TECHNOLOGY BRIEF

June 2002

Compaq Computer Corporation

Prepared by
ISS Technology Communications

CONTENTS

Executive Summary.....1

Introduction 3

RAID Levels: Functions and Limitations.....3

RAID ADG Provides Optimum Fault Tolerance 5

RAID ADG is a Cost-Effective Solution 7

Performance Considerations.....8

Choosing a RAID level 8

Summary.....9

Compaq RAID Advanced Data Guarding: A Cost-effective, Fault-tolerant Solution

EXECUTIVE SUMMARY

Although most customers are not experiencing sharp increases in business revenue, almost all are experiencing an explosion in the volume of enterprise data they are accruing. As customers evaluate storage solutions for this data, they are focusing on three very important needs: fault tolerance, storage capacity efficiency (cost effectiveness), and performance. Compaq RAID Advanced Data Guarding (ADG) is a cost-effective, fault tolerant solution for storing large volumes of enterprise data. Its performance, like that of other RAID levels, depends on the nature of the application. RAID ADG is supported on the Compaq Smart Array 5300 Controller and the Smart Array Cluster Storage and Modular SAN Array 1000 controllers.

Organizations implementing a large array should consider RAID ADG because it can tolerate up to two simultaneous drive failures without downtime or data loss. It can safely protect an array of up to 56 hard disk drives with greater fault tolerance than RAID 5 and with a lower implementation cost than RAID 1.

Please direct comments regarding this communication to the ISSG Technology Communications Group at this Internet address: TechCom@hp.com

COMPAQ

NOTICE

Compaq RAID Advanced Data Guarding: A Cost-effective, Fault-tolerant Solution

First Edition (June 2002)
Document Number TC020604TB

INTRODUCTION

RAID - Redundant array of independent disks.

Customers are being flooded by a rising tide of data that is the result of e-business and traditional applications: transaction processing, enterprise resource planning, decision analysis, etc. As the data level rises, customers are focusing on cost-effective storage technologies to protect the data they are amassing on an increasing number of disk drives. For these customers, RAID provides better performance and fault tolerance than storing data on separate hard drives. RAID enables a group of disk drives to function, from the operating system perspective, like a single physical disk drive.

RAID schemes, called levels, are differentiated by the method each uses to provide fault tolerance, but there is no correlation between the level numbers and the degree of fault protection. RAID 1, 1+0, 5, and Compaq RAID ADG are described in this paper because they are the levels best suited for arrays that store enterprise data. RAID 0 does not provide fault tolerance; however, it is described here in relation to RAID 1+0.

This paper first describes the functions and limitations of various RAID levels in protecting data in large storage volumes. Then, it describes three of the most important factors customers should consider in their decision-making process: fault tolerance, cost effectiveness, and performance.

RAID LEVELS: FUNCTIONS AND LIMITATIONS

Customers needing to create large arrays with a high number of disk drives, or with high-capacity disk drives, should consider the limitations of current RAID schemes in protecting data during a single- or multiple-drive failure. This section will help customers distinguish between the different RAID levels, including RAID ADG. Table 1 on the following page summarizes these RAID technologies in regards to function, best-suited applications, and limitations.

Striping – The spreading of data over multiple disk drives to improve performance. Data is interleaved by bytes or by sectors across the drives.

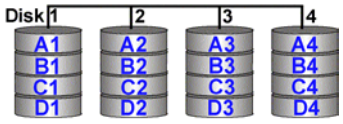
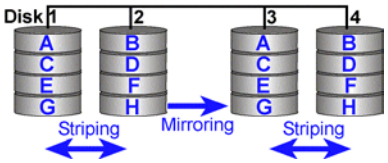
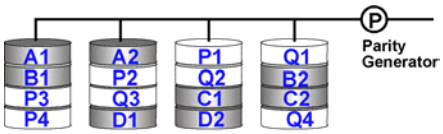
In a **RAID 0** implementation, a file is split into blocks and each block is *striped* across all the drives in the array. For large files, reading this data in parallel from the separate drives is faster than reading the file from a single drive. However, this RAID scheme offers no fault tolerance; the entire array will fail if one drive fails.

