INTRODUCING THE DELL PERC 6 FAMILY OF SAS RAID CONTROLLERS

By Bhanu Prakash Dixit
Sanjay Tiwari
Kedar Vaze
Joe H. Trickey III

The Dell™ PowerEdge™ Expandable RAID Controller (PERC) 6 family of enterprise-class Serial Attached SCSI (SAS) RAID controllers is designed for enhanced performance, increased reliability and fault tolerance, and simplified management—providing a powerful, easy-to-manage way to create a robust infrastructure and help maximize server uptime.

Related Categories:

Dell PowerEdge RAID Controller (PERC)
PERC 6
RAID
Serial Attached SCSI (SAS)
Storage

Visit DELL.COM/PowerSolutions for the complete category index.

As enterprise data requirements continue to increase, deploying powerful, easy-to-manage RAID solutions can become essential for organizations seeking to simplify their IT environments while increasing controller and storage reliability and server uptime. The Dell PowerEdge Expandable RAID Controller (PERC) line of storage controllers has evolved to address these data storage requirements, providing innovative features to help support complex server-based storage environments, increase server uptime, and reduce administrative burdens.

The PERC 6 family of Serial Attached SCSI (SAS) RAID controllers supports SAS devices and Dell-qualified Serial ATA (SATA) devices. It consists of three PCI Express (PCIe)–based, 3 Gbps SAS RAID controllers: the PERC 6/E adapter (which fits in a standard x8 PCIe slot supporting Dell PowerVault™ direct attach storage), the PERC 6/i adapter (which fits in a standard x8 PCIe slot supporting internal storage for Dell PowerEdge servers and Dell Precision™ workstations), and the PERC 6/i integrated card (which supports internal storage for Dell PowerEdge servers).

Each PERC 6 controller offers a 256 MB double data rate 2 (DDR2) cache memory, with the PERC 6/E supporting upgrades to 512 MB to allow quick access to additional data and help increase performance. PERC 6 controllers also include a battery backup unit to help maintain data consistency and protect data in the controller cache during a system power and/or controller failure. PERC 6 controllers support RAID levels 0, 1, 5, 6, 10, 50, and 60.

The evolution of the PERC 6 from the PERC 5 reflects the Dell focus on simplifying IT, advancing green technology, and designing total data solutions. PERC 6 controllers take advantage of the PCIe architecture and include an enhanced controller chip design to help substantially increase throughput and remove the controller as a performance bottleneck (see Figure 1). The optimization of the RAID design, which incorporates the I/O processor and I/O controller on a single RAID-on-a-chip (ROC) solution, enables a significant reduction in power consumption for PERC 6 controllers compared with equivalent PERC 5 controllers. In addition, using a true PCIe solution by removing the PCI Extended (PCI-X) bridge helps significantly increase performance under sequential workloads. Figure 2 provides a comprehensive comparison of PERC 5 and PERC 6 features and specifications.

PERC 6 controllers are designed to provide enhanced performance, simplified management, increased reliability and fault tolerance, optimized

