MegaRAID Elite 1600

Hardware Guide

MAN-493
6/12/01

LSI Logic Corporation
6145-D Northbelt Parkway
Norcross, GA 30071

Revision History

3/17/00	Initial release.
3/16/01	Added Chapter 7 Cluster Installation and Configuration, and corrected the RAID 0 graphic in Chapter 3. Added Appendix C Cluster Configuration with a Crossover Cable.
6/12/01	Make corrections, such as cache size (16 MB is smallest option), and the number of physical disk drives supported at various RAID levels.

Table of Contents

Preface

The MegaRAID Elite 160M I2O PCI Disk Array Controller supports two Ultra and Wide SCSI channels with data transfer rates up to 160 MB/s. This manual describes the MegaRAID Elite 1600 64-bit 160M controller.

Limited Warranty The buyer agrees if this product proves to be defective, that LSI Logic Corporation is obligated only to repair or replace this product at LSI Logic's discretion according to the terms and conditions of the warranty registration card that accompanies this product. LSI Logic shall not be liable in tort or contract for any loss or damage, direct, incidental or consequential resulting from the use of this product. Please see the Warranty Registration Card shipped with this product for full warranty details.

Limitations of Liability LSI Logic Corporation shall in no event be held liable for any loss, expenses, or damages of any kind whatsoever, whether direct, indirect, incidental, or consequential (whether arising from the design or use of this product or the support materials provided with the product). No action or proceeding against LSI Logic may be commenced more than two years after the delivery of product to Licensee of Licensed Software.

Licensee agrees to defend and indemnify LSI Logic Corporation from any and all claims, suits, and liabilities (including attorney's fees) arising out of or resulting from any actual or alleged act or omission on the part of Licensee, its authorized third parties, employees, or agents, in connection with the distribution of Licensed Software to end-users, including, without limitation, claims, suits, and liability for bodily or other injuries to end-users resulting from use of Licensee's product not caused solely by faults in Licensed Software as provided by LSI Logic Corporation to Licensee.

Package Contents You should have received:

- a MegaRAID Elite 1600 64-Bit 160M Controller
- a CD with drivers, utilities, and documentation
- a *MegaRAID Elite 1600 Hardware Guide*
- a *MegaRAID Configuration Software Guide*
- a *MegaRAID Operating System Drivers Guide*
- software license agreement
- a warranty registration card

Technical Support If you need help installing, configuring, or running the MegaRAID Controller, call LSI Logic Technical Support at 678-728-1250. Before you call, please complete the ***MegaRAID Problem Report*** form on the next page.

Web Site We invite you to access the LSI Logic Corporation world wide web site at:

http://www.lsil.com.

MegaRAID Problem Report Form

Customer Information	MegaRAID Information
Name	Today's Date
Company	Date of Purchase
Address	Invoice Number
City/State	Serial Number
Country	Number of Channels
email address	Cache Memory
Phone	Firmware Version
Fax	BIOS Version
System Information	
Motherboard:	BIOS manufacturer:
Operating System:	BIOS Date:
Op. Sys. Ver.:	Video Adapter:
MegaRAID Driver Ver.:	CPU Type/Speed:
Network Card:	System Memory:
Other disk controllers installed:	Other adapter cards installed:
Description of problem:	
Steps necessary to re-create problem: 1. 2. 3. 4.	

Logical Drive Configuration

Logical Drive	RAID Level	Stripe Size	Logical Drive Size	Cache Policy	Read Policy	Write Policy	# of Physical Drives
LD0							
LD1							
LD2							
LD3							
LD4							
LD5							
LD6							
LD7							
LD8							
LD9							
LD10							
LD11							
LD12							
LD13							
LD14							
LD15							

Logical Drive	RAID Level	Stripe Size	Logical Drive Size	Cache Policy	Read Policy	Write Policy	# of Physical Drives
LD16							
LD17							
LD18							
LD19							
LD20							
LD21							
LD22							
LD23							
LD24							
LD25							
LD26							
LD27							
LD28							
LD29							
LD30							
LD31							
LD32							
LD33							
LD34							
LD35							
LD36							
LD37							
LD38							
LD39							

Physical Device Layout

	Channel A	Channel B
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		

	Channel A	Channel B
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		

	Channel A	Channel B
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		

Disclaimer

Disclaimer

This manual describes the operation of the LSI Logic MegaRAID Controller. Although efforts have been made to assure the accuracy of the information contained here, LSI Logic Corporation expressly disclaims liability for any error in this information, and for damages, whether direct, indirect, special, exemplary, consequential or otherwise, that may result from such error, including but not limited to the loss of profits resulting from the use or misuse of the manual or information contained therein (even if LSI Logic has been advised of the possibility of such damages). Any questions or comments regarding this document or its contents should be addressed to LSI Logic at the address shown on the cover.

LSI Logic provides this publication “as is” without warranty of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability or fitness for a specific purpose.

Some states do not allow disclaimer of express or implied warranties or the limitation or exclusion of liability for indirect, special, exemplary, incidental or consequential damages in certain transactions; therefore, this statement may not apply to you. Also, you may have other rights which vary from jurisdiction to jurisdiction.

This publication 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. LSI Logic may make improvements and/or revisions in the product(s) and/or the program(s) described in this publication at any time.

Requests for technical information about LSI Logic products should be made to your LSI Logic authorized reseller or marketing representative.

FCC Regulatory Statement

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.

Warning: **Changes or modifications to this unit not expressly approved by the party responsible for compliance could void the user's authority to operate the equipment.**

Note: This equipment has been tested and found to comply with the limits for a Class B digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference in a residential installation. This equipment generates, uses and can radiate radio frequency energy and, if not installed and used in accordance with the instructions, may cause harmful interference to radio communications. However, there is no guarantee that interference will not occur in a specific installation. If this equipment does cause harmful interference to radio or television reception, which can be determined by turning the equipment off and on, try to correct the interference by one or more of the following measures:

1) Reorient or relocate the receiving antenna.
2) Increase the separation between the equipment and the receiver.
3) Connect the equipment into an outlet on a circuit different from that to which the receiver is connected.
4) Consult the dealer or an experienced radio/TV technician for help.

Shielded interface cables must be used with this product to ensure compliance with the Class B FCC limits. **LSI Logic MegaRAID Elite 1600 64-Bit 160M PCI SCSI Disk Array Controller**
Model Number: Series 493
FCC ID Number: IUESER493

Disclaimer

LSI Logic certifies only that this product will work correctly when this product is used with the same jumper settings, the same system configuration, the same memory module parts, and the same peripherals that were tested by LSI Logic with this product. The complete list of tested jumper settings, system configurations, peripheral devices, and memory modules are documented in the LSI Logic Compatibility Report for this product. Call your LSI Logic sales representative for a copy of the Compatibility Report for this product.

1 Overview

The MegaRAID® Elite 1600 LVD (Low Voltage Differential SCSI) PCI RAID controller adapter card provides two SCSI channels. Using LVD, you can use cables up to 25 meters long. Throughput on each SCSI channel can be as high as 160 MB/s. MegaRAID supports both a low voltage differential SCSI bus or a single ended SCSI bus.

MegaRAID Elite 1600 64-Bit LVD is a high performance intelligent PCI-to-SCSI host adapter with RAID control capabilities. MegaRAID Elite 1600 64-Bit LVD requires no special motherboard PCI expansion slot. The MegaRAID Elite 1600 card includes an Intel i960RN processor. MegaRAID provides reliability, high performance, and fault-tolerant disk subsystem management.

SCSI Channels MegaRAID Elite 1600 has two 160M SCSI channels. The two channels are supported by one Qlogic dual SCSI controller. Each SCSI channel supports up to fifteen 160M SCSI devices.

NVRAM and Flash ROM A 32 KB x 8 NVRAM stores RAID system configuration information. The firmware is stored in flash memory for easy upgrade.

SCSI Connectors MegaRAID has two ultra high density 68-pin external SCSI connectors and two 68-pin internal SCSI connectors for internal SCSI drives.

Single Ended and Differential SCSI Buses

The SCSI standard defines two electrical buses:

- a single ended bus
- a differential bus

Maximum Cable Length for SCSI Standards

Standard	Single ended	LVD	Maximum Number of Drives
SCSI I	6 m	12 m	7
Fast SCSI	6 m	12 m	7
Fast Wide SCSI	6 m	12 m	15
Ultra SCSI	1.5 m	12 m	7
Ultra SCSI	3 m	12 m	3
Wide Ultra SCSI		12 m	15
Wide Ultra SCSI	1.5 m	12 m	7
Wide Ultra SCSI	3 m	12 m	3
Ultra2 SCSI		25 m	1
Ultra2 SCSI		12 m	7
Wide Ultra2 SCSI		25 m	1
Wide Ultra2 SCSI		12 m	15

SCSI Bus Widths and Maximum Throughput

SCSI Standard	SCSI Bus Width	SCSI Throughput
SCSI I	8 bits	5 MB/s
Fast SCSI	8 bits	10 MB/s
Fast Wide SCSI	16 bits	20 MB/s
Ultra SCSI	8 bits	20 MB/s
Wide Ultra SCSI	16 bits	40 MB/s
Ultra2 SCSI	8 bits	40 MB/s
Wide Ultra2 SCSI	16 bits	80 MB/s
160M SCSI	8 bits	80 MB/s
Wide 160M SCSI	16 bits	160 MB/s

Documentation Set

The MegaRAID Elite 1600 64-Bit LVD technical documentation set includes:

- the *MegaRAID Elite 1600 Hardware Guide*
- the *MegaRAID Configuration Software Guide*
- the *WebBIOS Guide*
- the *MegaRAID Operating System Drivers Guide*

Using MegaRAID Elite 1600 Manuals The *MegaRAID Elite 1600 Hardware Guide* includes a RAID overview, RAID planning, and RAID system configuration information. Read it first.

MegaRAID Configuration Software Guide This manual describes the MegaRAID software utilities that configure and modify RAID systems. The software utilities include:

- MegaRAID Configuration Utility
- MegaRAID Manager
- Power Console Plus

WebBIOS Guide This manual explains the operation of the WebBIOS Configuration Utility. WebBIOS allows you to configure and manager RAID systems running in remote servers.

MegaRAID Operating System Drivers Guide This manual provides detailed information about the operating system drivers.

2 Introduction to RAID

RAID (Redundant Array of Independent Disks) is an array of multiple independent hard disk drives that provide high performance and fault tolerance. A RAID disk subsystem improves I/O performance. The RAID array appears to the host computer as a single storage unit or as multiple logical units. I/O is faster because drives can be accessed simultaneously. RAID improve data storage reliability and fault tolerance. You can prevent data loss caused by drive failure by reconstructing missing data from the remaining data and parity drives.

RAID Overview

The following topics are discussed:

RAID Levels

RAID (Redundant Array of Independent Disks) is a collection of specifications that describe a system for ensuring the reliability and stability of data stored on large disk subsystems. A RAID system can be implemented in a number of different versions (or RAID Levels). The standard RAID levels are 0, 1, 3, and 5. MegaRAID supports all standard RAID levels and RAID levels 10, 30, and 50, special RAID versions supported by MegaRAID.

Consistency Check

In RAID, check consistency verifies the correctness of redundant data in an array. For example, in a system with dedicated parity, checking consistency means computing the parity of the data drives and comparing the results to the contents of the dedicated parity drive.

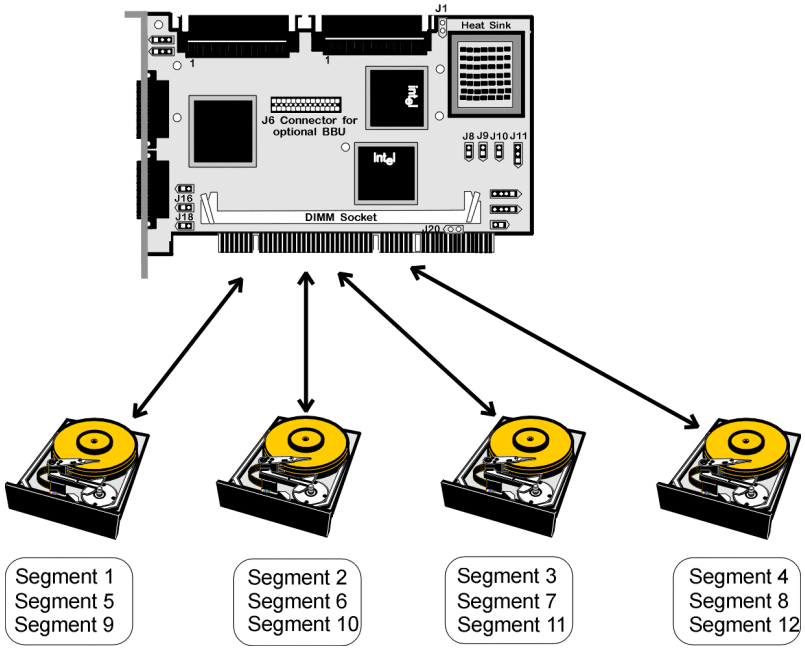
Fault Tolerance

Fault tolerance is achieved through cooling fans, power supplies, and the ability to hot swap drives. MegaRAID provides hot swapping through the hot spare feature. A hot spare drive is an unused online available drive. MegaRAID can use a hot spare to instantly rebuild a logical drive.

After the hot spare is automatically moved into the RAID subsystem, the failed drive can be automatically rebuilt. The RAID disk array continues to handle requests while the rebuild occurs.

Disk Striping

Disk striping writes data across multiple disk drives instead of just one disk drive. Disk striping involves partitioning each drive storage space into stripes that can vary in size from 2 KB to 128 KB. These stripes are interleaved in a repeated sequential manner. The combined storage space is composed of stripes from each drive. MegaRAID supports stripe sizes of 2 KB, 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, or 128 KB. For example, in a four-disk system using only disk striping (as in RAID level 0), segment 1 is written to disk 1, segment 2 is written to disk 2, and so on. Disk striping enhances performance because multiple drives are accessed simultaneously, but disk striping does not provide data redundancy.



Stripe Width Stripe width is the number of disks involved in an array where striping is implemented. For example, a four-disk array with disk striping has a stripe width of four.

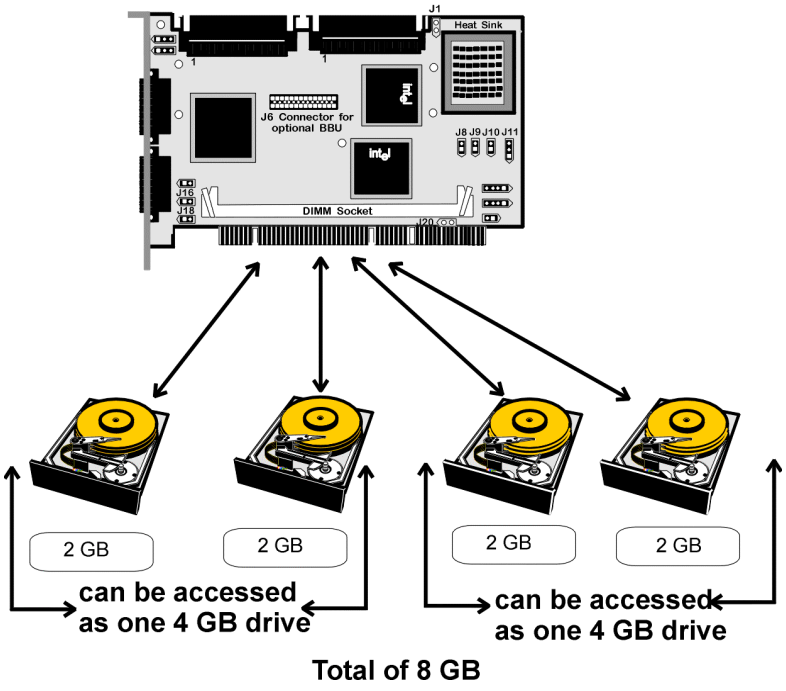
Stripe Size The stripe size is the length of the interleaved data segments that MegaRAID writes across multiple drives. MegaRAID supports stripe sizes of 2 KB, 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, or 128 KB.

Disk Spanning

Disk spanning allows multiple disk drives to function like one big drive. Spanning overcomes lack of disk space and simplifies storage management by combining existing resources or adding relatively inexpensive resources. For example, four 400 MB disk drives can be combined to appear to the operating system as one single 1600 MB drive.

Spanning alone does not provide reliability or performance enhancements. Spanned logical drives must have the same stripe size and must be contiguous. In the following graphic, RAID 1 array is turned into a RAID 10 array.

This controller supports a span depth of eight. That means that eight RAID 1, 3 or 5 arrays can be spanned to create one logical drive.



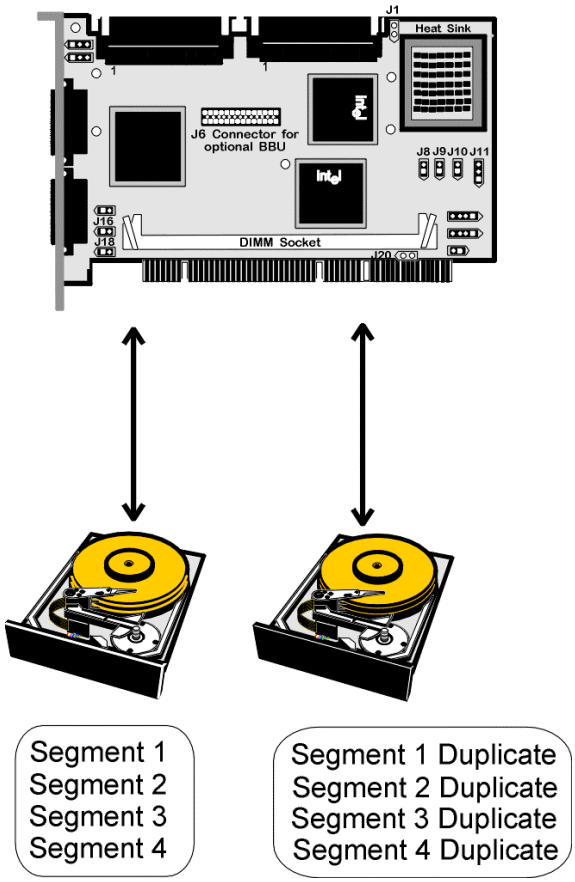
Spanning for RAID 10, RAID 30, or RAID 50

Level	Description
10	Configure RAID 10 by spanning two contiguous RAID 1 logical drives. The RAID 1 logical drives must have the same stripe size.
30	Configure RAID 30 by spanning two contiguous RAID 3 logical drives. The RAID 3 logical drives must have the same stripe size.
50	Configure RAID 50 by spanning two contiguous RAID 5 logical drives. The RAID 5 logical drives must have the same stripe size.
Note:	Spanning two contiguous RAID 0 logical drives does not produce a new RAID level or add fault tolerance. It does increase the size of the logical volume and improves performance by doubling the number of spindles.

Disk Mirroring

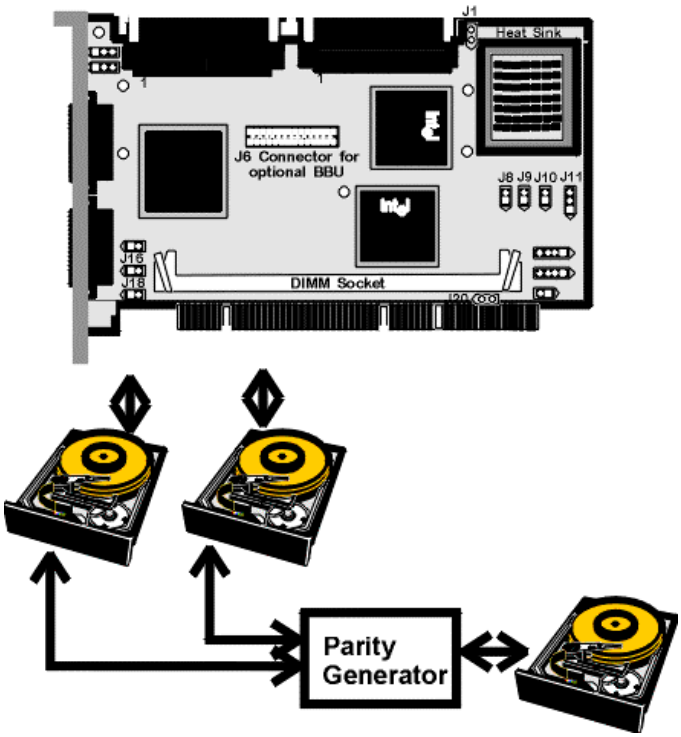
With mirroring (used in RAID 1), data written to one disk drive is simultaneously written to another disk drive. If one disk drive fails, the contents of the other disk drive can be used to run the system and reconstruct the failed drive. The primary advantage of disk mirroring is that it provides 100% data redundancy. Since the contents of the disk drive are completely written to a second drive, it does not matter if one of the drives fails. Both drives contain the same data at all times. Either drive can act as the operational drive.

Disk mirroring provides 100% redundancy, but is expensive because each drive in the system must be duplicated.



Parity

Parity generates a set of redundancy data from two or more parent data sets. The redundancy data can be used to reconstruct one of the parent data sets. Parity data does not fully duplicate the parent data sets. In RAID, this method is applied to entire drives or stripes across all disk drives in an array. A dedicated parity scheme during normal read/write operations is shown below.



The types of parity are:

Type	Description
Dedicated Parity	The parity of the data on two or more disk drives is stored on an additional disk.
Distributed Parity	The parity data is distributed across all drives in the system.

If a single disk drive fails, it can be rebuilt from the parity and the data on the remaining drives.

RAID level 3 combines dedicated parity with disk striping. The parity disk in RAID 3 is the last physical drive in a RAID set.

RAID level 5 combines distributed parity with disk striping. Parity provides redundancy for one drive failure without duplicating the contents of entire disk drives, but parity generation can slow the write process.

Hot Spares

A hot spare is an extra, unused disk drive that is part of the disk subsystem. It is usually in standby mode, ready for service if a drive fails. Hot spares permit you to replace failed drives without system shutdown or user intervention.

MegaRAID implements automatic and transparent rebuilds using hot spare drives, providing a high degree of fault tolerance and zero downtime. MegaRAID RAID Management software allows you to specify physical drives as hot spares. When a hot spare is needed, the MegaRAID controller assigns the hot spare that has a capacity closest to and at least as great as that of the failed drive to take the place of the failed drive.

Important

Hot spares are employed only in arrays with redundancy, for example, RAID levels 1, 3, 5, 10, 30, and 50.

A hot spare connected to a specific MegaRAID controller can only be used to rebuild a drive that is connected to the same controller.

Disk Rebuild

You rebuild a disk drive by recreating the data that had been stored on the drive before the drive failed.

Rebuilding can be done only in arrays with data redundancy such as RAID level 1, 3, 5, 10, 30, and 50.

Standby (warm spare) rebuild is employed in a mirrored (RAID 1) system. If a disk drive fails, an identical drive is immediately available. The primary data source disk drive is the original disk drive.

A hot spare can be used to rebuild disk drives in RAID 1, 3, 5, 10, 30, or 50 systems. If a hot spare is not available, the failed disk drive must be replaced with a new disk drive so that the data on the failed drive can be rebuilt.

Using hot spares, MegaRAID can automatically and transparently rebuild a failed drive with user-defined rebuild rates. If a hot spare is available, the rebuild starts automatically when a drive fails. MegaRAID automatically restarts the system and the rebuild if the system goes down during a rebuild.

Rebuild Rate The rebuild rate is the fraction of the compute cycles dedicated to rebuilding failed drives. A rebuild rate of 100 percent means the system is totally dedicated to rebuilding the failed drive.

The rebuild rate can be configured between 0% and 100%. At 0%, the rebuild is done only if the system is not doing anything else. At 100%, the rebuild has a higher priority than any other system activity.

Physical Array A RAID array is a collection of physical disk drives governed by the RAID management software. A RAID array appears to the host computer as one or more logical drives.

Logical Drive

A logical drive is a partition in a physical array of disks that is made up of contiguous data segments on the physical disks. A logical drive can consist of any of the following:

- an entire physical array
- more than one entire physical array
- a part of an array
- parts of more than one array
- a combination of any two of the above conditions

Hot Swap

A hot swap is the manual replacement of a defective physical disk unit while the computer is still running. When a new drive has been installed, you must issue a command to rebuild the drive. MegaRAID can be configured to detect the new disks and to rebuild the contents of the disk drive automatically.

SCSI Drive States

A SCSI disk drive can be in one of these states:

State	Description
Online (ONLIN)	The drive is functioning normally and is a part of a configured logical drive.
Ready (READY)	The drive is functioning normally but is not part of a configured logical drive and is not designated as a hot spare.
Hot Spare (HOTSP)	The drive is powered up and ready for use as a spare in case an online drive fails.
Fail (FAIL)	A fault has occurred in the drive placing it out of service.
Rebuild (REB)	The drive is being rebuilt with data from a failed drive.

Logical Drive States

State	Description
Optimal	The drive operating condition is good. All configured drives are online
Degraded	The drive operating condition is not optimal. One of the configured drives has failed or is offline.
Failed	The drive has failed.
Offline	The drive is not available to MegaRAID.

Disk Array Types

The RAID disk array types are:

Type	Description
Software-Based	The array is managed by software running in a host computer using the host CPU bandwidth. The disadvantages associated with this method are the load on the host CPU and the need for different software for each operating system.
SCSI to SCSI	The array controller resides outside of the host computer and communicates with the host through a SCSI adapter in the host. The array management software runs in the controller. It is transparent to the host and independent of the host operating system. The disadvantage is the limited data transfer rate of the SCSI channel between the SCSI adapter and the array controller.
Bus-Based	The array controller resides on the bus (for example, a PCI or EISA bus) in the host computer and has its own CPU to generate the parity and handle other RAID functions. A bus-based controller can transfer data at the speed of the host bus (PCI, ISA, EISA, VL-Bus) but is limited to the bus it is designed for. MegaRAID resides on a PCI bus, which can handle data transfer at up to 266 MB/s. With MegaRAID, each channel can handle data transfer rates up to 160 MB/s per SCSI channel.

Enclosure Management

Enclosure management is the intelligent monitoring of the disk subsystem by software and/or hardware.

The disk subsystem can be part of the host computer or separate from it. Enclosure management helps you stay informed of events in the disk subsystem, such as a drive or power supply failure. Enclosure management increases the fault tolerance of the disk subsystem.

3 RAID Levels

There are six official RAID levels (RAID 0 through RAID 5). MegaRAID supports RAID levels 0, 1, 3, and 5. LSI Logic has designed three additional RAID levels (10, 30, and 50) that provide additional benefits. The RAID levels that MegaRAID supports are:

RAID Level	Type	turn to
0	Standard	page 17
1	Standard	page 18
3	Standard	page 19
5	Standard	page 21
10	MegaRAID only	page 22
30	MegaRAID only	page 23
50	MegaRAID only	page 24

Select RAID Level To ensure the best performance, you should select the optimal RAID level when you create a system drive. The optimal RAID level for your disk array depends on a number of factors:

- the number of drives in the disk array
- the capacity of the drives in the array
- the need for data redundancy
- the disk performance requirements

Selecting a RAID Level The factors you need to consider when selecting a RAID level are listed on the next page.

Selecting a RAID Level

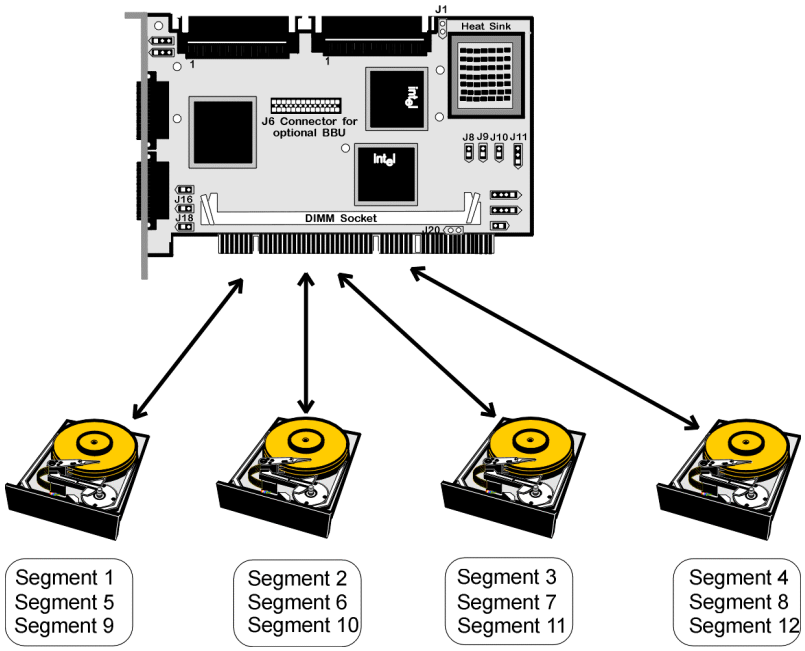
Level	Description and Use	Pros	Cons	Max. Drives	Fault Tolerant
0	Data divided in blocks and distributed sequentially (pure striping). Use for non-critical data that requires high performance.	High data throughput for large files	No fault tolerance. All data lost if any drive fails.	One to 30	No
1	Data duplicated on another disk (mirroring). Use for read-intensive fault-tolerant systems.	100% data redundancy	Doubles disk space. Reduced performance during rebuilds.	Two	Yes
3	Disk striping with a dedicated parity drive. Use for non-interactive apps that process large files sequentially.	Achieves data redundancy at low cost	Performance not as good as RAID 1	Three to 30	Yes
5	Disk striping and parity data across all drives. Use for high read volume but low write volume, such as transaction processing.	Achieves data redundancy at low cost	Performance not as good as RAID 1	Three to 30	Yes
10	Data striping and mirrored drives.	High data transfers, complete redundancy	More complicated	Four to 30 (must be a multiple of two)	Yes
30	Disk striping with a dedicated parity drive.	High data transfers, redundancy	More complicated	Six to 30	Yes
50	Disk striping and parity data across all drives.	High data transfers, redundancy	More complicated	Six to 30	Yes

RAID 0

RAID 0 provides disk striping across all drives in the RAID subsystem. RAID 0 does not provide any data redundancy, but does offer the best performance of any RAID level. RAID 0 breaks up data into smaller blocks and then writes a block to each drive in the array. The size of each block is determined by the stripe size parameter, set during the creation of the RAID set. RAID 0 offers high bandwidth.

