Appendix A

System Features

A.1 Overview

The Suma RAID Subsystem comes with many different features. Some of these features enhance the performance of the system, other features add configuration flexibility and other features simplify the installation, maintenance and upgrade procedures of the system. This section highlights some of the features of the Suma RAID Subsystem.

A.2 Flexible Configuration Options

A.2.1 Single and Redundant Models

The Suma series described in this manual comes as a single or redundant RAID subsystem. The controller modules in the redundant controller subsystems come with three SFP connectors. Two SFP connectors are pre-set as host channels and the third SFP connector is used for system expansion. The controller module in the single controller subsystem comes with either two SFP connectors or two mini SCSI connectors that are used to connect the subsystem to a host.

A.2.2 Rear Panel Variations

The rear panels of the single and redundant controller modules do not appear the same. The single controller model has a pre-attached sheet placed over the location where the second controller module would be in the redundant model. This sheet should not be removed

as it is used to normalize the internal airflow and allow the system to be properly ventilated.

NOTE
If the metal sheet covering the second controller module bay in the single controller model is removed the subsystem may sustain irreparable damage.

A.2.3 Fibre Channel Configuration

All the Suma RAID Subsystems come with two 2GBps fibre Host channels. The redundant models also come with a 2GBps fibre channel expansion port.

The host channels support point-to-point, fibre switch and fibre channel arbitrated loop (FC-AL) connectivity, and are therefore able to connect to the Host Bus Adapter (HBA) through either a Hub or a Fabric Switch.

A.3 RAID Support and RAID Levels

The RAID controllers that come with the Suma RAID subsystems are designed to provide RAID level 0, 1 (0+1), 3, 5, 10, 30, 50, or JBOD RAID storage.

A.3.1 JBOD

JBOD stands for Just a Bunch of Drives. The controller treats each drive as a stand-alone disk, therefore each drive is an independent logical drive. JBOD does not provide data redundancy. (See ***Figure A- 1***)

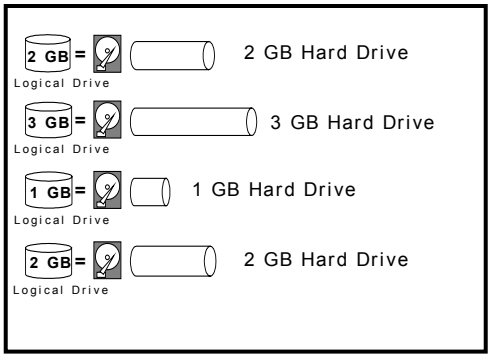

Figure A- 1: JBOD

A.3.2 RAID 0

RAID 0 implements *block striping* where data is broken into logical blocks and striped across several drives. Although called "RAID 0," this is not a true implementation of RAID because there is no facility for redundancy. In the event of a disk failure, data is lost.

In block striping, the total disk capacity is equivalent to the sum of the capacities of all drives in the array. This combination of drives appears to the system as a single logical drive.

RAID 0 provides the highest performance without redundancy. It is fast because data can be simultaneously transferred to/from multiple disks. Furthermore, read/writes to different drives can be processed concurrently.

RAID 0 block striping is illustrated in ***Figure A- 2***.

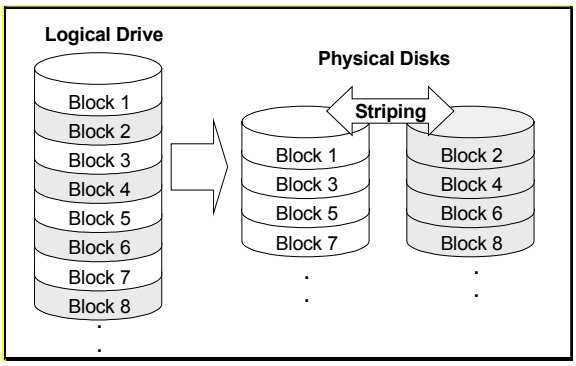

Figure A- 2: RAID 0

A.3.3 RAID 1

RAID 1 implements *disk mirroring* where a copy of the same data is recorded onto two sets of striped drives. By keeping two copies of data on separate disks or arrays, data is protected against a disk failure. If, at any time, a disk on either side fails, the remaining good disk (copy) can provide all of the data needed, thus preventing downtime.