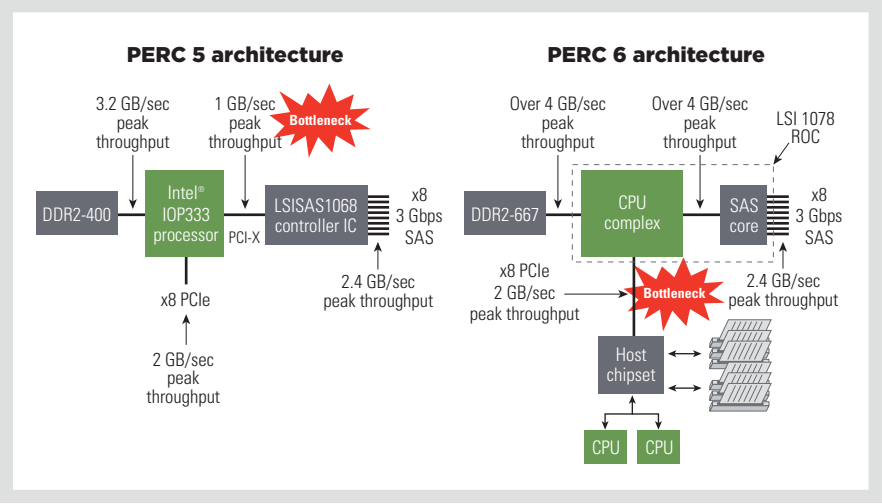

Figure 1. *Dell PERC 5 and PERC 6 controller architectures*

deployment with Dell PowerVault storage, and flexible RAID configuration and management tools. The controllers support Microsoft® Windows Server® 2003 with Service Pack 1 (SP1) (32-bit and 64-bit), Microsoft Windows Server 2008 (32-bit and 64-bit), Red Hat® Enterprise Linux® 4 Update 5 and Red Hat Enterprise Linux 5 (32-bit and 64-bit), Novell® SUSE® Linux Enterprise Server 10 (64-bit), Microsoft Windows® XP with SP2 (32-bit and 64-bit), and Microsoft Windows Vista® (32-bit and 64-bit) operating systems.

ENHANCED PERFORMANCE

The PERC 6 is designed to increase performance in each of the four key areas that can affect controller performance: hardware features, caching policies, read and I/O policies, and software features. Each is important and should be considered when evaluating a RAID controller in a host server along with other server features that may affect performance, such as the OS and hard drives.

Organizations should also keep in mind that different applications display different workload characteristics, and that these workload footprints can affect server performance. However, many applications can be categorized as having a *sequential* read/write profile or a *random* read/write profile. Sequential read/write applications typically include media streaming, backup and restore, and network attached storage and near-line storage applications. Workstation and file servers also handle large data file read/write requests, making sequential data performance a critical performance requirement in enterprise environments. For these applications, storage subsystem performance is typically measured by sequential throughput, in megabytes per second.

Random read/write applications typically include transaction processing applications such as databases, Web-based e-commerce applications, and other multitasking business solutions. For these applications, storage subsystem performance is typically measured in I/Os per second.

Hardware features

The PERC 6 continues the Dell focus on designing RAID solutions based on the latest interface technologies, from PCI to PCIe, from parallel to serial. Employing a ROC integrated circuit (IC) featuring a high-speed, 64-bit, error-correcting code (ECC)–protected DDR2 SDRAM interface; a hardware RAID assist engine for parity

	PERC 5	PERC 6
RAID levels	0, 1, 5, 10, and 50	0, 1, 5, 6, 10, 50, and 60
Enclosures per port	Up to three	Up to three
Ports	2 x 4 wide ports	2 x 4 wide ports
Write policies	Write-through and write-back	Write-through and write-back
Read policies	Normal, read-ahead, and adaptive	Normal, read-ahead, and adaptive
Virtual disks per controller	Up to 64	Up to 64
Cache memory size	256 MB	256 MB (up to 512 MB for PERC 6/E)
PCIe link width	x8	x8
256 KB, 512 KB, and 1,024 KB stripe sizes		✔
SATA NCQ support		✔
Global, dedicated, and affinity hot spares	Global and dedicated hot spares only	✔
Online capacity expansion	✔	✔
Hot-swappable drives	✔	✔
Mixed-capacity physical disk support	✔	✔
Hardware exclusive OR (XOR) assistance	✔	✔

Figure 2. *Feature comparison of Dell PERC 5 and PERC 6 controllers*

calculations; and support for an Intelligent I/O (I2O) message unit helps maximize performance. The controller is powered by a state-of-the-art LSI 1078 ROC controller, which includes both instruction and data cache to help provide an additional performance boost. This solid hardware foundation is designed to simplify integration of new RAID firmware features and performance increases in future release updates.

Caching policies

Flexible caching policies in PERC 6 controllers allow administrators to tune cache writing schemes to help optimize performance or maximize data protection, including choosing between *write-through* caching and *write-back* caching. With write-through caching enabled, the controller transfers a block of data directly to the disk, and does not send an acknowledgment of data transfer completion to the host system until the data is committed to the disk subsystem. With write-back caching enabled, the controller sends a data transfer completion signal to the host when the controller cache has received all data in a transaction, and the controller then writes the cached data to the storage device when system activity is low or when the write buffer approaches capacity; the cached data is not written to the storage device immediately.