RAID 1 is a mirroring scheme that stores identical data on two sets of drives. It is used in applications that require very high availability. RAID 1 has high fault tolerance, but it has low storage efficiency because it requires twice the number of drives. **RAID 1+0** is implemented as a striped array of mirrored drives. It is best suited for sites that need high performance and maximum reliability, but are willing to sacrifice storage efficiency. RAID 1+0 can withstand the failure of half the drives as long as no two drives in a mirrored pair fail; however, it sacrifices storage capacity efficiency.

RAID 5 can be implemented on arrays of three or more drives. Parity information is calculated for each stripe of data and is placed on a drive other than the drive used to store the stripe of data. The parity information is striped across all drives in the array, and it occupies the equivalent capacity of one physical drive. Overall, RAID 5 provides good performance, but it can only withstand the loss of one drive without failure of the array.

Compaq RAID ADG is an extension of RAID 5 for implementation on arrays of four or more drives. Data and two sets of parity information are striped across all drives in the array. The additional set of parity improves the fault tolerance of the array but results in lower write performance. The two parity sets occupy the equivalent capacity of two drives in the array. RAID ADG protects against the simultaneous failure of two drives in the array.

Table 1. Summary of RAID Technologies for Large Arrays

RAID LEVELS	Function/Applications	Limitations
RAID 0  Requires a minimum of one drive.	Files are split up and striped across separate disk drives. Image Editing • Video Production • Pre-Press Applications	No fault tolerance and highly vulnerable to failure. The entire array will fail if one drive fails.
RAID 1  Requires a minimum of two drives.	Mirroring - Identical data stored on two drives, high fault tolerance, very good performance. Accounting • Payroll • Financial	Requires 50% of capacity to be dedicated to fault protection. Doubles the number of drives required.
RAID 1+0  Requires a minimum of four drives.	Implemented as a stripe of mirrored disks. Database applications requiring high performance and fault tolerance, but willing to sacrifice storage efficiency.	
RAID 5  Requires a minimum of three drives. P*n* represents one set of parity.	One set of parity data is distributed across all drives. Protects against the failure of any one drive in an array. Transaction processing • File and application servers • ERP • Internet and Intranet servers	Relatively low fault tolerance. It can only withstand the loss of one drive without total array failure so it is highly risky for large arrays.
Compaq RAID Advanced Data Guarding  Requires a minimum of four drives. P*n* and Q*n* represent two sets of parity.	Two sets of parity data are distributed across all drives. Protects against the failure of two drives in an array. Provides higher fault tolerance than RAID 5. For 24x7 applications that require a higher level of fault tolerance than RAID 5.	Lower write performance than other RAID methods.

RAID ADG PROVIDES OPTIMUM FAULT TOLERANCE

Often, the terms "reliability" and "fault tolerance" are used interchangeably in describing RAID schemes; however, there is a distinction between them. Reliability refers to the likelihood that an individual drive or drive array will continue to function without experiencing a failure. Reliability is typically measured over some period of time. Although reliability is a very important issue for customers, it is a function of the reliability of the individual array components rather than the RAID technology implemented.

Fault tolerance, on the other hand, is the ability to withstand and recover from a failure despite disk drive reliability issues. Fault tolerance is provided by some sort of redundancy—mirroring, parity, or a combination of both—and it is typically measured by the number of drives that can fail without bringing down the entire array. The fault tolerance of various RAID levels is as follows:

- **RAID 0** has no fault tolerance because it uses no type of redundancy. The array will fail if one physical drive fails.
- With **RAID 1 or 1+0**, the maximum number of hard drives that can fail without failure of the array is n/2, assuming none of the failed drives are mirrored to each other. In practice, logical drive failure will usually occur before this maximum number is reached, since the probability of a newly failed drive not being mirrored to a previously failed drive becomes increasingly small as the number of failed drives increases. A RAID 1+0 array will fail if a drive and its mirrored drive fail, but the probability of this decreases as the number of mirrored pairs increase.
- **RAID 5** can withstand the failure of one physical drive. The array will fail if two drives fail simultaneously.
- **RAID ADG** can withstand the failure of two physical drives. Three hard drives must fail before the entire array will fail. RAID ADG also protects against the loss of data if a drive fails and a defect occurs in a single sector of another drive. This is important if data is being rebuilt after a drive failure and a media defect occurs in one of the good drives.