In disk mirroring, the total disk capacity is equivalent to half the sum of the capacities of all drives in the combination. Thus, combining four 1GB SATA drives, for example, would create a single logical drive with a total disk capacity of 2GB. This combination of drives appears to the system as a single logical drive.

NOTE
One drawback to RAID 1 is that it does not allow running expansion. Once a RAID 1 array has been created, to expand it, the data must be backed up elsewhere before a new drive can be added. Other RAID levels permit running expansion.

RAID 1 is simple and easy to implement; however, it is more expensive as it doubles the investment required for a non-redundant disk array implementation.

RAID 1 mirroring is illustrated in ***Figure A- 3***.

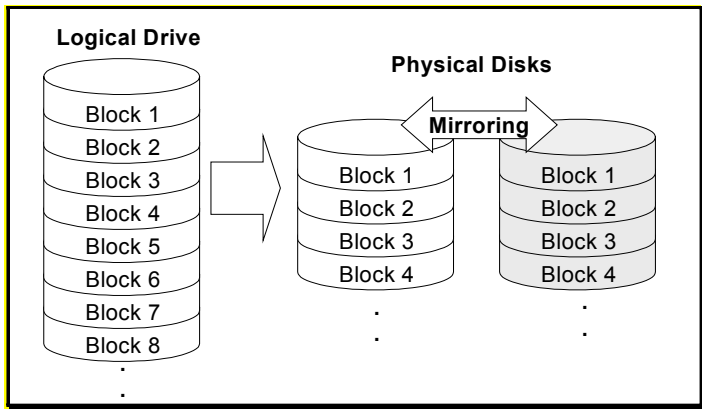

Figure A- 3: RAID 1

In addition to the data protection RAID 1 provides, this RAID level also improves performance. In cases where multiple concurrent I/Os are occurring, these I/Os can be distributed between disk copies, thus reducing total effective data access time.

A.3.4 RAID 1(0+1)

RAID 1 (0+1) combines RAID 0 and RAID 1 – *mirroring and disk striping.* RAID (0+1) allows multiple drive failure because of the full redundancy of the hard disk drives. If more than two hard disk drives are chosen for RAID 1, RAID (0+1) will be performed automatically.

RAID 1(0 + 1) mirroring and striping is illustrated in ***Figure A- 4***.

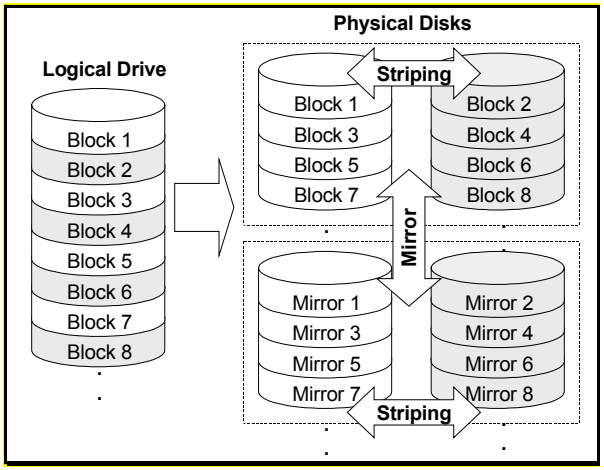

Figure A- 4*:* RAID 1(0 + 1)

IMPORTANT:
RAID (0+1) will not appear in the list of RAID levels supported by the controller. If you wish to perform RAID 1, the controller will determine whether to perform RAID 1 or RAID (0+1). This will depend on the number of drives selected for the logical drive.

