole or in part without the written permission of the manufacturer.

The manufacturer has prepared this information for use with the specific machines indicated. The manufacturer makes no representations that it is suitable for any other purpose.

The manufacturer's computer systems contain mechanisms designed to reduce the possibility of undetected data corruption or loss. This risk, however, cannot be eliminated. Users who experience unplanned outages, system failures, power fluctuations or outages, or component failures must verify the accuracy of operations performed and data saved or transmitted by the system at or near the time of the outage or failure. In addition, users must establish procedures to ensure that there is independent data verification before relying on such data in sensitive or critical operations. Users should periodically check the manufacturer's support websites for updated information and fixes applicable to the system and related software.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at Copyright and trademark information at www.ibm.com/legal/copytrade.shtml.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

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

Electronic emission notices

Class A Notices

The following Class A statements apply to the IBM servers that contain the POWER7 processor and its features unless designated as electromagnetic compatibility (EMC) Class B in the feature information.

Federal Communications Commission (FCC) statement

Note: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this

equipment in a residential area is likely to cause harmful interference, in which case the user will be required to correct the interference at his own expense.

Properly shielded and grounded cables and connectors must be used in order to meet FCC emission limits. IBM is not responsible for any radio or television interference caused by using other than recommended cables and connectors or by unauthorized changes or modifications to this equipment. Unauthorized changes or modifications could void the user's authority to operate the equipment.

This device complies with Part 15 of the FCC rules. Operation is subject to the following two conditions: (1) this device may not cause harmful interference, and (2) this device must accept any interference received, including interference that may cause undesired operation.

Industry Canada Compliance Statement

This Class A digital apparatus complies with Canadian ICES-003.

Avis de conformité à la réglementation d'Industrie Canada

Cet appareil numérique de la classe A est conforme à la norme NMB-003 du Canada.

European Community Compliance Statement

This product is in conformity with the protection requirements of EU Council Directive 2004/108/EC on the approximation of the laws of the Member States relating to electromagnetic compatibility. IBM cannot accept responsibility for any failure to satisfy the protection requirements resulting from a non-recommended modification of the product, including the fitting of non-IBM option cards.

This product has been tested and found to comply with the limits for Class A Information Technology Equipment according to European Standard EN 55022. The limits for Class A equipment were derived for commercial and industrial environments to provide reasonable protection against interference with licensed communication equipment.

European Community contact:
IBM Deutschland GmbH
Technical Regulations, Department M456
IBM-Allee 1, 71139 Ehningen, Germany
Tele: +49 7032 15-2937
e-mail: tjahn@de.ibm.com

Warning: This is a Class A product. In a domestic environment, this product may cause radio interference, in which case the user may be required to take adequate measures.

VCCI Statement - Japan

この装置は、クラスA 情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

VCCI-A

The following is a summary of the VCCI Japanese statement in the box above:

This is a Class A product based on the standard of the VCCI Council. If this equipment is used in a domestic environment, radio interference may occur, in which case, the user may be required to take corrective actions.

**Japanese Electronics and Information Technology Industries Association (JEITA)
Confirmed Harmonics Guideline (products less than or equal to 20 A per phase)**

高調波ガイドライン適合品

**Japanese Electronics and Information Technology Industries Association (JEITA)
Confirmed Harmonics Guideline with Modifications (products greater than 20 A per phase)**

高調波ガイドライン準用品

Electromagnetic Interference (EMI) Statement - People's Republic of China

声 明

此为 A 级产品,在生活环境中,
该产品可能会造成无线电干扰。
在这种情况下,可能需要用户对其
干扰采取切实可行的措施。

Declaration: This is a Class A product. In a domestic environment this product may cause radio interference in which case the user may need to perform practical action.

Electromagnetic Interference (EMI) Statement - Taiwan

警告使用者：

這是甲類的資訊產品，在
居住的環境中使用時，可
能會造成射頻干擾，在這
種情況下，使用者會被要
求採取某些適當的對策。

The following is a summary of the EMI Taiwan statement above.