Although RAID 1 and RAID 1+0 provide a higher level of fault tolerance than RAID 5, this protection comes at a very high price because 50 percent of the drives are dedicated to fault protection. For RAID 5 configurations, we recommend using no more than 14 physical drives be used per array due to the increased likelihood of drive array failure with more hard drives. For arrays with more than 14 drives, we recommend RAID ADG due to its fault tolerance and storage efficiency. RAID ADG supports the use of up to 56 physical drives per array because that is the maximum number of hard drives that can be physically connected to a four-channel Compaq Smart Array Controller.

Note: RAID ADG is standard with the four-channel Compaq Smart Array 5300 (SA 5304/256) Controllers and is sold as an option for the two-channel Smart Array models (SA 5302/128). RAID ADG requires a minimum of 64 MB read/write cache.

Figure 1 shows the relative probability of logical drive failure for different RAID settings and different logical drive sizes, assuming no online spares are present. RAID ADG is less likely to experience a logical drive failure than RAID 0, 5, and 1+0. An online spare (hot spare) can be added to any of the fault-tolerant RAID levels to further decrease the probability of logical drive failure. As soon as drive failure occurs, missing data can be automatically rebuilt from parity data onto the online spare. Without an online spare there is a chance that more hard drives will fail before the failed drive is replaced, leading to failure of the complete drive array and consequent loss of data. This is less likely with RAID ADG than with RAID 5, since RAID ADG can sustain failure of two drives. RAID ADG supports online spare drives and Online RAID Level Migration from RAID 1 or RAID 5.[1]

[1] For more information about online spare drives and Online RAID Level Migration from RAID 1 or RAID 5, please refer to the *Advanced Data Guarding for the Smart Array 5300 Controller Reference Guide*, part number 188270-001 at www.compaq.com.