By breaking up a large file into smaller blocks, MegaRAID can use multiple SCSI channels and drives to read or write the file faster. RAID 0 involves no parity calculations to complicate the write operation. This makes RAID 0 ideal for applications that require high bandwidth but do not require fault tolerance.

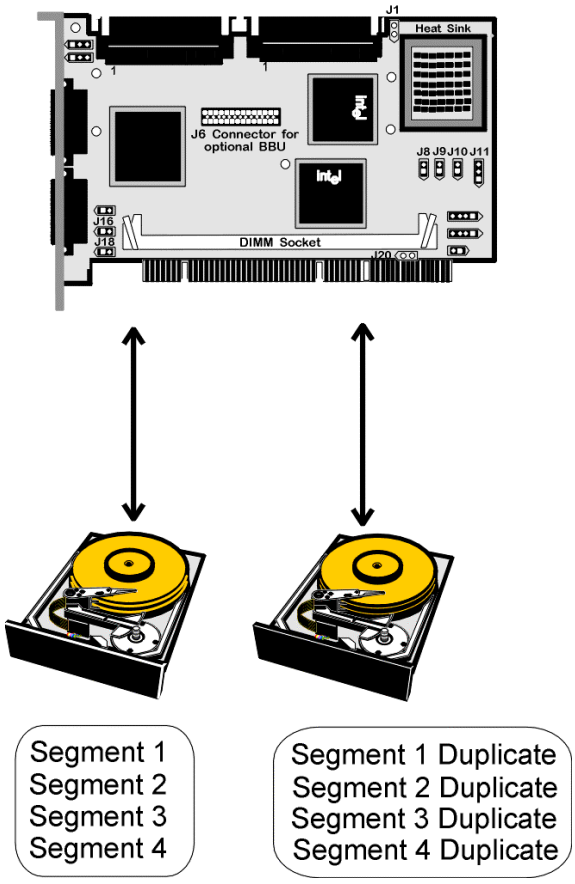
Uses	RAID 0 provides high data throughput, especially for large files. Any environment hat does not require fault tolerance.
Strong Points	Provides increased data throughput for large files. No capacity loss penalty for parity.
Weak Points	Does not provide fault tolerance. All data lost if any drive fails.
Drives	One to 30 The initiator takes one ID per channel. This leaves 15 IDs available per channel, for a maximum of 30 for two channels.



RAID 1

In RAID 1, MegaRAID duplicates all data from one drive to a second drive. RAID 1 provides complete data redundancy, but at the cost of doubling the required data storage capacity.

Uses	Use RAID 1 for small databases or any other environment that requires fault tolerance but small capacity.
Strong Points	RAID 1 provides complete data redundancy. RAID 1 is ideal for any application that requires fault tolerance and minimal capacity.
Weak Points	RAID 1 requires twice as many disk drives. Performance is impaired during drive rebuilds.
Drives	Two



RAID 3

RAID 3 provides disk striping and complete data redundancy though a dedicated parity drive. The stripe size must be 64 KB if RAID 3 is used. RAID 3 handles data at the block level, not the byte level, so it is ideal for networks that often handle very large files, such as graphic images.

RAID 3 breaks up data into smaller blocks, calculates parity by performing an exclusive-or on the blocks, and then writes the blocks to all but one drive in the array. The parity data created during the exclusive-or is then written to the last drive in the array. The size of each block is determined by the stripe size parameter, which is set during the creation of the RAID set.

If a single drive fails, a RAID 3 array continues to operate in degraded mode. If the failed drive is a data drive, writes will continue as normal, except no data is written to the failed drive. Reads reconstruct the data on the failed drive by performing an exclusive-or operation on the remaining data in the stripe and the parity for that stripe. If the failed drive is a parity drive, writes will occur as normal, except no parity is written. Reads retrieve data from the disks.

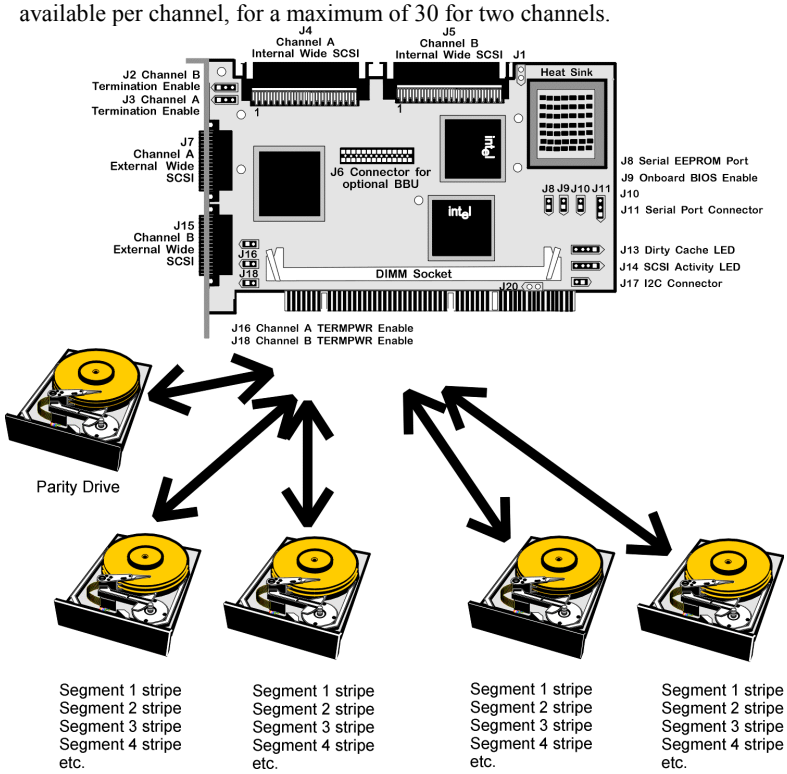
Uses Best suited for applications such as graphics, imaging, or video, or any application that calls for reading and writing huge, sequential blocks of data.

Strong Points Provides data redundancy and high data transfer rates.

Weak Points The dedicated parity disk is a bottleneck with random I/O.

Drives Three to 30

The initiator takes one ID per channel. This leaves 15 IDs available per channel, for a maximum of 30 for two channels.



Cont'd

RAID 5 vs RAID 3 You may find that RAID 5 is preferable to RAID 3 even for applications characterized by sequential reads and writes, because MegaRAID has very robust caching algorithms and hardware based exclusive-or assist.

The benefits of RAID 3 disappear if there are many small I/O operations scattered randomly and widely across the disks in the logical drive. The RAID 3 fixed parity disk becomes a bottleneck in such applications. For example: The host attempts to make two small writes and the writes are widely scattered, involving two different stripes and different disk drives. Ideally both writes should take place at the same time. But this is not possible in RAID 3, since the writes must take turns accessing the fixed parity drive. For this reason, RAID 5 is the clear choice in this scenario.

RAID 5

RAID 5 includes disk striping at the byte level and parity. In RAID 5, the parity information is written to several drives. RAID 5 is best suited for networks that perform a lot of small I/O transactions simultaneously.

RAID 5 addresses the bottleneck issue for random I/O operations. Since each drive contains both data and parity numerous writes can take place concurrently. In addition, robust caching algorithms and hardware based exclusive-or assist make RAID 5 performance exceptional in many different environments.

Uses

RAID 5 provides high data throughput, especially for large files. Use RAID 5 for transaction processing applications because each drive can read and write independently. If a drive fails, MegaRAID uses distributed parity to recreate all missing information. Use also for office automation and online customer service that requires fault tolerance. Use for any application that has high read request rates but low write request rates.

Strong Points

Provides data redundancy and good performance in most environments

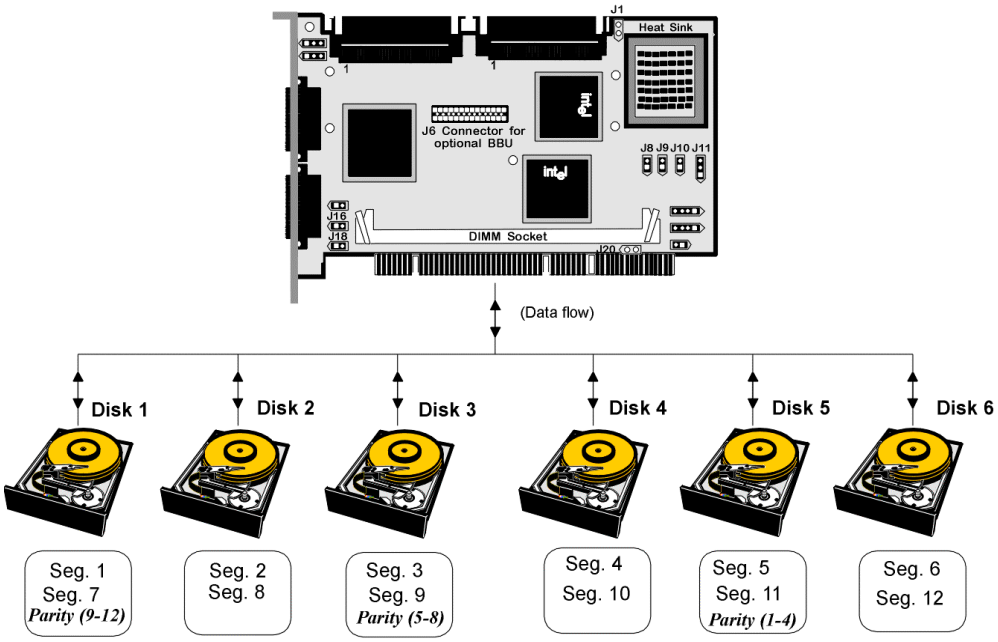
Weak Points

Disk drive performance will be reduced if a drive is being rebuilt. Environments with few processes do not perform as well because the RAID overhead is not offset by the performance gains in handling simultaneous processes.

Drives

Three to 30

The initiator takes one ID per channel. This leaves 15 IDs available per channel, for a maximum of 30 for two channels.



Parity is distributed across all drives in array.

RAID 10

RAID 10 is a combination of RAID 0 and RAID 1. RAID 10 has mirrored drives. RAID 10 breaks up data into smaller blocks, and then stripes the blocks of data to each RAID 1 raid set. Each RAID 1 raid set then duplicates its data to its other drive. The size of each block is determined by the stripe size parameter, which is set during the creation of the RAID set. RAID 10 can sustain one to four drive failures while maintaining data integrity if each failed disk is in a different RAID 1 array.

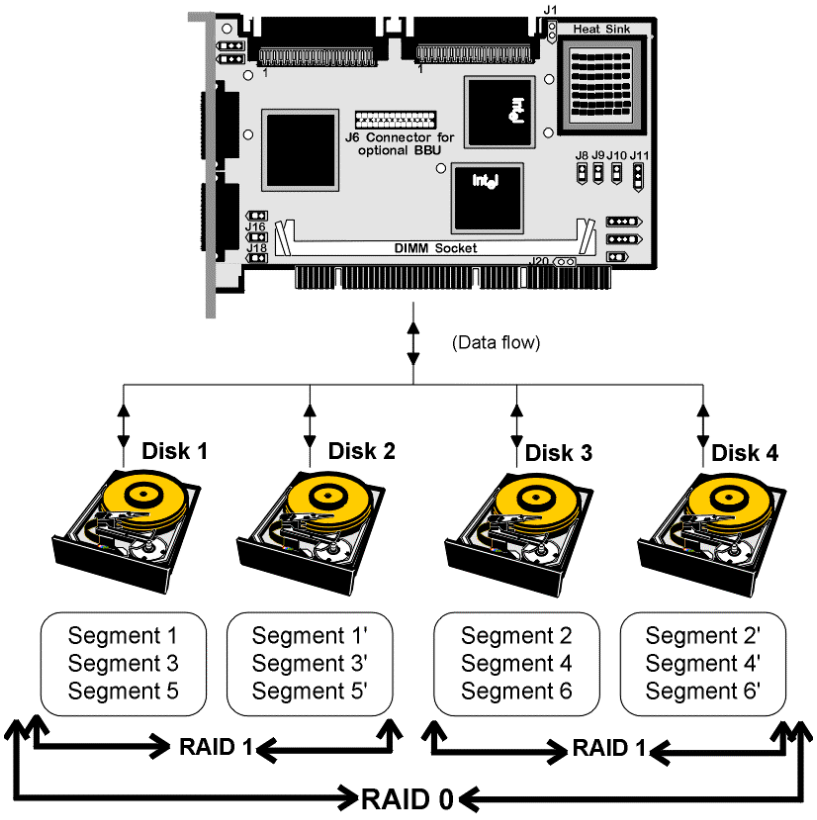
Uses RAID 10 works best for data storage that must have 100% redundancy of mirrored arrays and that also needs the enhanced I/O performance of RAID 0 (striped arrays). RAID 10 works well for medium-sized databases or any environment that requires a higher degree of fault tolerance and moderate to medium capacity.

Strong Points RAID 10 provides both high data transfer rates and complete data redundancy.

Weak Points RAID 10 requires twice as many drives as all other RAID levels except RAID 1.

Drives Four to 30 (must be a multiple of two)

The initiator takes one ID per channel. This leaves 15 IDs available per channel, for a maximum of 30 for two channels.



RAID 30

RAID 30 is a combination of RAID 0 and RAID 3. RAID 30 provides high data transfer speeds and high data reliability. RAID 30 is best implemented on two RAID 3 disk arrays with data striped across both disk arrays. RAID 30 breaks up data into smaller blocks, and then stripes the blocks of data to each RAID 3 raid set. RAID 3 breaks up data into smaller blocks, calculates parity by performing an exclusive-or on the blocks, and then writes the blocks to all but one drive in the array. The parity data created during the exclusive-or is then written to the last drive in each RAID 3 array. The size of each block is determined by the stripe size parameter, which is set during the creation of the RAID set.

RAID 30 can sustain one drive failure per RAID 3 array and still maintain data integrity. For example, the RAID 30 configuration in the graphic below has two RAID 3 arrays. It can survive two drive failures, as long as the failed drives are in different RAID 3 arrays.

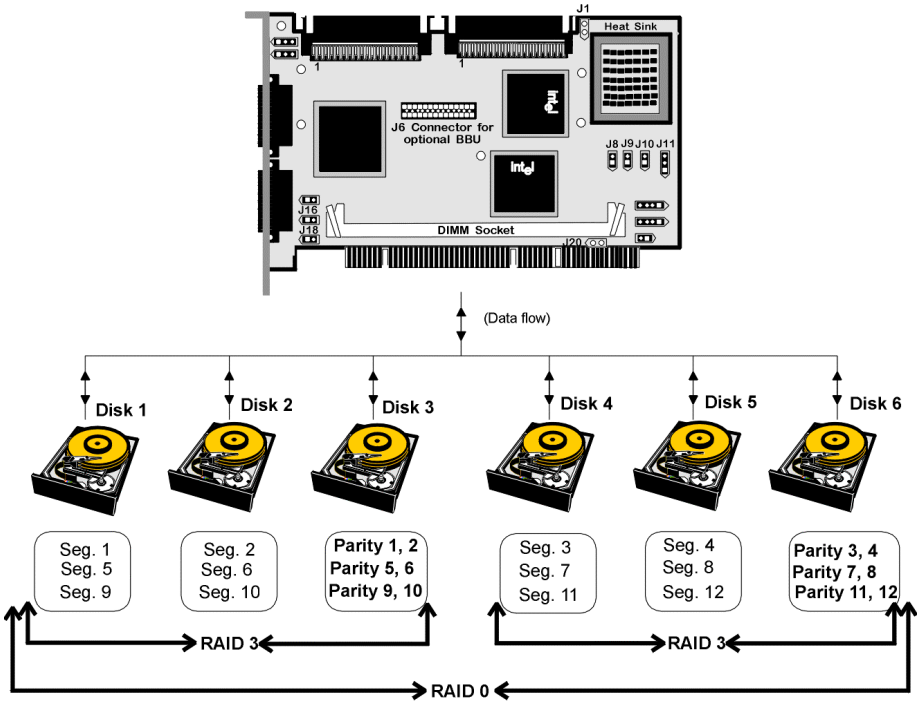
Uses Use RAID 30 for sequentially written and read data, pre-press and video on demand that requires a higher degree of fault tolerance and medium to large capacity.

Strong Points Provides data reliability and high data transfer rates.

Weak Points Requires 2 – 4 times as many parity drives as RAID 3.

Drives Six to 30

The initiator takes one ID per channel. This leaves 15 IDs available per channel, for a maximum of 30 for two channels.



RAID 50

RAID 50 provides the features of both RAID 0 and RAID 5. RAID 50 includes both parity and disk striping across multiple drives. RAID 50 is best implemented on two RAID 5 disk arrays with data striped across both disk arrays. RAID 50 breaks up data into smaller blocks, and then stripes the blocks of data to each RAID 5 raid set. RAID 5 breaks up data into smaller blocks, calculates parity by performing an exclusive-or on the blocks, and then writes the blocks of data and parity to each drive in the array. The size of each block is determined by the stripe size parameter, which is set during the creation of the RAID set.

RAID 50 can sustain one drive failure per RAID 5 array and still maintain data integrity. For example, the RAID 50 configuration in the graphic below has two RAID 5 arrays. It can survive two drive failures, as long as the failed drives are in different RAID 5 arrays.

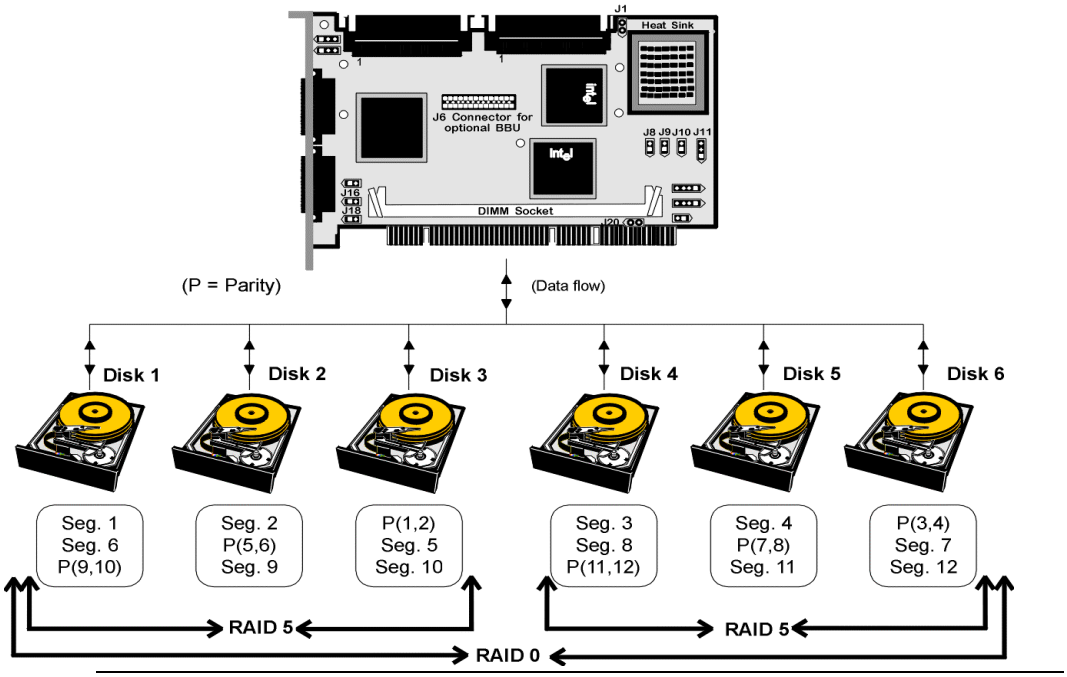
Uses RAID 50 works best when used with data that requires high reliability, high request rates, and high data transfer and medium to large capacity.

Strong Points RAID 50 provides high data throughput, data redundancy, and very good performance.

Weak Points Requires 2 to 4 times as many parity drives as RAID 5.

Drives Six to 30

The initiator takes one ID per channel. This leaves 15 IDs available per channel, for a maximum of 30 for two channels.



4 Features

MegaRAID is a family of high performance intelligent PCI-to-SCSI host adapters with RAID control capabilities. MegaRAID Elite 1600 64-Bit LVD has two SCSI channels that support 160M and Wide SCSI, with data transfer rates of up to 160 MB/s per SCSI channel. Each SCSI channel supports up to 15 Wide devices and up to seven non-Wide devices.

Features

MegaRAID features include:

- remote configuration and array management through MegaRAID WebBIOS
- high performance I/O migration path while preserving existing PCI-SCSI software
- SCSI data transfers up to 160 MB/s
- synchronous operation on a wide LVD SCSI bus
- up to 15 LVD SCSI devices on the wide bus
- up to 128 MB of 3.3V SDRAM cache memory in one single-sided or double-sided DIMM socket (Cache memory is used for read and write-back caching and for RAID 5 parity generation.)
- NVRAM storage for RAID configuration data
- audible alarm
- DMA chaining support
- separate DRAM bus
- support for differential or single ended SCSI with active termination
- up to 12 MegaRAID Elite 1600 adapter cards per system
- support for up to 15 SCSI devices per channel
- support for RAID levels 0, 1, 3, 5, 10, 30, and 50
- span depth of eight for RAID 1, 3 or 5 arrays
- support for scatter/gather and tagged command queuing
- ability to multithread up to 256 commands simultaneously
- support for multiple rebuilds and consistency checks with transparent user-definable priority setting
- support for variable stripe sizes for all logical drives
- automatic detection of failed drives
- automatic and transparent rebuild of hot spare drives
- ability to hot swap new drives without taking the system down
- optional battery backup
- server clustering support
- optional firmware provides multi-initiator support
- support for server failover
- software drivers for major operating systems

SMART Technology

The MegaRAID Self-Monitoring Analysis and Reporting Technology (SMART) detects up to 70% of all predictable drive failures. SMART monitors the internal performance of all motors, heads, and drive electronics. You can recover from drive failures through RAID remapping and online physical drive migration.

Configuration on Disk

Configuration on Disk (drive roaming) saves configuration information both in NVRAM on MegaRAID and on the disk drives connected to MegaRAID. If MegaRAID is replaced, the new MegaRAID controller can detect the actual RAID configuration, maintaining the integrity of the data on each drive, even if the drives have changed channel and/or target ID.

Hardware Requirements

MegaRAID can be installed in an IBM AT®-compatible or EISA computer with a motherboard with PCI expansion slots. The computer must support PCI version 2.1 or later. The computer should have an Intel Pentium or later CPU, a floppy drive, color monitor and VGA adapter card, a keyboard, and mouse.

Configuration Features

Specification	Feature
RAID Levels	0, 1, 3, 5, 10, 30, and 50.
SCSI Channels	2
Maximum number of drives per channel	15
Array interface to host	64-bit PCI
PCI bus master	Supports write invalidate
Drive interface	Wide 160M
Upgradable cache memory sizes	16 MB, 32 MB, 64 MB, or 128 MB,
Cache Function	Write-through, write-back, ARA, NRA, RA
Multiple logical drives/arrays per controller	Up to 40 logical drives per controller
Maximum number of MegaRAID controllers per system	12
Online capacity expansion	Yes
Dedicated and pool hot spare	Yes
Flashable firmware	Yes
Hot swap devices supported	Yes
Non-disk devices supported	Yes
Mixed capacity hard disk drives	Yes
Number of 16-bit internal SCSI connectors	2
Number of external SCSI connectors	2
Support for hard disk drives with capacities of more than 8 GB.	Yes
Clustering support (Failover control)	Yes
Online RAID level migration	Yes
RAID remapping	Yes
No reboot necessary after expansion	Yes
More than 200 Qtags per physical drive	Yes
Hardware clustering support on the board	Yes
User-specified rebuild rate	Yes

Hardware Architecture Features

Specification	Feature
Processor	Intel i960RN
SCSI Controller	one Q-Logic 12160 Dual SCSI controller
memory type	One 64-bit 168-pin SDRAM DIMM socket provides write-through or write-back caching on a logical drive basis. It also provides adaptive readahead.
Size of Flash ROM	1 MB
Amount of NVRAM	32 KB
Hardware XOR assistance	Yes
Direct I/O	Yes
Removable battery-backed cache memory module	Yes
SCSI bus termination	Active, LVD and SE
Double-sided DIMMs	Yes
Direct I/O bandwidth	266 MB/s

Array Performance Features

The MegaRAID array performance features include:

Specification	Feature
Host data transfer rate	266 MB/s
Drive data transfer rate	160 MB/s
Maximum Scatter/Gathers	26 elements
Maximum size of I/O requests	6.4 MB in 64 KB stripes
Maximum Queue Tags per drive	211
Stripe Sizes	2 KB, 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, or 128 KB
Maximum number of concurrent commands	255
Support for multiple initiators	Yes

RAID Management Features

The MegaRAID RAID management features include:

Specification	Feature
Support for SNMP	Yes
Performance Monitor provided	Yes
Remote control and monitoring	Yes
Event broadcast and event alert	Yes
Hardware connector	3-pin serial
Drive roaming	Yes
Support for concurrent multiple stripe sizes	Yes
Windows NT and NetWare server support via GUI client utility	Yes
SCO Unix, OS/2, and UnixWare server support via GUI client utility	Yes
DMI support	Yes
Management through an industry-standard browser	Yes

Fault Tolerance Features

The MegaRAID fault tolerance features include:

Specification	Feature
Support for SMART	Yes
Enclosure management	SAF-TE compliant
Drive failure detection	Automatic
Drive rebuild using hot spares	Automatic and transparent
Parity Generation and checking	Software and hardware

Software Utilities

The MegaRAID software utility features include:

Specification	Feature
FlexRAID reconfiguration on the fly	Yes
FlexRAID RAID level migration on the fly	Yes
FlexRAID online capacity expansion	Yes
Remote configuration and management over the Internet	Yes
Graphical user interface	Yes
Diagnostic utility	Yes
Management utility	Yes
Bootup configuration via MegaRAID Manager	Yes
Online Read, Write, and cache policy switching	Yes
Internet and intranet support through TCP/IP	Yes

Operating System Software Drivers

Operating System Drivers MegaRAID includes a DOS software configuration utility and drivers for all major operating systems. See the *MegaRAID Operating System Drivers Guide* for additional information.

The DOS drivers for MegaRAID are contained in the firmware on MegaRAID except the DOS ASPI and CD-ROM drivers. Call LSI Logic technical support at 770-246-8600 or access the LSI Logic web site at www.lsil.com for information about drivers for other operating systems.

MegaRAID Specifications

Parameter	Specification
Card Size	6.875" x 4.2" (half length PCI)
Processor	Intel i960RN @ 100 MHz
Bus Type	PCI 2.2
Bus Data Transfer Rate	Up to 266 MB/s
BIOS	MegaRAID BIOS
Cache Configuration	16, 32, 64, or 128 MB through a single bank using 66 MHz, 3.3V unbuffered ECC SDRAM in a single-sided or double-sided 168-pin DIMM.
Firmware	1 MB × 8 flash ROM
Nonvolatile RAM	32 KB × 8 for storing RAID configuration
Operating Voltage	5.00 V ± 0.25 V and 3.30V +/- 0.3V
SCSI Controller	1 SCSI controller for 160M and Wide support.
SCSI Data Transfer Rate	Up to 160 MB/s.
SCSI Bus	low voltage differential or SE
SCSI Termination	Active
Termination Disable	Automatic through cable detection
Devices per SCSI Channel	Up to 15 wide or seven non-wide SCSI devices. Up to 6 non-disk SCSI drives per MegaRAID controller.
SCSI Device Types Supported	Synchronous or Asynchronous. Disk and non-disk.
RAID Levels Supported	0, 1, 3, 5,10, 30, and 50
SCSI Connectors	Two 68-pin internal high-density connectors for 16-bit SCSI devices. Two ultra-high density 68-pin external connectors
SCSI cables	Up to 25 meters
Serial Port	3-pin RS232C-compatible berg

Components

CPU The MegaRAID controller uses the 64-bit Intel i960RN Intelligent I/O processor with an embedded 64-bit 80960 Jx RISC processor that runs at 100 MHz. This processor directs all functions of the controller including command processing, PCI and SCSI bus transfers, RAID processing, drive rebuilding, cache management, and error recovery.

Cache Memory Cache memory resides in a single 64-bit DIMM socket that requires one X8 or X16 unbuffered 3.3V SDRAM single-sided or double-sided DIMM. Possible configurations are 16, 32, 64, or 128 MB.

MegaRAID supports write-through or write-back caching, which can be selected for each logical drive. To improve performance in sequential disk accesses, MegaRAID does not use read-ahead caching for the current logical drive. The default setting for the read policy is *Normal*, meaning no read-ahead caching. You can disable read-ahead caching.

Warning!

Write caching is not recommended for the physical drives. When write cache is enabled, loss of data can occur when power is interrupted.

MegaRAID BIOS The BIOS resides on a 1 MB or 2 MB × 8 flash ROM for easy upgrade. The MegaRAID BIOS supports INT 13h calls to boot DOS without special software or device drivers. The MegaRAID BIOS provides an extensive setup utility that can be accessed by pressing <Ctrl> <M> at BIOS initialization. MegaRAID Configuration Utility is described in the *MegaRAID Configuration Software Guide.*

Cont'd

Components, Continued

Onboard Speaker MegaRAID has an onboard tone generator for audible warnings when system errors occur. Audible warnings can be generated through this speaker. The audible warnings are listed on page 127.

Serial Port MegaRAID includes a 3-pin RS232C-compatible serial port berg connector, which can connect to communications devices and external storage devices.