Write-through caching typically provides a data security advantage over write-back caching, while write-back caching typically provides a performance advantage. The risk of using write-back caching is that the cached data can be lost if a power failure occurs before the data is written to the storage device. This risk can be mitigated by using the battery backup units on PERC 6 controllers.

Read and I/O policies

Read policies determine the type of read option for a logical drive and can be set to normal (the default setting), read-ahead, or adaptive. The normal setting specifies that the controller reads only the requested data and does not read ahead for the current logical drive. The read-ahead setting specifies that the controller reads sequentially ahead of requested data and stores the additional data in cache memory, anticipating that the data will be needed soon; this policy typically supplies sequential data faster than the normal (no-read-ahead) setting, but is not as effective when accessing random data. The adaptive setting specifies that the controller begins using read-ahead if the two most recent disk accesses occurred in sequential sectors. If all read requests are random, the algorithm reverts to the normal setting; however, requests are still evaluated for possible sequential operation.

I/O policies are used to enable read buffering in cache memory and apply to reads on a specific logical drive. This setting does not affect the read-ahead cache. The cached I/O setting specifies that the controller buffers reads in cache memory, while the direct I/O setting specifies that reads and writes are not buffered in cache memory. This setting does not override the read policy settings; data is transferred to cache and the host concurrently. If the same data block is read again, it comes from cache memory.

"As enterprise data requirements continue to increase, deploying powerful, easy-to-manage RAID solutions can become essential for organizations seeking to simplify their IT environments while increasing controller and storage reliability and server uptime."

Software features

Enabling disk caching can have a significant impact on system performance, but the potential for data loss increases if power to the disk drives is lost while data written to the disk is still in the cache. Native command queuing (NCQ) for SATA and tagged command queuing (TCQ) for SAS hard drives helps increase performance by allowing individual hard disks to optimize the order in which read and write commands are executed.

SIMPLIFIED MANAGEMENT

PERC 6 controllers integrate technologies and applications designed to simplify management for IT administrators. For example, SAS disk connectivity capabilities have made it increasingly common to connect a large number of disks to a PERC controller, and PERC 6 controllers include features that can automatically track disk order and organization through disk roaming. PERC 6 controllers also support moving physical disks from one cable connection or backplane slot to another on the same controller; the controllers can automatically recognize the relocated physical disks and logically place them in the appropriate virtual disks that are part of the disk group.

Other advanced online capabilities allow administrators to migrate disks from one controller to another. The Advanced Import feature enables administrators to import, or *migrate*, disk arrays created by a different PERC controller while the server remains online. The controllers can even import partial, or *degraded*, arrays and provide array preview capabilities before import to further enhance data serviceability.

Dell PERC 6 controllers utilize Smart Cache Technology that offers administrators pre-tuned cache ratios for read and write operations that automate cache

"PERC 6 controllers are designed to provide enhanced performance, simplified management, increased reliability and fault tolerance, optimized deployment with Dell PowerVault storage, and flexible RAID configuration and management tools."

allocations based on the characteristics of the storage environment. This auto-tuning feature helps administrators avoid the time-consuming effort of manually tuning the controller cache in a test environment.

Administrators can also perform real-time RAID-level migration by converting one RAID level to a different RAID level. Figure 3 lists possible RAID-level migrations and the number of drives required at the beginning and end of the process. RAID-level migration and expansion cannot be performed on RAID-10, RAID-50, or RAID-60 configurations.

INCREASED RELIABILITY AND FAULT TOLERANCE

Beyond industry-standard RAID levels, PERC 6 controllers offer myriad features to help protect data, including advanced media error protection, enhanced rebuild capabilities, consistency check and background initialization, and other key features such as Self-Monitoring, Analysis, and Reporting Technology (SMART), failed physical disk detection, and hot spares.

