



By André Dumouchelle
Joe H. Trickey III
John Seward

RETHINKING THE DELL PERC 6 RAID CONTROLLER

The Dell™ PowerEdge™ Expandable RAID Controller (PERC) 6.1 firmware release introduces multiple features designed to enhance reliability, performance, and ease of use to help organizations meet the challenges of ongoing data growth.

As data continues to grow at an exponential rate, many organizations find they must repeatedly update their storage systems to keep up with increasing data demands. The Dell PowerEdge Expandable RAID Controller (PERC) 6 is designed to help these organizations meet the challenges of data growth. The PERC 6 features an LSI 1078 RAID-on-a-chip (ROC) solution, 256 MB of double data rate 2 (DDR2) cache memory (upgradable to 512 MB on the PERC 6/E), and flexible support for RAID levels 0, 1, 5, 6, 10, 50, and 60. Figure 1 summarizes the key PERC 6 features.¹

Related Categories:

Dell PowerEdge RAID Controller (PERC)

Dell PowerEdge servers

PERC 6

RAID

Storage

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

The PERC 6.1 firmware update introduces multiple features designed to further enhance reliability, performance, and ease of use—including enhanced support for RAID-10 and an Intelligent Mirroring feature, redundant paths and load balancing, revertible hot spares, simplified storage administration, and increased battery shelf life. Organizations currently using PERC 6 controllers with previous firmware versions can download the 6.1 firmware release from the Dell support Web site at support.dell.com.

ENHANCED RAID-10 SUPPORT AND INTELLIGENT MIRRORING

IT organizations often need to balance performance and redundancy. Although traditionally they have used RAID-5 to help provide an appropriate balance,

random I/O applications such as the Microsoft® SQL Server® and Exchange platforms can take advantage of the speed and protection offered by RAID-10. Because RAID-10 combines striping and mirroring to help produce large virtual disks with both high performance and fault tolerance—remaining functional for up to one drive failure per mirrored pair—it can help increase performance by striping across mirror sets without requiring parity calculations. RAID-10 is well suited for environments requiring high levels of performance and redundancy as well as rapid data recovery following a drive failure.

As with previous PERCs, the PERC 6 includes support for RAID-10. However, previous PERCs limited the number of hard drives in a single RAID-10 volume to 16, supporting only 8 mirrored spans (see Figure 2). The PERC 6.1 firmware now enables administrators to configure up to 144 drives into a single RAID-10 volume to help them take advantage of the increased capacity of the 2U, 24-drive Dell PowerVault™ MD1120 storage enclosure (see Figure 3). The RAID-10 features also enable administrators to mirror enclosures such that data on the hard drives of one enclosure is mirrored to drives on another enclosure in the array.

In addition, the PERC 6.1 firmware now supports use of the Dell OpenManage™ 5.5 suite to help simplify redundancy and capacity maximization on external enclosures. The Intelligent Mirroring feature introduced

¹ For more information, see “Introducing the Dell PERC 6 Family of SAS RAID Controllers,” by Bhanu Prakash Dixit, Sanjay Tiwari, Kedar Vaze, and Joe H. Trickey III, in *Dell Power Solutions*, May 2008, DELL.COM/Downloads/Global/Power/ps2q08-20080255-Dixit.pdf.

