

Figure 1: Two sets of parity data, P+Q, are striped across disks. RAID 6 safeguards against data loss during rebuild mode by allowing up to two consecutive drive failures for maximum fault tolerance.

- High capacity systems with eight or more high capacity drives
- For enterprise/mission critical installations

3ware RAID 6 Highlights

- **Red Hot RAID 6**
Simultaneous parity calculations and dedicated silicon-based stripe handlers radically reduce the RAID 6 write penalty
- **Higher data availability**
Data is safeguarded against up to two consecutive drive failures
- **2-Drive Parity**
Data from two failed drives can be rebuilt with assured data accessibility
- **RAID Protection in Degraded Mode**
Data is protected against another drive failure during rebuild or from latent media errors in degraded mode
- **Less than 10% RAID 6 performance penalty vs. RAID 5**

RAID 6

Double-parity RAID, commonly known as RAID 6, safeguards against data loss by allowing up to two consecutive drive failures

Background

An understanding of RAID 5 is critical to the understanding of RAID 6. In a RAID 5 array, data is striped across all drives in the array. Parity information is distributed and stored across all the disks. If an individual drive fails, the surviving array operates in degraded mode until the failed drive is replaced and its data is rebuilt from the parity information retained on the surviving disks.

The vulnerability of RAID 5 arrays is manifested within two types of error conditions. The first condition is an obvious one; all data will be lost in the unlikely event of a second drive failure during a rebuild. Latent media errors, the second error condition, is not so obvious, but is equally important to protect against. A latent media error is triggered by an undetected media defect that causes a read error during a rebuild. This type of error can lead to data loss due to the controller's inability to rebuild a stripe during RAID 5 degraded mode operations. Today's increased hard disk capacities are causing longer rebuild times. Longer rebuild times widen the window of likelihood that a second drive will fail during rebuild which would necessarily result in data loss without RAID 6 protection.

RAID 5 Performance Considerations

RAID 5 write performance is influenced by the number of disk accesses that are required during the write process. While there is no adverse effect on RAID 5 read performance, write performance can drop by almost 50% between RAID 0 (data striping across multiple drives) and RAID 5 (data striping across multiple drives with distributed parity). The effect on overall RAID 5 performance will depend on the ratio between reads and writes for a given application; in general, more writes mean lower performance.

What is RAID 6?

RAID 6 eliminates the risk of data loss if a second hard disk drive fails or an unrecoverable read error occurs while the RAID array is rebuilding. In a RAID 6 enabled system, a second set of parity is calculated, written, and distributed across all the drives. This second parity

calculation provides significantly more robust fault tolerance because two drives can fail without resulting in data loss. A RAID 6 implementation is diagramed above (Figure 1).

RAID 6 Performance Considerations

RAID 6 requires a second set of parity calculations to be made so that data from two failed drives can be rebuilt from the parity information retained on the surviving disks. This additional parity calculation step adversely affects write performance. Performance benchmarks show that a RAID controller can suffer more than a 30% drop in overall write performance in RAID 6 compared to a RAID 5 implementation. RAID 5 and RAID 6 read performance is unaffected.

AMCC has implemented sophisticated Galois field algorithms and improved StreamFusion (advanced stripe handling) technology to maximize RAID 5 and RAID 6 performance by dramatically minimizing the write performance penalty. Simultaneous parity calculations and dedicated silicon-based stripe handlers radically reduce the write penalty on 3ware controllers to less than 15% between RAID 5 and RAID 0 and to less than 10% between RAID 6 and RAID 5.

Capacity Considerations

RAID 5 implementations require a minimum of three drives and have the storage capacity of N-1 drives because the equivalent capacity of one drive is exclusively dedicated to holding parity data. For example in a 4 drive, 200GB per drive array, the total available storage capacity is 600GB out of a total of 800GB.

RAID 6 implementations require a minimum of 4 drives and have the storage capacity of N-2 drives because the equivalent capacity of two drives is exclusively dedicated to holding parity data. The available storage capacity, using the same example as above, is 400GB out of 800GB. The percentage of usable system capacity is greater in larger RAID 5 and RAID 6 configurations. In a typical 8-drive SATA RAID array, 25% of the total drive capacity will be used for RAID 6 parity, compared to 12.5% of a RAID 5 array (see figure 2).

