Benefits of Intel® Matrix Storage Technology

White Paper

December 2005

Document Number: 310855-001

Contents

Figures

Tables

Revision History

Document Number	Revision Number	Description	Revision Date
310855	-001	Initial Release.	December 2005

§

1 Abstract

We live in a digital world. We digitally create, record, edit, share, and save practically everything, from the movies we watch, to the pictures we take, to the presentations we create. The amount and value of data that is being stored on a PC for personal and office use has undergone a dramatic increase. A hard drive failure can result in unacceptable loss of home or office files. This paper provides information on the unique benefits offered by the Intel® Matrix Storage Technology. By combining hardware capabilities, a streamlined user interface, and proven RAID technology, Intel Matrix Storage Technology can help protect the user from losing important files when a hard drive failure occurs.

§

2 Introduction

2.1 Purpose and Scope

The amount of data we store on a PC has increased tremendously over the years as the world goes digital. People continue to store more and more important information in both their business and personal worlds. This information ranges from wedding and vacation photos to tax returns and business files. Often, most of us do not plan backup storage well in advance. A hard drive failure, resulting in the loss of all important information, can be traumatic because it results in content being irretrievable.

Intel Matrix Storage Technology offers unique benefits and, through the use of RAID technology, can help prevent users from losing important data. RAID technology uses two to four low cost hard drives that are linked together to form a single large capacity storage device, offering superior performance, storage capability and reliability over older storage solutions. Intel Matrix Storage Technology offers the following benefits: protection, fault tolerance, performance, ease of use, and newer usage models.

2.2 Glossary

Term	Definition
AHCI	Advanced Host Controller Interface
ATA	Advance Technology Attachment
BIOS	Basic Input/Output System
HDD	Hard Disk Drive
I/O	Input/Output
NCQ	Native Command Queuing
OROM	Option ROM
OS	Operating System
RAID	Redundant Array Of Independent Disks
ROM	Read Only Memory
SATA	Serial Advance Technology Attachment
SMART	Self-Monitoring, Analysis and Reporting Technology
SSU	Staggered Spin Up
UI	User Interface

§

3 Benefits of Intel Matrix Storage Technology

Intel Matrix Storage Technology provides a number of benefits for both mainstream and enthusiast end-users. These benefits are primarily targeted at providing high levels of data availability and integrity while maintaining ease of use. The benefits of Intel Matrix Storage Technology are as follows:

Protection: Since a lot of data needs to be stored reliably in today's environment, different RAID levels can be used to provide redundancy. Redundancy provides data protection when one of the disks fails.

- The Intel solution uses RAID 1 to provide seamless data protection against a single hard drive failure. RAID 1 mirrors data across two drives, allowing the system to function normally even if one drive fails.

- If a SMART event occurs on a member disk of a RAID volume and a spare is present, the Intel solution copies the data from the drive with the SMART event to the spare drive.

Performance: Performance is improved by allowing the controller to use the capabilities of multiple hard disks joined together with RAID or by taking advantage of the benefits of AHCI to access the data faster.

- The Intel solution provides support for AHCI (“Native Serial ATA”). One of the benefits of AHCI is NCQ support. NCQ allows multiple commands to be sent to the hard drive simultaneously and dynamically re-orders drive commands for more efficient hard drive access. This improves single-drive performance, and can also be used on systems with multiple hard drives.
- The Intel solution provides a high performance workstation storage solution by taking advantage of the benefits of RAID.
- RAID 0 can be scaled for up to four drives for high-performance media creation.

Protection & Performance

- The Intel solution supports high performance RAID 5 for protection of large amounts of critical data
- RAID 10 is also supported, allowing the benefits of both RAID 0 and RAID 1 to be combined. Refer to Appendix A for more information on RAID 10.

Ease of use: Intel Matrix Storage Technology provides a number of features that allow efficient and easy use of the Intel solution.

- Users can upgrade from a single drive to RAID capability by adding a second drive at a later date. No OS re-install is required and the system will remain usable while the migration is in progress.

- The Intel solution provides allows users to migrate a RAID 0, RAID 1 or RAID 10 volume to RAID 5. The system will remain available to the user while the migration is in progress. Appendix C lists the exact migrations supported.
- A streamlined user interface helps users manage their RAID systems more easily. The user interface has self monitoring tools that indicate RAID failures and SMART events, as well as features that allow users to configure RAID volumes, initiate migrations, and check the status of devices. Note: Refer to Appendix B for more information on SMART.
- For normal usage, character-based OROM user interface interaction is required only on initial setup.
- The Intel solution supports the hot plug of a device plugged directly into the SATA controller.