SCSI Bus MegaRAID Elite 1600 has two 160M Wide SCSI channels that support low voltage differential SCSI devices with active termination. Both synchronous and asynchronous devices are supported. MegaRAID provides automatic termination disable via cable detection. Each channel supports up to 15 wide or seven non-wide SCSI devices at speeds up to 160 MB/s per SCSI channel. MegaRAID supports up to six non-disk devices per controller. The SCSI bus mode defaults to LVD for each SCSI channel. If a single ended device is attached to a SCSI channel, MegaRAID automatically switches to SE mode for that SCSI channel.

SCSI Connectors MegaRAID has two types of SCSI connectors:

- two 68-pin high density internal SCSI connectors (Channels A and B)
- two 68-pin external ultra-high-density external SCSI connectors (Channels A and B)

SCSI Termination MegaRAID uses active termination on the SCSI bus conforming to Alternative 2 of the SCSI-2 specifications. Termination enable/disable is automatic through cable detection.

Cont'd

SCSI Firmware The firmware handles all RAID and SCSI command processing and also supports:

Feature	Description
Disconnect/Reconnect	Optimizes SCSI Bus seek.
Tagged Command Queuing	Multiple tags to improve random access
Scatter/Gather	Multiple address/count pairs
Multi-threading	Up to 255 simultaneous commands with elevator sorting and concatenation of requests per SCSI channel
Stripe Size	Variable for all logical drives: 2 KB, 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, or 128 KB.
Rebuild	Multiple rebuilds and consistency checks with user-definable priority.

RAID Management The RAID utilities manage and configure the RAID system and MegaRAID, create and manage multiple disk arrays, control and monitor multiple RAID servers, provide error statistics logging and online maintenance:

- MegaRAID Configuration Utility
- WebBIOS Configuration Utility
- Power Console
- MegaRAID Manager

MegaRAID Configuration Utility It configures and maintains RAID arrays, formats disk drives, and manages the RAID system. It is independent of any operating system.

WebBIOS Configuration Utility It allows you to configure and manage a RAID system on a remote server over the Internet.

Power Console Plus It configures, monitors, and manages RAID servers from any Windows NT network node or remote server.

MegaRAID Manager A character-based utility for DOS, Linux, Solaris, SCO Unix, SCO UnixWare, OS/2, and Novell NetWare.

Cont'd

Components, Continued

Fault-Tolerance The MegaRAID fault-tolerance features are:

- built-in 3-pin berg connector that provides an RS-232C serial communication interface
- automatic failed drive detection
- automatic failed drive rebuild with no user intervention required
- hot swap manual replacement without bringing the system down
- SAF-TE compliant enclosure management,
- cache memory

Detect Failed Drive The MegaRAID firmware automatically detects and rebuilds failed drives. This can be done transparently with hot spares.

Hot Swap MegaRAID supports the manual replacement of a disk unit in the RAID subsystem without system shutdown.

Compatibility MegaRAID compatibility issues include:

- server management
- SCSI device compatibility
- software compatibility

Server Management As an SNMP agent, MegaRAID supports all SNMP managers and RedAlert from Storage Dimensions.

SCSI Device Compatibility MegaRAID supports SCSI hard disk drives, CD-ROMs, tape drives, optical drives, DAT drives and other SCSI peripheral devices.

Cont'd

Components, Continued

Software All SCSI backup and utility software should work with MegaRAID. Software that has been tested and approved for use with MegaRAID includes Cheyenne®, CorelSCSI®, Arcserve®, and Novaback®. This software is not provided with MegaRAID.

Clustering Support LSI Logic provides OEM-optional firmware with multi-initiator support. This software provides high system availability by permitting server failover.

Summary

MegaRAID features were discussed in this chapter. In the next chapter, MegaRAID configuration is described.

5 Configuring MegaRAID

Configuring SCSI Physical Drives

SCSI Channels Physical SCSI drives must be organized into logical drives. The arrays and logical drives that you construct must be able to support the RAID level that you select.

Your MegaRAID adapter has two SCSI channels.

Distributing Drives Distribute the disk drives across all channels for optimal performance. It is best to stripe across channels instead of down channels. Performance is most affected for sequential reads and writes. MegaRAID supports SCSI CD-ROM drives, SCSI tape drives, and other SCSI devices as well as SCSI hard disk drives. For optimal performance, all non-disk SCSI devices should be attached to one SCSI channel.

Basic Configuration Rules You should observe the following guidelines when connecting and configuring SCSI devices in a RAID array:

- attach non-disk SCSI devices to a single SCSI channel that does not have any disk drives
- distribute the SCSI hard disk drives equally among all available SCSI channels except any SCSI channel that is being reserved for non-disk drives
- you can place up to 30 physical disk drives in a logical array
- an array can contain SCSI devices that reside on any channel
- include all drives that have the same capacity to the same array
- make sure any hot spare has a capacity that is at least as large as the largest drive that may be replaced by the hot spare
- when replacing a failed drive, make sure that the replacement drive has a capacity that is at least as large as the drive being replaced

Current Configuration

SCSI ID	Device Description	Termination?
SCSI Channel A		
0		
1		
2		
3		
4		
5		
6		
8		
9		
10		
11		
12		
13		
14		
15		
SCSI Channel B		
0		
1		
2		
3		
4		
5		
6		
8		
9		
10		
11		
12		
13		
14		
15		

Logical Drive Configuration

Logical Drive	RAID Level	Stripe Size	Logical Drive Size	Cache Policy	Read Policy	Write Policy	# of Physical Drives
LD0							
LD1							
LD2							
LD3							
LD4							
LD5							
LD6							
LD7							
LD8							
LD9							
LD10							
LD11							
LD12							
LD13							
LD14							
LD15							
LD16							
LD17							
LD18							
LD19							
LD20							
LD21							
LD22							
LD23							
LD24							
LD25							
LD26							
LD27							
LD28							
LD29							
LD30							
LD31							
LD32							
LD33							
LD34							
LD35							
LD36							
LD37							
LD38							
LD39							

Physical Device Layout

	Channel A	Channel B
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		

	Channel A	Channel B
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		
Target ID		
Device Type		
Logical Drive Number/ Drive Number		
Manufacturer/Model Number		
Firmware level		

Configuring Arrays

Connect the physical drives to MegaRAID, configure the drives, then initialize them. The number of physical disk drives that an array can support depends on the firmware version.

For MegaRAID Elite 1600, an array can consist of up to 30 physical disk drives, depending on the RAID level. Elite 1600 supports up to 40 logical drives per controller. The number of drives in an array determines the RAID levels that can be supported.

Arranging Arrays You must arrange the arrays to provide additional organization for the drive array. You must arrange arrays so that you can create system drives that can function as boot devices.

You can sequentially arrange arrays with an identical number of drives so that the drives in the group are spanned. Spanned drives can be treated as one large drive. Data can be striped across multiple arrays as one logical drive.

You can create spanned drives by using the MegaRAID Configuration utility or the MegaRAID Manager. See the *MegaRAID Configuration Software Guide* for additional information.

Creating Hot Spares Any drive that is present, formatted, and initialized but not included in a array or logical drive is automatically designated as a hot spare.

You can also designate drives as hot spares by using the MegaRAID Configuration Utility, MegaRAID Manager, or Power Console. See the *MegaRAID Configuration Software Guide* for additional information.

Creating Logical Drives Logical drives are arrays or spanned arrays that are presented to the operating system. You must create one or more logical drives.

The logical drive capacity can include all or any portion of a array. The logical drive capacity can also be larger than an array by using spanning. MegaRAID Elite 1600 supports up to 40 logical drives.

Configuration Strategies

The most important factors in RAID array configuration are: drive capacity, drive availability (fault tolerance), and drive performance. You cannot configure a logical drive that optimizes all three factors, but it is easy to choose a logical drive configuration that maximizes one factor at the expense of the other two factors, although needs are seldom that simple.

Maximize Capacity RAID 0 achieves maximum drive capacity, but does not provide data redundancy. Maximum drive capacity for each RAID level is shown below. OEM level firmware that can span up to 4 logical drives is assumed.

RAID Level	Description	Drives Required	Capacity
0	Striping without parity	1 – 30	(Number of disks) X capacity of smallest disk
1	Mirroring	2	(Capacity of smallest disk) X (1)
3	Striping with fixed parity drive	3 – 30	(Number of disks) X (capacity of smallest disk) - (capacity of 1 disk)
5	Striping with floating parity drive	3 – 30	(Number of disks) X (capacity of smallest disk) - (capacity of 1 disk)
10	Mirroring and Striping	4 – 30 (Must be a multiple of 2)	(Number of disks) X (capacity of smallest disk) / (2)
30	RAID 3 and Striping	6 – 30 (Must be a multiple of arrays)	(Number of disks) X (capacity of smallest disk) – (capacity of 1 disk X number of Arrays)
50	RAID 5 and Striping	6 – 30 (Must be a multiple of arrays)	(Number of disks) X (capacity of smallest disk) – (capacity of 1 disk X number of Arrays)

Cont'd

Maximize Drive Availability You can maximize the availability of data on the physical disk drive in the logical array by maximizing the level of fault tolerance. The levels of fault tolerance provided by the RAID levels are:

RAID Level	Fault Tolerance Protection
0	No fault tolerance.
1	Disk mirroring, which provides 100% data redundancy.
3	100% protection through a dedicated parity drive.
5	100% protection through striping and parity. The data is striped and parity data is written across a number of physical disk drives.
10	100% protection through data mirroring.
30	100% protection through data striping. All data is striped across all drives in two or more arrays.
50	100% protection through data striping and parity. All data is striped and parity data is written across all drives in two or more arrays.

Maximizing Drive Performance You can configure an array for optimal performance. But optimal drive configuration for one type of application will probably not be optimal for any other application. A basic guideline of the performance characteristics for RAID drive arrays at each RAID level is:

RAID Level	Performance Characteristics
0	Excellent for all types of I/O activity, but provides no data security.
1	Provides data redundancy and good performance.
3	Provides data redundancy.
5	Provides data redundancy and good performance in most environments.
10	Provides data redundancy and excellent performance.
30	Provides data redundancy and good performance in most environments.
50	Provides data redundancy and very good performance.

Assigning RAID Levels

Only one RAID level can be assigned to each logical drive. The drives required per RAID level is shown below:

RAID Level	Minimum Number of Physical Drives	Maximum Number of Physical Drives
0	One	30
1	Two	Two
3	Three	30
5	Three	30
10	Four	30
30	Six	30
50	Six	30

Configuring Logical Drives

After you have installed the MegaRAID controller in the server and have attached all physical disk drives, perform the following actions to prepare a RAID disk array:

Step	Action
1	Optimize the MegaRAID controller options for your system. See Chapter 3 for additional information.
2	Press <Ctrl> <M> to run MegaRAID Manager.
3	Perform a low-level format of the SCSI drives to be used in the array, and the drives to be used as hot spares.
4	Define and configure one or more logical drives. Select Easy Configuration in MegaRAID Manager or select New Configuration to customize the RAID array.
5	Create and configure one or more system drives (logical drives). Select the RAID level, cache policy, read policy, and write policy.
6	Save the configuration.
7	Initialize the system drives. After initialization, you can install the operating system.

Optimizing Data Storage

Data Access Requirements Each type of data stored in the disk subsystem has a different frequency of read and write activity. If you know the data access requirements, you can more successfully determine a strategy for optimizing the disk subsystem capacity, availability, and performance.

Servers that support Video on Demand typically read the data often, but write data infrequently. Both the read and write operations tend to be long. Data stored on a general-purpose file server involves relatively short read and write operations with relatively small files.

Array Functions You must first define the major purpose of the disk array. Will this disk array increase the system storage capacity for general-purpose file and print servers? Does this disk array support any software system that must be available 24 hours per day? Will the information stored in this disk array contain large audio or video files that must be available on demand? Will this disk array contain data from an imaging system?

You must identify the purpose of the data to be stored in the disk subsystem before you can confidently choose a RAID level and a RAID configuration.

Planning the Array Configuration

Answer the following questions about this array:

Question	Answer
Number of MegaRAID SCSI channels	2
Number of physical disk drives in the array	
Purpose of this array. Rank the following factors:	
Maximize drive capacity	
Maximize the safety of the data (fault tolerance)	
Maximize hard drive performance and throughput	
How many hot spares?	
Amount of cache memory installed on the MegaRAID	
Are all of the disk drives and the server that MegaRAID is installed in protected by a UPS?	

Using the Array Configuration Planner The following table lists the possible RAID levels, fault tolerance, and effective capacity for all possible drive configurations for an array consisting of one to eight drives.

The following table does not take into account any hot spare (standby) drives. You should always have a hot spare drive in case of drive failure.

RAID 1 requires 2 drives. RAID 10 requires at least 4 drives. RAID 30 and RAID 50 require at least 6 drives.

Array Configuration Planner

Number of Drives	Possible RAID Levels	Relative Performance	Fault Tolerance	Effective Capacity
1	None	Excellent	No	100%
1	RAID 0	Excellent	No	100%
2	None	Excellent	No	100%
2	RAID 0	Excellent	No	100%
2	RAID 1	Good	Yes	50%
3	None	Excellent	No	100%
3	RAID 0	Excellent	No	100%
3	RAID 3	Good	Yes	67%
3	RAID 5	Good	Yes	67%
4	None	Excellent	No	100%
4	RAID 0	Excellent	No	100%
4	RAID 3	Good	Yes	75%
4	RAID 5	Good	Yes	75%
4	RAID 10	Good	Yes	50%
5	None	Excellent	No	100%
5	RAID 0	Excellent	No	100%
5	RAID 3	Good	Yes	80%
5	RAID 5	Good	Yes	80%
6	None	Excellent	No	100%
6	RAID 0	Excellent	No	100%
6	RAID 3	Good	Yes	83%
6	RAID 5	Good	Yes	83%
6	RAID 10	Good	Yes	50%
6	RAID 30	Good	Yes	67%
6	RAID 50	Good	Yes	67%
7	None	Excellent	No	100%
7	RAID 0	Excellent	No	100%
7	RAID 3	Good	Yes	86%
7	RAID 5	Good	Yes	86%
8	None	Excellent	No	100%
8	RAID 0	Excellent	No	100%
8	RAID 3	Good	Yes	87%
8	RAID 5	Good	Yes	87%
8	RAID 10	Good	Yes	50%
8	RAID 30	Good	Yes	75%
8	RAID 50	Good	Yes	75%

6 Hardware Installation

Requirements You must have the following items before installing the MegaRAID controller in a server:

- a MegaRAID Elite 1600 64-Bit 160M RAID Controller
- a host computer with an available PCI expansion slot
- the MegaRAID Elite 1600 Installation CD
- the necessary SCSI cables and terminators (depends on the number and type of SCSI devices to be attached)
- an Uninterruptible Power Supply (UPS) for the entire system
- 160M SCSI hard disk drives and other SCSI devices, as desired

> ***Important***
>
> The MegaRAID Elite 1600 controller must be installed in a PCI expansion slot.

Optional Equipment You may also want to install SCSI cables that interconnect MegaRAID Elite 1600 to external SCSI devices.

Checklist

Perform the steps in the installation checklist:

Check	Step	Action
	1	Turn all power off to the server and all hard disk drives, enclosures, and system, components.
	2	Prepare the host system. See the host system technical documentation.
	3	Determine the SCSI ID and SCSI termination requirements.
	4	Make sure the jumper settings on the MegaRAID controller are correct. Install the cache memory.
	5	Connect the battery pack harness on Series 492 or Series 495 (optional.)
	6	Install the MegaRAID card in the server and attach the SCSI cables and terminators as needed. Make sure Pin 1 on the cable matches Pin 1 on the connector. Make sure that the SCSI cables you use conform to all SCSI specifications.
	7	Perform a safety check. Make sure all cables are properly attached. Make sure the MegaRAID card is properly installed. Turn power on after completing the safety check. Connect the battery pack.
	8	Install and configure the MegaRAID software utilities and drivers.
	9	Format the hard disk drives as needed.
	10	Configure system drives (logical drives).
	11	Initialize the logical drives.
	12	Install the appropriate MegaRAID drivers for your operating system.

Installation Steps

MegaRAID provides extensive customization options. If you need only basic MegaRAID features and your computer does not use other adapter cards with resource settings that may conflict with MegaRAID settings, even custom installation can be quick and easy.

Step	Action	Additional Information
1	Unpack the MegaRAID controller and inspect for damage. Make sure all items are in the package.	If damaged, call LSI Logic Technical Support at 678-728-1250.
2	Turn the computer off and remove the cover.	
3	Make sure the motherboard jumper settings are correct.	
4	Install cache memory on the MegaRAID card.	16 MB minimum cache memory is required.
5	Check the jumper settings 16 – 32 on the MegaRAID controller.	See page 55 for the MegaRAID jumper settings.
6	Set SCSI termination.	
7	Set SCSI terminator power (TermPWR).	
8	Connect the battery backup.	Optional
9	Install the MegaRAID card.	
10	Connect the SCSI cables to SCSI devices.	
11	Set the target IDs for the SCSI devices.	Do not use the same target ID for different SCSI devices. The devices should be non-cluster. Do not use 7 as the target ID.
12	Replace the computer cover and turn the power on.	Be sure the SCSI devices are powered up before or at the same time as the host computer.
13	Run MegaRAID Configuration Utility.	Optional.
14	Install software drivers for the desired operating systems.	

Each step is described in detail in the following pages.

Step 1 Unpack

Unpack and install the hardware in a static-free environment. The MegaRAID controller card is packed inside an anti-static bag between two sponge sheets. Remove the controller card and inspect it for damage. If the card appears damaged, or if any of items listed below are missing, contact LSI Logic Technical Support at 678-728-1250. The MegaRAID Controller is also shipped with the following items that are on CD:

- the *MegaRAID Configuration Software Guide,*
- the *MegaRAID Operating System Drivers Guide*
- the *MegaRAID Elite 1600 Hardware Guide*
- the software license agreement
- the MegaRAID Configuration Utilities for DOS
- the warranty registration card

Step 2 Power Down

Turn off the computer and remove the cover. Make sure the computer is turned off and disconnected from any networks before installing the controller card.

Step 3 Configure Motherboard

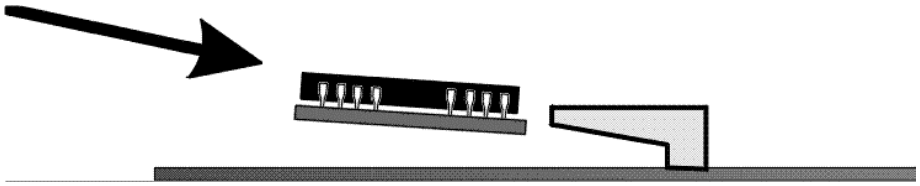
Make sure the motherboard is configured correctly for MegaRAID. MegaRAID is essentially a SCSI Controller. Each MegaRAID card you install will require an available PCI IRQ; make sure an IRQ is available for each controller you install.

Step 4 Install Cache Memory

> ***Important***
>
> *A minimum of 16 MB of cache memory is required. The cache memory must be installed before MegaRAID is operational.*

Memory Specifications Insert one in the cache memory socket.

DIMM Specifications Install cache memory DIMMs on the MegaRAID controller card in the cache memory socket. Use a 64-bit 3.3V single-sided or double-sided 168-pin unbuffered DIMM. Lay the controller card component-side up on a clean static-free surface. The memory socket is mounted flush with the MegaRAID card, so the DIMM is parallel to the MegaRAID card when properly installed. The DIMM clicks into place, indicating proper seating in the socket. The MegaRAID card is shown lying on a flat surface below.



Cont'd

Step 4 Install Cache Memory, Continued

Installing or Changing Memory Perform the following steps to install cache memory.

Step	Action
1	Bring down the operating system properly. Make sure that cache memory has been flushed. You must perform a system reset if operating under DOS. When the computer reboots, the MegaRAID controller will flush cache memory.
2	Turn the computer power off. Disconnect the power cables from the computer.
3	Remove the computer cover.
4	Disconnect the Series 492 battery backup card from the MegaRAID controller. **Note:** If you use a Series 495 memory module (it contains battery backup), you do not need to use the 492 battery backup card.
5	Remove the MegaRAID controller.
6	You can now add or remove DRAM modules from the MegaRAID controller. Follow the instructions on page 53.
7	Reinstall the MegaRAID controller in the computer. Follow the instructions in this chapter.
8	Reattach the Series 492 battery backup card to the MegaRAID controller.
9	Replace the computer cover and turn the computer power on.

Recommended Memory Vendors Call LSI Logic Technical Support at 678-728-1250 for a current list of recommended memory vendors.

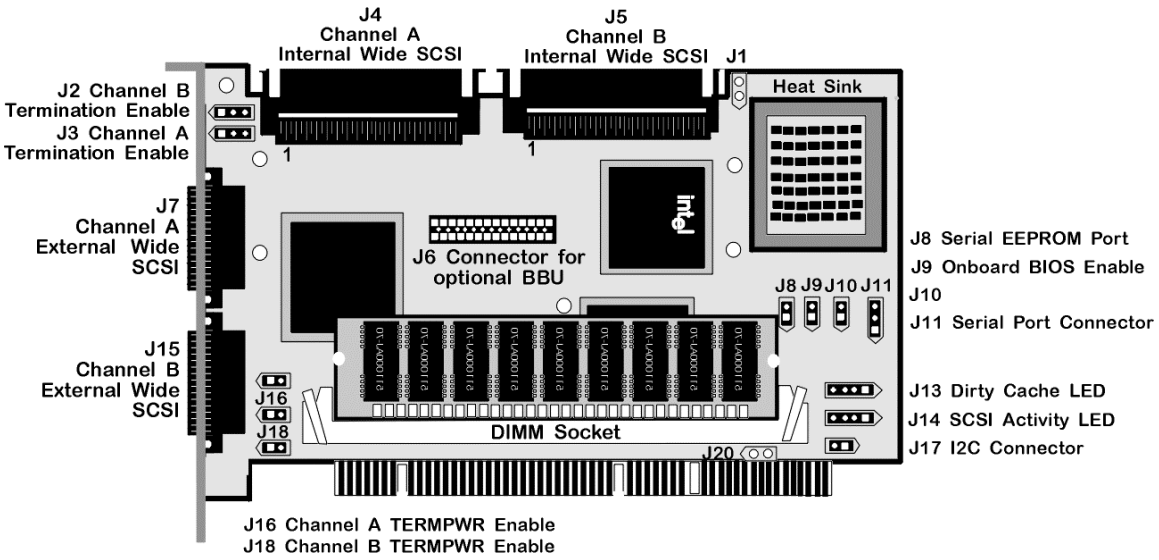
Step 5 Set Jumpers

Make sure the jumper settings on the MegaRAID card are correct. The jumpers and connectors are:

Connector	Description	Type
J2	Channel B Termination Enable	3-pin header
J3	Channel A Termination Enable	3-pin header
J4	Channel A Internal Wide SCSI	68-pin connector
J5	Channel B Internal Wide SCSI	68-pin connector
J6	Connector for optional BBU (battery backup unit)	28-pin connector
J7	Channel A External Wide SCSI	68-pin connector
J8	Serial EEPROM Port	2-pin header
J9	Onboard BIOS Enable	2-pin header
J11	Serial port connector	3-pin header
J13	Dirty Cache (Write Pending) LED	2-pin header
J14	SCSI activity LED	4-pin header
J15	Channel B External Wide SCSI	68-pin connector
J16	Channel A TERMPWR Enable	2-pin header
J17	I2C connector	4-pin header
J18	Channel B TERMPWR Enable	2-pin header

Cont'd

MegaRAID Elite 1600 64-Bit 160M Card Layout



Cont'd

Step 5 Set Jumpers, Continued

J2, and J3 Termination Enable J2, and J3 are 3-pin bergs that set the SCSI termination for each SCSI channel:

Jumper	SCSI Channel	SCSI Termination Controlled by Software	SCSI Termination Always Disabled	SCSI Termination Always Enabled
J2	A	Short Pins 1-2	Short Pins 2-3	OPEN
J3	B	Short Pins 1-2	Short Pins 2-3	OPEN

J16, and J18 TERMPWR Enable J16, and J18 are 2-pin bergs that enable TERMPWR to the SCSI bus for each SCSI channel:

Jumper	Term. Power Channel	Settings
J16	A	Short Pins 1-2 to have the PCI bus on the host computer provide TermPWR. This is the factory setting. Leave Open to let the SCSI bus provide TermPWR.
J18	B	Short Pins 1-2 to have the PCI bus on the host computer provide TermPWR. This is the factory setting. Leave Open to let the SCSI bus provide TermPWR.

J6 Connector for optional BBU J6 is 28-pin connector that is used to mount the Series 492 battery backup card. Battery backup is optional. An alternative to using Series 492 for battery backup is to use a Series 495 battery-backed DIMM.

Cont'd

J11 Serial Port J11 is a 3-pin header that attaches to a serial cable. The pinout is:

Pin	Signal Description
1	Receive Data
2	Transmit Data
3	Ground

J9 Onboard BIOS Enable J9 is a 2-pin berg that enables or disables MegaRAID onboard BIOS. The onboard BIOS should be enabled (J9 unjumpered) for normal board position.

J9 Setting	Onboard BIOS Status
Unjumpered	Enabled
Jumpered	Disabled

J13 Dirty Cache LED J13 is a two-pin connector for an LED mounted on the computer enclosure. The LED indicates when the data in the cache has yet to be written to the storage devices.

Pin	Description
1	High
2	Dirty Cache Signal

J14 SCSI Activity LED J14 is a four-pin connector for an LED mounted on the computer enclosure.

Pin	Description
1	VCC
2	SCSI Activity Signal
3	SCSI Activity Signal
4	VCC

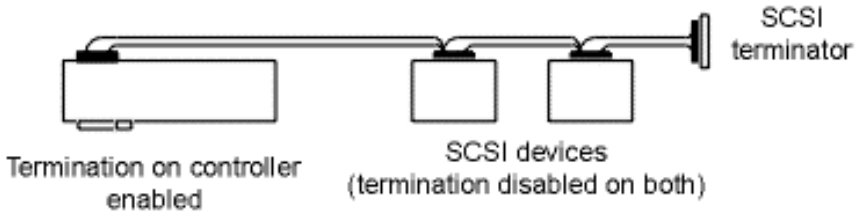
J17 I2C connector J17 is a 4-pin header.

Pin	Description
1	Data
2	GND
3	Clock
4	Power (fused)

Step 6 Set Termination

Each MegaRAID SCSI channel can be individually configured for termination enable mode by setting the J2, and J3 jumpers (see page 57 for jumper information).

You must terminate the SCSI bus properly. Set termination at both ends of the SCSI cable. The SCSI bus is an electrical transmission line and must be terminated properly to minimize reflections and losses. Termination should be set at each end of the SCSI bus, as shown below.



Setup using one connector for one channel

For a disk array, set SCSI bus termination so that removing or adding a SCSI device does not disturb termination. An easy way to do this is to:

- Connect the MegaRAID card to one end of the SCSI cable for each channel.
- Connect an external terminator module at the other end of each cable

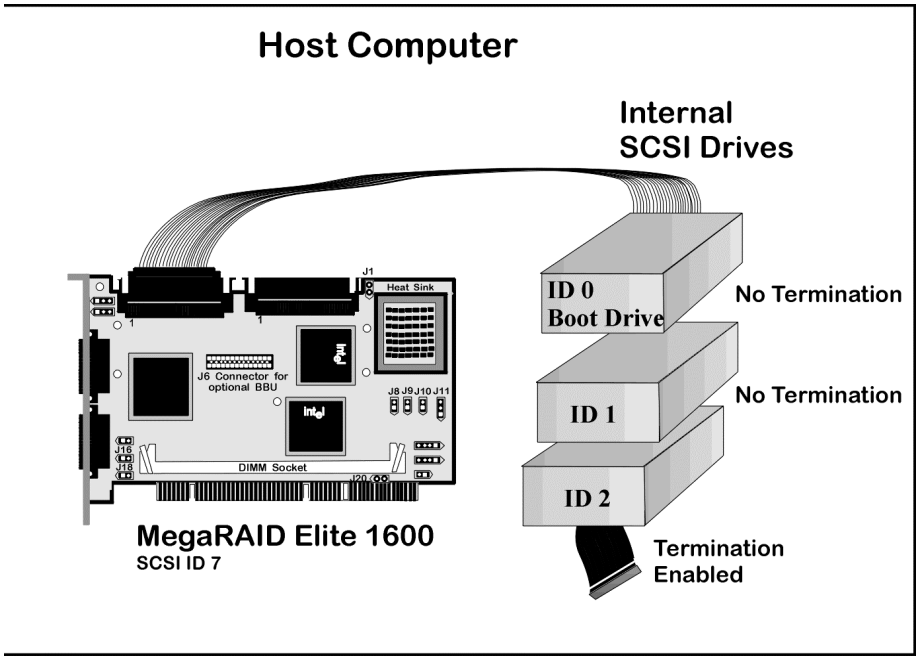
The connectors between the two ends can connect SCSI devices. Disable termination on the SCSI devices. See the manual for each SCSI device to disable termination.

SCSI Termination

The SCSI bus on a SCSI channel is an electrical transmission line. It must be terminated properly to minimize reflections and losses. You complete the SCSI bus by setting termination at both ends. MegaRAID automatically provides SCSI termination at one end of the SCSI bus for each channel. Terminate the other end of the bus by attaching an external SCSI terminator module to the end of the cable for each channel or by attaching a SCSI device that internally terminates the SCSI bus at the end of each SCSI channel.

Use standard external SCSI terminators on SCSI channels operating at 10 MB/s or higher synchronous data transfer. Termination should support the highest speed available to the bus.