Advanced media error monitoring	When the controller encounters media errors during normal read and write operations, it can attempt to repair the data immediately. This on-the-fly repair capability helps increase data reliability by fixing data written to bad media sectors, thus helping reduce the potential for data loss.
Advanced RAID sets	The PERC 6 supports advanced RAID-6, RAID-50, and RAID-60 configurations.
Auto-import protection	Auto-import protection enables administrators to import disk arrays created by a different PERC while the server remains online. It can import degraded arrays and offers array preview capabilities before import, helping further enhance data serviceability. This innovative feature also prompts administrators to accept foreign configurations rather than automatically importing the drives to help ensure optimized data availability.
Automatic drive rebuild on replacement	A drive rebuild can begin automatically when administrators replace a failed drive with a new drive that has the necessary capacity and uses the same drive technology as the failed drive (SAS or SATA).
Consistency check	This check helps verify that all stripes in a virtual disk with a redundant RAID level are consistent. Consistent RAID-1 virtual disks have matched data across each stripe; consistent RAID-5 virtual disks have matched data and parity across each stripe.
Background initialization	Background initialization is a consistency check that runs within five minutes of the creation of a new logical drive. This operation checks for media errors on physical disks and helps ensure that striped data segments are accurate and complete. Accurate and complete parity enables the RAID firmware to use RAID algorithms that enhance I/O performance.
Hot spares	The PERC 6 supports global hot spares, dedicated hot spares, and enclosure affinity configurations.
Online capacity expansion	Administrators can expand the virtual disk size without an OS reboot for supported RAID levels other than RAID-10, RAID-50, and RAID-60.
Patrol Read	The Patrol Read feature checks each sector of configured hard drives for media errors and can recover data if an error is detected.
RAID-level migration	Administrators can migrate systems from one RAID level to another as long as the necessary capacity and number of disks are available. (RAID-level migration cannot be performed on RAID-10, RAID-50, and RAID-60 configurations, however.)
Rebuild check pointing	Rebuild operations can resume at the last point following a restart.
Selectable rebuild rate	Administrators can manually adjust the percentage of time the controller spends rebuilding rather than handling I/O. Reduced rates help minimize performance impact, while high rates help minimize exposure to a second drive failure.
SMART	SMART monitors the internal performance of motors, heads, and physical disk electronics to help detect predictable physical disk failures.
Smart Cache Technology	Smart Cache Technology provides automatic cache allocation based on the characteristics of the storage configuration and I/O workload.
Transportable battery backup unit (TBBU)	The TBBU is a cache memory module with an integrated battery pack that facilitates transport of the cache module with the battery into a different controller.

Figure 1. Key features of the Dell PERC 6

in this firmware update provides administrators with the tools to select disks, enclosures, or an entire RAID subsystem and have this selection automatically configured into a RAID-10 virtual disk. Administrators can enable the Intelligent Mirroring feature in the Create Virtual Disk Advanced Wizard options screen in Dell OpenManage Server Administrator (OMSA) Storage Management (see Figure 4). Once enabled, this feature helps maximize redundancy by ordering the selected disks based first on mirroring between PERC 6/E connectors, then ordering them between enclosures, then matching drives within the enclosures based on size and slot.