Productivity Enhancements: Intel Matrix Storage Technology allows for the “hot swap” of any member disk of a redundant RAID array. This eliminates the need to power down the system in the event of a drive hardware failure. After the failed drive is replaced, the system will remain operational during the subsequent rebuild. Note: hot swap must be enabled in the system BIOS by the OEM/motherboard manufacturer.

Fault Tolerance: Fault tolerance is the capability of a system to tolerate hard drive faults depending on the RAID level implemented. Fault tolerance is increased by redundancy and offers a lower chance for the whole subsystem to fail due to hardware failures

- The Intel solution supports RAID 1, 5 and 10.
- The Intel solution provides a seamless upgrade to a RAID level that offers fault tolerance via data migration.

Power savings: The Intel solution has the ability to spin up hard disks one at a time to reduce stress on small power supplies. This is called SSU capability.

Matrix RAID: Intel allows the user to combine the benefits of two RAID volumes in a single RAID array. The two RAID volumes can consist of the same or different RAID levels. Refer to Appendix A for a detailed description of matrix RAID.

§

Appendix A

Background on RAID

RAID technology was formally defined by a group of computer scientists at the University of California at Berkeley in the 1980s. They studied the possibility of using two or more disks that appear as a single device to the host system. Even though the array's performance was better than that of large single-disk storage systems, reliability was unacceptably low. To address this, the scientists proposed redundant architectures to provide ways of achieving storage fault tolerance.

In RAID, several drives are combined into a single RAID volume or array. The drives are partitioned into “strips” that can range in size from 1 sector (512 bytes) to 256 sectors (128K). Data is assigned to these strips in a round robin fashion across each of the drives in the array (not in RAID 1). The array management software controls where the data is to be written so that this data can be retrieved. Because data is read from each of the drives, performance gains can be achieved by retrieving portions of the data from multiple drives in parallel.

A “strip” is the grouping of data on a single disk. In Figure1 below, D0, D1, D2….DF individually are called strips. A “stripe” is the sum of all strips on a horizontal axis. So D0-D3 combined is a stripe.

RAID Levels

The different RAID levels supported are given below:

RAID 0: This is called "striping". This is a performance-oriented data mapping technique since there is a near-linear performance scaling with the addition of connected hard drive(s). Data written to the array is divided into stripes and written across the disks in the array. This procedure enables high I/O performance at a low cost. The fact that individual disks can be accessed in parallel improves performance.

In a two disk RAID 0 volume, the performance is roughly twice that of each component disk. Three and four disk RAID 0 volumes also have near linear scaling in performance. RAID 0 offers no data redundancy. If we have a four disk RAID 0 volume, the chance of a hardware disk drive breakdown is four times greater than that of a single disk.

Volume capacity = (# of HDDs) * (capacity of smallest HDD)