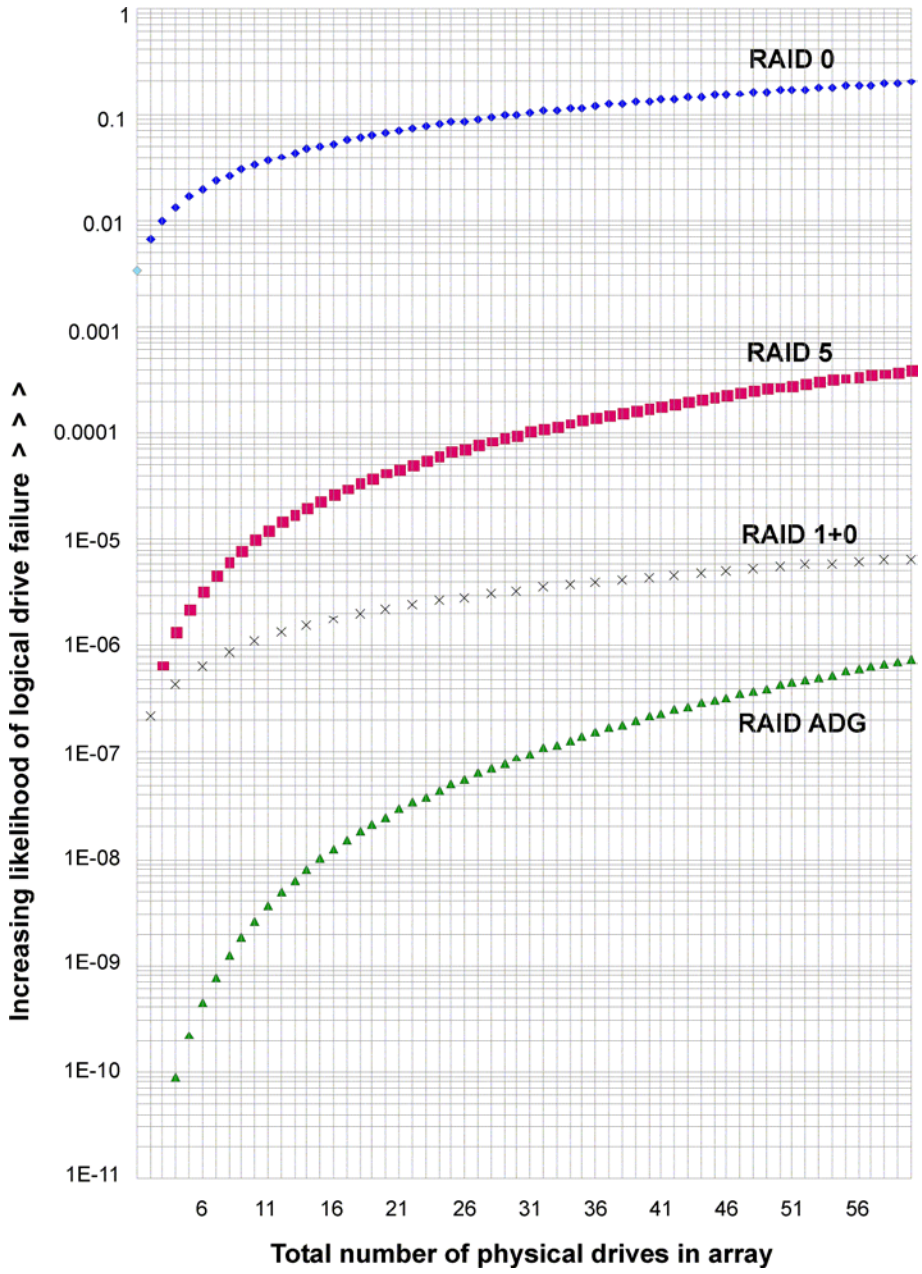


Figure 1. Probability of drive array failure for different RAID levels. RAID ADG is less likely to experience a complete drive array failure than RAID 0, 1+0, and 5.

RAID ADG IS A COST-EFFECTIVE SOLUTION

The cost effectiveness of each RAID solution is a balance between the total cost of the array and its usable capacity. While the total cost includes all the drives in the array, the usable capacity includes only the drives that store non-redundant (not parity or mirrored) data. One way to evaluate cost effectiveness is to compare the cost per gigabyte of usable capacity of various RAID levels. Another useful way to evaluate cost effectiveness is to compare storage efficiency—the usable capacity divided by the total of capacity of all the drives.

An important factor to note is that the usable capacity of any RAID array is limited by the size of the smallest hard drive in the array; the extra capacity on larger drives goes unused. For example, an array with four drives (40 GB, 60 GB, 60 GB, and 60 GB) would have a usable capacity of 4 x 40 GB, or 160 GB. To maximize storage efficiency, all RAID array drives should have the same capacity. If drives with different capacities are attached to the same controller, it is possible to create multiple arrays that contain only drives of the same capacity.

Table 2 lists the storage efficiencies of the various RAID levels. The storage efficiency of RAID 1 and RAID 1+0 is constant, but the storage efficiency of RAID 5 and RAID ADG varies with the number of drives. The number of parity drives in RAID 5 and RAID ADG schemes is fixed (one parity drive for RAID 5 and two parity drives for RAID ADG), so their storage efficiency increases with the number of drives.

As shown in Table 2, RAID 1 and 1+0 have the lowest storage efficiency at 50 percent; therefore, they are less cost-effective solutions for large arrays. A comparison of the storage efficiencies of RAID 5 and RAID ADG requires closer scrutiny. For a given number of drives, RAID 5 will have higher storage efficiency than RAID ADG, although this difference shrinks as the number of drives increases. The storage efficiency of a RAID 5 array varies from 67 percent for three drives to 93 percent for 14 drives (the maximum recommended by Compaq). The storage efficiency of RAID ADG varies from 50 percent for four drives to 96 percent for a maximum of 56 drives. A maximum of 56 drives can be physically connected to a four-channel controller using four external Compaq storage systems.

The bottom line is that RAID 1 and 1+0 have very low storage efficiency. RAID 5 and RAID ADG have much higher storage efficiencies, but the level of efficiency depends on the number of drives in the array.