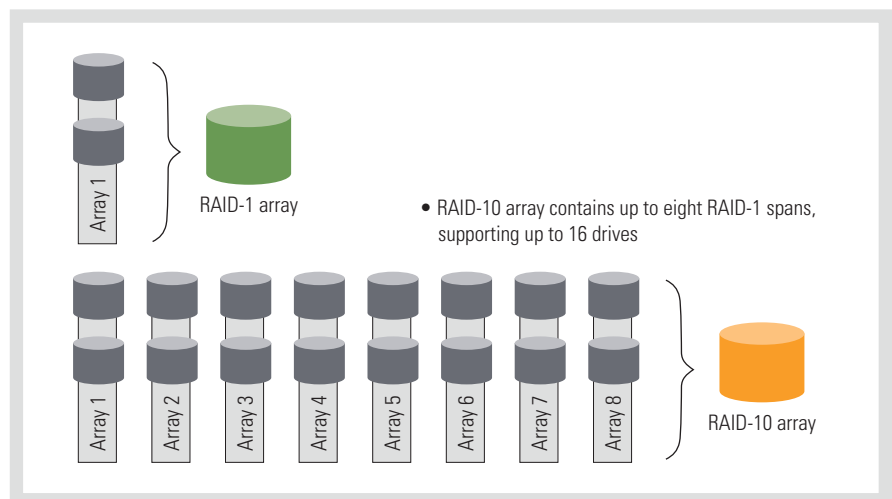


Figure 2. Maximum RAID-10 configuration supported by previous Dell PERCs

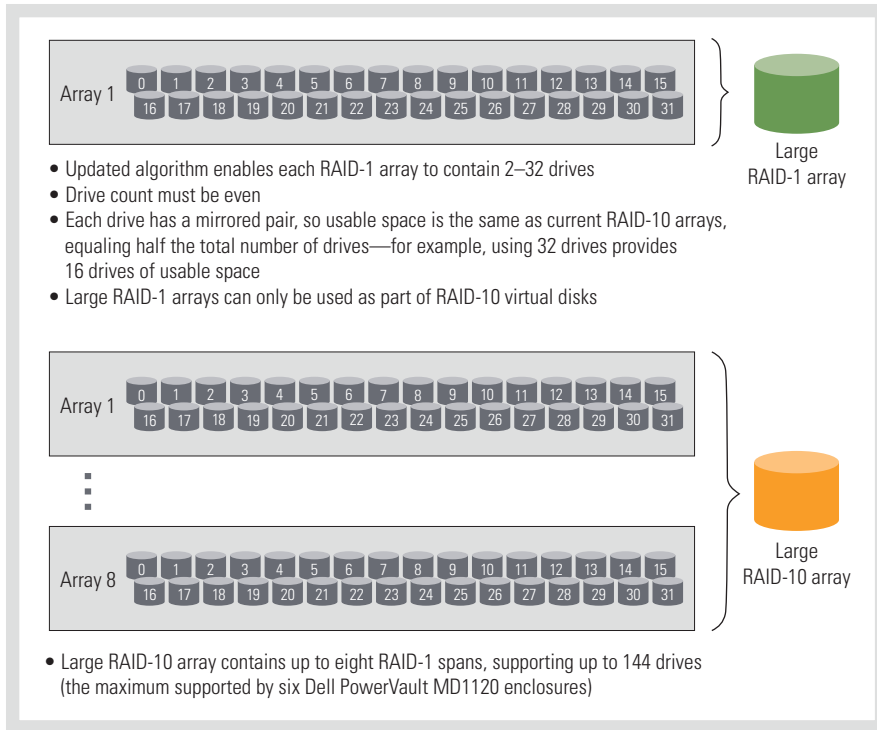


Figure 3. Expanded maximum RAID-10 configuration supported by the Dell PERC 6.1 firmware

Administrators can then review the selection and make configuration changes to help optimize the mirrored pairs, after which the array is created and ready for use. This feature enables administrators to easily create an array that can tolerate a pulled cable, enclosure power loss, or hardware failure. Administrators can combine the Intelligent Mirroring feature with redundant paths to help maximize data redundancy.

REDUNDANT PATHS AND LOAD BALANCING

The PERC 6.1 firmware enables administrators to redundantly connect Dell PowerVault MD1000 and PowerVault MD1120 external storage enclosures to both external connectors on the PERC 6/E when in unified mode. The connection not only offers failover protection, but can also provide static load balancing and performance enhancements compared with systems using nonredundant paths.

The PERC 6/E is designed to independently discover the hard drives on each of the controller’s two ports along with their Serial Attached SCSI (SAS) addresses.

If a drive has two SAS addresses, it is redundantly pathed. This feature works for SAS disks (because they have two addresses) and for Serial ATA (SATA) disks with interposers. After the controller discovers the drives, the controller ports

are merged into a single redundant path connector in the OMSA Storage Management graphical user interface (GUI) (see Figure 5).