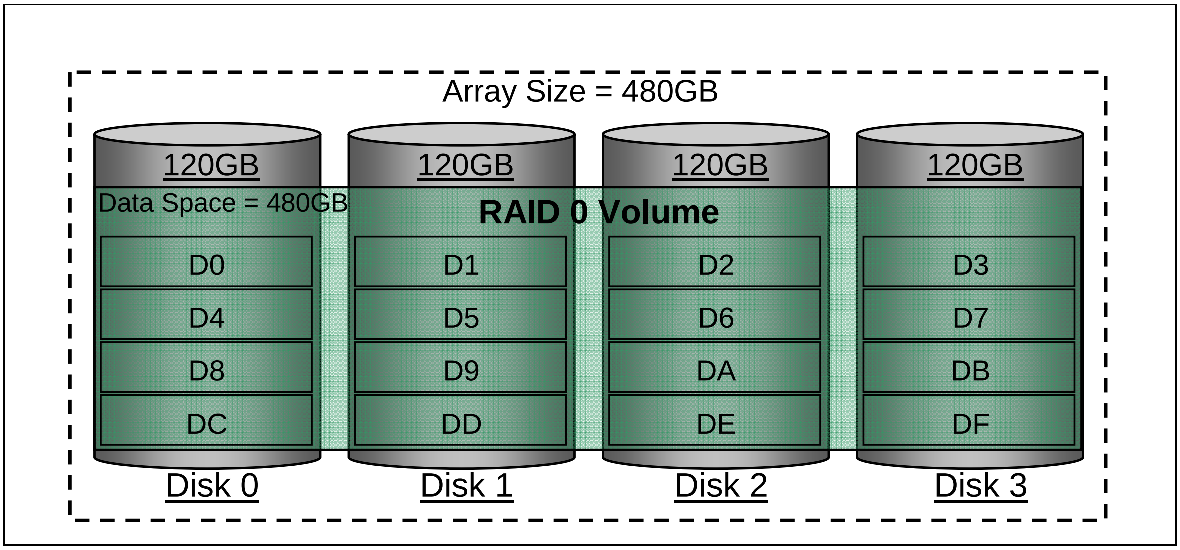

Figure 1: Four Disk SATA RAID Array

RAID 1: This is called "mirroring". It is a data redundancy technique. All data in the system is simultaneously written to two hard disks instead of one. The redundancy provides full protection against data loss in the event of the failure of either of the disks. When a hard drive failure occurs, all data is immediately available on the second hard drive. In Figure 2, if Disk 0 fails, all data is immediately available on Disk 1.

The performance is higher for read-intensive applications because only one read request needs to be issued, but the data can be read in parallel from either disk. In the case of read requests, the software uses an algorithm to effectively choose the more optimal disk to read the data from. If the disks are both busy, the requests are sent to the disk that is the least busy. If both are not busy, the request is sent to the disk which is closest to the location of the requested data.

The performance is lower for write-intensive applications since two write requests need to be issued, one to each disk.

Volume capacity = capacity of smallest HDD

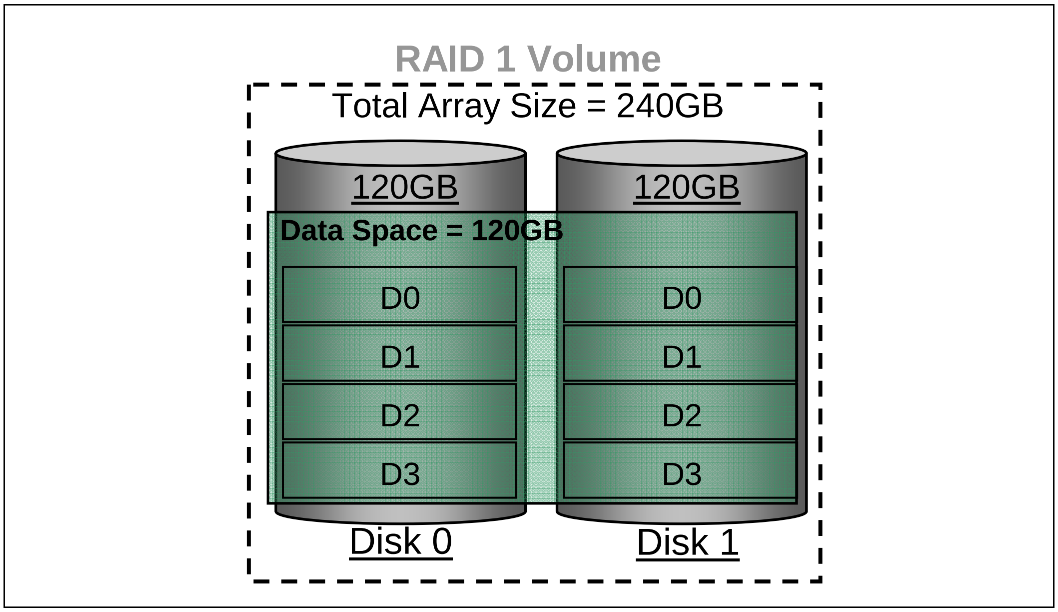

Figure 2: RAID 1 volume

RAID 5: RAID 5 stripes both data and parity information across three or more drives. RAID 5 has a distributed parity algorithm, where data and parity blocks are written across all drives in the array. Fault tolerance is maintained by ensuring that the parity information for any given block is placed on a drive separate from those used to store the data itself. Read operations outperform write operations. Read requests can be read from any of the disks, but in the case of write operations, the data must be written, parity must be calculated, and then parity must be written. RAID 5 requires a minimum of three hard drives.

Parity is the XOR of all the data in the stripe. In Figure 3 below, the data from drives D0, D1 and D2 are XOR'd and the result is stored in the fourth drive. This is indicated as P0 in the figure. A failed drive can be hot swapped with a new one, and the RAID software automatically rebuilds the lost data. In addition, RAID systems may be built using a spare drive (hot spare) ready and waiting to be the replacement for a drive that fails.