A.3.5 RAID 3

RAID 3 implements *block striping with dedicated parity.* This RAID level breaks data into logical blocks, the size of a disk block, and then stripes these blocks across several drives. One drive is dedicated to parity. In the event a disk fails, the original data can be reconstructed from the parity information.

In RAID 3, the total disk capacity is equivalent to the sum of the capacities of all drives in the combination, excluding the parity drive. Thus, combining four 1GB SATA drives, for example, would create a single logical drive with a total disk capacity of 3GB. This combination appears to the system as a single logical drive.

RAID 3 provides increased data transfer rates when data is being accessed in large chunks or sequentially. However, in write operations that do not span multiple drives, performance is reduced since the information stored in the parity drive needs to be re-calculated and re-written every time new data is written to any of the data disks.

RAID 3 striping with dedicated parity is shown in ***Figure A- 5***.

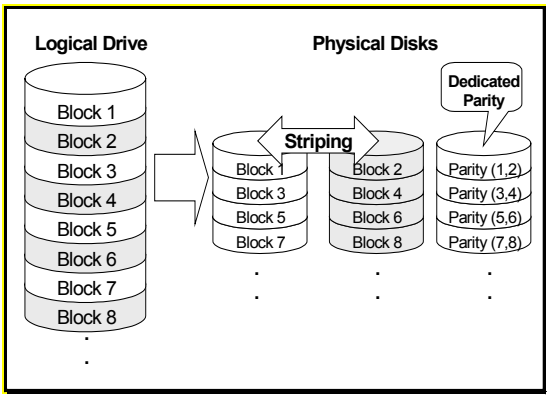

Figure A- 5: RAID 3

A.3.6 RAID 5

RAID 5 implements *multiple-block striping with distributed parity*. This RAID level offers the same redundancy available in RAID 3; though the parity information this time is distributed across all disks in the array. Data and relative parity are never stored on the same disk. In the event a disk fails, original data can be reconstructed using the available parity information.

An illustration of RAID 5 striping with non-dedicated parity is shown in ***Figure A- 6***.

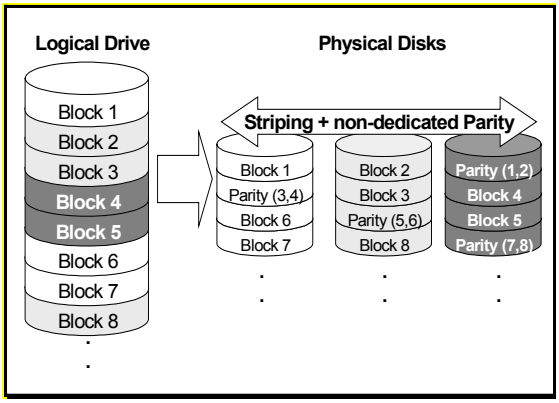

Figure A- 6: RAID 5

RAID 5 offers increased data transfer rates when data is accessed in large chunks (i.e., sequentially) and reduced data access time for many simultaneous I/O's when they do not span more than one drive.

A.3.7 RAID 30 and RAID 50

RAID 30 is a logical volume with RAID 3 logical drives. RAID 50 is a logical volume with RAID 5 logical drives.

A.3.8 Non-RAID Storage

One common option for expanding disk storage capacity is simply to install multiple disk drives into the system and then combine them end to end. This method is called *disk spanning*.

In disk spanning, the total disk capacity is equivalent to the sum of the capacities of all drives in the combination. This combination appears to the system as a single logical drive. Thus, combining four 1GB SATA drives in this way, for example, would create a single logical drive with a total disk capacity of 4GB.

Disk spanning is considered non-RAID due to the fact that it provides neither redundancy nor improved performance. Disk spanning is inexpensive, flexible, and easy to implement; however, it does not improve the performance of the drives and any single disk failure will result in total data loss.

