Understanding Server-Based RAID Using Ultra ATA Technology

By Steve Hogge and Bo Zhou

Although SCSI has been the dominant technology for server-based RAID, higher performance standards such as Ultra ATA/100 have led to the emergence of other options. This article introduces the Dell® CERC ATA100/4CH, a new ATA-based PCI controller that offers a low-cost, high-performance alternative to SCSI RAID.

SCSI I/O and drive technology are today's leading choices for fault-tolerant, server-based data storage using RAID (redundant array of independent disks). Advanced Technology Attachment (ATA) RAID is becoming a viable alternative to SCSI RAID because of continued improvements in ATA bus speed and the availability of higher density, low-cost ATA drives.

Dell now offers its first intelligent peripheral component interconnect (PCI) host bus adapter (HBA), the Dell® Cost-Effective RAID Controller, with 100 MB/sec throughput for each of its four ATA channels (CERC ATA100/4CH). The CERC HBA, which includes many of the features offered in the SCSI-based Dell PowerEdge® Expandable RAID Controller (PERC) family, introduces a lower cost, high-performance option that is easy to deploy and maintain.

Why use ATA for RAID?

Cost is a significant reason, although not the only one, for considering CERC ATA100/4CH to implement RAID. First deployed on Dell PowerEdge 500SC and PowerEdge 1400SC servers with up to three and four ATA drives respectively, CERC ATA100/4CH offers similar features (see sidebar "CERC ATA100/4CH Highlights") at a low price compared to SCSI options.

Cost compared to SCSI

The price difference between a SCSI and an ATA hard disk drive (HDD) system can be substantial depending on the amount of data storage required. For example, Figure 1 shows a potential cost savings of approximately 60 percent when comparing the retail cost of ATA-based RAID hardware to that of the SCSI equivalent.

Advantages over non-RAID ATA options

By adding a CERC ATA100/4CH card, users can realize the benefits of RAID for a small incremental cost.

Significantly reduce the chance of data loss. By using any of the available RAID levels (except RAID-0), users can eliminate data

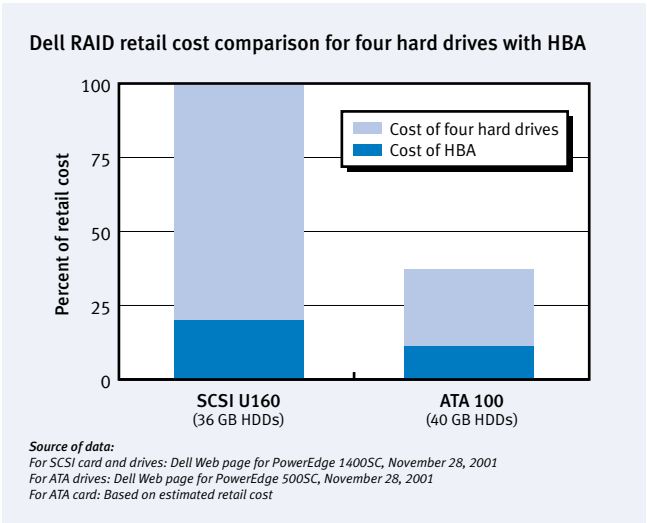

Figure 1. Retail price comparison of Dell RAID alternatives

Deployment consideration	CERC ATA100/4CH	PERC 3/SC (SCSI)
Life expectancy of drives	Good	Better
Performance	Good	Better

Figure 2. Differences between CERC ATA100/4CH and PERC 3/SC

AN OVERVIEW OF THE IDE/ATA INTERFACE

The Integrated Drive Electronics (IDE), or ATA, interface was introduced when the hard disk controller was integrated onto the hard disk. The first IDE drives appeared in PCs that used the Industry Standard Architecture (ISA) bus. The name ATA reflects the fact that the IBM® PC/AT was the first PC to implement the ISA bus. In this article, ATA refers to both IDE and ATA.