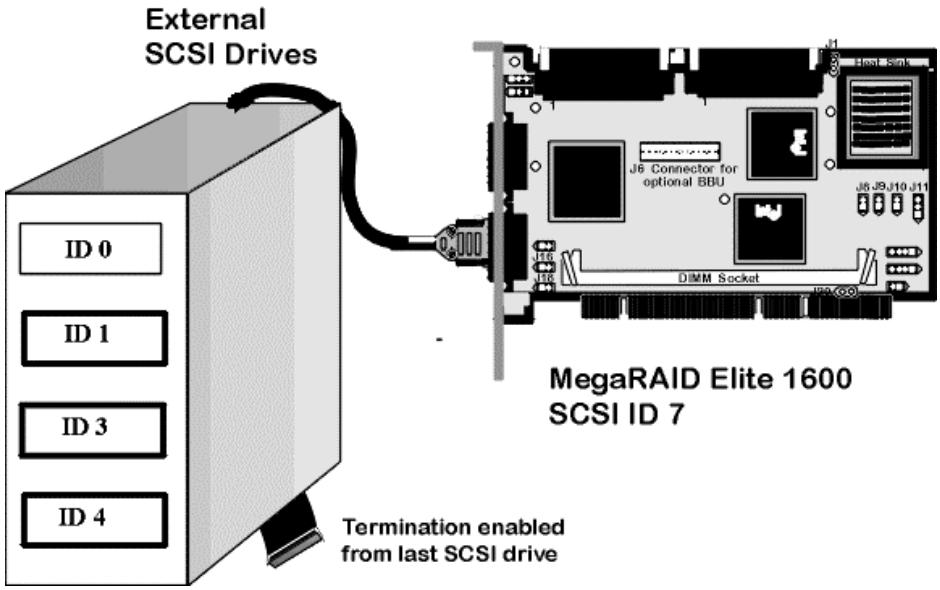
Terminating Internal SCSI Disk Arrays Set the termination so that SCSI termination and termination power are intact when any disk drive is removed from a SCSI channel, as shown below. MegaRAID termination should always be enabled or controlled by software. Make sure J2, and J3 are either always open (termination always enabled), or Pins 1-2 are shorted (termination controlled by software).



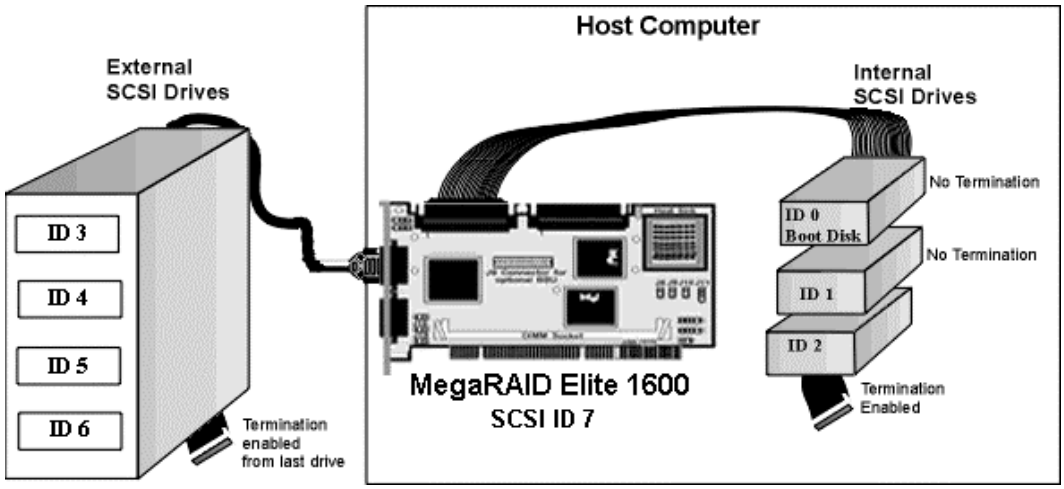
Cont'd

SCSI Termination, Continued

Terminating External Disk Arrays In most array enclosures, the end of the SCSI cable has an independent SCSI terminator module that is not part of a SCSI drive. In this way, SCSI termination is not disturbed when a drive is removed. MegaRAID termination should always be enabled or controlled by software. Make sure J2, and J3 are either always open (termination always enabled), or Pins 1-2 are shorted (termination controlled by software).



Terminating Internal and External Disk Arrays You can use both internal and external drives with MegaRAID. You still must make sure that the proper SCSI termination and termination power is preserved. MegaRAID termination should always be enabled or controlled by software.

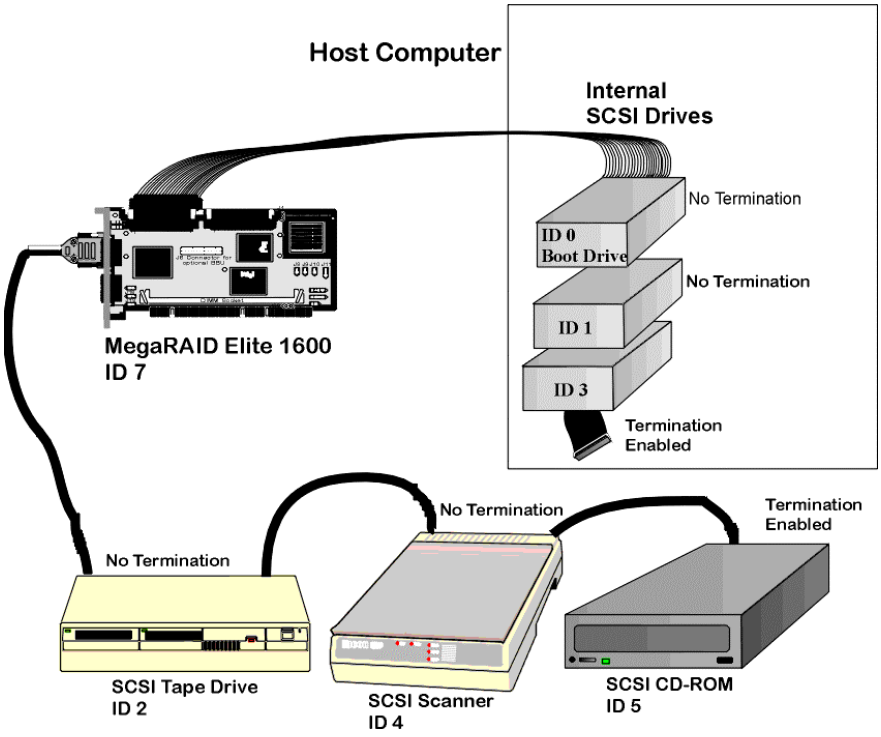


Cont'd

Connecting Non-Disk SCSI Devices SCSI Tape drives, scanners, CD-ROM drives, and other non-disk drive devices must each have a unique SCSI ID regardless of the SCSI channel they are attached to. The general rule for operating systems is:

- tape drive set to SCSI ID 2
- CD-ROM drive set to SCSI ID 5
- all non-disk SCSI devices attached to SCSI channel A

Make sure that no hard disk drives are attached to the same SCSI channel as the non-disk SCSI devices. Drive performance will be significantly degraded if SCSI hard disk drives are attached to this channel.



Step 7 Set SCSI Terminator Power

J16, and J18 These jumpers control TermPWR for the MegaRAID SCSI channels. See the documentation for each SCSI device for information about enabling TermPWR. The factory settings supply TermPWR from the PCI bus.

Important

The SCSI channels need Termination power to operate. If a channel is not being used, make sure the jumper setting for that channel is set to supply TermPWR from the PCI bus.

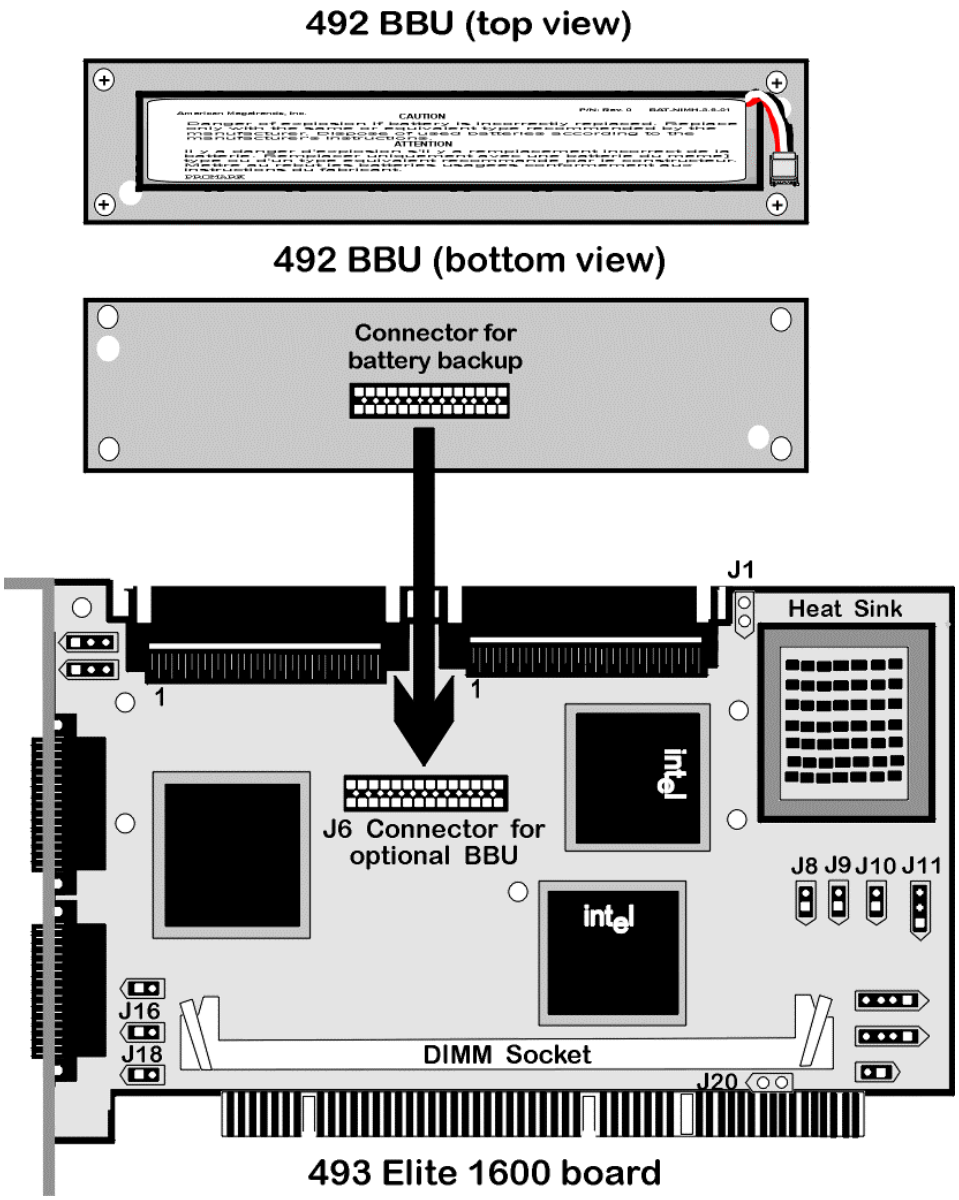
J16 SCSI Channel A – Short Pins 1-2 for PCI power.
J18 SCSI Channel B – Short Pins 1-2 for PCI power.

Step 8 Install Battery Pack (Optional)

You can install a battery pack onto the Series 493 MegaRAID Elite 1600 160M RAID controller. There are two ways to install a battery pack on the controller:

- install a Series 492 battery backup board on the Elite 1600 board
- install a Series 495 battery backed DIMM module into the DIMM socket

Series 492 battery backup You can connect the 492 battery backup card to the Elite 1600 board to provide up to 72-hour battery backup in case of power failure. The Elite 1600 board has a 28-pin connector, J6, which accepts the BBU board, as shown below:



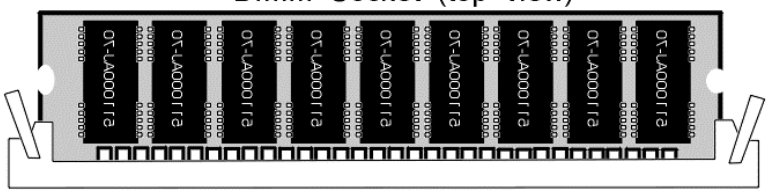
Cont'd

Step 8 Install Battery Pack (Optional), Continued

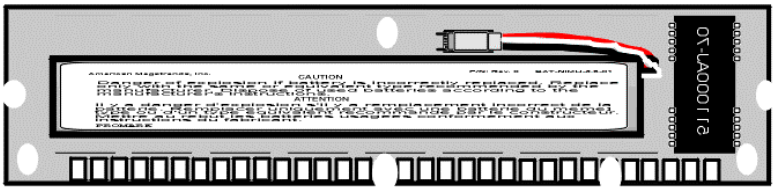
DIMM with battery backup (Series 495) The battery pack is shown in the bottom view of the DIMM socket below. Connect the 3-wire cable from the battery pack to the connector as shown below. Pin 1 on the cable from the battery pack is usually denoted by a red wire. The caution information is also shown below.

Install the DIMM as described in Step 4 on page 53.

DIMM Socket (top view)



DIMM Socket (bottom view)



CAUTION

Danger of explosion if battery is incorrectly replaced. Replace only with the same or equivalent type recommended by the manufacturer. Dispose of used batteries according to the manufacturer's instructions.

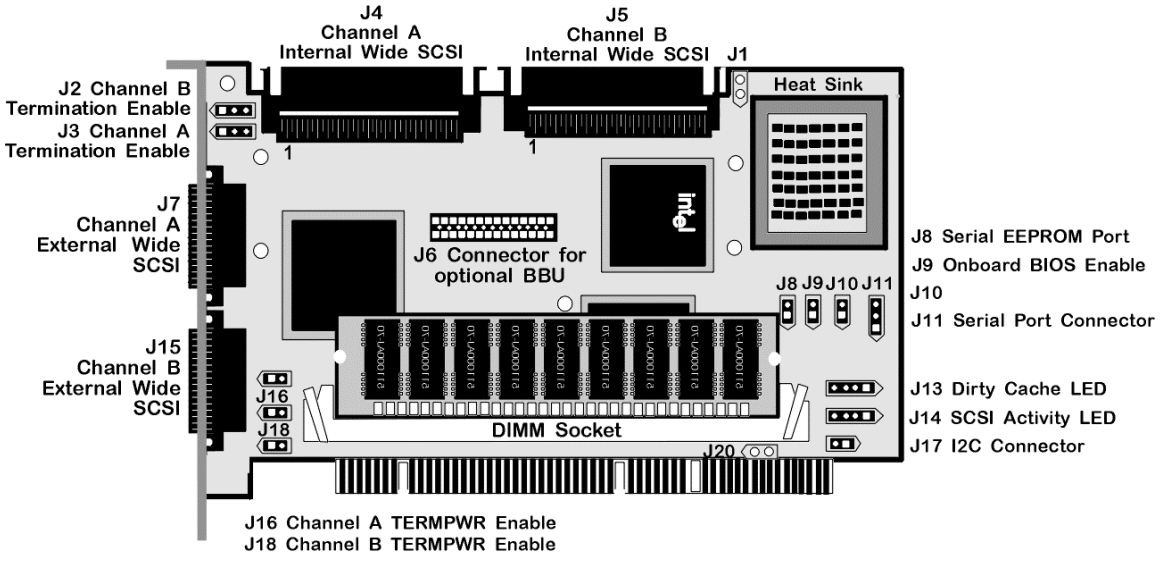
ATTENTION

Il y a danger d'explosion s'il y a remplacement incorrect de la batterie. Remplacer uniquement avec une batterie du meme] type ou d'un type equivalent recommande par le constructeur. Mettre au rebut les batteries usagees conformement aux instructions du fabricant.

Cont'd

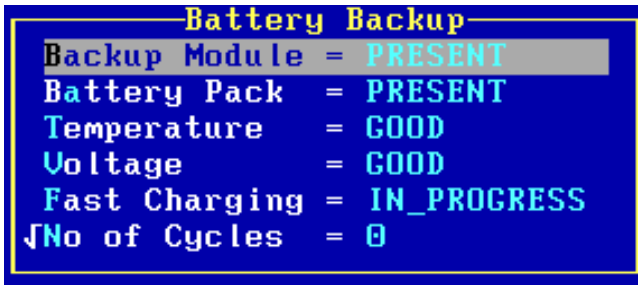
Step 8 Install Battery Pack (Optional), Continued

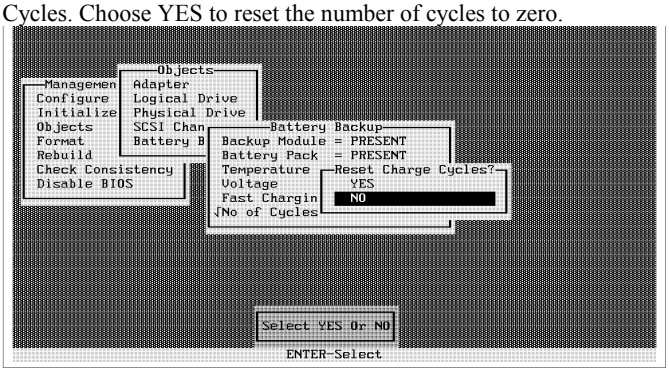
Board with battery A drawing of the MegaRAID Elite 1600 160M RAID Controller with a battery backed DIMM (Series 495) is shown below. The battery pack is on the backside of the DIMM in the DIMM socket.



Cont'd

Configure Battery Backup After installing the MegaRAID controller and booting, press <Ctrl> <M>. Choose the Objects menu. Select Battery Backup. The battery backup menu displays, as shown below:



Menu Item	Explanation
Battery Pack	PRESENT will appear if the battery pack is properly installed; ABSENT if it is not.
Temperature	GOOD appears if the temperature is within the normal range. HIGH appears if the module is too hot.
Voltage	GOOD appears if the voltage is within the normal range. BAD appears if the voltage is out of range.
Fast Charging	COMPLETED appears if the fast charge cycle is done. CHARGING appears if the battery pack is charging.
No. of Cycles	This must be configured. When first installing a battery pack, set the Charge Cycle to 0. The screen below appears when you select No. of Cycles. Choose YES to reset the number of cycles to zero.  After 1100 charge cycles, the life of the battery pack is assumed to be over and you must replace the battery pack.

Cont'd

Step 8 Install Battery Pack (Optional), Continued

Charging the Battery Pack The battery pack is shipped uncharged. You must charge the battery pack before you can use it. The minimum time that the battery must be charged is:

LSI Logic Part Number	Description	Time to Charge
BAT-NIMH-3.6-01	Battery, NIMH, 3.6V, 650mA onboard battery pack with mounting brackets	4 hours

The full data retention time is not available until the battery pack is fully charged. It is a good idea to set the MegaRAID controller cache write policy option *to Write-Through* during the battery pack charging period. After the battery pack is fully charged, you can change the cache write policy to *Write-Back*.

Cont'd

Step 8 Install Battery Pack (Optional), Continued

Changing the Battery Pack The MegaRAID configuration software warns when the battery pack must be replaced. A new battery pack should be installed every 1 to 5 years.

Step	Action
1	Bring down the operating system properly. Make sure that cache memory has been flushed. You must perform a system reset if operating under DOS. When the computer reboots, the MegaRAID Elite 1600 160M controller flushes cache memory. Turn the computer power off. Remove the computer cover. Remove the MegaRAID controller.
2	Unscrew the screws holding the battery pack on the back of the Series 492 battery backup card or the back of the Series 495 battery backup DIMM.
3	Install a new battery pack and connect the new battery pack to the connector.
4	Disable write-back caching using MegaRAID Manager or Power Console Plus.

Disposing of a Battery Pack

Warning

Do not dispose of the MegaRAID battery pack by fire. Do not mutilate the battery pack. Do not damage it in any way. Toxic chemicals can be released if it is damaged. Do not short-circuit the battery pack.

The material in the battery pack contains heavy metals that can contaminate the environment. Federal, state, and local laws prohibit disposal of some rechargeable batteries in public landfills. These batteries must be sent to a specific location for proper disposal. Call the Rechargeable Battery Recycling Corporation at 352-376-6693 (FAX: 352-376-6658) for an authorized battery disposal site near you. For a list of battery disposal sites, write to:

Rechargeable Battery Recycling Corporation
2293 NW 41st Street
Gainesville FL 32606
Voice: 352-376-6693
FAX: 352-376-6658

Cont'd

Battery Disposal Laws

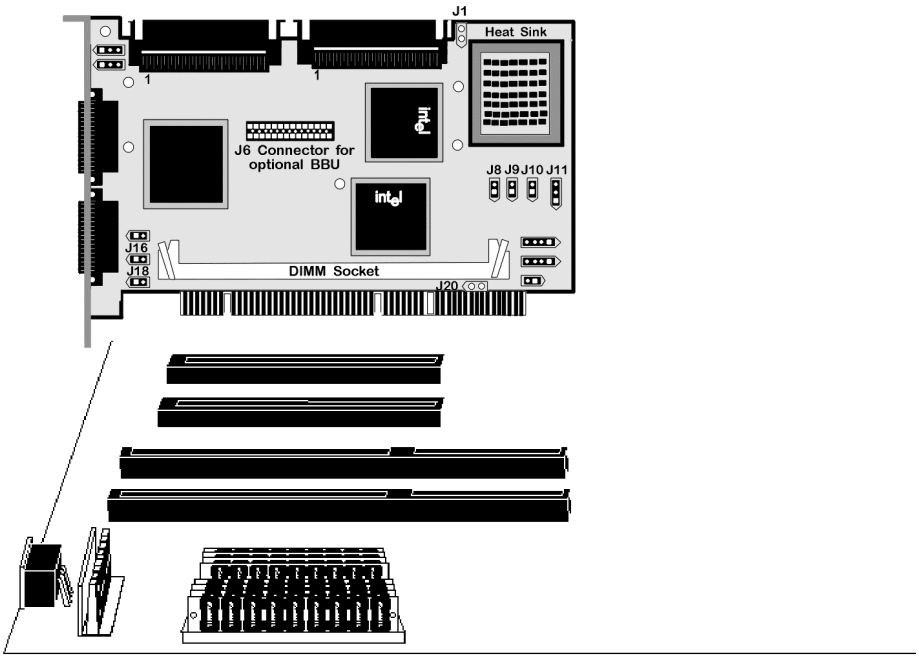
> ***Important***
>
> *Most used Nickel-Metal Hydride batteries are not classified as hazardous waste under the federal RCRA (Resource Conservation and Recovery Act). Although Minnesota law requires that Nickel-Metal Hydride batteries be labeled "easily removable" from consumer products, and that Nickel-Metal Hydride batteries must be collected by manufacturers, the Minnesota Pollution Control Agency (MPCA) has granted a temporary exemption from these requirements.*

Other Laws in Other Areas LSI Logic reminds you that you must comply with all applicable battery disposal and hazardous material handling laws and regulations in the country or other jurisdiction where you are using an optional battery pack on the MegaRAID Elite 1600 160M controller.

Step 9 Install MegaRAID Card

The MegaRAID card can plug into a 32-bit or 64-bit PCI slot that receives 5 V, and, optionally, 3.3 V through the motherboard. Choose a PCI slot and align the MegaRAID controller card edge connector to the slot. Note the key on the PCI slot and look for any obstructions. Press down gently, but firmly to make sure that the card is properly seated in the slot. The bottom edge of the controller card should be flush with the slot.

Insert the MegaRAID card in a PCI slot as shown below:



Screw the bracket to the computer frame.

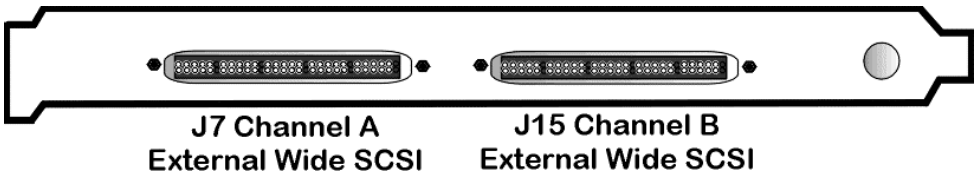
Step 10 Connect SCSI Cables

SCSI Connectors Connect the SCSI cables to the SCSI devices. MegaRAID provides two types of SCSI connectors:

- external
- internal

External Connectors J7 provides one ultra high density external connector for SCSI channels A.

J15 provides one ultra high density connector for SCSI channels B.



Internal Connectors Internal connectors are provided for channels A and B only.

J4 is the internal connector for channel A.

J1 is the internal connector for channel B.

See the board layout for the location of J4 and J1.

Cont'd

Step 10 Connect SCSI Cables, Continued

J7 Channel A External Connector J7 is a 68-pin ultra-high density external SCSI connector. It is on the MegaRAID mounting bracket.

J15 Channel B External Connector J15 is a 68-pin ultra-high density external SCSI connector. It is on the MegaRAID mounting bracket.

Connect SCSI Devices When connecting SCSI devices:

Action	Description
1	Disable termination on any SCSI device that does *not* sit at the end of the SCSI bus.
2	Configure all SCSI devices to supply TermPWR.
3	Set proper target IDs (TIDs) for all SCSI devices.
4	Distribute SCSI devices evenly across the SCSI channels for optimum performance.
5	The cable length should not exceed three meters for Fast SCSI (10 MB/s) devices or 1.5 meters for Ultra SCSI devices.
6	The cable length should not exceed six meters for non-Fast SCSI devices.
7	Try to connect all non-disk SCSI devices to a SCSI channel that has no SCSI disk drives connected to it.

Cable Suggestions System throughput problems can occur if the SCSI cable used is not optimized. You should:

- use the shortest SCSI cables (in SE mode, no more than 3 meters for Fast SCSI, no more than 1.5 meters for an 8-drive Ultra SCSI system and no more than 3 meters for a 6-drive Ultra SCSI system)
- note that LVD mode cable lengths should be no more than 25 meters with two devices and no more than 12 meters with 15 devices
- use active termination
- avoid clustering the stubs
- cable stub length should be no more than 0.1 meter (4 inches), depending on impedance of the cable
- route SCSI cables carefully
- use high impedance cables
- do not mix cable types (choose either flat or rounded and shielded or non-shielded)
- note that ribbon cables have fairly good cross-talk rejection characteristics

Step 11 Set Target IDs

Set target identifiers (TIDs) on the SCSI devices. Each device in a specific SCSI channel must have a unique TID in that channel. Non-disk devices (CD-ROM or tapes) should have unique SCSI IDs *regardless of the channel where they are connected*. See the documentation for each SCSI device to set the TIDs.

The MegaRAID controller automatically occupies TID 7 in each SCSI channel. Eight-bit SCSI devices can use the TIDs from 0 to 6 only. 16-bit devices can use the TIDs from 0 to 15. The arbitration priority for a SCSI device depends on its TID.

Priority	Highest											Lowest
TID	7	6	5	…	2	1	0	15	14	…	9	8

Important

Non-disk devices (CD-ROM or tapes) should have unique SCSI IDs regardless of the channel they are connected to. ID 0 cannot be used for non-disk devices because they are limited to IDs 1 through 6. There is a limit of six IDs for non-disk devices per controller.

Device Identification on MegaRAID Controllers

Example of MegaRAID ID Mapping

ID	Channel A	Channel B
0	A1-1	A1-2
1	A2-1	Scanner
2	CD	A2-3
3	A2-5	A2-6
4	CD	A3-1
5	A4-1	Tape
6	Optical	A5-1
7	Reserved	Reserved
8	A5-2	A5-3
9	A5-6	A5-7
10	A6-1	A6-2
11	A6-4	A6-5
12	A6-7	A6-8
13	A7-2	A7-3
14	A7-5	A7-6
15	A7-8	A8-1

As Presented to the Operating System

ID	LUN	Device	ID	LUN	Device
0	0	Disk (A1-X)	1	0	Scanner
0	1	Disk (A2-X)	2	0	CD
0	2	Disk (A3-X)	3	0	Tape
0	3	Disk (A4-X)	4	0	CD
0	4	Disk (A5-X)	5	0	Tape
0	5	Disk (A6-X)	6	0	Optical
0	6	Disk (A7-X)			
0	7	Disk (A8-X)			

Step 12 Power Up

Replace the computer cover and reconnect the AC power cords. Turn power on to the host computer. Set up the power supplies so that the SCSI devices are powered up at the same time as or before the host computer. If the computer is powered up before a SCSI device, the device might not be recognized.

During boot, the MegaRAID BIOS message appears:

```
MegaRAID Elite 1600 Disk Array Adapter BIOS Version x.xx date
Copyright (c) LSI Logic Corporation
Firmware Initializing... [ Scanning SCSI Device ...(etc.)... ]
```

The firmware takes several seconds to initialize. During this time the adapter will scan each SCSI channel. When it is ready, the following lines appear:

```
Host Adapter-1 Firmware Version x.xx DRAM Size 16 MB
0 Logical Drives found on the Host Adapter
0 Logical Drives handled by BIOS
Press <Ctrl><M> to run MegaRAID Elite BIOS Configuration Utility
```

The <Ctrl> <M> prompt times out after several seconds.

The MegaRAID Elite 1600 host adapter (controller) number, firmware version, and cache DRAM size are displayed in the second portion of the BIOS message. The numbering of the controllers follows the PCI slot scanning order used by the host motherboard.

Step 13 Run MegaRAID Configuration Utility

Press <Ctrl> <M> to run the MegaRAID Configuration Utility. See the *MegaRAID Configuration Software Guide* for information about running MegaRAID Configuration Utility.