Impact of Parity Calculation on Arrays

# Drives	Array capacity used for parity (%)		Storage efficiency (%)	
	RAID 5	RAID 6	RAID 5	RAID 6
3	33.3	N/A	66.7	N/A
4	25.0	50.0	75.0	50.0
8	12.5	25.0	87.5	75.0

Figure 2: Usable system storage capacity is greater in larger RAID 5 and RAID 6 systems. RAID 6 uses more capacity for additional parity storage.



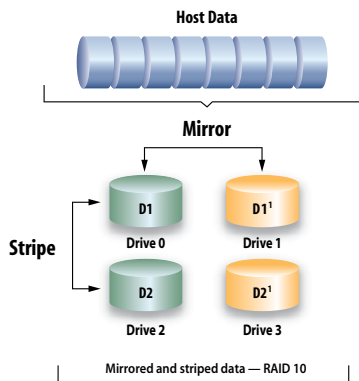


Figure 3: RAID 10 first mirrors, then stripes data across the drives to provide double drive failure protection without the parity calculation performance penalty. The mirror provides the redundancy.

RAID 6 or RAID 10 for 4 drives?

RAID 6 can technically be used with as few as four drives. However, a higher performing configuration that still tolerates two drive failures in a four drive array is RAID 10. RAID 10 mirrors and stripes data to maximize both redundancy and performance (see figure 3). A RAID mirror does not require a read-modify-write operation as RAID 5 and RAID 6 arrays do. A performance benefit is realized by leveraging the mirrored read benefit (read from a track closest to either the mirror A set or the mirror B set) and the ability to recover from a drive failure in either mirrored set.

RAID 6 is recommended over RAID 10 for RAID implementations with more than four drives.

Avoiding a Second Drive Failure

There are several steps to safeguard against a second drive failure that should be considered with any RAID implementation, these include:

1. **Hot sparing with automatic rebuilds.** This does not speed up the rebuild, but does remove the time delay between drive failures and drive replacement. Multiple arrays on a single controller can share a single hot spare for automatic rebuild purposes.
2. **Set the rebuild priority to its highest level.** This will slow the application down during rebuilds but will minimize the exposure time between degraded mode and optimal RAID mode operations.
3. **Balance capacity requirements and performance expectations against the number of arrays in a single volume.**
 - a. Minimize the number of drives per array if capacity is king. The greater the number of high capacity drives in a single array, the higher the probability of a second drive failure.
 - b. Use a greater quantity of smaller capacity drives if performance is the priority. The smaller the drives the shorter the rebuild times will be.
4. **Always look for the highest MTBF rated drives for your RAID 5 and RAID 6 arrays.** The higher the MTBF (Mean Time Between Failure), the lower the probability of a drive failure.

Summary

RAID 6 represents a sea change in the RAID landscape. It provides the industry with higher levels of data protection, data availability and fault tolerance than RAID 5. By assuring data availability following a second drive failure, users can rest assured that they are enjoying maximum data protection, both in normal and degraded modes. And, with RAID 6 enabled 3ware controllers, RAID 6 will not cripple performance.

It is important to remember that RAID 6 does not come without costs. RAID 6 requires the equivalent capacity of two drives in the array to be dedicated to storing only parity information. Also, most RAID 6 systems carry a heavy write performance burden due to the additional parity calculation and the additional memory interrupts. AMCC's simultaneous parity calculations mitigate these performance impediments to provide the fastest RAID 6 solution available.

AMCC

Sales Offices for 3ware Products:

USA: +1-877-883-9273
+1-408-542-8771
Europe: +44-161-232-6360
Asia/Pacific: +65-6826-3381
Japan: +0800-123-9273
China: +86-10-6581-4916

3waresales@amcc.com
www.3ware.com
www.amcc.com

AMCC reserves the right to make changes to its products, or to discontinue any product or service without notice, and advises its customers to obtain the latest version of relevant information to verify, before placing orders, that the information being relied upon is current.

AMCC is a registered trademark of Applied Micro Circuits Corporation. 3ware, SwitchedRAID and 3DM are registered trademarks in the United States and StorSwitch is a trademark in the United States, of Applied Micro Circuits Corporation. All other trademarks are the property of their respective holders. Copyright © 2005 Applied Micro Circuits Corporation.

All Rights Reserved. TBR6_112206