Warning: This is a Class A product. In a domestic environment this product may cause radio interference in which case the user will be required to take adequate measures.

IBM Taiwan Contact Information:

台灣IBM 產品服務聯絡方式：
台灣國際商業機器股份有限公司
台北市松仁路7號3樓
電話：0800-016-888

Electromagnetic Interference (EMI) Statement - Korea

이 기기는 업무용으로 전자파적합등록을 한 기기이오니 판매자 또는 사용자는 이점을 주의하시기 바라며, 만약 잘못 판매 또는 구입하였을 때에는 가정용으로 교환하시기 바랍니다.

Please note that this equipment has obtained EMC registration for commercial use. In the event that it has been mistakenly sold or purchased, please exchange it for equipment certified for home use.

Germany Compliance Statement

Deutschsprachiger EU Hinweis: Hinweis für Geräte der Klasse A EU-Richtlinie zur Elektromagnetischen Verträglichkeit

Dieses Produkt entspricht den Schutzanforderungen der EU-Richtlinie 2004/108/EG zur Angleichung der Rechtsvorschriften über die elektromagnetische Verträglichkeit in den EU-Mitgliedsstaaten und hält die Grenzwerte der EN 55022 Klasse A ein.

Um dieses sicherzustellen, sind die Geräte wie in den Handbüchern beschrieben zu installieren und zu betreiben. Des Weiteren dürfen auch nur von der IBM empfohlene Kabel angeschlossen werden. IBM übernimmt keine Verantwortung für die Einhaltung der Schutzanforderungen, wenn das Produkt ohne Zustimmung von IBM verändert bzw. wenn Erweiterungskomponenten von Fremdherstellern ohne Empfehlung von IBM gesteckt/eingebaut werden.

EN 55022 Klasse A Geräte müssen mit folgendem Warnhinweis versehen werden:

"Warnung: Dieses ist eine Einrichtung der Klasse A. Diese Einrichtung kann im Wohnbereich Funk-Störungen verursachen; in diesem Fall kann vom Betreiber verlangt werden, angemessene Maßnahmen zu ergreifen und dafür aufzukommen."

Deutschland: Einhaltung des Gesetzes über die elektromagnetische Verträglichkeit von Geräten

Dieses Produkt entspricht dem "Gesetz über die elektromagnetische Verträglichkeit von Geräten (EMVG)". Dies ist die Umsetzung der EU-Richtlinie 2004/108/EG in der Bundesrepublik Deutschland.

Zulassungsbescheinigung laut dem Deutschen Gesetz über die elektromagnetische Verträglichkeit von Geräten (EMVG) (bzw. der EMC EG Richtlinie 2004/108/EG) für Geräte der Klasse A

Dieses Gerät ist berechtigt, in Übereinstimmung mit dem Deutschen EMVG das EG-Konformitätszeichen - CE - zu führen.

Verantwortlich für die Einhaltung der EMV Vorschriften ist der Hersteller:
International Business Machines Corp.

New Orchard Road
Armonk, New York 10504
Tel: 914-499-1900

Der verantwortliche Ansprechpartner des Herstellers in der EU ist:
IBM Deutschland GmbH
Technical Regulations, Abteilung M456
IBM-Allee 1, 71139 Ehningen, Germany
Tel: +49 7032 15-2937
e-mail: tjahn@de.ibm.com

Generelle Informationen:

Das Gerät erfüllt die Schutzanforderungen nach EN 55024 und EN 55022 Klasse A.

Electromagnetic Interference (EMI) Statement - Russia

**ВНИМАНИЕ! Настоящее изделие относится к классу А.
В жилых помещениях оно может создавать
радиопомехи, для снижения которых необходимы
дополнительные меры**

Terms and conditions

Permissions for the use of these publications is granted subject to the following terms and conditions.

Personal Use: You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative works of these publications, or any portion thereof, without the express consent of the manufacturer.

Commercial Use: You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of the manufacturer.

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

The manufacturer reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by the manufacturer, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

THE MANUFACTURER MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THESE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.