An illustration of the non-RAID storage capacity is shown in ***Figure A- 7***

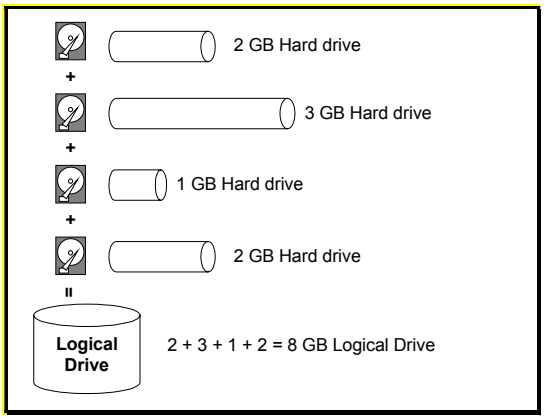

Figure A- 7: Non-RAID

A.3.9 Spares

RAID implementations include one other basic concept that needs to be introduced at this point: spare drives. RAID levels that have redundancy, levels 1, 3, and 5, all allow users to include a drive as a "spare." Spare drives are installed, fully functioning, "hot-ready" hard drives which a RAID controller will use to replace a failed drive as soon as the failure is detected. The purpose of this, obviously, is to enhance the existing fault-tolerant capabilities of a RAID array.

A.4 Redundant Features

A.4.1 Dual-Active Redundant Controllers

The RAID controllers in the EonStor subsystem can be configured in a dual-active redundant mode. If one controller fails the other controller will automatically take over the workload of the failed controller and manage the entire storage system. Both the cache writes and configuration data are fully synchronized. Failover and failback operations are completely transparent to the host and IOs are moved between the controllers without any user intervention.

A.4.2 Redundant Data Paths

Dual data paths can be connected to the HBA from the EonStor RAID Subsystem. If one data path should fail the data can be transmitted through the alternative data path. Full host side redundancy can be achieved if two host channels are connected to two separate HBAs and then to the Host computer.

A.5 Fault Tolerance

A.5.1 Intelligent Drive Handling

Hard drives can fail and bad blocks may occur simultaneously on two member drives of an array. The occurrence of bad blocks on more than one drive can cause loss of data. To prevent data loss two options can be implemented: "Media Scan" and "Bad Block Handling in Degrade Mode."

The Media Scan can be performed regularly to examine drives and, if any bad blocks are found during the process, data can be reconstructed onto good sectors automatically.

If bad blocks are encountered on yet another drive during the rebuild process, the block LBA (Logical Block Address) of those bad blocks will be shown and the rebuild process of the unaffected sectors will continue, salvaging most of your precious data.

Intelligent drive handling will occur in both the degraded mode and during the rebuild process. Optional write-verify for normal writes, rebuild writes and LD intialization is also available. Further low quality drive handling comes in the transparent resetting of hung hard drives. Power-failure management and bad drive handling during LD expansion provide further data security.

A.5.2 Hot-swappable active components

All the active components; including the controller modules, power supply units (PSU), the battery back up units (BBU), and the hard-drives are hot-swappable. If any of these components fail, they can be replaced without turning off the system or disrupting the smooth operation of the system.

A.5.3 Global and Local Spares

Both Global and Local (dedicated) spares are supported. The controller(s) will automatically disconnect from a failed drive and start to rebuild data on the spare drive. The spare drive will then replace the failed drive.

A.5.4 Hot-Swapping of Drives

A failed drive in the EonStor subsystem can be exchanged without turning off the system or interrupting the smooth operation of the system. Once the failed drive is replaced the data will be rebuilt in the background. Hot-swapping is supported through the automatic disconnection from a failed drive and the detection of a reserve drive. All these failure recovery procedures are completely transparent to the host.

A.5.5 S.M.A.R.T. Support

S.M.A.R.T (Self Monitoring Analysis and Reporting Technology) is supported with configurable reaction schemes. Users may select different reaction schemes for immediate prevention against S.M.A.R.T. detected errors. Available options include: detect only, clone and replace, and perpetual clone. A faulty drive can be cloned to an active spare upon the discovery of errors.