Step 14 Install the Operating System Driver

> ***Important***
>
> *When booting the system from a drive connected to a MegaRAID controller and using EMM386.EXE, MEGASPI.SYS must be loaded in CONFIG.SYS before EMM386.EXE is loaded. If you do not do this, you cannot access the boot drive after EMM386 is loaded.*

DOS ASPI Driver The MegaRAID DOS ASPI driver can be used under DOS, Windows 3.*x*, and Windows 95. The DOS ASPI driver supports:

- up to six non-disk SCSI devices (each SCSI device must use a unique SCSI ID regardless of the SCSI channel it resides on. SCSI IDs 1 through 6 are valid.)
- up to six MegaRAID adapters (you should configure only one MegaRAID adapter per system if possible)

ASPI Driver The ASPI driver is MEGASPI.SYS. It supports disk drives, tape drives, CD-ROM drives, etc. You can use it to run CorelSCSI, Novaback, PC Tools, and other software that requires an ASPI driver. *CorelSCSI, Novaback, and PC Tools are not provided with MegaRAID.* Copy MEGASPI.SYS to your hard disk drive. Add the following line to CONFIG.SYS. *MEGASPI.SYS must be loaded in CONFIG.SYS before EMM386.EXE is loaded.*

```
device=<path>\MEGASPI.SYS /v
```

Parameters The MEGASPI.SYS parameters are:

Parameter	Description
/h	INT 13h support is not provided.
/v	Verbose mode. All message are displayed on the screen.
/a	Physical drive access mode. Permits direct access to physical drives.
/q	Quiet mode. All message except error message are suppressed.

Cont'd

Step 14 Install Operating System Driver, Continued

CD-ROM Driver A device driver is provided with MegaRAID for CD-ROM drives operating under DOS, Windows 3.*x*, and Windows 95. The driver filename is AMICDROM.SYS.

The MEGASPI.SYS ASPI manager must be added to the CONFIG.SYS file before you can install the CD-ROM device driver. See the instructions on the previous page for adding the MEGASPI.SYS driver. Copy AMICDROM.SYS to the root directory of the C: drive. Add the following line to CONFIG.SYS, making sure it is preceded by the line for MEGASPI.SYS:

```
DEVICE=C:\AMICDROM.SYS
```

Add the following to AUTOEXEC.BAT. Make sure it precedes the SMARTDRV.EXE line.

```
MSCDEX /D:MSCD001
```

MSCDEX is the CD-ROM drive extension file that is supplied with MS-DOS® and PC-DOS® Version 5.0 or later. See your DOS manual for the command line parameters for MSCDEX.

Summary

This chapter discussed hardware installation. See the MegaRAID Configuration Software Guide for information about the MegaRAID software utilities. You configure the RAID system via software configuration utilities. The utility programs for configuring MegaRAID are:

Configuration Utility	Operating System
MegaRAID Configuration Utility	independent of the operating system
MegaRAID Manager	DOS SCO UNIX SVR3.2 Novell NetWare 3.x, 4.x UnixWare
Power Console	Microsoft Windows NT

7 Cluster Installation and Configuration

Overview This chapter contains the procedures for installing Cluster Service for servers running the Windows 2000 server operating system.

Clusters Physically, a cluster is a grouping of two independent servers that can access the same data storage and provide services to a common set of clients. With current technology, this usually means servers connected to common I/O buses and a common network for client access.

Logically, a cluster is a single management unit. Any server can provide any available service to any authorized client. The servers must have access to the same data and must share a common security model. Again, with current technology, this generally means that the servers in a cluster will have the same architecture and run the same version of the same operating system.

The Benefits of Clusters Clusters provide three basic benefits:

- improved application and data availability
- scalability of hardware resources
- simplified management of large or rapidly growing systems

Software Requirements

The software requirements for cluster installation are:

- MS Windows 2000 Advanced Server or Windows 2000 Datacenter Server must be installed.
- You must use a name resolution method, such as Domain Naming System (DNS), Windows Internet Naming System (WINS), or HOSTS.
- Using a Terminal Server for remote cluster administration is recommended.

Hardware Requirements

The hardware requirements for the Cluster Service node can be found at the following web site: http://www.microsoft.com/windows2000/upgrade/compat/default.asp.

- The cluster hardware must be on the Cluster Service Hardware Compatibility List (HCL). To see the latest version of the Cluster Service HCL, go to the following web site:

  http://www.microsoft.com/hcl/default.asp

  and search using the word "Cluster."

- Two HCL-approved computers, each with the following:

  - A boot disk that has Windows 2000 Advanced Server or Windows 2000 Datacenter Server installed. You cannot put the boot disk on the shared storage bus described below.

  - A separate PCI storage host adapter (SCSI or Fibre Channel) is required for the shared disks. This is along with the boot disk adapter.

  - Each machine in the cluster needs two PCI network adapters.

  - An HCL-approved external disk storage unit connected to all the computers in the cluster. This is used as the clustered disk. RAID (redundant array of independent disks) is recommended for this storage unit.

  - Storage cables are needed to attach the shared storage device to all the computers in the cluster.

  - Make sure that all hardware is identical, slot for slot, card for card, for all nodes. This will make it easier to configure the cluster and eliminate potential compatibility problems.

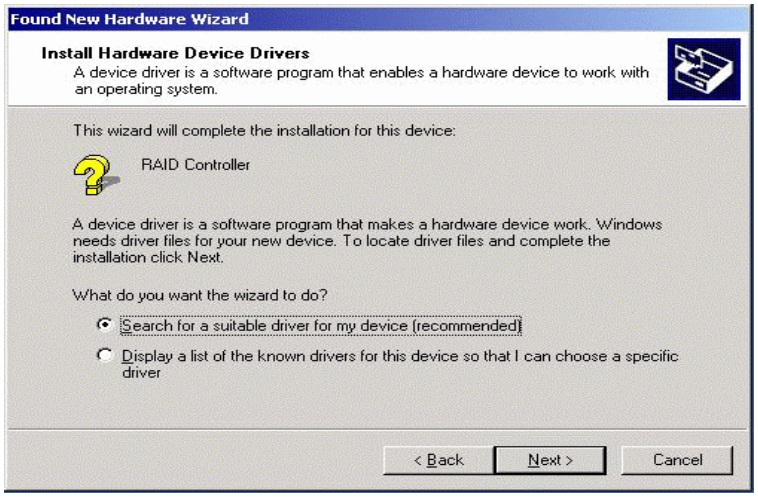
Installation and Configuration

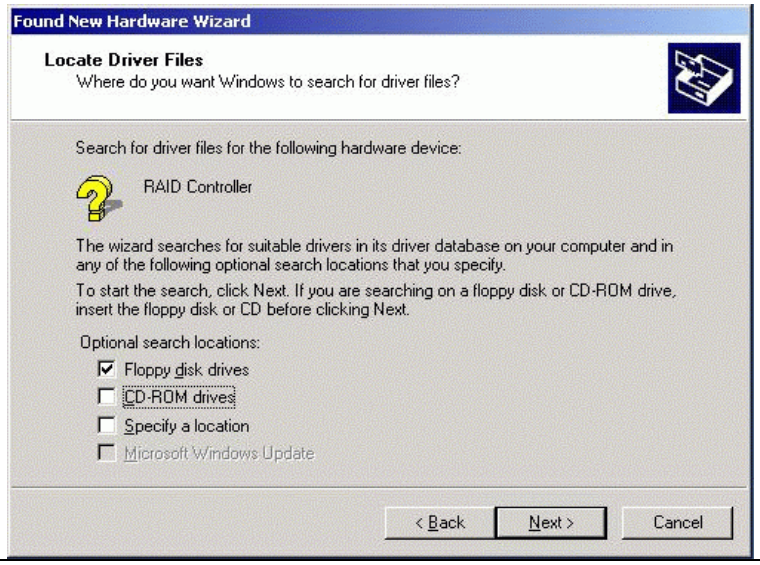
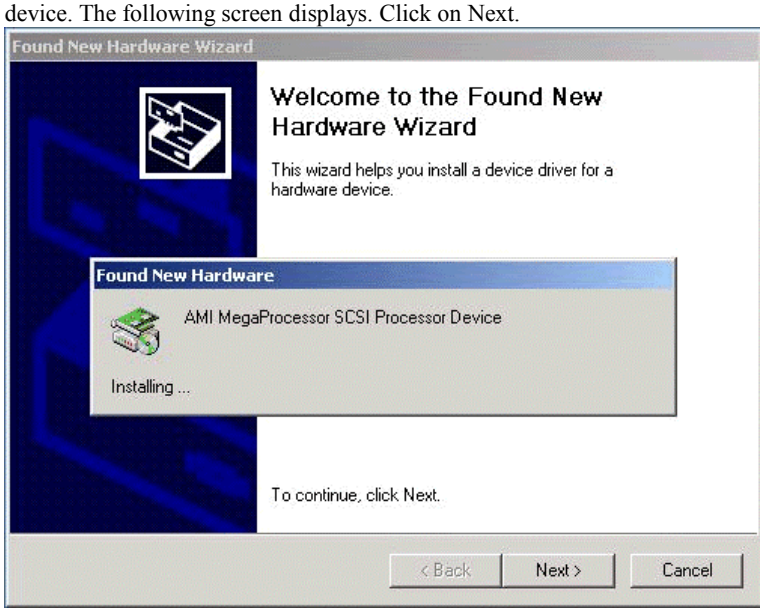
Use the following procedures to install and configure your system as part of a cluster.

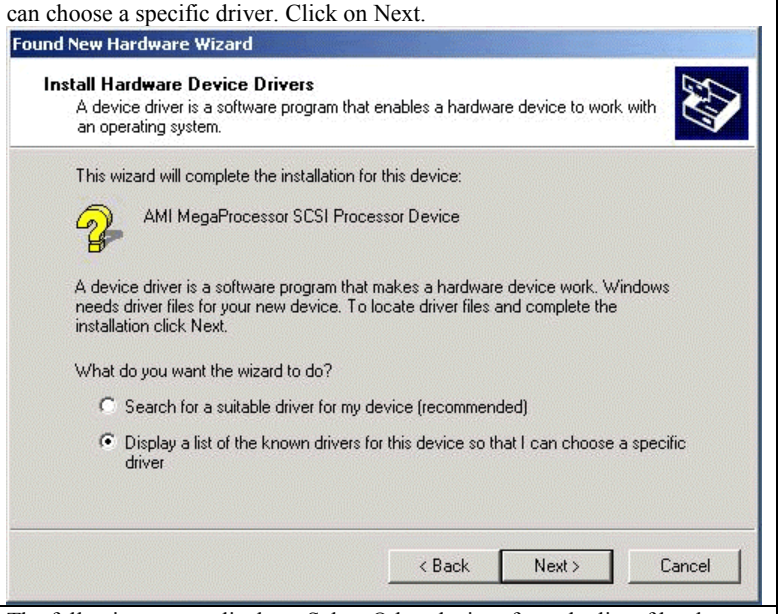
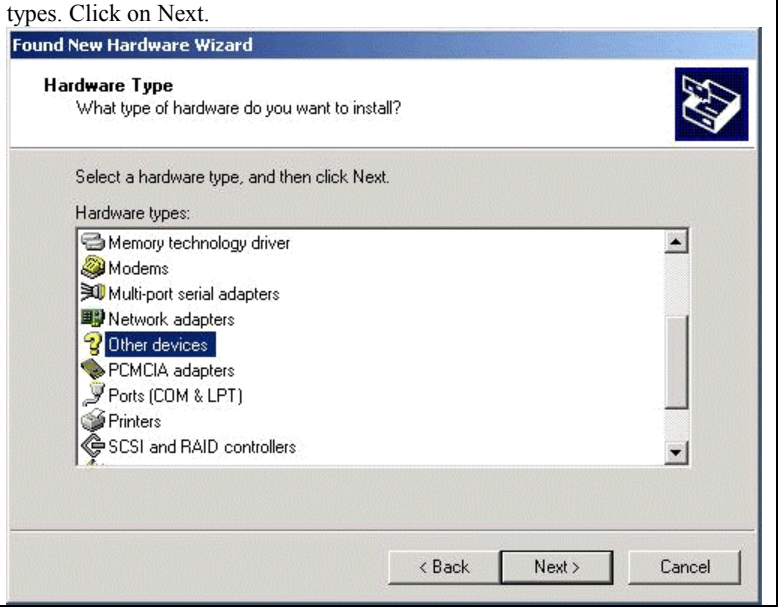
Step	Action
1	Unpack the controller following the instructions on page 52.
2	Set the hardware termination for the controller as “always on”. Refer to the J2, and J3 Termination Enable jumper settings on page 57 for more information.
3	Configure the IDs for the drives in the enclosure. See the enclosure configuration guide for information.
4	Install one controller at a time. Press <Ctrl> <M> at BIOS initialization to configure the options in the steps 5 – 11. Do not attach the disks yet.
5	Set the controller to Cluster Mode in the Objects > Adapter > Cluster Mode menu.
6	Disable the BIOS in the Objects > Adapter > Enable/Disable BIOS menu.
7	Change the initiator ID in the Objects > Adapter > Initiator ID menu.
8	Power down the first system.
9	Attach the controller to the shared array.
10	Configure the first controller to the desired arrays using the Configure > New Configuration menu.
11	Follow the on-screen instructions to create arrays and save the configuration. Initialize the logical drives before powering off the system.
12	Power down the first system.
13	Repeat steps 4 – 7 for the second controller. **Note:** Do not have the cables for the second controller attached to the shared enclosure yet.
14	Power down the second server.
15	Attach the cables for the second controller to the shared enclosure and power up the second system.
16	If a configuration mismatch occurs, enter the <Ctrl> <M> utility. Go to the Configure > View/Add Configuration > View Disk menu to view the disk configuration. Save the configuration.
17	Proceed to the driver installation for a Microsoft cluster environment.

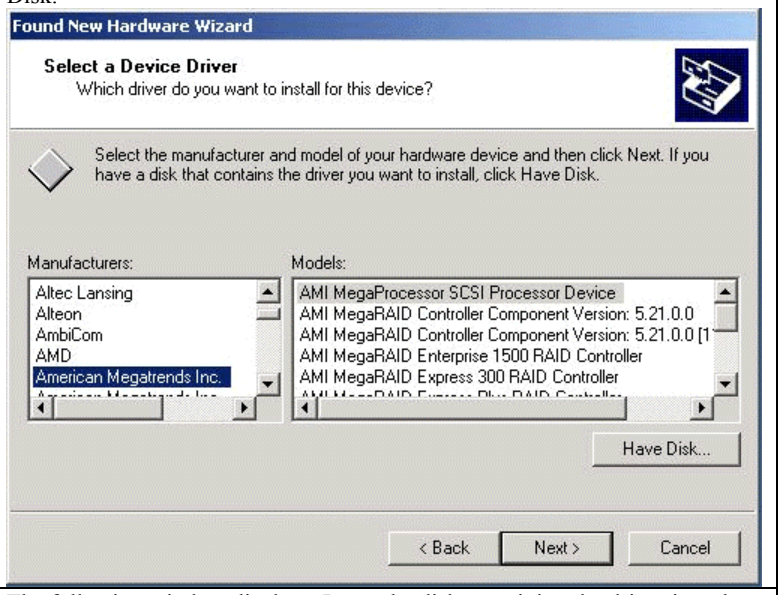
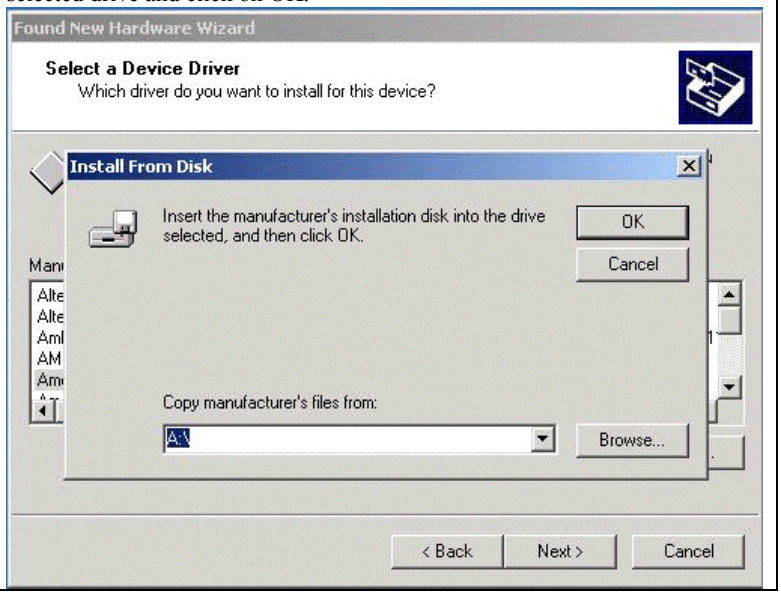
Driver Installation Instructions under Microsoft Windows 2000 Advanced Server

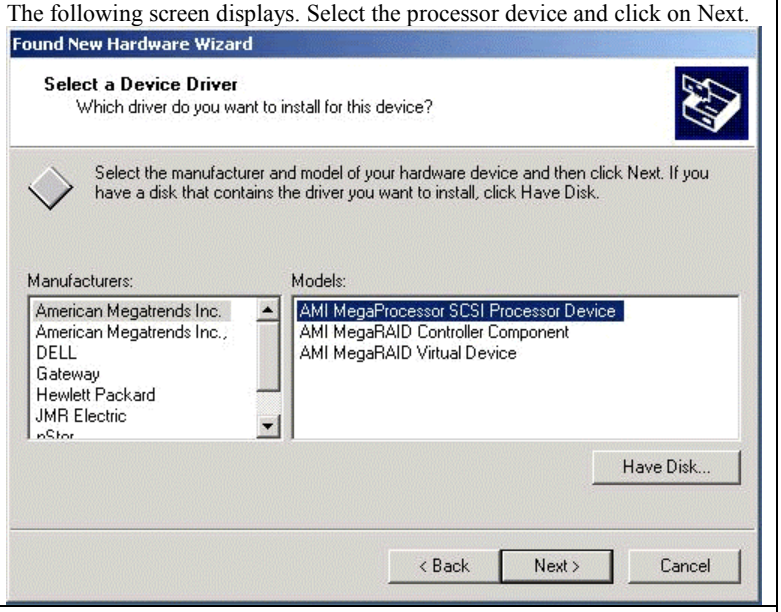
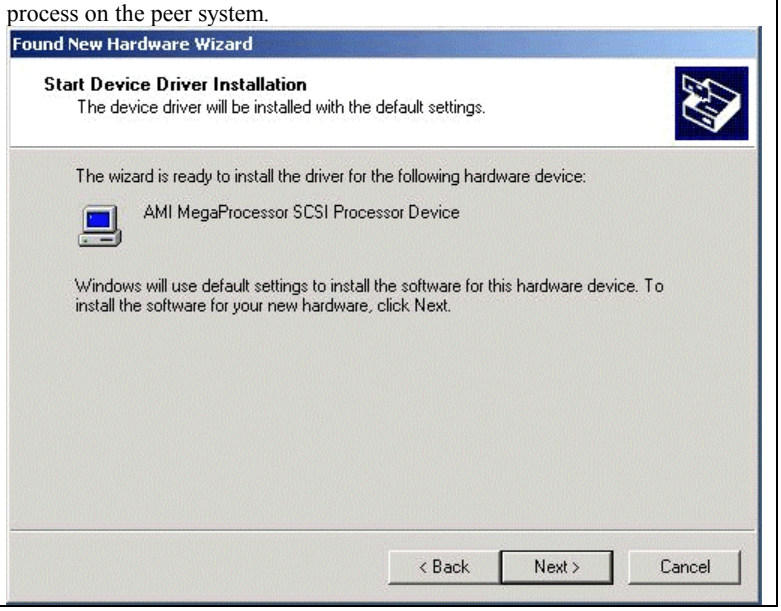
After the hardware is set up for the MS cluster configuration, perform the following procedure to configure the driver.

Step	Action
1	When the controller is added to an existing Windows 2000 Advanced Server installation, the operating system detects the controller.
2	The following screen displays the detected hardware device. Click on Next. 
3	The following screen appears. This screen is used to locate the device driver for the hardware device. Select Search for a suitable driver… and click on Next. 

Step	Action
4	The following screen displays. Insert the floppy diskette with the appropriate driver disk for Windows 2000. Select Floppy disk drives in the screen below and click on Next. 
5	The Wizard detects the device driver on the diskette and the "Completing the upgrade device driver" wizard displays the name of the controller. Click on Finish to complete the installation.
6	Repeat steps 1 – 5 to install the device driver on the second system.
7	After the cluster is installed, and both nodes are booted to the Microsoft Windows 2000 Advanced Server, installation will detect a SCSI processor device. The following screen displays. Click on Next. 

Step	Action
8	On the screen below, choose to display a list of the known drivers, so that you can choose a specific driver. Click on Next. 
9	The following screen displays. Select Other devices from the list of hardware types. Click on Next. 

Step	Action
10	The following screen displays. Select the driver that you want to install for the device. If you have a disk with the driver you want to install, click on Have Disk. 
11	The following window displays. Insert the disk containing the driver into the selected drive and click on OK. 

Step	Action
12	The following screen displays. Select the processor device and click on Next. 
13	On the final screen, click on Finish to complete the installation. Repeat the process on the peer system. 

Network Requirements

The network requirements for clustering are:

- A unique NetBIOS cluster name
- Five unique, static IP addresses:
  - two are for the network adapters on the internal network
  - two are for the network adapters on the external network
  - one is for the cluster itself
- A domain user account for Cluster Service (all nodes must be part of the same domain.)
- Two network adapters for each node—one for connection to the external network and the other for the node-to-node internal cluster network. If you do not use two network adapters for each node, your configuration is unsupported. HCL certification requires a separate private network adapter.

Shared Disk Requirements

Disks can be shared by the nodes. The requirements for sharing disks are as follows:

- Physically attach all shared disks, including the quorum disk, to the shared bus.
- Make sure that all disks attached to the shared bus are seen from all nodes. You can check this at the setup level in <Ctrl> <M> (the BIOS configuration utility.) See page 83 for installation information.
- Assign unique SCSI identification numbers to the SCSI devices and terminate the devices properly. Refer to the storage enclosure manual about installing and terminating SCSI devices.
- Configure all shared disks as basic (not dynamic.)
- Format all partitions on the disks as NTFS.

It is best to use fault-tolerant RAID configurations for all disks. This includes RAID levels 1, 5, 10, 30 or 50.

Cluster Installation

Installation Overview During installation, some nodes are shut down, and other nodes are rebooted. This is necessary to ensure uncorrupted data on disks attached to the shared storage bus. Data corruption can occur when multiple nodes try to write simultaneously to the same disk, if that disk is not yet protected by the cluster software.

The table below shows which nodes and storage devices should be powered on during each step.

Step	Node 1	Node 2	Storage	Comments
Set Up Networks	On	On	Off	Make sure that power to all storage devices on the shared bus is turned off. Power on all nodes.
Set up Shared Disks	On	Off	On	Power down all nodes. Next, power on the shared storage, then power on the first node.
Verify Disk Configuration	Off	On	On	Shutdown the first node. Power on the second node.
Configure the First Node	On	Off	On	Shutdown all nodes. Power on the first node.
Configure the Second Node	On	On	On	Power on the second node after the first node was successfully configured.
Post-installation	On	On	On	All nodes should be active.

Before installing the Cluster Service software you must follow the steps below:

- Install Windows 2000 Advanced Server or Windows 2000 Datacenter Server on each node
- Setup networks
- Setup disks

Note: These steps must be completed on every cluster node before proceeding with the installation of Cluster Service on the first node.

To configure the Cluster Service on a Windows 2000-based server, you must be able to log on as administrator or have administrative permissions on each node. Each node must be a member server, or be domain controllers inside the same domain. A mix of domain controllers and member servers in a cluster is not acceptable.

Installing the Windows 2000 Operating System

Install Microsoft Windows 2000 to each node. See your Windows 2000 manual on how to install the Operating System.

Log on as administrator before you install the Cluster Services.

Setting Up Networks

Note: Do not allow both nodes to access the shared storage device before the Cluster Service is installed. In order to prevent this, power down any shared storage devices and then power up nodes one at a time. Install the Clustering Service on at least one node and make sure it is online before you power up the second node.

Install at least two network card adapters per each cluster node. One network card adapter card is used to access the public network. The second network card adapter is used to access the cluster nodes.

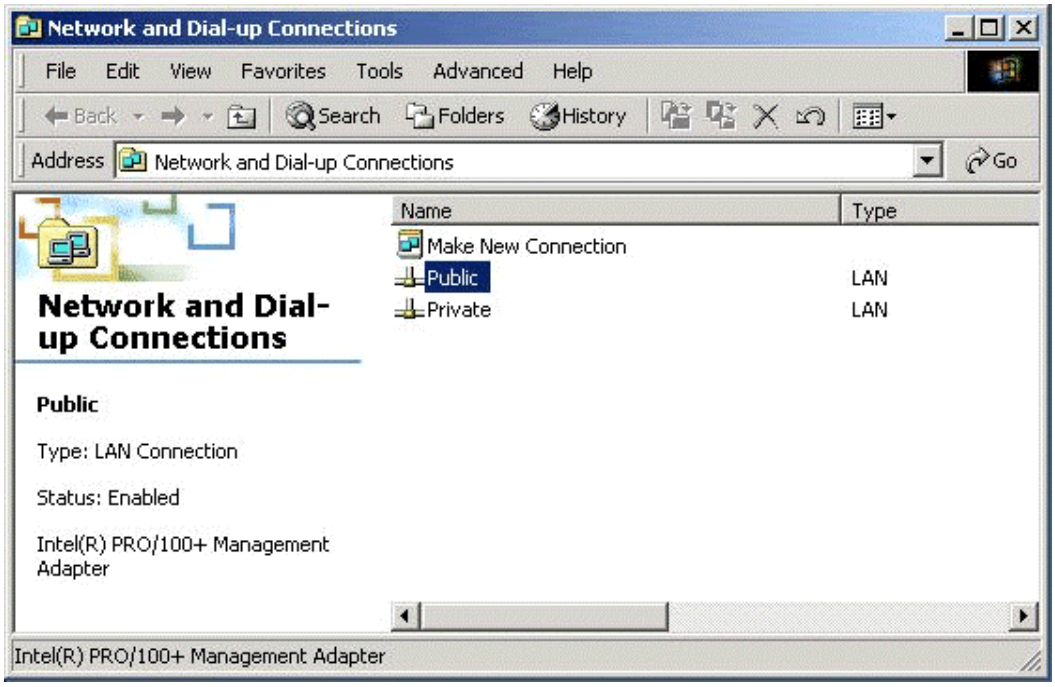
The network card adapter that is used to access the cluster nodes establishes the following:

- Node to node communications
- Cluster status signals
- Cluster Management

Check to make sure that all the network connections are correct. Network cards that access the public network must be connected to the public network. Network cards that access the cluster nodes must connect to each other.

Cont'd

Verify that all network connections are correct, with private network adapters connected to other private network adapters only, and public network adapters connected to the public network. View the Network and Dial-up Connections screen to check the connections.



Note: Use crossover cables for the network card adapters that access the cluster nodes. If you do not use the crossover cables properly, the system will not detect the network card adapter that accesses the cluster nodes. If the network card adapter is not detected, then you cannot configure the network adapters during the Cluster Service installation.

However, if you install Cluster Service on both nodes, and both nodes are powered on, you can add the adapter as a cluster resource and configure it properly for the cluster node network in Cluster Administrator.

Configuring the Cluster Node Network Adapter

Note: Which network adapter is private and which is public depends upon your wiring. For the purposes of this chapter, the first network adapter (Local Area Connection) is connected to the public network, and the second network adapter (Local Area Connection 2) is connected to the private cluster network. This may not be the case in your network.

Renaming the Local Area Connections In order to make the network connection more clear, you can change the name of the Local Area Connection (2). Renaming it will help you identify the connection and correctly assign it. Follow the steps below to change the name:

Step	Description
1	Right-click on the Local Area Connection 2 icon.
2	Click on Rename.
3	Type Private Cluster Connection into the textbox, then press Enter.
4	Repeat steps 1-3 to change the name of the public LAN network adapter to Public Cluster Connection.
5	The renamed icons should look like those in the picture above. Close the Networking and Dial-up Connections window. The new connection names automatically replicate to other cluster servers as the servers are brought online.

Setting up the First Node in your Cluster Follow the steps below to setup the first node in your cluster:

Step	Description
1	Right-click on My Network Places, then click on Properties.
2	Right-click the Private Connection icon.
3	Click on Status. The Private Connection Status window shows the connection status, as well as the speed of connection. If the window shows that the network is disconnected, examine cables and connections to resolve the problem before proceeding.
4	Click on Close
4	Right-click Private Connection again
5	Click on Properties.
6	Click on Configure.
7	Click on Advanced. The network card adapter properties window displays.
8	You should set network adapters on the private network to the actual speed of the network, rather than the default automated speed selection. Select the network speed from the drop-down list. Do not use “Auto-select” as the setting for speed. Some adapters can drop packets while determining the speed. Set the network adapter speed by clicking the appropriate option, such as Media Type or Speed.
9	Configure identically all network adapters in the cluster that are attached to the same network, so they use the same Duplex Mode, Flow Control, Media Type, and so on. These settings should stay the same even if the hardware is different.
10	Click on Transmission Control Protocol/Internet Protocol (TCP/IP).
11	Click on Properties.
12	Click on the radio-button for Use the following IP address.
13	Enter the IP addresses you want to use for the private network.
14	Type in the subnet mask for the network.
15	Click the Advanced radio button, then select the WINS tab.
16	Select Disable NetBIOS over TCP/IP.
17	Click OK to return to the previous menu. *Perform this step for the private network adapter only*.

Configuring the Public Network Adapter

Note: It is strongly recommended that you use static IP addresses for all network adapters in the cluster. This includes both the network adapter used to access the cluster nodes and the network adapter used to access the LAN (Local Area Network). If you must use a dynamic IP address through DHCP, access to the cluster could be terminated and become unavailable if the DHCP server goes down or goes offline.

The use of long lease periods is recommended to assure that a dynamically assigned IP address remains valid in the event that the DHCP server is temporarily lost. In all cases, set static IP addresses for the private network connector. Note that Cluster Service will recognize only one network interface per subnet.

Verifying Connectivity and Name Resolution

In order to verify that the network adapters are working properly, perform the following steps.