Table 2. Summary of RAID Array Storage Efficiency*

	RAID 1	RAID 1+0	RAID 5	Advanced Data Guarding
Usable Capacity (C= capacity of smallest drive; n= number of drives)	C*(n/2)	C*(n/2)	C*(n-1)	C*(n-2)
Minimum number of drives	2	4	3	4
Maximum recommended number of drives*	N/A	N/A	14	N/A
Storage efficiency from minimum to maximum recommended no. of drives**	50%	50%	67% to %93	50% to 96%

*Compaq recommends not exceeding these maximum figures (excluding any allowable online spares) when configuring a drive array, due to the increased likelihood of drive array failure if more hard drives are added.

**The value for storage efficiency is calculated assuming all drives in the array have the same capacity with no online spares.

PERFORMANCE CONSIDERATIONS

The key to RAID performance is parallelism—the ability to access multiple disks simultaneously. This parallelism allows data to be written to or read from a RAID array faster than would be possible with a single drive. Analyzing RAID performance can be very complicated because several factors must be considered (sequential versus random reads and writes, block size, data transfer rate, etc).

The performance of a RAID array can be subdivided into read performance and write performance, which vary based on the RAID level. For example, in RAID 1 (mirroring), the data and redundant data must be written to two separate drives. The data is normally read from one drive, so the read performance is much faster than the write performance; however, data can be read from both drives to increase the read performance. RAID 0 uses striping to improve performance by splitting up files into smaller portions (stripes) and distributing them to multiple hard disks; however, RAID 0 has no fault tolerance. RAID 5 and RAID ADG also use striping, but their write performance is significantly affected by the multiple reads and writes needed to perform the parity calculations prior to updating the array. The write performance of RAID ADG is less than that of RAID 5 because RAID ADG has dual parity overhead. The read performance of RAID 5 and RAID ADG is very good and may be improved by tweaking the stripe size.

In the final analysis, RAID array performance boundaries are largely predetermined by the intended application. Applications that require a large capacity array, yet have a relatively low ratio of writes to reads (ERP, transaction processing, and web servers) may benefit from striping with parity—RAID 5 or RAID ADG. On the other hand, applications that have a much higher ratio of writes to reads (file servers, database servers, and media development), may benefit from a RAID 1+0 array; but eventually cost becomes an issue as capacity requirements grow. The following section will help guide customers through this complicated decision-making process.

CHOOSING A RAID LEVEL

To choose the optimum RAID level for data protection in large arrays, customers should consider a variety of factors, including:

- Fault tolerance (based on availability requirements)
- Cost effectiveness (based on storage efficiency or cost per gigabyte of usable capacity)
- Performance (based on application requirements)

Customers can use the decision chart in Table 3 on the next page to determine which RAID level provides the best solution. For example, if cost effectiveness is of primary importance and fault tolerance is of secondary importance, or vice versa, the best choice is RAID ADG.

Table 3: Important Factors in Choosing the Optimum RAID Level for Large Arrays

Most Important	Secondary Importance	RAID Level Choice
Cost Effectiveness	Fault Tolerance →	RAID ADG
	Performance →	RAID 5 (RAID 0 if fault tolerance is not needed)
Fault Tolerance	Cost Effectiveness →	RAID ADG
	Performance →	RAID 1+0
Performance	Cost Effectiveness →	RAID 5 (RAID 0 if fault tolerance is not needed)
	Fault Tolerance →	RAID 1+0

SUMMARY

Compaq RAID ADG provides an advanced level of data protection for customers who need a higher level of fault tolerance than RAID 5 and a lower implementation cost than RAID 1. RAID ADG is best implemented when customers need to protect enterprise data at a lower cost than RAID 1 arrays, and when performance is not an overriding factor.

RAID ADG is supported on the Compaq Smart Array 5300 Controller and can effectively protect an array of up to 56 total drives. A RAID ADG array can tolerate up to two simultaneous drive failures without downtime or data loss. RAID ADG supports Online Spare Drives and Online RAID Level Migration from RAID 1 or RAID 5.