A.5.6 Other Fault Tolerant Features

Other comprehensive failure management features on the EonStor RAID Subsystem include:

- Automatic Bad Block Assignment
- Background rebuilding
- Verify-after-Write is supported on normal writes, rebuild writes and/or RAID initialization writes.
- Regeneration of parity of logical drives in the background.

A.6 SAN Features

A.6.1 Logical Unit Numbers

Up to 1024 Logical Unit Numbers (LUNs) are supported. Each LUN can be mapped to a logical unit (drive or volume). The LUN provides the logical unit with a unique signifier which enables the controllers to identify it.

A.6.2 LUN Masking

The RAID Controllers in the EonStor subsystem support LUN Masking. This enables a specific LUN to be uniquely assigned to a

specific host. The host will then only be able to access the LUNs which are assigned to it, all the other LUNs will be “hidden”.

A.7 Mechanical Features

A.7.1 Modular Design

The modular design of the EonStror simplifies the installation process and makes these systems easy to maintain. All the active components are modular and are therefore easy to replace.

A.7.2 Cableless Design

All the active components are cableless. This simplifies the installation of the system. Users do not have to be concerned with connecting any cables.

This page is intentionally left blank

APPENDIX A: SYSTEM FEATURES A-1

A.1 OVERVIEW A-1
A.2 FLEXIBLE CONFIGURATION OPTIONS A-1
A.2.1 Single and Redundant Models A-1
A.2.2 Rear Panel Variations A-1
A.2.3 Fibre Channel Configuration A-2
A.3 RAID SUPPORT AND RAID LEVELS A-2
A.3.1 JBOD A-2
A.3.2 RAID 0 ..... A-3
A.3.3 RAID 1 ..... A-4
A.3.4 RAID 1(0+1) A-5
A.3.5 RAID 3 ..... A-6
A.3.6 RAID 5 ..... A-7
A.3.7 RAID 30 and RAID 50 ..... A-8
A.3.8 Non-RAID Storage A-8
A.3.9 Spares A-9
A.4 REDUNDANT FEATURES A-10
A.4.1 Dual-Active Redundant Controllers A-10
A.4.2 Redundant Data Paths A-10
A.5 FAULT TOLERANCE A-10
A.5.1 Intelligent Drive Handling A-10
A.5.2 Hot-swappable active components A-11
A.5.3 Global and Local Spares A-11
A.5.4 Hot-Swapping of Drives A-11
A.5.5 S.M.A.R.T. Support A-12
A.5.6 Other Fault Tolerant Features A-12
A.6 SAN FEATURES A-12
A.6.1 Logical Unit Numbers A-12
A.6.2 LUN Masking A-12
A.7 MECHANICAL FEATURES A-13
A.7.1 Modular Design A-13
A.7.2 Cableless Design A-13

Automatic Bad Block Assignment, A-12

Background rebuilding, A-12
Bad Block Handling in Degrade Mode, A-10
block striping, A-3

controller, A-2

disk failure, A-3, A-4, A-9
disk mirroring, A-4
disk spanning, A-8
dual data paths, A-10
dual-active, A-10

failback, A-10
failed drive, A-11
failover, A-10
FC-AL, A-2
fibre switch, A-2

HBA, A-2
hot-swappable, A-11

Intelligent drive handling, A-11

JBOD, A-2

LUN, A-12
LUN Masking, A-12

Media Scan, A-10
modular, A-13
multiple-block striping with distributed parity, A-7

parity, A-6
point-to-point, A-2

RAID 0, A-3
RAID 1, A-4, A-5
RAID 1(0+1), A-5
RAID 3, A-6
RAID 5, A-7, A-8
RAID level, A-2
redundancy, A-2, A-3, A-9
Regeneration of parity, A-12

S.M.A.R.T, A-12
spares, A-11
Spares, A-9

Verify-after-Write, A-12