Note: Before proceeding, you must know the IP address for each network card adapter in the cluster. You can obtain it by using the IPCONFIG command on each node.

Step	Description
1	Click on Start.
2	Click on Run.
3	Type `cmd` in the text box.
4	Click on OK.
5	Type ipconfig /all and press Enter. IP information displays for all network adapters in the machine.
6	If you do not already have the command prompt on your screen, click on Start.
7	Click on Run.
8	Type `cmd` in the text box.
9	Click on OK.

10	Type
	
	`ping ipaddress`
	
	where *ipaddress* is the IP address for the corresponding network adapter in the other node. For example, assume that the IP addresses are set as follows:
	
	Node / **Network Name** / **Network Adapter IP Address**
	1 / Public Cluster Connection / 192.168.0.171
	1 / Private Cluster Connection / 10.1.1.1
	2 / Public Cluster Connection / 192.168.0.172
	2 / Private Cluster Connection / 10.1.1.2
	
	In this example, you would type
	
	`Ping 192.168.0.172`
	
	and
	
	`Ping 10.1.1.1`
	
	from Node 1.
	
	They you would type

	Ping 192.168.0.172 and 10.1.1.1 from Node 2. To confirm name resolution, ping each node from a client using the node's machine name instead of its IP number.

Verifying Domain Membership

All nodes in the cluster have to be members of the same domain and capable of accessing a domain controller and a DNS Server. You can configure them as either member servers or domain controllers. If you plan to configure one node as a domain controller, you should configure all other nodes as domain controllers in the same domain as well.

Setting Up a Cluster User Account

The Cluster Service requires a domain user account that the Cluster Service can run under. You must create the user account before installing the Cluster Service. The reason for this is that setup requires a user name and password. This user account should not belong to a user on the domain.

Step	Description
1	Click on Start.
2	Point to Programs, then point to Administrative Tools.
3	Click on Active Directory Users and Computers.
4	Click the plus sign (+) to expand the domain name (if it is not already expanded.)
5	Click on Users.
6	Right-click on Users.
7	Point to New and click on User.
8	Type in the cluster name and click on Next.
9	Set the password settings to User Cannot Change Password and Password Never Expires.
10	Click on Next, then click on Finish to create this user. **Note:** If your company's security policy does not allow the use of passwords that never expire, you must renew the password on each node before password expiration. You must also update the Cluster Service configuration
11	Right-click on Cluster in the left pane of the Active Directory Users and Computers snap-in.
12	Select Properties from the context menu.
13	Click on Add Members to a Group.
14	Click on Administrators and click on OK. This gives the new user account administrative privileges on this computer.
15	Close the Active Directory Users and Computers snap-in.

Setting Up Shared Disks

 **Warning**: Make sure that Windows 2000 Advanced Server or Windows 2000 Datacenter Server and the Cluster Service are installed and running on one node before you start an operating system on another node. If the operating system is started on other nodes before you install and configure Cluster Service and run it on at least one node, the cluster disks will have a high chance of becoming corrupted.

To continue, power off all nodes. Power up the shared storage devices. Once the shared storage device is powered up, power up node one.

Quorum Disk The quorum disk stores cluster configuration database checkpoints and log files that help manage the cluster. Windows 2000 makes the following quorum disk recommendations:

- Create a small partition [Use a minimum of 50 megabytes (MB) as a quorum disk. Windows 2000 generally recommends a quorum disk to be 500 MB.]
- Dedicate a separate disk for a quorum resource. The failure of the quorum disk would cause the entire cluster to fail; therefore, Windows 2000 strongly recommends that you use a volume on a RAID disk array.

During the Cluster Service installation, you have to provide the drive letter for the quorum disk.

Note: For our example, we use the letter **E** for the quorum disk drive letter.

Configuring Shared Disks

Perform the following procedure to configure the shared disks.

Step	Description
1	Right-click on My Computer.
2	Click on Manage, then click on Storage.
3	Double-click on Disk Management.
4	Make sure that all shared disks are formatted as NTFS and are designated as Basic. If you connect a new drive, the Write Signature and Upgrade Disk Wizard starts automatically. If this occurs, click on Next to go through the wizard. The wizard sets the disk to dynamic, but you can uncheck it at this point to set it to basic. To reset the disk to Basic, right-click on Disk # (where # identifies the disk that you are working with) and click on Revert to Basic Disk.
5	Right-click on unallocated disk space.
6	Click on Create Partition…
7	The Create Partition Wizard begins. Click on Next twice.
8	Enter the desired partition size in MB and click on Next.
9	Accept the default drive letter assignment by clicking on Next.
10	Click on Next to format and create a partition.

Assigning Drive Letters

After you have configured the bus, disks, and partitions, you must assign drive letters to each partition on each clustered disk.


Note: Mountpoints is a feature of the file system that lets you mount a file system using an existing directory without assigning a drive letter. Mountpoints is not supported on clusters. Any external disk that is used as a cluster resource must be partitioned using NTFS partitions and have a drive letter assigned to it. Use the procedure below to assign driver letters.

Step	Description
1	Right-click on the desired partition and select Change Drive Letter and Path.
2	Select a new drive letter.
3	Repeat steps 1 and 2 for each shared disk.
4	Close the Computer Management window.

Verifying Disk Access and Functionality

Perform the steps below to verify disk access and functionality.

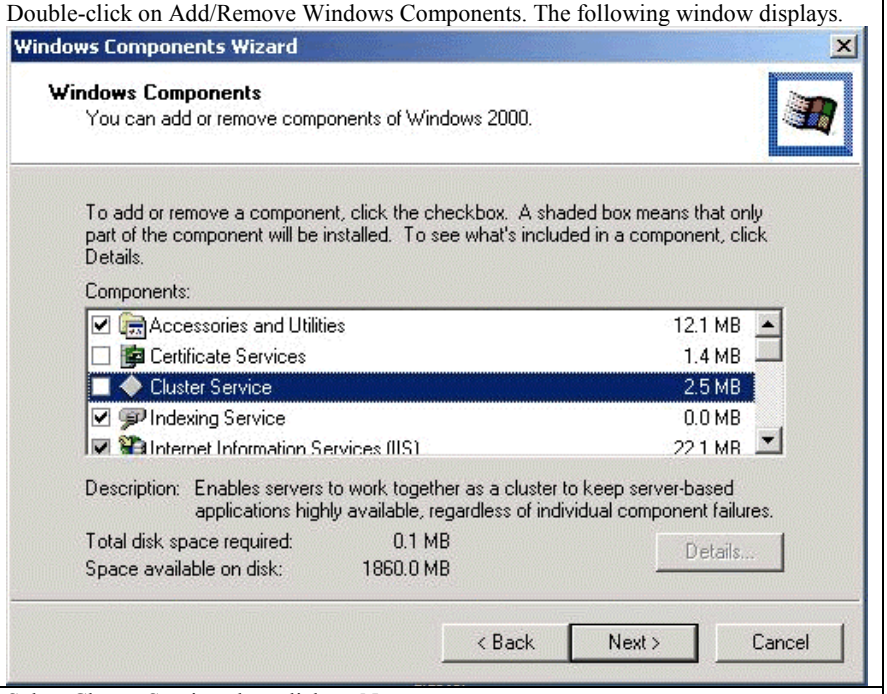
Step	Description
1	Click on Start.
2	Click on Programs. Click on Accessories, then click on Notepad.
3	Type some words into Notepad and use the File/Save As command to save it as a test file called test.txt. Close Notepad.
4	Double-click on the My Documents icon.
5	Right-click on test.txt and click on Copy.
6	Close the window.
7	Double-click on My Computer.
8	Double-click on a shared drive partition.
9	Click on Edit and click on Paste.
10	A copy of the file should now exist on the shared disk.
11	Double-click on test.txt to open it on the shared disk.
12	Close the file.
13	Highlight the file and press the Del key to delete it from the clustered disk.
14	Repeat the process for all clustered disks to make sure they can be accessed from the first node.

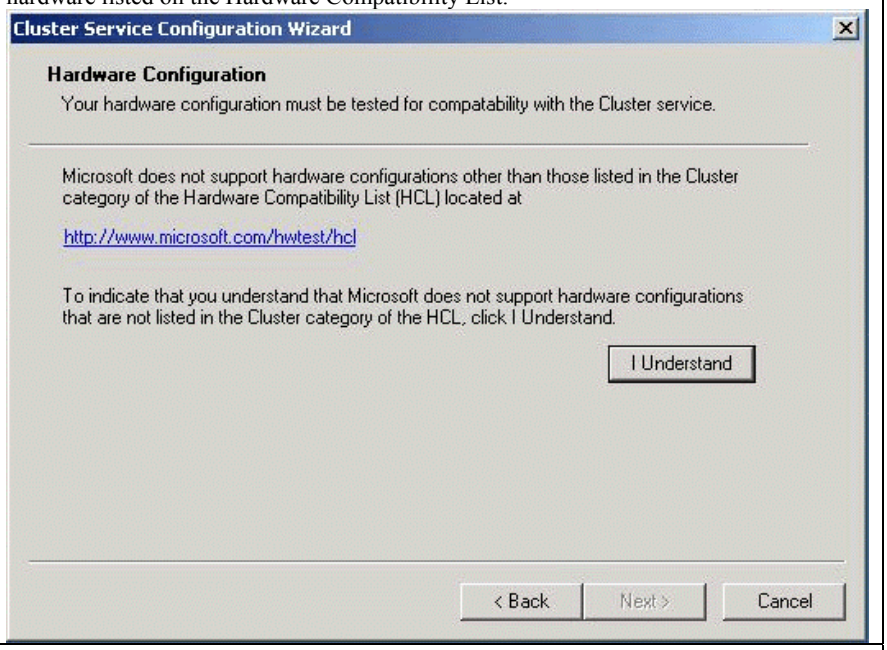
After you complete the procedure, shut down the first node, power on the second node and repeat the procedure above. Repeat again for any additional nodes. After you have verified that all nodes can read and write from the disks, turn off all nodes except the first, and continue with this guide.

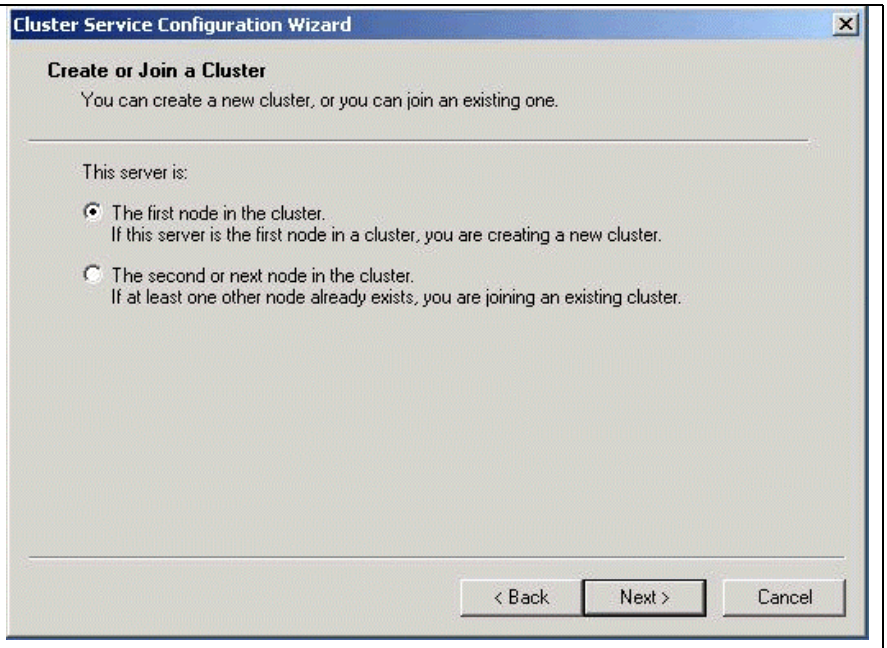
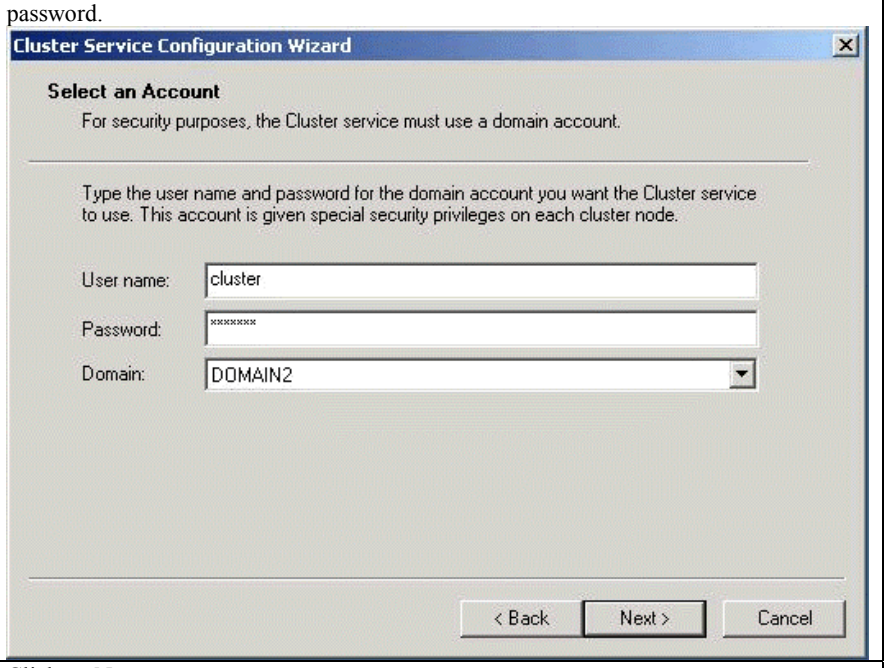
Cluster Service Software Installation

Before you begin the Cluster Service Software installation on the first node, make sure that all other nodes are either powered down or stopped and that all shared storage devices are powered on.

Cluster Configuration Wizard To create the cluster, you must provide the cluster information. The Cluster Configuration Wizard will allow you to input this information.

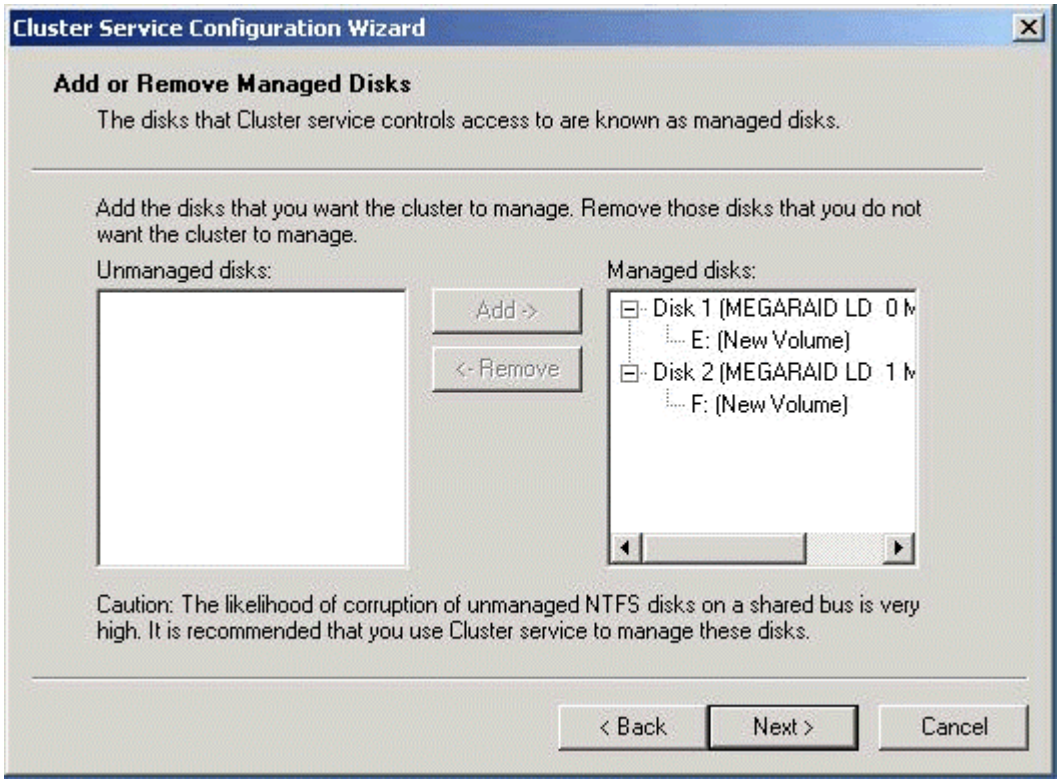
Step	Description
1	Click on Start.
2	Click on Settings, then click on Control Panel.
3	Double-click on Add/Remove Programs.
4	Double-click on Add/Remove Windows Components. The following window displays. 
5	Select Cluster Service, then click on Next.
6	Cluster Service files are located on the Windows 2000 Advanced Server or Windows 2000 Datacenter Server CD-ROM. Enter x:\i386 (where *x* is the drive letter of your CD-ROM). If you installed Windows 2000 from a network, enter the appropriate network path instead. (If the Windows 2000 Setup flashscreen displays, close it.)
7	Click on OK. The following screen displays.

	
8	Click on Next.
9	The Hardware Configuration Certification window appears. Click on I Understand to accept the condition that Cluster Service is supported only on hardware listed on the Hardware Compatibility List. 
10	This is the first node in the cluster; therefore, you must create the cluster itself. Select The first node in the cluster, as shown below and then click on Next.

	
11	Enter a name for the cluster (up to 15 characters), and click on Next. (In our example, the cluster is named ClusterOne.)
12	Type the user name of the Cluster Service account that you created during the pre-installation. (In our example, the user name is cluster.) Do not enter a password. Type the domain name, then click on Next. At this point the Cluster Service Configuration Wizard validates the user account and password. 
13	Click on Next. The Add or Remove Managed Disks screen displays next. This screen is in the following section about configuring cluster disks.

Configuring Cluster Disks

Windows 2000 Managed Disks displays all SCSI disks, as shown on the screen below. It displays SCSI disks that do not reside on the same bus as the system disk. Because of this, a node that has multiple SCSI buses will list SCSI disks that are not to be used as shared storage. You must remove any SCSI disks that are internal to the node and not to be shared storage.



In production clustering scenarios, you need to use more than one private network for cluster communication to avoid having a single point of failure. Cluster Service can use private networks for cluster status signals and cluster management. This provides more security than using a public network for these roles. In addition, you can use a public network for cluster management, or you can use a mixed network for both private and public communications.

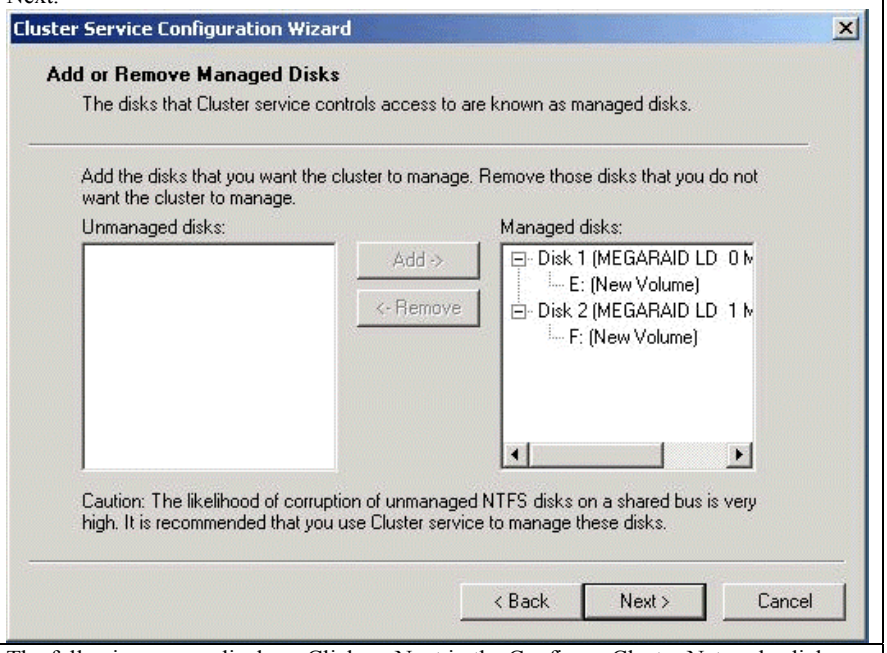
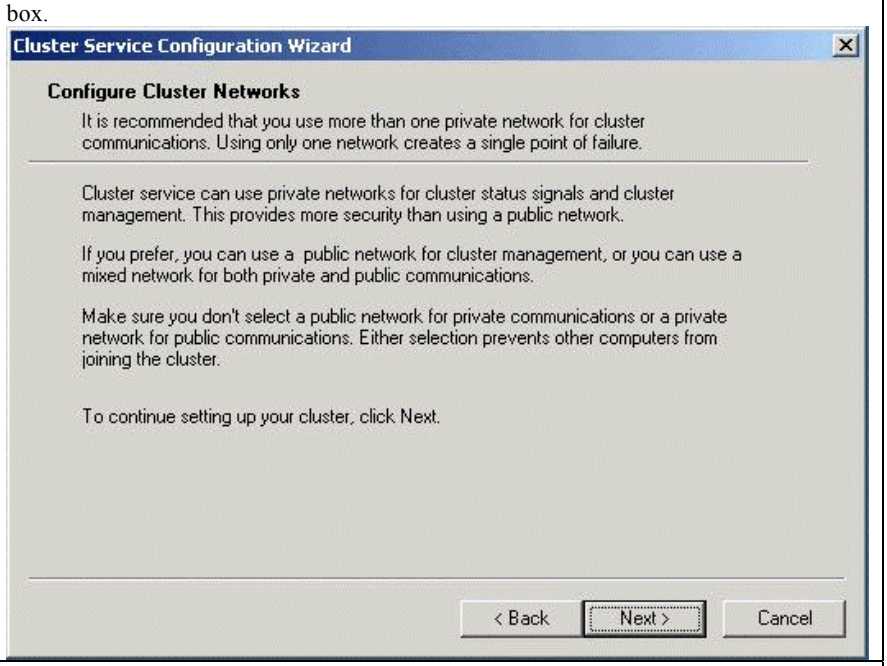
In any case, verify that at least two networks are used for cluster communication; using a single network for node-to-node communication creates a potential single point of failure. We recommend that you use multiple networks, with at least one network configured as a private link between nodes and other connections through a public network. If you use more than one private network, make sure that each uses a different subnet, as Cluster Service recognizes only one network interface per subnet.

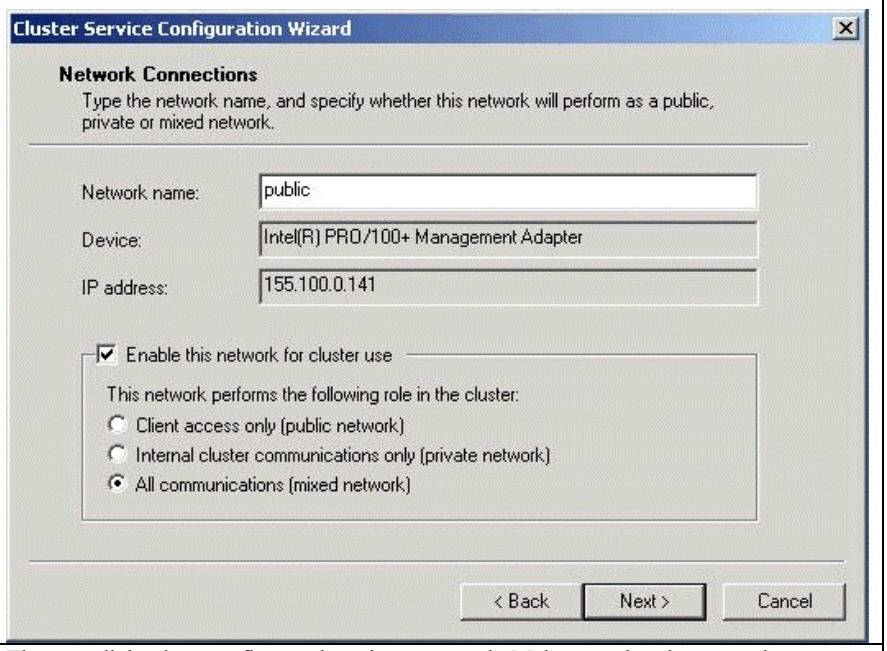
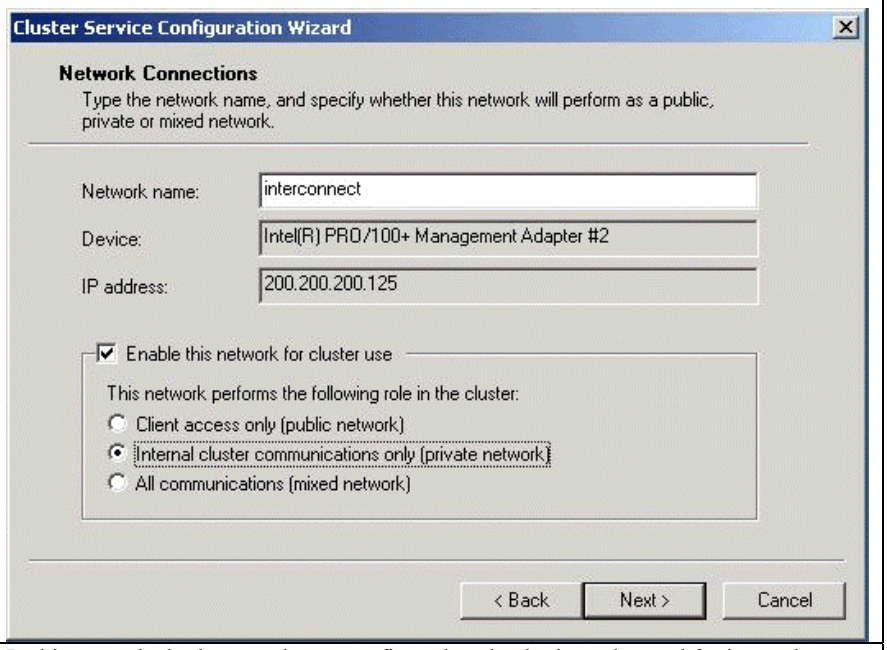
This document assumes that only two networks are in use. It describes how you can configure these networks as one mixed and one private network.

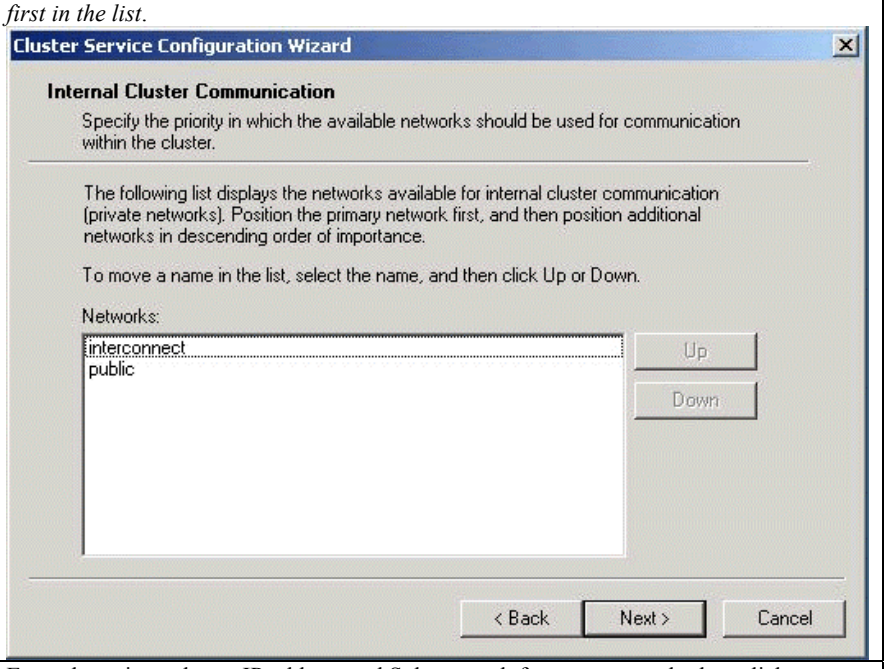
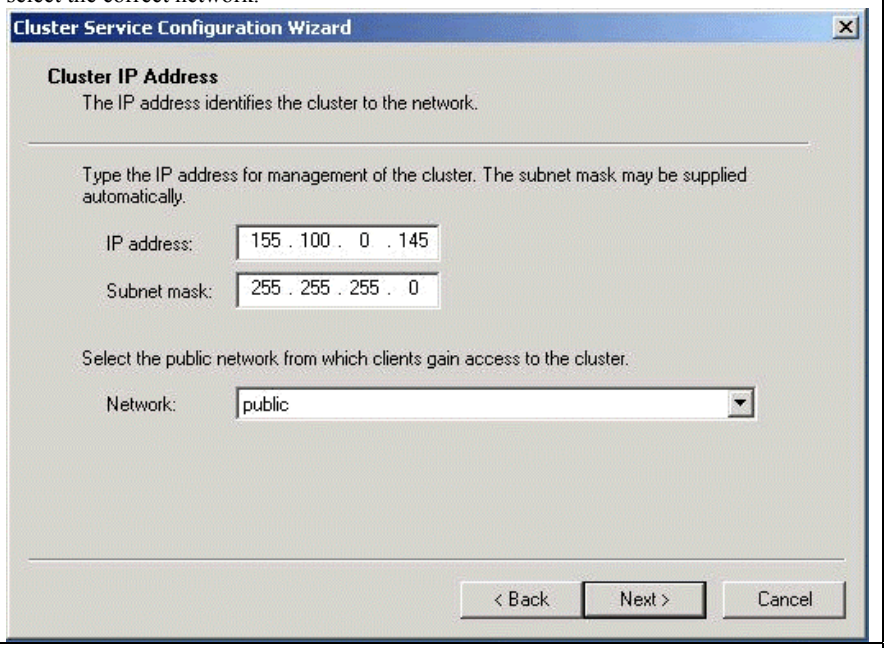
The order in which the Cluster Service Configuration Wizard presents these networks can vary. In this example, the public network is presented first.