The user can choose to enable or disable the RAID 5 volume write cache. When it is enabled, it can operate in write-back or write through mode. Write-back cache means that when data is written by the OS, it is kept in memory and may not immediately be written to the disk. If power is lost, all data in the cache is lost. Write-through means that data is stored in the cache and is also written to the disk. Data is still in memory and there may be no need to access the disk for data access. Users can enable the volume write cache through the Intel Matrix Storage Manager even if there is no UPS system connected. In general, if there is a UPS connected to the system, then if a loss of power occurs, the data in the cache will be flushed to disk before the system shuts down.

Volume capacity = ((# of HDDs) – 1) * (capacity of smallest HDD)

Writes are coalesced to reduce the number of I/Os per write for parity calculations for sequential data streams. This means that if four sequential 16K access requests have been made, in the next I/O access, the 16K I/Os are combined into a single 64K access. RAID 5 utilizes a cache for write buffering and to improve coalescing for desktop usage models. ~4MB of main memory is allocated for the cache at boot time.

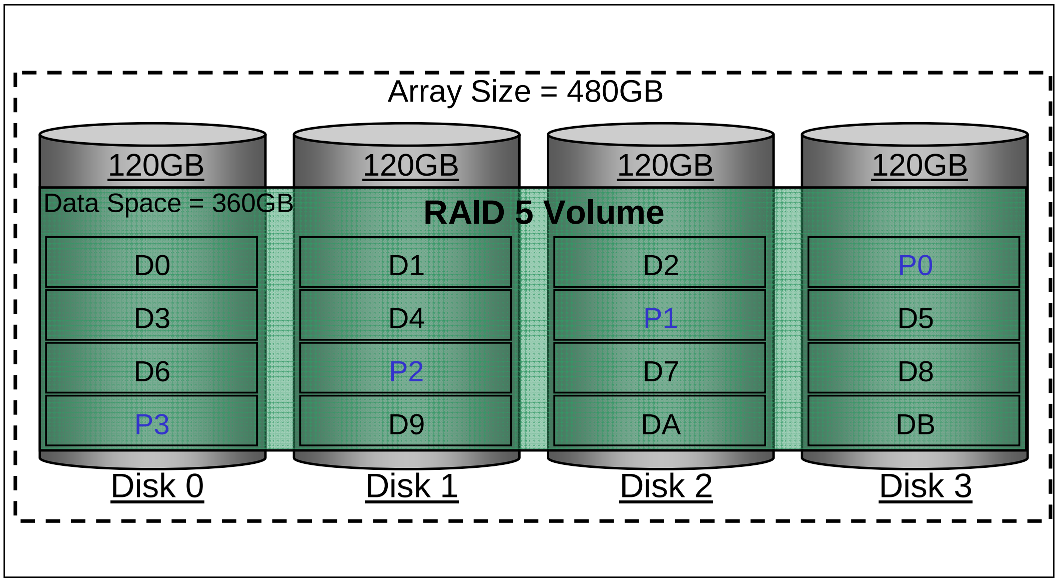

Figure 3: RAID 5 volume

RAID 10: This is also known as RAID 1+0 (striped mirrors) and is similar to RAID 0+1 (mirrored stripes). It uses a combination of RAID 0 and RAID 1. This combines the best features of striping and mirroring to yield large arrays with high performance and superior fault tolerance. RAID 10 provides excellent overall performance by combining the speed of RAID 0 with the redundancy of RAID 1. Four drives are required so it is more expensive in terms of redundancy information.

Volume capacity = 2 * (capacity of smallest HDD)