The original ATA interface was designed for slow data transfer rates in the 5 MB/sec range. Because hard disk performance increased over the years, manufacturers began using the direct memory access (DMA) transfer protocol on the ATA interface to increase data transfer rates. Manufacturers further improved I/O performance with the introduction of Ultra DMA, a DMA transfer mode that uses techniques such as double transition clocking and cyclical redundancy check (CRC). Double transition clocking doubles the data throughput of the interface for any given clock speed, and CRC improves the integrity of data.

The Ultra DMA mode 5 implementation provides up to 100 MB/sec of throughput. ATA drives that use this mode are often called Ultra ATA/100 drives.

Each ATA channel can support up to two drives: a master and a slave (see Figure A). Most modern systems have two ATA channels that are integrated onto the motherboard.

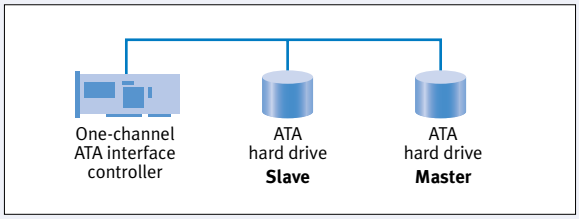

Figure A. ATA channel hard drive connections

loss caused by the failure of one drive (in the case of RAID-1 or RAID-5) or two (in the case of RAID-10 if both failed drives are in different RAID-1 sets).

Increase server uptime. Data storage using RAID (except RAID-0) can withstand the loss of at least one disk drive without forced server downtime. Consequently, administrators can schedule downtime for repair during non-peak hours of operation. The hot spare feature provides greater fault tolerance because administrators can assign one or more drives to automatically replace a failed drive. For example, on the PowerEdge 1400SC, an administrator can assign three drives to a RAID-5 array and the fourth drive as a hot spare.

Decrease mean time to repair. Instead of reloading from tape or other backup media, administrators can repair a degraded disk array under control of a CERC card by bringing down the server, replacing the failed drive, and restarting the server. The card rebuilds the array and restores it to full health in a matter of hours.

Increase potential for higher system performance. The CERC card is an intelligent I/O controller with its own microprocessor to perform RAID functions on data for storage to disk. By off-loading RAID tasks to a dedicated processor, CERC enables the main server CPU to focus on other tasks, which helps maximize overall system performance. For example, RAID-5 parity calculations can be CPU intensive because each write to disk requires recalculating parity. Additionally, RAID arrays perform a higher number of I/Os per second by distributing I/O to multiple drives.

Deployment considerations

Although the lower cost of an ATA system is an advantage, cost-feature trade-offs are also a factor. Figure 2 highlights some differences between CERC ATA100/4CH and PERC 3/SC (PERC version 3, Single Channel) based on their deployment in PowerEdge 500SC and PowerEdge 1400SC servers.

Currently, SCSI hard drives have longer life expectancies than ATA drives. For example, the Seagate® Cheetah® X15 36LP 15K RPM has a mean time between failure (MTBF) of 1.2 million hours, whereas the Seagate Barracuda® ATA IV 7.5K RPM has an MTBF of 0.6 million hours.[1] In addition, higher I/O bus speeds and faster hard drive access, as well as other factors, contribute to better SCSI performance.

An alternative to SCSI RAID

Advances in I/O technology will continue to offer RAID alternatives to customers seeking to optimize their systems for price and performance. Today, CERC ATA100/4CH, which is deployed

[1] Data from Seagate *Cheetah X15 36LP Product Manual, Rev. A* and Seagate *Barracuda ATA IV Family Product Manual, Rev. A*, respectively.

CERC ATA100/4CH HIGHLIGHTS

CERC ATA100/4CH is based on proven PERC 3 core firmware. It uses the same set of SCSI drivers as PERC 3 because the drivers talk to the same SCSI firmware interface. The following highlights describe some key features.

