Implementing Cost-Effective RAID

with CERC SATA 2s Disk Technology

Dell recently introduced a low-cost RAID implementation known as the Cost Effective RAID Controller (CERC) SATA 2s, a software-based RAID controller designed for use with Serial ATA (SATA) disk technology. This article examines the architecture, configuration, and management of CERC SATA 2s, including an overview of how CERC SATA 2s interacts with the server’s BIOS and how it compares with other RAID implementations.

BY HARISH JAYAKUMAR AND ED MATTHEWS

A mainstay for organizations deploying servers, RAID techniques are primarily used to provide protection from disk failures. This protection is achieved by storing data redundantly on multiple disks; if one disk fails, the other disks can still service incoming I/O requests. The various RAID levels represent the different algorithms used to provide data redundancy. Each level has its own cost and performance advantages.

Figure 1 shows a theoretical calculation1 for the maximum throughput per dollar relative to RAID-0 for different types of RAID.2 For the purposes of this example, the cost of the RAID hardware is directly proportional to the number of disks the RAID system uses in the disk array. This comparison shows that, given equivalently priced RAID-0 and RAID-1 systems with comparable performance, a RAID-1 system is expected to sustain half the number of small writes per second that a RAID-0 system is expected to sustain. Likewise, a RAID-5 system is expected to sustain a maximum of only one-fourth to one-Nth (where N is the number of disks) the number of small writes per second that a RAID-0 system is expected to sustain.

To limit the CPU burden placed on a system, the Dell® Cost Effective RAID Controller (CERC) SATA 2s implements only RAID-0 and RAID-1.

Introduction to SATA and CERC SATA 2s

Serial ATA (SATA) disk drives are based on a low-cost technology that is replacing Parallel ATA (PATA) disk drives in low-end servers. SATA implements a serial data transfer mechanism, and the serial interface is designed to provide higher throughput rates than PATA technology.

CERC SATA 2s supports two SATA disk drives and is a driver-based software RAID implementation.3 To understand

---

1 The source of this calculation is "The RAID Tutorial: Cost & Performance Issues" by Israel Koren, Department of Electrical and Computer Engineering, University of Massachusetts, Amherst; for more information, visit www.ecs.umass.edu/ece/koren/architecture/Raid/cp.html.

2 RAID-0 is actually a misnomer because it does not provide redundancy or protection from disk failures. In a RAID-0 implementation, the data is stored across multiple disks. The RAID-0 approach enables increased performance because multiple drives can concurrently service an I/O request.

3 CERC SATA 2s is also referred to as CERC SATA 1.5/2s—the number 1.5 stands for the interface speed.
the behavior and advantages of CERC SATA 2s, administrators can compare different types of RAID implementations (see Figure 2).

As Figure 2 shows, CERC SATA 2s can be a cost-effective alternative when the more advanced capabilities of a hardware implementation are not needed. As in a hardware RAID implementation, CERC SATA 2s configures RAID through option ROM.

### Role of option ROM in RAID functionality

The role of option ROM is to extend the functionality of the BIOS so that the OS can use vendor-specific hardware or special devices. For CERC SATA 2s, option ROM provides Int 13h support and configures metadata and RAID.

Enabling Int 13h support

Int 13h is a software interrupt that the BIOS provides to the OS to perform disk I/Os. In the Dell CERC SATA 2s implementation, the Int 13h interrupt is not a direct part of the BIOS. Instead, option ROM is loaded at system startup and is invoked to handle disk I/O requests made through Int 13h.

Configuring metadata and RAID

The metadata stores disk configuration information including the type of RAID in use. This data is stored in reserved areas on a disk, which are not available to user data (hardware-based RAID implementations also store metadata in nonvolatile memory).

Option ROM is responsible for configuring this metadata on the disks, and system administrators can choose between a RAID-0 and a RAID-1 implementation during system configuration: turn RAID on in the BIOS settings, and then choose the type of RAID using the option ROM interface, which is invoked through the key sequence Ctrl + M. Option ROM then stores the metadata in the reserved areas on the disks, and the drivers read it at system startup.

Historically, hardware vendors have formatted disk metadata differently. The Storage Networking Industry Association (SNIA) Common RAID Disk Data Format (DDF) Technical Working Group was chartered to define a standard data structure describing how data is formatted across the disks in a RAID group. The Common RAID DDF specification is designed to provide standardization and interoperability among different suppliers of RAID technology.