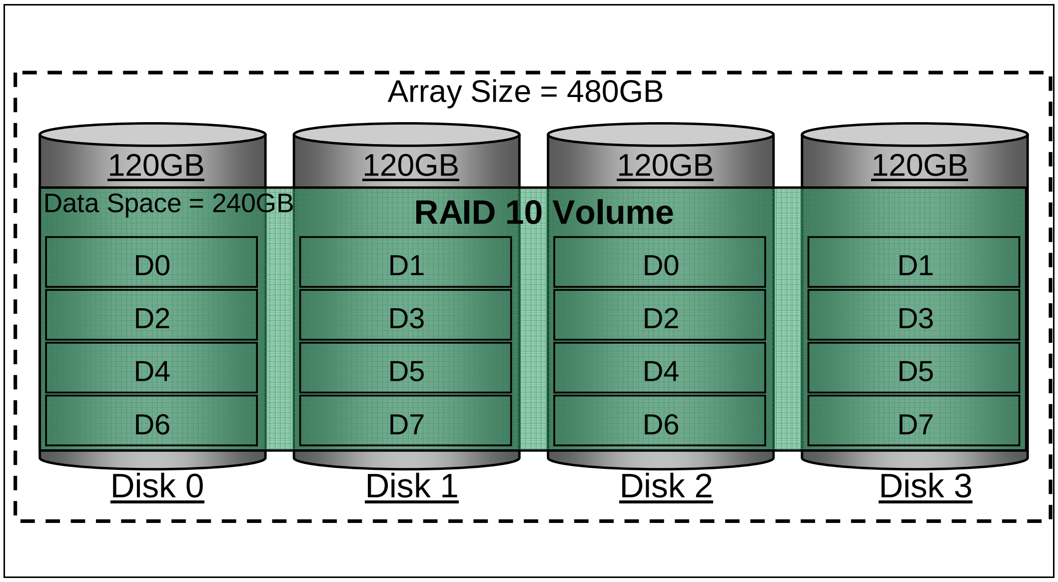

Figure 4: RAID 10 volume

Matrix RAID: Matrix RAID combines the benefits of two RAID levels as it allows two RAID volumes to be created on a single RAID array. As an example, a RAID 0 and a RAID 1 volume can be created on two hard drives; critical files can be stored on the RAID 1 volume and non-critical items can be stored on the RAID 0 volume.

S0- S5 refers to the RAID 1 volume in Figure 5 below. For the RAID 1 and RAID 0 combination, a RAID volume can be created using the option ROM or RAIDCfg.exe that spans half of the array. During Windows* setup, partitions and file systems can be created within the volume. The RAID 1 volume can be used to store critical data like pictures, home movies, financial records, and research papers, etc.

S0-SA and S1-SB refer to the RAID 0 volume in Figure 5 below. The RAID 0 volume can be created on the remainder of the array using the option ROM or Intel Matrix Storage Console. Partitions and file systems can be created within the volume. The RAID 0 volume can be used to store less critical items such as applications, games and video editing workspace.