Supported by the Dell OpenManage™ Array Manager. CERC is configured and managed through the same software used with the PERC products. This feature allows current Dell PERC users to apply their existing expertise. Additionally, administrators can perform pre-operating-system configuration through the same BIOS utility available for PERC 3.

Common device driver and card BIOS utilities. By leveraging the core firmware and device driver used on PERC 3, CERC ATA100/4CH enhances the ease of use for server updates. Administrators do not need to track different names and versions of drivers; one driver update works for all PERC 3 HBAs and the CERC HBA. These HBAs are also managed by the same BIOS configuration utility.

Subset of PERC 3 features. CERC ATA100/4CH supports most of the RAID features offered by PERC 3.

- **Support for RAID levels 0, 1, 10, and 5.** CERC supports a full set of RAID levels, including RAID-5.
- **Online capacity expansion.** Administrators can add disk capacity to the existing virtual disk without taking the server out of service.
- **Support for up to 40 virtual disks.** Administrators can mix virtual disks with different RAID levels for different redundancy requirements.
- **Online RAID-level migration.** Administrators can change the RAID type of a virtual disk, such as RAID-0 to RAID-1, or RAID-5 to RAID-0 (certain restrictions apply) without taking the server out of service.
- **Fast RAID array initialization.** CERC reduces array initialization from hours to seconds.
- **Minimal utilization of host-server CPU.** An onboard Intel® i960® microprocessor performs all I/O processes and RAID calculations in local memory. The card also provides 16 MB of onboard cache for fast I/O access.

High performance. CERC ATA100/4CH has a peak sequential transfer rate of more than 100 MB/sec, as measured for sequential reads using Intel IOmeter on a PowerEdge 1400SC with four Ultra ATA/100 hard drives in a RAID-0 array.

Low price per GB of RAID with RAID-5. CERC ATA100/4CH support for RAID-5 lets users obtain a lower cost per gigabyte of protected data than options that offer only mirroring, which doubles the amount of storage required for data protection.

Universal half-length PCI adapter. On a half-length PCI adapter, CERC ATA100/4CH provides four independent Ultra ATA/100 channels, one per hard drive, which are backward compatible with Ultra ATA/33 and ATA/66. Each channel supports one master drive (see Figure B).

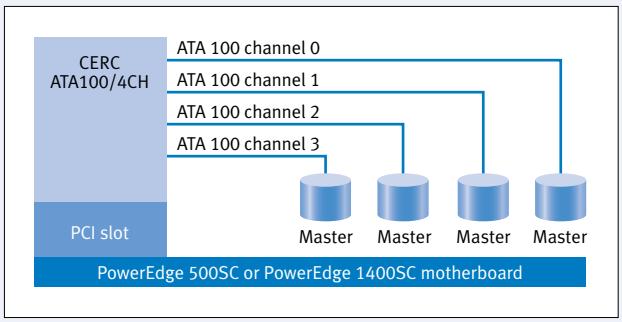

Figure B. CERC ATA100/4CH hard drive configuration

in the Dell PowerEdge 500SC and PowerEdge 1400SC, allows customers to realize the benefits of fault-tolerant data storage at a price for performance once limited to the world of SCSI.

Steve Hogge *(steve_hogge@dell.com) is a server-based RAID product manager at Dell. Prior to Dell, he worked at GE Industrial Systems, where he spent five years in hardware and firmware design engineering and five years in global product management. Steve received a B.S. in Computer Engineering from Old Dominion University.*

Bo Zhou *(bo_zhou@dell.com) is a systems engineer working on ATA RAID controller development in the Dell RAID II Development Group. Bo has several years of experience in computer architecture modeling, simulation, and PC storage subsystem design. He earned a B.S. in Computer Engineering from the Huazhong University of Science and Technology and an M.S. in Computer Science from Louisiana Tech University.*

FOR MORE INFORMATION

For more information, please visit http://www.dell.com