Advanced media error protection

PERC 6 controllers provide advanced media error monitoring and repair technologies designed to prevent data loss and enable safe data retrieval. For example, a typical response to a media error in a RAID controller would be to read the data from the redundant disk and continue on. However, when a PERC 6 encounters media errors during normal read and write operations, it can immediately attempt to repair the data. This on-the-fly media error repair capability helps increase data reliability by fixing data written to bad media sectors, thus helping reduce the potential for data loss.

Preemptive media error monitoring through disk scanning services is typically essential for enterprise-class RAID controllers. While a typical approach to this monitoring would employ a background consistency check to scan for media errors, the Patrol Read feature of PERC 6 controllers implements verify commands to help validate media—a method that can help significantly increase efficiency, particularly when encountering multiple media errors on drives in an array.

Enhanced rebuild capabilities

When an array is in degraded mode following a disk failure, the risk for potential data loss increases—particularly if the controller encounters media errors on optimal disks in the array during a rebuild. In contrast to controllers that might immediately stop the rebuild, the PERC 6 continues the rebuild to the end, allowing administrators to access all valid data on the virtual disk even though data has been lost on the virtual disk stripe where the media error occurred.

Consistency check and background initialization

The consistency check operation helps verify the correctness of data in logical drives that use RAID levels 1, 5, 6, 10, 50, and 60 (RAID level 0 does not provide data redundancy). For example, in a system with parity, checking consistency means computing the data on one drive and comparing the results to the contents of the parity drive. Dell recommends that administrators perform a consistency check at least once a month to help maintain optimum array status.

Background initialization is a consistency check that is forced when creating a logical drive, and automatically begins five minutes after the drive has been created. This operation checks for media errors on physical disks and helps ensure that striped data segments are the same on all physical drives in an array.

Source RAID level	Target RAID level	Required number of source physical disks	Required number of target physical disks
RAID-0	RAID-1	1	2
RAID-0	RAID-5	1	3
RAID-0	RAID-6	1	4
RAID-1	RAID-0	2	1
RAID-1	RAID-5	2	3
RAID-1	RAID-6	2	4
RAID-5	RAID-0	3	3
RAID-5	RAID-6	3	4
RAID-6	RAID-0	4	4
RAID-6	RAID-5	4	3

Figure 3. *Possible RAID migrations supported with Dell PERC 6 controllers*

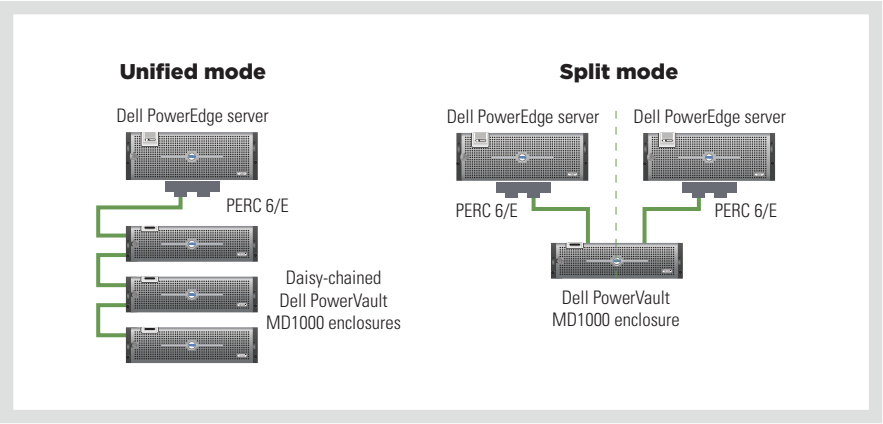

Figure 4. *Unified and split modes for Dell PowerVault MD1000 enclosures with Dell PERC 6/E controllers*

The background initialization rate is controlled by the BGI rate set using Dell OpenManage™ Server Administrator (OMSA) Storage Management.