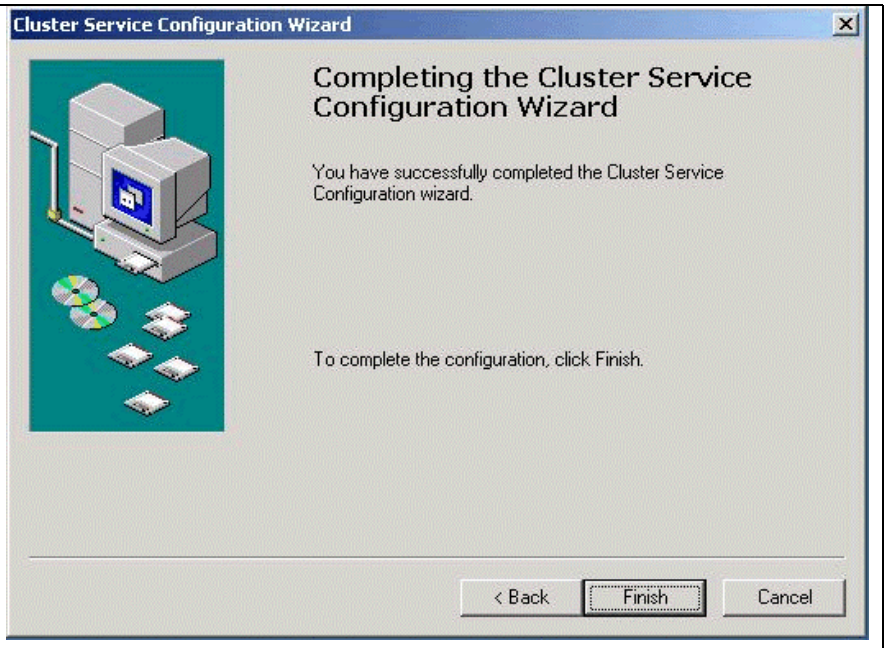
Cont'd

Use the following procedure to configure the clustered disks.

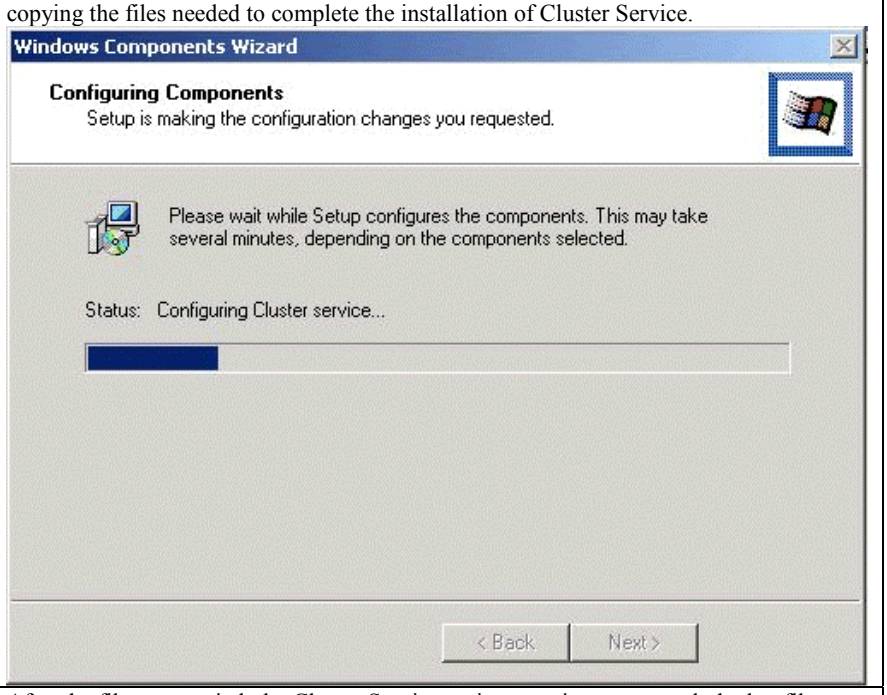
Step	Description
1	The Add or Remove Managed Disks dialog box specifies disks on the shared SCSI bus that will be used by Cluster Service. Add or remove disks as necessary, then click on Next. 
2	The following screen displays. Click on Next in the Configure Cluster Networks dialog box. 
3	Verify that the network name and IP address correspond to the network interface for the *public* network.
4	Check the box Enable this network for cluster use.
5	Select the option All communications (mixed network), as shown below, and click on

	Next. 
6	The next dialog box configures the private network. Make sure that the network name and IP address correspond to the network interface used for the *private* network. Check the box Enable this network for cluster use. Select the option Internal cluster communications only, then click on Next. 
7	In this example, both networks are configured so that both can be used for internal cluster communication. The next dialog window offers an option to modify the order in which the networks are used. Because Private Cluster Connection represents a direct connection between nodes, it remains at the top of the list. In normal operation, this connection is used for cluster communication. In case of the Private Cluster Connection failure, Cluster Service automatically switches to the next

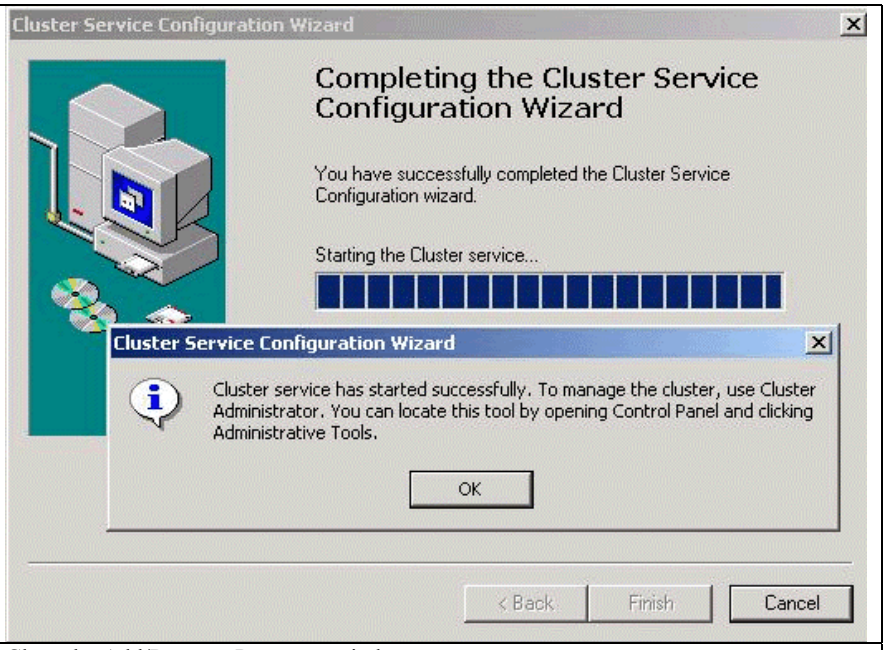
	network on the list—in this case Public Cluster Connection. Verify that the first connection in the list is the Private Cluster Connection, then click on Next. **Note:** *Always set the order of the connections so that the Private Cluster Connection is first in the list.* 
8	Enter the unique cluster IP address and Subnet mask for your network, then click on Next. The Cluster Service Configuration Wizard shown below automatically associates the cluster IP address with one of the public or mixed networks. It uses the subnet mask to select the correct network. 
9	Click Finish to complete the cluster configuration on the first node.



The Cluster Service Setup Wizard completes the setup process for the first node by copying the files needed to complete the installation of Cluster Service.



10	After the files are copied, the Cluster Service registry entries are created, the log files on the quorum resource are created, and the Cluster Service is started on the first node. A dialog box appears telling you that Cluster Service has started successfully. Click on OK.

	
11	Close the Add/Remove Programs window.

Validating the Cluster Installation

Use the Cluster Administrator snap-in to validate the Cluster Service installation on the first node.

Step	Description
1	Click on Start.
2	Click on Programs.
3	Click on Administrative Tools.
4	Click on Cluster Adminstrator.
5	The following screen displays. If your snap-in window is similar to that shown above below, your Cluster Service was successfully installed on the first node. You are now ready to install Cluster Service on the second node.

Configuring the Second Node

Note: For this procedure, have node one and all shared disks powered on, then power up the second node.

Installation of Cluster Service on the second node takes less time than on the first node. Setup configures the Cluster Service network settings on the second node based on the configuration of the first node.

Installation of Cluster Service on the second node begins the same way as installation on the first node. The first node must be running during installation of the second node.

Follow the same procedures used to install Cluster Service on the first node, with the following differences:

1. In the Create or Join a Cluster dialog box, select The second or next node in the cluster, then click Next.
2. Enter the cluster name that was previously created (it is MyCluster in this example**)**, and click Next.
3. Leave Connect to cluster as unchecked. The Cluster Service Configuration Wizard automatically supplies the name of the user account selected when you installed the first node. Always use the same account you used when you set up the first cluster node.
4. Enter the password for the account (if there is one), then click Next.
5. At the next dialog box, click Finish to complete configuration.
6. The Cluster Service will start. Click OK.
7. Close Add/Remove Programs.

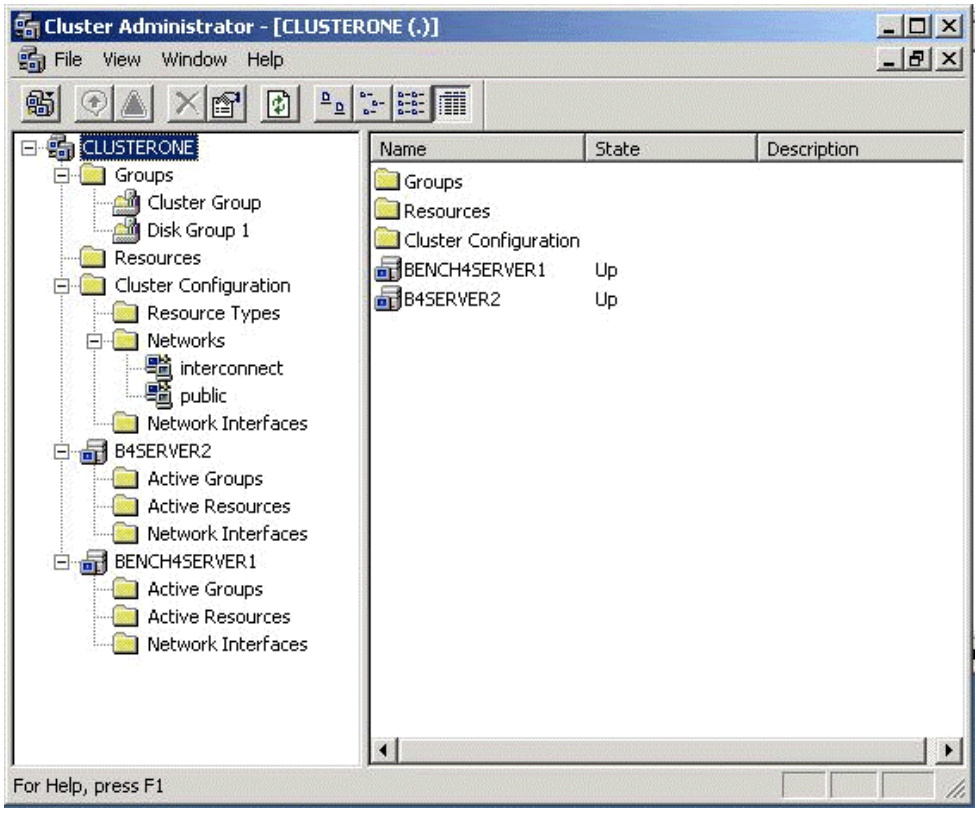
If you install additional nodes, repeat these steps to install Cluster Service on all other nodes.

Verify Installation

There are several ways to verify that Cluster Service was successfully installed. Here is a simple one:

1. Click Start, click Programs, click Administrative Tools, then click Cluster Administrator.

   The presence of two nodes (pictured below) shows that a cluster exists and is in operation.



2. Right-click the group Disk Group 1 and select the option Move. This option moves the group and all its resources to another node. After a short period of time, the Disk F: G: will be brought online on the second node. If you watch the screen, you will see this shift. Close the Cluster Administrator snap-in.

Congratulations! You have completed installing Cluster Service on all nodes. The server cluster is fully operational. Now, you are ready to install cluster resources, such as file shares, printer spoolers, cluster aware services like IIS, Message Queuing, Distributed Transaction Coordinator, DHCP, WINS, or cluster aware applications like Exchange or SQL Server.

SCSI Drive Installations

This information is provided as a generic instruction set for SCSI drive installations. If the SCSI hard disk vendor's instructions conflict with the instructions in this section, always use the instructions supplied by the vendor.

The SCSI bus listed in the hardware requirements must be configured prior to installation of Cluster Services. This includes:

- Configuring the SCSI devices.
- Configuring the SCSI controllers and hard disks to work properly on a shared SCSI bus.
- Properly terminating the bus. The shared SCSI bus must have a terminator at each end of the bus. It is possible to have multiple shared SCSI buses between the nodes of a cluster.

In addition to the information on the next page, refer to the documentation from the SCSI device manufacturer or the SCSI specifications, which can be ordered from the American National Standards Institute (ANSI). The ANSI web site contains a catalog that you can search for the SCSI specifications.

Configuring the SCSI Devices

Each device on the shared SCSI bus must have a unique SCSI ID. Since most SCSI controllers default to SCSI ID 7, part of configuring the shared SCSI bus will be to change the SCSI ID on one controller to a different SCSI ID, such as SCSI ID 6. If there is more than one disk that will be on the shared SCSI bus, each disk must also have a unique SCSI ID.

Some SCSI controllers reset the SCSI bus when they initialize at boot time. If this occurs, the bus reset can interrupt any data transfers between the other node and disks on the shared SCSI bus. Therefore, SCSI bus resets should be disabled if possible.

Terminating the Shared SCSI Bus

You can connect ***Y*** cables to devices if the device is at the end of the SCSI bus. You can then attach a terminator to one branch of the ***Y*** cable to terminate the SCSI bus. This method of termination requires either disabling or removing any internal terminators the device has.

Note: Any devices that are not at the end of the shared bus must have their internal termination disabled.

8 Troubleshooting

Problem	Suggested Solution
Some operating systems do not load in a computer with a MegaRAID adapter.	Check the system BIOS configuration for PCI interrupt assignments. Make sure some Interrupts are assigned for PCI. Initialize the logical drive before installing the operating system.
One of the hard drive in the array fails often	Check the drive error counts using Power Console. Format the drive. Rebuild the drive If the drive continues to fail, replace the drive with another drive with the same capacity.
Pressed <Ctrl> <M>. Ran Megaconf.exe and tried to make a new configuration. The system hangs when scanning devices.	Check the drives IDs on each channel to make sure each device has a different ID. Check the termination. The device at the end of the channel must be terminated. Replace the drive cable.
Multiple drives connected to MegaRAID using the same power supply. There is a problem spinning the drives all at once.	Set the drives to spin on command. This will allow MegaRAID to spin two devices simultaneously.
Pressing <Ctrl> <M> or running megaconf.exe does not display the Management Menu.	These utilities require a color monitor.
At system power-up with the MegaRAID installed, the screen display is garbled.	At least 16 MB of memory must be installed before power-up.
Cannot flash or update the EEPROM.	You may need a new EEPROM.
`Firmware Initializing...` appears and remains on the screen.	Make sure that TERMPWR is being properly provided to each peripheral device populated channel. Make sure that each end of the channel chain is properly terminated using the recommended terminator type for the peripheral device. The channel is automatically terminated at the MegaRAID card if only one cable is connected to a channel. Make sure that memory modules are rate at 60 ns or faster. Make sure that the MegaRAID controller is properly seated in the PCI slot.
What is the maximum number of MegaRAID adapters per computer?	Currently, all the utilities and drivers support up to 12 MegaRAID adapters per system.
What SCSI IDs can a non-hard disk device have and what is maximum number allowed per adapter?	Non-hard disk devices can accommodate only SCSI IDs 1, 2, 3, 4, 5 or 6, regardless of the channel used. A maximum of six non-hard disk devices are supported per MegaRAID adapter.
Why does a failed logical array still get a drive assignment?	To maintain the DOS Path statement integrity.

BIOS Boot Error Messages

Message	Problem	Suggested Solution
Adapter BIOS Disabled. No Logical Drives Handled by BIOS	The MegaRAID BIOS is disabled. Sometimes the BIOS is disabled to prevent booting from the BIOS.	Enable the BIOS via the MegaRAID Configuration Utility utility.
Host Adapter at Baseport xxxx Not Responding	The BIOS cannot communicate with the adapter firmware.	Make sure MegaRAID is properly installed. Rty moving the MegaRAID card to another PCI slot. Run the MegaRAID Manager Diagnostics to verify that MegaRAID is functioning properly.
No MegaRAID Adapter	The BIOS cannot communicate with the adapter firmware.	Make sure MegaRAID is properly installed. Move the MegaRAID card to another PCI slot. Run the MegaRAID Manager Diagnostics to verify that MegaRAID is functioning properly.
Configuration of NVRAM and drives mismatch. Run View/Add Configuration option of Configuration Utility. Press any key to run the Configuration Utility.	The configuration stored in the MegaRAID adapter does not match the configuration stored in the drives.	Press a key to run MegaRAID Manager. Choose View/Add Configuration from the Configure menu. Use View/Add Configuration to examine both the configuration in NVRAM and the configuration stored on the disk drives. Resolve the problem by selecting one of the configurations.
Configuration of NVRAM and drives mismatch for Host Adapter. Run View/Add Configuration option of Configuration Utility. Press any key to run the Configuration Utility.	The configuration stored in the MegaRAID adapter does not match the configuration stored in the drives.	Press a key to run MegaRAID Manager. Choose View/Add Configuration from the Configure menu. Use View/Add Configuration to examine both the configuration in NVRAM and the configuration stored on the disk drives. Resolve the problem by selecting one of the configurations.
1 Logical Drive Failed	A logical drive failed to sign on.	Make sure all physical drives are properly connected and are powered on. Run MegaRAID Manager to find out if any physical drives are not responding. Reconnect, replace, or rebuild any drive that is not responding.
X Logical Drives Degraded	x number of logical drives signed on in a degraded state.	Make sure all physical drives are properly connected and are powered on. Run MegaRAID Manager to find out if any physical drives are not responding. Reconnect, replace, or rebuild any drive that is not responding.
1 Logical Drive Degraded	A logical drive signed on in a degraded state.	Make sure all physical drives are properly connected and are powered on. Run MegaRAID Manager to find out if any physical drives are not responding. Reconnect, replace, or rebuild any drive that is not responding.
Insufficient memory to run BIOS. Press any key to continue…	Not enough MegaRAID memory to run MegaRAID BIOS.	Make sure MegaRAID memory has been properly installed.

Message	Problem	Suggested Solution
Insufficient Memory	Not enough memory on the MegaRAID adapter to support the current configuration.	Make sure MegaRAID memory has been properly installed.
The following SCSI IDs are not responding: Channel x:a.b.c	The physical drives with SCSI IDs a, b, and c are not responding on SCSI channel x.	Make sure the physical drives are properly connected and are powered on.

Other BIOS Error Messages

Message	Problem	Suggested Solution
Following SCSI disk not found and no empty slot available for mapping it	The physical disk roaming feature did not find the physical disk with the displayed SCSI ID. No slot is available to map the physical drive. MegaRAID cannot resolve the physical drives into the current configuration.	Reconfigure the array.
Following SCSI IDs have the same data y, z Channel x: a, b, c	The physical drive roaming feature found the same data on two or more physical drive on channel x with SCSI IDs a, b, and c. MegaRAID cannot determine the drive that has the duplicate information.	Remove the drive or drives that should not be used.
Unresolved configuration mismatch between disks and VRAM on the adapter	The configuration stored in the MegaRAID NVRAM does not match the configuration stored on the drives.	Press a key to run MegaRAID Manager. Choose View/Add Configuration from the Configure menu. Use View/Add Configuration to examine both the configuration in NVRAM and the configuration stored on the disk drives. Resolve the problem by selecting one of the configurations.

DOS ASPI Driver Error Messages

Message	Corrective Action
LSI Logic ASPI Manager has NOT been loaded.	The ASPI manager is not loaded. One of the failure codes listed below is displayed next.
Controller setup FAILED error code=[0xab]	Correct the condition that caused the failure. The failure codes are: 0x40 No MegaRAID adapters found 0x80 Timed out waiting for interrupt to be posted 0x81 Timed out waiting for the MegaRAID Response command. 0x82 Invalid command completion count. 0x83 Invalid completion status received. 0x84 Invalid command ID received. 0x85 No MegaRAID adapters found or no PCI BIOS support. 0x90 Unknown Setup completion error
No non-disk devices were located	The driver did not find any non-hard drive devices during scanning. A SCSI device that is not a hard disk drive, such as a tape drive or CD-ROM drive, must be attached to this SCSI channel. The SCSI ID must be unique for each adapter and cannot be SCSI ID 0. The supported SCSI IDs are 1, 2, 3, 4, 5, and 6.
'ERROR: VDS support is *INACTIVE* for MegaRAID logical drives	The /h option is appended to driver in CONFIG.SYS or this driver is used with a BIOS that is earlier than v1.10, or no logical drives are configured.

Other Potential Problems

Topic	Information
DOS ASPI	MEGASPI.SYS, the MegaRAID DOS ASPI manager, uses 6 KB of system memory once it is loaded.
CD-ROM drives under DOS	At this time, copied CDs are not accessible from DOS even after loading MEGASPI.SYS and AMICDROM.SYS.
Physical Drive Errors	To display the MegaRAID Manager Media Error and Other Error options, press <F2> after selecting a physical drive under the Physical Drive menu, selected from the Objects menu. A Media Error is an error that occurred while actually transferring data. An Other Error is an error that occurs at the hardware level because of a device failure, poor cabling, bad termination, signal loss, etc.
Virtual Sizing	The FlexRAID Virtual Sizing option enables RAID expansion. FlexRAID Virtual Sizing must be enabled to increase the size of a logical drive or add a physical drive to an existing logical drive. Run MegaRAID Manager by pressing <Ctrl> <M> to enable FlexRAID Virtual Sizing. Select the Objects menu, then select the Logical Drive menu. Select View/Update Parameters. Set FlexRAID Virtual Sizing to Enabled.
BSD Unix	We do not provide a driver for BSDI Unix. MegaRAID does not support BSDI Unix.
Multiple LUNs	MegaRAID supports one LUN per each target ID. No multiple LUN devices are supported.
MegaRAID Power Requirements	The Maximum MegaRAID power requirements are 15 watts at 5V and 3 Amps.
SCSI Bus Requirements	The ANSI specification dictates the following: The maximum signal path length between terminators is 3 meters when using up to 4 maximum capacitance (25 pF) devices and 1.5 meters when using more than 4 devices. SCSI devices should be uniformly spaced between terminators, with the end devices located as close as possible to the terminators. The characteristic impedance of the cable should be 90 +/- 6 ohms for the /REQ and /ACK signals and 90 +/- 10 ohms for all other signals. The stub length(the distance from the controller's external connector to the mainline SCSI bus) shall not exceed.1m (approximately 4 inches). The spacing of devices on the mainline SCSI bus should be at least three times the stub length. All signal lines shall be terminated once at both ends of the bus powered by the TERMPWR line. For more detailed information, refer to SPI (SCSI Parallel Interface) documentation.

Topic	Information

Windows NT Installation	When Windows NT is installed via a bootable CD, the devices on the MegaRAID will not be recognized until after the initial reboot. The Microsoft documented workaround is in SETUP.TXT:
	
	SETUP.TXT is on the CD.
	
	To install drivers when Setup recognizes one of the supported SCSI host adapters without making the devices attached to it available for use:
	
	1 Restart Windows NT Setup.
	2 When Windows NT Setup displays
	
	`Setup is inspecting your computer's hardware Configuration...,`
	
	press <F6> to prevent Windows NT Setup from performing disk controller detection. This allows you to install the driver from the Drivers disk you created. All SCSI adapters must be installed manually.
	
	When Windows NT Setup displays
	
	`Setup could not determine the type of one or more mass storage devices installed in your system, or you have chosen to manually specify an adapter,`
	
	press S to display a list of supported SCSI host adapters.
	
	4 Select Other from the bottom of the list.
	
	5 Insert the Drivers Disk you made when prompted to do so and select MegaRAID from this list. In some cases, Windows NT Setup repeatedly prompts to swap disks. Windows NT will now recognize any devices attached to this adapter. Repeat this step for each host adapter not already recognized by Windows NT Setup.

A SCSI Cables and Connectors

SCSI Connectors

MegaRAID provides several different types of SCSI connectors for each channel. The connectors are:

- 68-pin high density internal connectors,
- 68-pin ultra high density external connectors.

68-Pin High Density SCSI Internal Connectors

Each of the SCSI channels on the MegaRAID has a 68-pin high density 0.050 inch pitch unshielded connector.

These connectors provide all signals needed to connect MegaRAID to wide SCSI devices. The connector pinouts are for a single-ended primary bus (P-CABLE) as specified in SPI documentation.

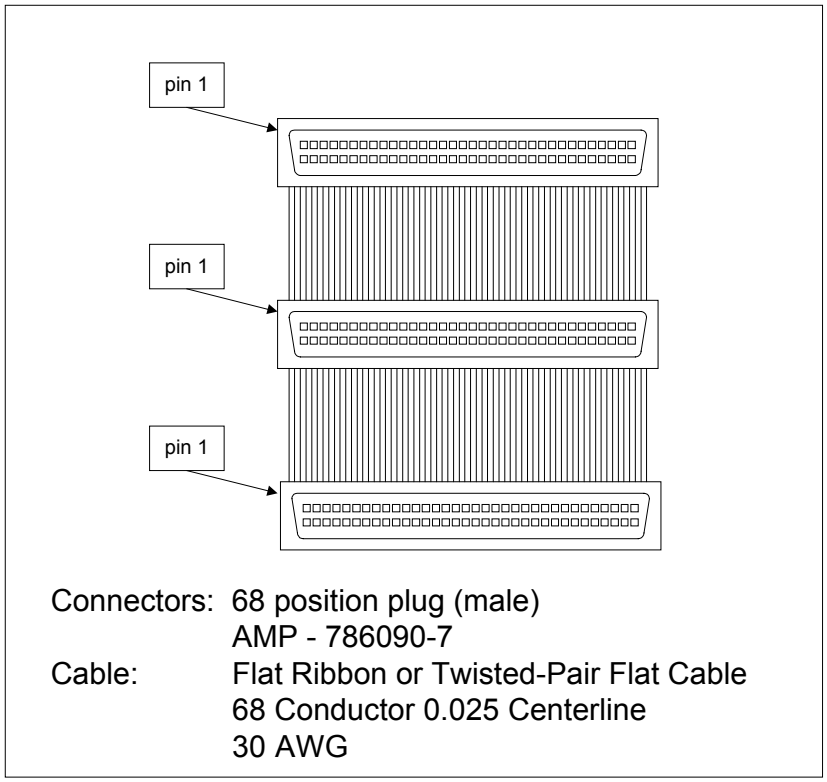
The cable assemblies that interface with this 68-pin connector are:

- flat ribbon or twisted pair cable for connecting internal wide SCSI devices
- flat ribbon or twisted pair cable for connecting internal and external wide SCSI devices
- cable assembly for converting from internal wide SCSI connectors to internal non-wide (Type 2) connectors
- cable assembly for converting from internal wide to internal non-wide SCSI connectors (Type 30), and
- cable assembly for converting from internal wide to internal non-wide SCSI connectors

Cont'd

68-Pin High Density Connectors, Continued

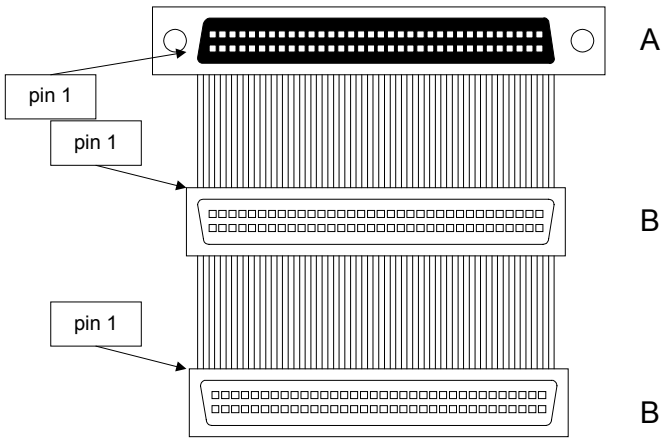
Cable Assembly for Internal Wide SCSI Devices The cable assembly for connecting internal wide SCSI devices is shown below:



Cont'd

68-Pin High Density Connectors, Continued

Connecting Internal and External Wide Devices The cable assembly for connecting internal wide and external wide SCSI devices is shown below:



Connector A: 68 position panel mount receptacle
with 4-40 holes (female)
AMP - 786096-7

NOTE: To convert to 2-56 holes, use screwlock kit 749087-1, 749087-2, or 750644-1 from AMP

Connector B: 68 position plug (male)
AMP - 786090-7

Cable: Flat Ribbon or Twisted-Pair Flat Cable
68 Conductor 0.025 Centerline
30 AWG

Cont'd

Converting Internal Wide to Internal Non-Wide (Type 2) The cable assembly for converting internal wide SCSI connectors to internal non-wide SCSI connectors is shown below:

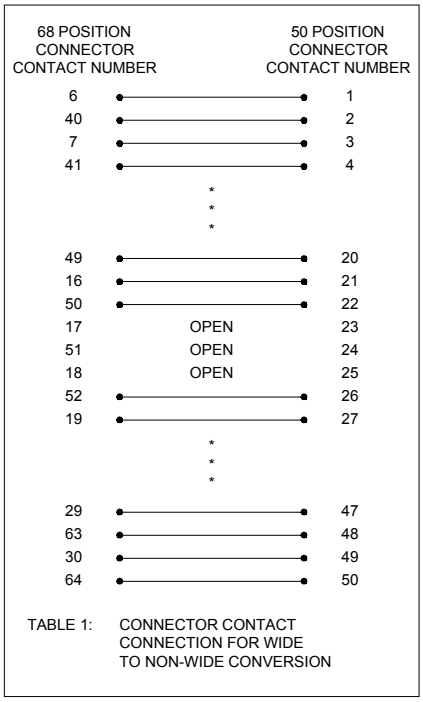
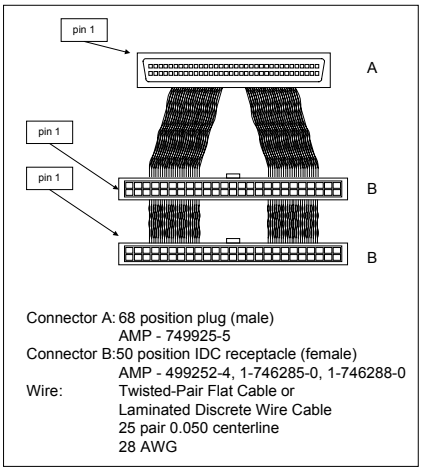


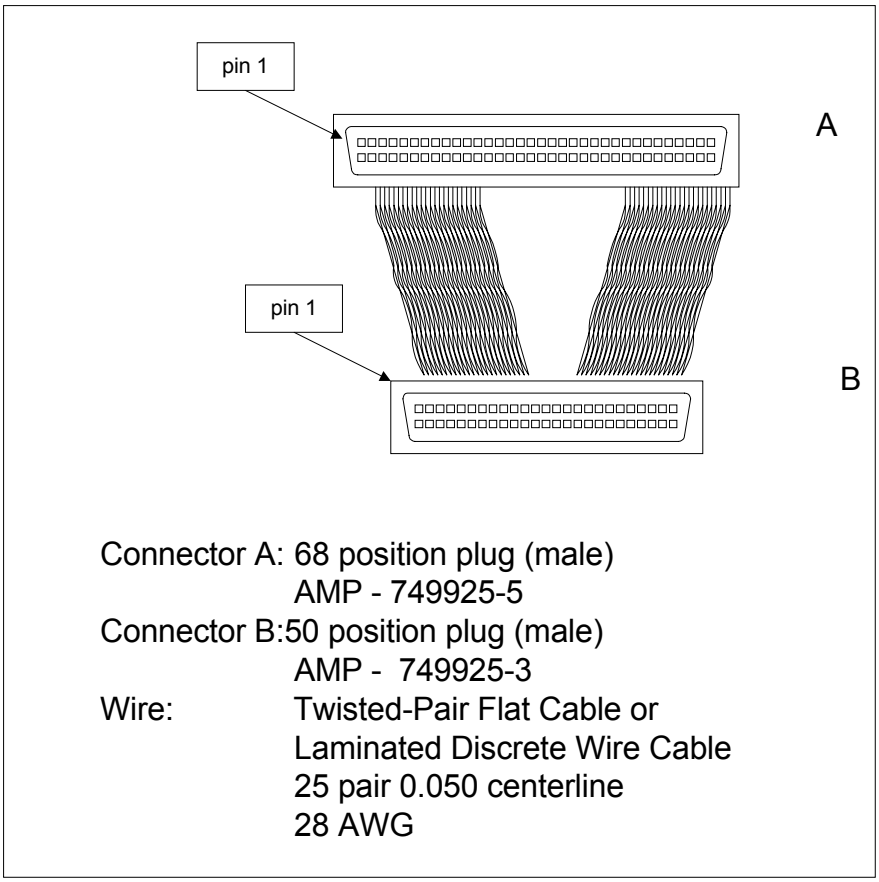
TABLE 1: CONNECTOR CONTACT CONNECTION FOR WIDE TO NON-WIDE CONVERSION



Connector A: 68 position plug (male)
AMP - 749925-5
Connector B: 50 position IDC receptacle (female)
AMP - 499252-4, 1-746285-0, 1-746288-0
Wire: Twisted-Pair Flat Cable or
Laminated Discrete Wire Cable
25 pair 0.050 centerline
28 AWG

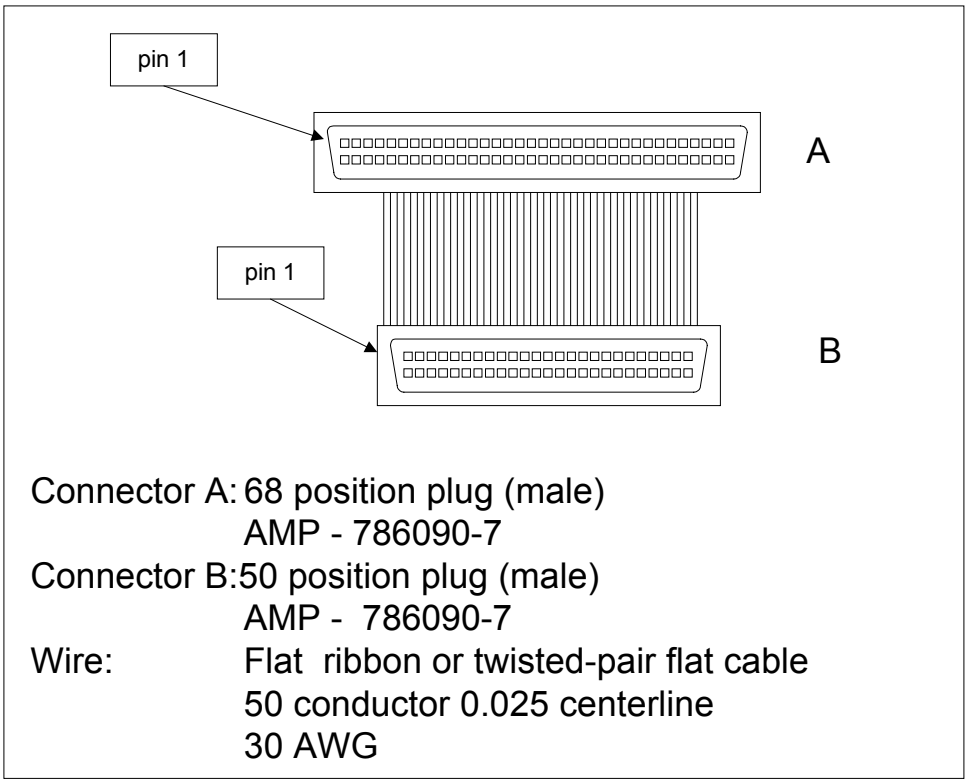
Cont'd

Converting Internal Wide to Internal Non-Wide (Type 30) The cable assembly for connecting internal wide SCSI devices to internal non-wide SCSI devices is shown below:



Cont'd

Converting from Internal Wide to Internal Non-Wide (Type 3) The cable assembly for connecting internal wide SCSI devices to internal non-wide (Type 3) SCSI devices is shown below:



SCSI Cable Vendors

Manufacturer	Telephone Number
Cables To Go	Voice: 800-826-7904 Fax: 800-331-2841
System Connection	Voice: 800-877-1985
Technical Cable Concepts	Voice: 714-835-1081
GWC	Voice: 800-659-1599

SCSI Connector Vendors

Manufacturer	Connector Part Number	Back Shell Part Number
AMP	749111-4	749193-1
Fujitsu	FCN-237R050-G/F	FCN-230C050-D/E
Honda	PCS-XE50MA	PCS-E50LA

68-Pin Connector Pinout for Single-Ended SCSI

Signal	Connector Pin	Cable Pin	Cable Pin	Connector Pin	Signal
Ground	1	1	2	35	-DB(12)
Ground	2	3	4	36	-DB(13)
Ground	3	5	6	37	-DB(14)
Ground	4	7	8	38	-DB(15)
Ground	5	9	10	39	-DB(P1)
Ground	6	11	12	40	-DB(0)
Ground	7	13	14	41	-DB(1)
Ground	8	15	16	42	-DB(2)
Ground	9	17	18	43	-DB(3)
Ground	10	19	20	44	-DB(4)
Ground	11	21	22	45	-DB(5)
Ground	12	23	24	46	-DB(6)
Ground	13	25	26	47	-DB(7)
Ground	14	27	28	48	-DB(P)
Ground	15	29	30	49	Ground
Ground	16	31	32	50	Ground
TERMPWR	17	33	34	51	TERMPWR
TERMPWR	18	35	36	52	TERMPWR
Reserved	19	37	38	53	Reserved
Ground	20	39	40	54	Ground
Ground	21	41	42	55	-ATN
Ground	22	43	44	56	Ground
Ground	23	45	46	57	-BSY
Ground	24	47	48	58	-ACK
Ground	25	49	50	59	-RST
Ground	26	51	52	60	-MSG
Ground	27	53	54	61	-SEL
Ground	28	55	56	62	-C/D
Ground	29	57	58	63	-REQ
Ground	30	59	60	64	-I/O
Ground	31	61	62	65	-DB(8)
Ground	32	63	64	66	-DB(9)
Ground	33	65	66	67	-DB(10)
Ground	34	67	68	68	-DB(11)

Cont’d

68-Pin SCSI Connector Pinout, Continued

High-Density Connector The following applies to the high-density SCSI connector table on the previous page:

- A hyphen before a signal name indicates that signal is active low.
- The connector pin refers to the conductor position when using 0.025 inch centerline flat ribbon cable with a high-density connector (AMPLIMITE.050 Series connectors).
- Eight-bit devices connected to the P-Cable must leave the following signals open: -DB (8), -DB (9), -DB (10), -DB (11), -DB(12), -DB (13), -DB (14), -DB 15), and -DB (P1).
- All other signals should be connected as defined.

> ***Caution***
>
> Lines labeled RESERVED should be connected to Ground in the bus terminator assemblies or in the end devices on the SCSI cable.
>
> RESERVED lines should be open in the other SCSI devices, but can be connected to Ground.

68-Pin Connector Pinout for Low-Voltage Differential SCSI

Signal	Connector Pin	Cable Pin	Cable Pin	Connector Pin	Signal
+DB(12)	1	1	2	35	-DB(12)
+DB(13)	2	3	4	36	-DB(13)
+DB(14)	3	5	6	37	-DB(14)
+DB(15)	4	7	8	38	-DB(15)
+DB(P1)	5	9	10	39	-DB(P1)
+DB(0)	6	11	12	40	-DB(0)
+DB(1)	7	13	14	41	-DB(1)
+DB(2)	8	15	16	42	-DB(2)
+DB(3)	9	17	18	43	-DB(3)
+DB(4)	10	19	20	44	-DB(4)
+DB(5)	11	21	22	45	-DB(5)
+DB(6)	12	23	24	46	-DB(6)
+DB(7)	13	25	26	47	-DB(7)
+DB(P)	14	27	28	48	-DB(P)
Ground	15	29	30	49	Ground
DIFFSENS	16	31	32	50	Ground
TERMPWR	17	33	34	51	TERMPWR
TERMPWR	18	35	36	52	TERMPWR
Reserved	19	37	38	53	Reserved
Ground	20	39	40	54	Ground
+ATN	21	41	42	55	-ATN
Ground	22	43	44	56	Ground
+BSY	23	45	46	57	-BSY
+ACK	24	47	48	58	-ACK
+RST	25	49	50	59	-RST
+MSG	26	51	52	60	-MSG
+SEL	27	53	54	61	-SEL
+C/D	28	55	56	62	-C/D
+REQ	29	57	58	63	-REQ
+I/O	30	59	60	64	-I/O
+DB(8)	31	61	62	65	-DB(8)
+DB(9)	32	63	64	66	-DB(9)
+DB(10)	33	65	66	67	-DB(10)
+DB(11)	34	67	68	68	-DB(11)

Notes The conductor number refers to the conductor position when using flat-ribbon cable.

B Audible Warnings

MegaRAID has an onboard tone generator that indicates events and errors.

Tone Pattern	Meaning	Examples
Three seconds on and one second off	A logical drive is offline.	One or more drives in a RAID 0 configuration failed. Two or more drives in a RAID 1, 3, or 5 configuration failed.
One second on and one second off	A logical drive is running in degraded mode.	One drive in a RAID 3 or 5 configuration failed.
One second on and three seconds off	An automatically initiated rebuild has been completed.	While you were away from the system, a disk drive in a RAID 1, 3, or 5 configuration failed and was rebuilt.

C Cluster Configuration with a Crossover Cable

When you are installing the Cluster Service on the first node in a server cluster, Setup may not detect the network adapter that is connected with a crossover cable. The icon in **Network and Dial-up Connections** that represents the network adapter connected to the crossover cable is displayed with a red X, and the **Network cable unplugged** icon in displayed on the taskbar.

You may also receive one of the following error messages:

During installation:

```
Only a singled Adapter is configured for internal cluster use. If
you have multiple adapters you may reconfigure them to avoid a
single point of failure.
```

Or, depending on the network role designated on other network adapters that are detected:

```
No network adapter was configured for internal cluster use.
```

The reason for this is because Media Sense is a default feature in Windows 2000 that removes bound protocols from an adapter sensed as "down" or "disconnected." Because the second node is powered off to avoid contention on the shared disk, Media Sense flags the network as "disconnected" because there is no end-to-end signal. During installation, the Cluster Service does not detect the adapter because there are no protocols bound to the adapter.

Solution

Note: Using Registry Editor incorrectly can cause serious problems that may require you to reinstall your operating system. Use Registry Editor at your own risk. You should back up the registry before you edit it. If you are running Windows NT or Windows 2000, you should also update your Emergency Repair Disk (ERD).

Disable the Media Sense feature:

1. Start Registry Editor (Regedt32.exe).

2. Locate the following key in the registry:

HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters

3. On the Edit menu, click Add Value, and then add the following registry value:

Value Name: DisableDHCPMediaSense
Data Type: REG_DWORD
Value: 1

4. Quit Registry Editor, and then restart the computer.

The network adapter still shows the "disconnected" status, but the cluster installation process can detect the adapter as available for cluster communication.

Alternatively, when you install the Cluster Service on the first node, you can have the second node powered up to the Control M (<Ctrl> <M>) menu. On the first node, a network connection will be detected for the private network.

Glossary

Array A grouping or array of disk drives combines the storage space on the disk drives into a single segment of contiguous storage space. MegaRAID can group disk drives on one or more SCSI channels into an array. A hot spare drive does not participate in an array.

Array Management Software Software that provides common control and management for a disk array. Array Management Software most often executes in a disk controller or intelligent host bus adapter, but can also execute in a host computer. When it executes in a disk controller or adapter, Array Management Software is often called firmware.

Array Spanning Array spanning by a logical drive combines storage space in two arrays of disk drives into a single, contiguous storage space in a logical drive. MegaRAID logical drives can span consecutively numbered arrays that each consist of the same number of disk drives. Array spanning promotes RAID levels 1, 3, and 5 to RAID levels 10, 30, and 50, respectively. See also *Disk Spanning*.

Asynchronous Operations Operations that bear no relationship to each other in time and can overlap. The concept of asynchronous I/O operations is central to independent access arrays in throughput-intensive applications.

Cache I/O A small amount of fast memory that holds recently accessed data. Caching speeds subsequent access to the same data. It is most often applied to processor-memory access, but can also be used to store a copy of data accessible over a network. When data is read from or written to main memory, a copy is also saved in cache memory with the associated main memory address. The cache memory software monitors the addresses of subsequent reads to see if the required data is already stored in cache memory. If it is already in cache memory (a cache hit), it is read from cache memory immediately and the main memory read is aborted (or not started.) If the data is not cached (a cache miss), it is fetched from main memory and saved in cache memory.

Channel An electrical path for the transfer of data and control information between a disk and a disk controller.

Cont'd

Consistency Check An examination of the disk system to determine whether all conditions are valid for the specified configuration (such as parity.)

Cold Swap A cold swap requires that you turn the power off before replacing a defective hard drive in a disk subsystem.

Data Transfer Capacity The amount of data per unit time moved through a channel. For disk I/O, bandwidth is expressed in megabytes per second (MB/s).

Degraded A drive that has become non-functional or has decreased in performance.

Disk A non-volatile, randomly addressable, rewritable mass storage device, including both rotating magnetic and optical disks and solid-state disks, or non-volatile electronic storage elements. It does not include specialized devices such as write-once-read-many (WORM) optical disks, nor does it include so-called RAM disks implemented using software to control a dedicated portion of a host computer volatile random access memory.

Disk Array A collection of disks from one or more disk subsystems combined with array management software. It controls the disks and presents them to the array-operating environment as one or more virtual disks.

Disk Duplexing A variation on disk mirroring where a second disk adapter or host adapter and redundant disk drives are present.

Disk Mirroring Writing duplicate data to more than one (usually two) hard disks to protect against data loss in the event of device failure. It is a common feature of RAID systems.

Disk Spanning Disk spanning allows multiple disk drives to function like one big drive. Spanning overcomes lack of disk space and simplifies storage management by combining existing resources or adding relatively inexpensive resources. For example, four 400 MB disk drives can be combined to appear to the operating system as one single 1600 MB drive. See also *Array Spanning* and *Spanning*.

Cont'd

Disk Striping A type of disk array mapping. Consecutive stripes of data are mapped round-robin to consecutive array members. A striped array (RAID Level 0) provides high I/O performance at low cost, but provides lowers data reliability than any of its member disks.

Disk Subsystem A collection of disks and the hardware that connects them to one or more host computers. The hardware can include an intelligent controller or the disks can attach directly to a host computer I/O a bus adapter.

Double Buffering A technique that achieves maximum data transfer bandwidth by constantly keeping two I/O requests for adjacent data outstanding. A software component begins a double-buffered I/O stream by issuing two requests in rapid sequence. Thereafter, each time an I/O request completes, another is immediately issued. If the disk subsystem is capable of processing requests fast enough, double buffering allows data to be transferred at the full-volume transfer rate.

Failed Drive A drive that has ceased to function or consistently functions improperly.

Fast SCSI A variant on the SCSI-2 bus. It uses the same 8-bit bus as the original SCSI-1, but runs at up to 10MB (double the speed of SCSI-1.)

Firmware Software stored in read-only memory (ROM) or Programmable ROM (PROM). Firmware is often responsible for the behavior of a system when it is first turned on. A typical example would be a monitor program in a computer that loads the full operating system from disk or from a network and then passes control to the operating system.

FlexRAID Power Fail Option The FlexRAID Power Fail option allows a reconstruction to restart if a power failure occurs. This is the advantage of this option. The disadvantage is, once the reconstruction is active, the performance is slower because an additional activity is added.

Cont'd

Format The process of writing zeros to all data fields in a physical drive (hard drive) to map out unreadable or bad sectors. Because most hard drives are factory formatted, formatting is usually only done if a hard disk generates many media errors.

GB Shorthand for 1,000,000,000 (10 to the ninth power) bytes. It is the same as 1,000 MB (megabytes).

Host-based Array A disk array with an Array Management Software in its host computer rather than in a disk subsystem.

Host Computer Any computer that disks are directly attached to. Mainframes, servers, workstations, and personal computers can all be considered host computers.

Hot Spare A stand-by drive ready for use if another drive fails. It does not contain any user data. Up to eight disk drives can be assigned as hot spares for an adapter. A hot spare can be dedicated to a single redundant array or it can be part of the global hot-spare pool for all arrays controlled by the adapter.

Hot Swap The substitution of a replacement unit in a disk subsystem for a defective one, where the substitution can be performed while the subsystem is running (performing its normal functions). Hot swaps are manual.

I/O Driver A host computer software component (usually part of the operating system) that controls the operation of peripheral controllers or adapters attached to the host computer. I/O drivers communicate between applications and I/O devices, and in some cases participates in data transfer.

Initialization The process of writing zeros to the data fields of a logical drive and generating the corresponding parity to put the logical drive in a Ready state. Initializing erases previous data and generates parity so that the logical drive will pass a consistency check. Arrays can work without initializing, but they can fail a consistency check because the parity fields have not been generated.

Cont'd

Logical Disk A set of contiguous chunks on a physical disk. Logical disks are used in array implementations as constituents of logical volumes or partitions. Logical disks are normally transparent to the host environment, except when the array containing them is being configured.

Logical Drive A virtual drive within an array that can consist of more than one physical drive. Logical drives divide the contiguous storage space of an array of disk drives or a spanned group of arrays of drives. The storage space in a logical drive is spread across all the physical drives in the array or spanned arrays. Configure at least one logical drive for each array.

Mapping The conversion between multiple data addressing schemes, especially conversions between member disk block addresses and block addresses of the virtual disks presented to the operating environment by Array Management Software.

MB (Megabyte) An abbreviation for 1,000,000 (10 to the sixth power) bytes. It is the same as 1,000 KB (kilobytes).

Multi-threaded Having multiple concurrent or pseudo-concurrent execution sequences. Used to describe processes in computer systems. Multi-threaded processes allow throughput-intensive applications to efficiently use a disk array to increase I/O performance.

Operating Environment The operating environment includes the host computer where the array is attached, any I/O buses and adapters, the host operating system, and any additional software required to operate the array. For host-based arrays, the operating environment includes I/O driver software for the member disks, but does not include Array Management Software, which is regarded as part of the array itself.

Cont'd

Parity Parity is an extra bit added to a byte or word to reveal errors in storage (in RAM or disk) or transmission. Parity is used to generate a set of redundancy data from two or more parent data sets. The redundancy data can be used to reconstruct one of the parent data sets. However, parity data does not fully duplicate the parent data sets. In RAID, this method is applied to entire drives or stripes across all disk drives in an array. Parity consists of dedicated parity, in which the parity of the data on two or more drives is stored on an additional drive, and distributed parity, in which the parity data are distributed among all the drives in the system. If a single drive fails, it can be rebuilt from the parity of the respective data on the remaining drives.

Partition An array virtual disk made up of logical disks rather than physical ones. Also known as logical volume.

Physical Disk A hard disk drive that stores data. A hard disk drive consists of one or more rigid magnetic discs rotating about a central axle with associated read/write heads and electronics.

Physical Disk Roaming The ability of some adapters to detect when hard drives have been moved to a different slots in the computer, for example, after a hot swap.

Protocol A set of formal rules describing how to transmit data, especially across a network. Low level protocols define the electrical and physical standards to be observed, bit- and byte-ordering, and the transmission and error detection and correction of the bit stream. High level protocols deal with the data formatting, including the message syntax, the terminal-to-computer dialogue, character sets, and sequencing of messages.

RAID Redundant Array of Independent Disks (originally Redundant Array of Inexpensive Disks) is an array of multiple small, independent hard disk drives that yields performance exceeding that of a Single Large Expensive Disk (SLED). A RAID disk subsystem improves I/O performance on a server using only a single drive. The RAID array appears to the host server as a single storage unit. I/O is expedited because several disks can be accessed simultaneously.

Cont'd

RAID Levels A style of redundancy applied to a logical drive. It can increase the performance of the logical drive and can decrease usable capacity. Each logical drive must have a RAID level assigned to it. The RAID level drive requirements are: RAID 0 requires one or more physical drives, RAID 1 requires exactly two physical drives, RAID 3 requires at least three physical drives, RAID 5 requires at least three physical drives. RAID levels 10, 30, and 50 result when logical drives span arrays. RAID 10 results when a RAID 1 logical drive spans arrays. RAID 30 results when a RAID 3 logical drive spans arrays. RAID 50 results when a RAID 5 logical drive spans arrays.

RAID Migration RAID migration is used to move between optimal RAID levels or to change from a degraded redundant logical drive to an optimal RAID 0. In Novell, the utility used for RAID migration is MEGAMGR and in Windows NT its Power Console. If a RAID 1 is being converted to a RAID 0, instead of performing RAID migration, one drive can be removed and the other reconfigured on the controller as a RAID 0. This is due to the same data being written to each drive.

Read-Ahead A memory caching capability in some adapters that allows them to read sequentially ahead of requested data and store the additional data in cache memory, anticipating that the additional data will be needed soon. Read-Ahead supplies sequential data faster, but is not as effective when accessing random data.

Ready State A condition in which a workable hard drive is neither online nor a hot spare and is available to add to an array or to designate as a hot spare.

Rebuild The regeneration of all data from a failed disk in a RAID level 1, 3, 4, 5, or 6 array to a replacement disk. A disk rebuild normally occurs without interruption of application access to data stored on the array virtual disk.

Rebuild Rate The percentage of CPU resources devoted to rebuilding.

Cont'd

Reconstruct The act of remaking a logical drive after changing RAID levels or adding a physical drive to an existing array.

Redundancy The provision of multiple interchangeable components to perform a single function to cope with failures or errors. Redundancy normally applies to hardware; a common form of hardware redundancy is disk mirroring.

Replacement Disk A disk available to replace a failed member disk in a RAID array.

Replacement Unit A component or collection of components in a disk subsystem that are always replaced as a unit when any part of the collection fails. Typical replacement units in a disk subsystem includes disks, controller logic boards, power supplies, and cables. Also called a hot spare.

SAF-TE SCSI Accessed Fault-Tolerant Enclosure. An industry protocol for managing RAID enclosures and reporting enclosure environmental information.

SCSI (Small Computer System Interface) A processor-independent standard for system-level interfacing between a computer and intelligent devices, including hard disks, floppy disks, CD-ROM, printers, scanners, etc. SCSI can connect up to 7 devices to a single adapter (or host adapter) on the computer's bus. SCSI transfers eight or 16 bits in parallel and can operate in either asynchronous or synchronous modes. The synchronous transfer rate is up to 160 MB/s (for 160M.) SCSI connections normally use single ended drivers, as opposed to differential drivers.

The original standard is now called SCSI-1 to distinguish it from SCSI-2 and SCSI-3, which include specifications of Wide SCSI (a 16-bit bus) and Fast SCSI (10 MB/s transfer). Ultra 160M SCSI is a subset of Ultra3 SCSI and allows a maximum throughput of 160 MB/s, which is more than twice as fast as Wide Ultra2 SCSI.

SCSI Channel MegaRAID controls the disk drives via SCSI-2 buses (channels) over which the system transfers data in either Fast and Wide or Ultra SCSI mode. Each adapter can control up to three SCSI channels.

Cont'd

Service Provider The Service Provider, (SP), is a program that resides in the desktop system or server and is responsible for all DMI activities. This layer collects management information from products (whether system hardware, peripherals or software) stores that information in the DMI's database and passes it to management applications as requested.

SMARTer Self-Monitoring, Analysis, and Reporting Technology with Error Recovery. An industry standard protocol for reporting server system information. Self-Monitoring, Analysis and Reporting Technology for disk drives is a specification designed to offer an early warning for some disk drive failures. These failures are predicted based upon actual performance degradation of drive components that are then reported through a graphical interface.

SNMP Simple Network Management Protocol is the most widely used protocol for communication management information between the managed elements of a network and a network manager. It focuses primarily on the network backbone. The Internet standard protocol developed to manage nodes on an Internet Protocol (IP) network.

Spanning Array spanning by a logical drive combines storage space in two arrays of disk drives into a single, contiguous storage space in a logical drive. MegaRAID logical drives can span consecutively numbered arrays that each consist of the same number of disk drives. Array spanning promotes RAID levels 1, 3, and 5 to RAID levels 10, 30, and 50, respectively. See also *Disk Spanning* and *Spanning*.

Cont'd

Spare A hard drive available to back up the data of other drives.

Stripe Size The amount of data contiguously written to each disk. You can specify stripe sizes of 2 KB, 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, and 128 KB for each logical drive. For best performance, choose a stripe size equal to or smaller than the block size used by the host computer.

Stripe Width The number of disk drives across which the data are striped.

Striping Segmentation of logically sequential data, such as a single file, so that segments can be written to multiple physical devices in a round-robin fashion. This technique is useful if the processor can read or write data faster than a single disk can supply or accept it. While data is being transferred from the first disk, the second disk can locate the next segment. Data striping is used in some modern databases and in certain RAID devices.

Terminator A resistor connected to a signal wire in a bus or network for impedance matching to prevent reflections, e.g., a 50-ohm resistor connected across the end of an Ethernet cable. SCSI chains and some LocalTalk wiring schemes also require terminators.

Ultra 160M A subset of Ultra3 SCSI, allows a maximum throughput of 160 MB/s, which is more than twice as fast as Wide Ultra2 SCSI. Ultra 160M allows the attachment of up to 15 SCSI devices (one SCSI ID is reserved for the controller), including a combination of LVD and older, single-end legacy devices, while maintaining backward compatibility with older versions of SCSI.

Ultra-SCSI An extension of SCSI-2 that doubles the transfer speed of Fast-SCSI, providing 20MB/s on an 8-bit connection and 40MB/s on a 16-bit connection.

Virtual Sizing FlexRAID Virtual Sizing is used to create a logical drive up to 80 GB. A maximum of eight logical drives can be configured on a RAID controller and RAID migration is possible for all logical drives except the eighth. Because it is not possible to do migration on the last logical drive, the maximum space available for RAID migration is 560 GB.

Wide SCSI A variant on the SCSI-2 interface. Wide SCSI uses a 16-bit bus, double the width of the original SCSI-1. Wide SCSI devices cannot be connected to a SCSI-1 bus. Wide SCSI supports transfer rates up to 20 MB/s, like Fast SCSI.

Write-Through/Write-Back When the processor writes to main memory, the data is first written to cache memory, assuming that the processor will probably read this data again soon. In write-through cache, data is written to main memory at the same time it is written to cache memory. In write-back cache, data is written only to main memory when it is forced out of cache memory.

Write-through caching is simpler than write-back because an entry to cache memory that must be replaced can be overwritten in cache memory because it will already have been copied to main memory. Write-back requires cache memory to initiate a main memory write of the flushed entry followed (for a processor read) by a main memory read. However, write-back is more efficient because an entry can be written many times to cache memory without a main memory access.

Index

D

E

F

G

H

I

J

L

M

N

O

P

R

S

T

U

V

W