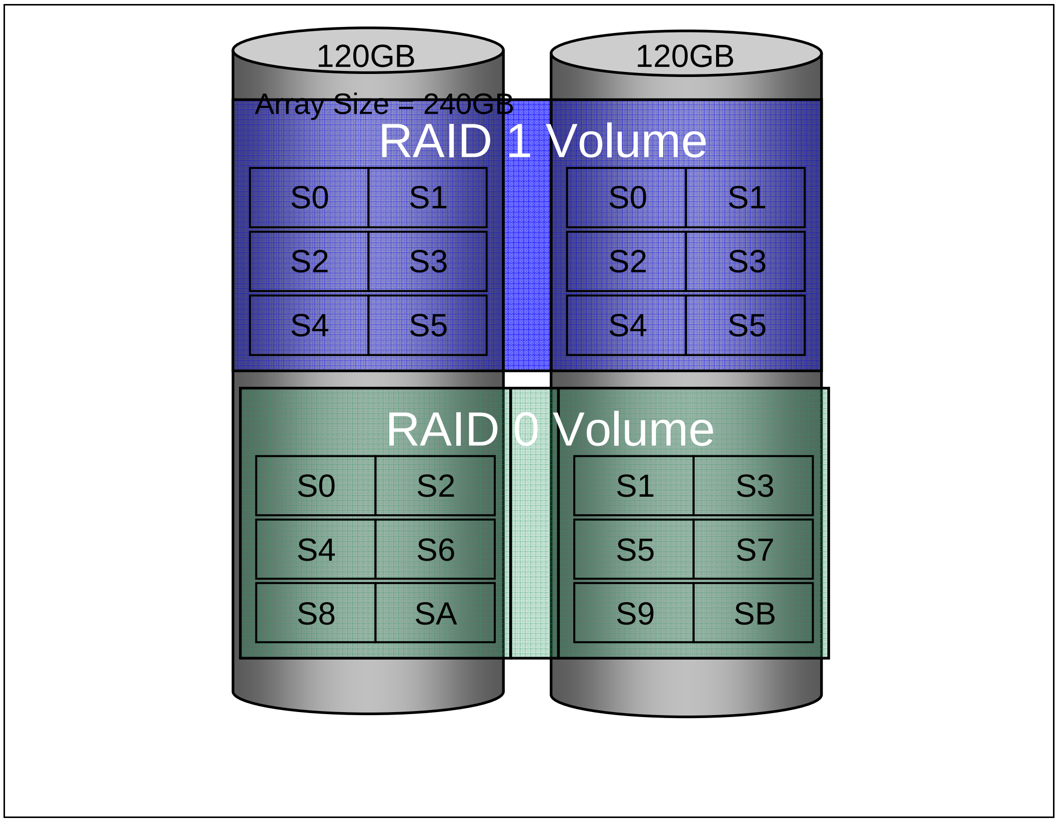

Figure 5: RAID 1 and RAID 0 combination

The top part of Figure 6 below refers to the RAID 5 volume. The bottom half of Figure 6 below refers to the RAID 0 volume. If RAID 5 and RAID 0 are combined, critical data/applications like the OS, business applications and personal media (e.g. photos, videos and financial records) can be stored on the RAID 5 volume. Non-critical data/applications like games, pagefile/scratch disks, and digital media workspace can be stored on the RAID 0 volume. This combination will help provide higher performance, capacity and fault tolerance and also provides up to 87.5% of usable total drive space.

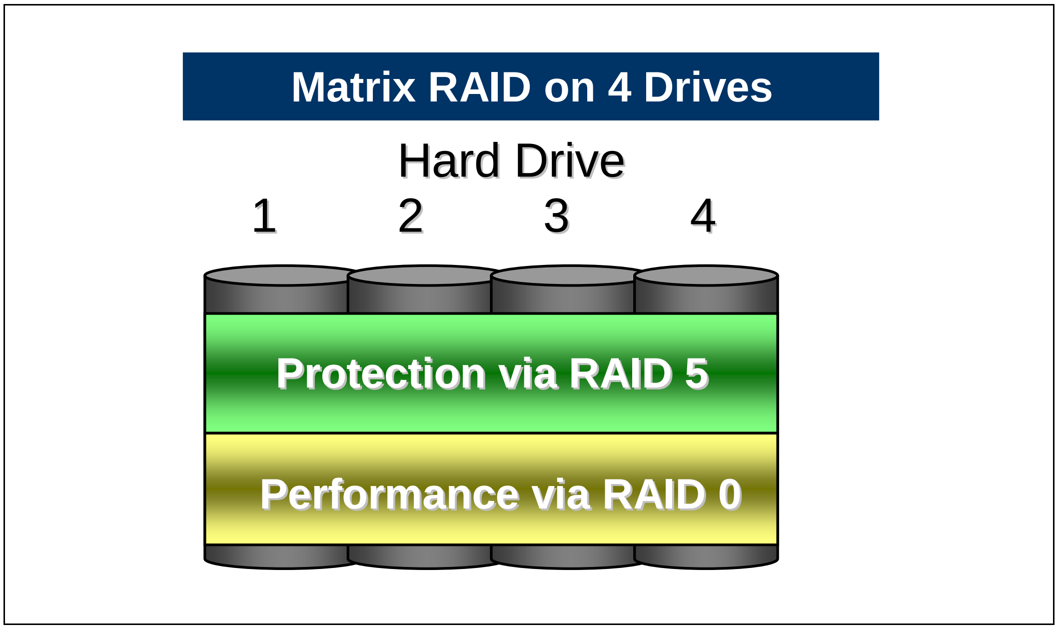

Figure 6: RAID 5 and RAID 0 combination

§

Appendix B

This section has additional information on SMART.

SMART

SMART is an open standard for developing hard drives and software systems that automatically monitors a hard drive's health and reports potential problems. The main principle behind SMART is that most problems with hard disks do not occur suddenly. They result from a slow degradation of some mechanical or electrical components. SMART serves as an "early warning system" for pending drive problems. This is important to the user because when there is a potential problem with the drive, the Intel Matrix Storage Manager can detect this event and provide a status in the user interface. Note: It is important that the user replace a drive on which a SMART event has taken place.

§

Appendix C

This section lists the exact migrations supported by the Intel solution.

Start\Finish	Single Drive	2 Disk RAID1	2 Disk RAID0	3 Disk RAID0	4 Disk RAID0	3 Disk RAID5	4 Disk RAID5	4 Disk RAID10
Single Drive	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2 Disk RAID1	No	N/A	No	No	No	Yes	Yes	No
2 Disk RAID0	No	No	N/A	No	No	Yes	Yes	No
3 Disk RAID0	No	No	No	N/A	No	No	Yes	No
4 Disk RAID 10	No	No	No	No	No	No	Yes	N/A

Table 1: Data Migration Matrix

§