Other key data protection features

Other key data protection features of PERC 6 controllers include the following:

- **SMART:** The SMART feature monitors the internal performance of motors, heads, and physical disk electronics to help detect predictable physical disk failures.
- **Failed physical disk detection:** The controllers can automatically detect and rebuild failed physical disks when a new drive is placed in the slot where the failed drive resided or when an applicable hot spare is present. Automatic rebuilds can be performed transparently with hot spares.
- **Hot spares:** The controllers support global hot spare, dedicated hot spare, and affinity configurations, which administrators can set up using the Dell BIOS Configuration Utility as well as OMSA Storage Management. Global hot spares can typically be used in any degraded RAID array when the hot spare has sufficient capacity to fit into the RAID container. Dedicated hot spares are reserved for a particular disk group. The affinity option prioritizes a hot spare to the enclosure that it is located in.
- **Transportable battery backup unit (TBBU):** The TBBU is a cache memory module with an integrated battery pack that enables administrators to transport the cache module with the battery into a different controller. The TBBU helps protect the integrity of cached data by providing backup power during a power outage. After power is restored to the controller, the cache is flushed and the protected data is written to the disk drives.
- **Physical disk hot swapping:** Administrators can manually substitute a replacement unit in a disk subsystem for a defective one while the subsystem is performing its normal functions.
- **Heartbeat status LED and dirty cache LED:** The heartbeat status LED on the controllers indicates activity on the chip, while the dirty cache LED indicates that the cache has data that has not yet been written to disk.

OPTIMIZED DEPLOYMENTS WITH DELL POWERVAULT STORAGE

Dell PowerVault storage is engineered to work optimally with Dell PowerEdge servers and PERC controllers. Combining the PERC 6/E with the Dell PowerVault MD1000 disk expansion enclosure can provide a high-performance, enterprise-class direct attach storage solution. Administrators can use OMSA Storage Management software to manage both the external array and internal storage within the server itself, helping streamline and simplify storage management through a single interface while helping reduce resource load on the system.

Administrators can customize the PowerVault MD1000 along with the PERC 6/E to meet different enterprise needs and help increase storage provisioning flexibility by using one of two backplane modes (see Figure 4):

- **Unified mode:** Unified mode allows administrators to access all 15 hard drives from a single controller, and supports daisy-chaining up to two additional enclosures per channel—supporting a total of up to 90 hard drives when using six enclosures and two channels in each controller.
- **Split mode:** Split mode allows two controllers to share the same enclosure, with one having access to eight hard drives and one having access to the other seven hard drives. This configuration does not support daisy-chaining additional enclosures.

"Integrating PERC 6 controllers into environments using Dell PowerEdge servers, Dell PowerVault storage, and Dell Precision workstations can help administrators create a robust and highly available environment designed to maximize server uptime."

By taking advantage of the support for daisy-chaining in unified mode, administrators can configure a specific combination of storage resources initially, and then add storage volumes quickly and easily as data requirements increase.

FLEXIBLE RAID CONFIGURATION AND MANAGEMENT TOOLS

In keeping with the Dell focus on simplifying IT, PERC 6 controllers share a common firmware code base, host drivers, and management software. The storage management software enables administrators to configure and manage multiple PERC 6 RAID systems, create and manage multiple disk groups, and perform online maintenance. By supporting multiple configuration and management options designed for flexibility and simplicity, PERC 6 controllers can help meet wide-ranging application and deployment requirements.

Administrators can choose from two primary management applications depending on their needs: OMSA Storage Management or the Dell BIOS Configuration Utility. OMSA Storage Management (see Figure 5), a storage management application for Dell PowerEdge servers, provides enhanced features for configuring locally attached RAID and non-RAID disk storage. It enables administrators to perform controller and enclosure functions—such as creating data redundancy, assigning hot spares, and rebuilding failed physical disks—for supported RAID and non-RAID controllers and enclosures from a single graphical user interface (GUI) or command-line interface (CLI) without requiring use of the controller BIOS utilities. The wizard-based GUI includes features for both novice and advanced administrators as well as detailed online help files, while the CLI is designed for comprehensive functionality and supports scripting.