The long-term goal is that drives from different systems, vendors, and controllers can be interchanged while preserving the data on those disks.

Key features of the DDF specification include:

- A standard data structure that defines how data should be formatted across the disks in a RAID group
- Interoperability between different suppliers of RAID technology
- A common format for storing RAID configuration information so that data on physical disks can be accessed independent of the RAID controller used
- No requirement for controllers to store this configuration information in the same format in their internal memory

### Role of drivers in CERC SATA 2s

Option ROM formats the RAID configuration and writes the metadata onto the SATA drives. After this is accomplished, the drives are ready for OS installation. First, administrators load the drivers for the OS, which provide the ability to read and interpret the metadata previously written by option ROM. The drivers then present the multiple disks as one logical drive to the OS.

![Figure 2. Comparative advantages and disadvantages of RAID implementations](image)

---

4 For more information about DDF, see “Solving the RAID Compatibility Puzzle” by Matthew Brisse and Bill Dawkins, Ph.D., in *User Power Solutions* May 2005.
Management of CERC SATA 2s

Management capabilities for CERC SATA 2s depend on the type of Dell server used. CERC SATA 2s is currently available on Dell PowerEdge™ SC420, PowerEdge SC1420, PowerEdge 800, and PowerEdge 1800 servers. On Dell PowerEdge SC420 and PowerEdge SC1420 servers, the RAID Storage Manager (RSM) utility is used. On Dell PowerEdge 800 and PowerEdge 1800 servers, Dell OpenManage™ Storage Services (OMSS) is used.

RSM on Dell PowerEdge SC platforms

For Dell PowerEdge SC platforms, which are entry-level servers, the RSM utility provides management functionality for the RAID configuration on the Microsoft Windows Server™ 2003 OS (see Figure 3). RSM features include:

• Task management with a scheduler
• Event management
• Multiple graphical user interface (GUI) views
• Documentation with a search facility

Dell OpenManage on Dell PowerEdge platforms

The Dell OpenManage infrastructure provides a comprehensive set of tools, including OMSA and Dell OpenManage Server Administrator (OMSA), to enable efficient systems management. OMSA is accessed through OMSA and can be used to manage the CERC SATA 2s RAID functionality on Dell PowerEdge servers.

OMSS consists of various components for the management of storage, firmware, and hardware. Using the storage component, administrators can implement controller functions without accessing the BIOS. Controller functions include configuring virtual disks, applying RAID levels, and creating hot spares for data protection. Many controller functions such as rebuilds and troubleshooting can be initiated using OMSS. Most of these functions can be implemented while the server remains online and continues to process requests.

The status of storage systems is reported through graphical displays and icons. When a change in status occurs, event traps are sent to the administrator and also maintained in an event log, which can be viewed from the console. In addition, most events generate Simple Network Management Protocol (SNMP) traps, which can be sent to a remote destination.

Figure 4 shows the OMSA/OMSS GUI for managing CERC SATA 2s. In this scenario, a RAID-1 array was created over two SATA disks, and the OS sees this array as a single virtual disk. The OMSS GUI has an Available Tasks drop-down menu, which provides the administrator with various options (delete, check consistency, rename, and so on) for managing the virtual disk.

Cost-effective RAID for low-cost servers

CERC SATA 2s is a low-cost RAID implementation that can provide effective RAID functionality on two SATA drives. It can be used for RAID-0 or RAID-1 and includes management tools for supported Dell PowerEdge and PowerEdge SC platforms. In this way, CERC SATA 2s can provide low-cost servers with basic RAID capabilities while helping cost-conscious organizations run an efficient data center.

Harish Jayakumar is a test engineer in the Dell OpenManage software development and testing organization. He has a master’s degree in Computer Science from Arizona State University and a bachelor’s degree in Computer Science from the University of Madras, India. His interests include software testing, networking, and operating systems.

Ed Matthews is a senior manager working in the Dell OpenManage software development and testing organization. He has been in the software/IT industry for nearly 30 years, and has prior experience working for semiconductor, telecommunications, storage, and computer manufacturers. He has a B.S. in Physics with Electronics from Liverpool University, England. He is a Chartered Information Technology Professional (CITP) and a Member of the British Computer Society.