The Dell BIOS Configuration Utility (see Figure 6) is embedded in PERC 6 controllers and enables administrators to configure and maintain RAID disk groups and virtual disks and manage RAID systems. It operates independently of the system OS. Administrators would typically use this utility for initial setup, and then configure advanced features through OMSA Storage Management.

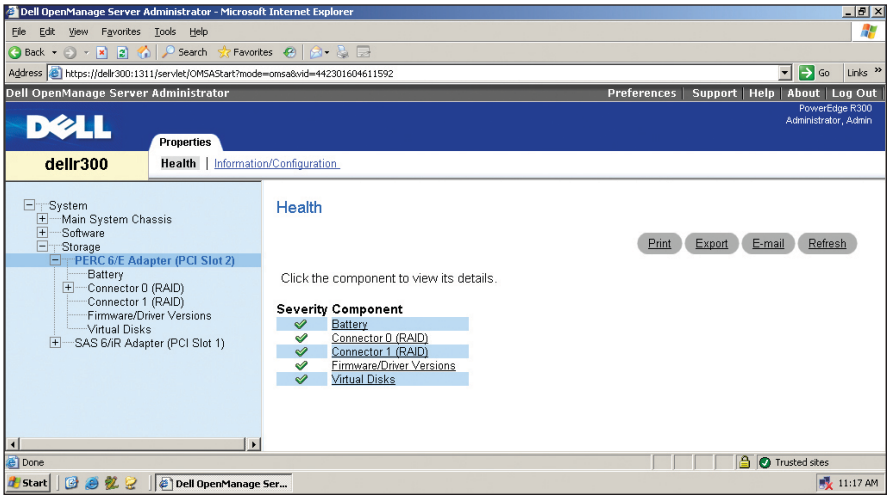

Figure 5. *Dell OpenManage Server Administrator Storage Management*

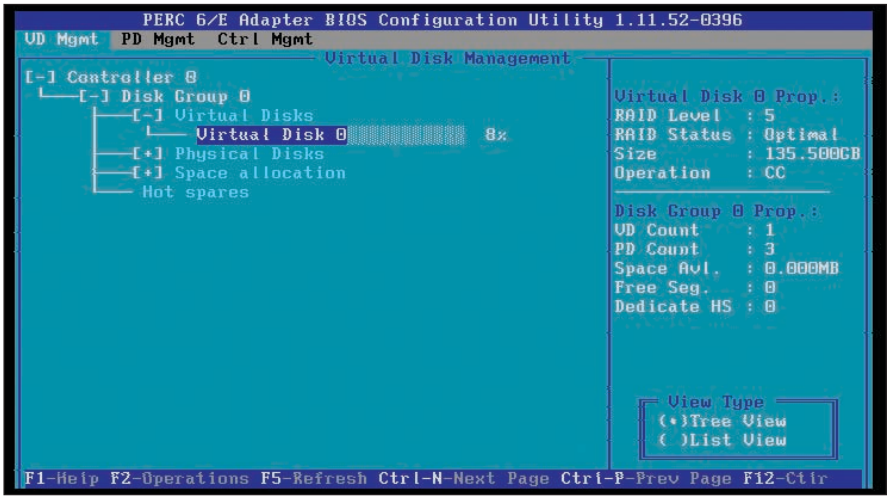

Figure 6. *Dell BIOS Configuration Utility*

SIMPLIFIED RAID CONFIGURATION AND MANAGEMENT

The Dell PERC 6 family of SAS RAID controllers is designed to provide enhanced performance, increased reliability and fault tolerance, and simplified management for supported Dell platforms. Integrating PERC 6 controllers into environments using Dell PowerEdge servers, Dell PowerVault storage, and Dell Precision workstations can help administrators create a robust and highly available environment designed to maximize server uptime.

Bhanu Prakash Dixit is a senior analyst in the Dell Enterprise Storage Group.

Sanjay Tiwari is an engineer adviser in the Dell Enterprise Storage Group.

Kedar Vaze is an engineering manager in the Dell Enterprise Storage Group.

Joe H. Trickey III is a storage product marketing consultant in the Dell Product Group.