After completing the discovery process, the controller counts the redundantly pathed disks, then assigns each disk to one of its two external ports, where the disk would then receive I/O traffic. The disks typically alternate between ports (slot 0 on port 0, slot 1 on port 1, and so on). Because the load balancing is performed based on disk count—not dynamically based on virtual disk or I/O load criteria—this feature helps avoid situations in which both ports attempt to communicate with a disk simultaneously, and helps prevent the possibility of overlapping read and write operations to a section of a disk (see Figure 6). Administrators can enable and disable load balancing through either the GUI within OMSA Storage Management or through the command-line interface (CLI). Virtual disks that use redundant paths are configured in the same way as those that use nonredundant paths; no administrator intervention is required during normal operation to enable this support.

Redundant paths also provide a failover method during a cable or

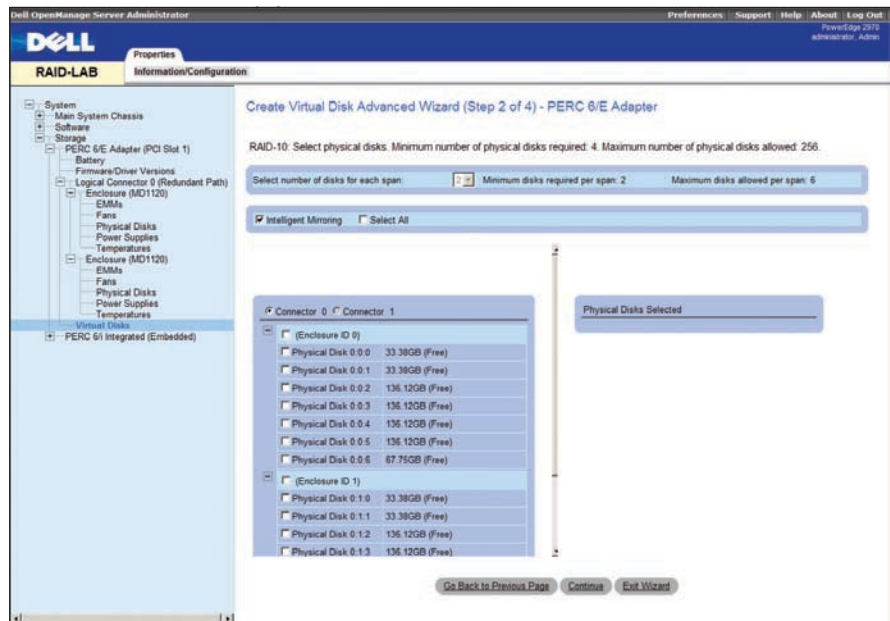


Figure 4. Create Virtual Disk Advanced Wizard options screen in OMSA Storage Management

enclosure management module disruption. Although this feature does not offer RAID controller redundancy, it does provide a means for queued I/Os to be completed by automatically transferring them down the other available data path if the primary data path is unavailable. Future I/Os can then continue down the available data path; administrators are notified of the data path disruption through OMSA Storage Management. When the data path is restored, the disks can be load balanced automatically so that an even number of drives are on each port.

The load-balancing architecture can also provide additional performance increases for random write and sequential read operations. In some workloads, the required bandwidth can exceed that of what a single x4 SAS link can provide (approximately 1,050 MB/sec). The additional path can help increase performance by nearly 50 percent for sequential read operations (see Figures 7 and 8).²

Administrators should keep in mind that because redundant paths utilize both PERC 6 ports, this feature can support only three Dell PowerVault MD1000 or PowerVault MD1120 storage enclosures, compared with six enclosures when using a single path. The Dell BIOS Configuration Utility (which administrators can access by pressing Ctrl+R during boot) does not indicate whether a subsystem is redundantly pathed; administrators can only view and manage redundant paths within the OMSA Storage Management GUI.

REVERTIBLE HOT SPARES

To help further simplify storage administration, the PERC 6.1 firmware introduces a revertible hot spare feature, which is designed to provide simple data protection when using hot spares in a RAID configuration. Traditionally, when a drive fails, the hot spare rebuilds into the array, and administrators replace the failed

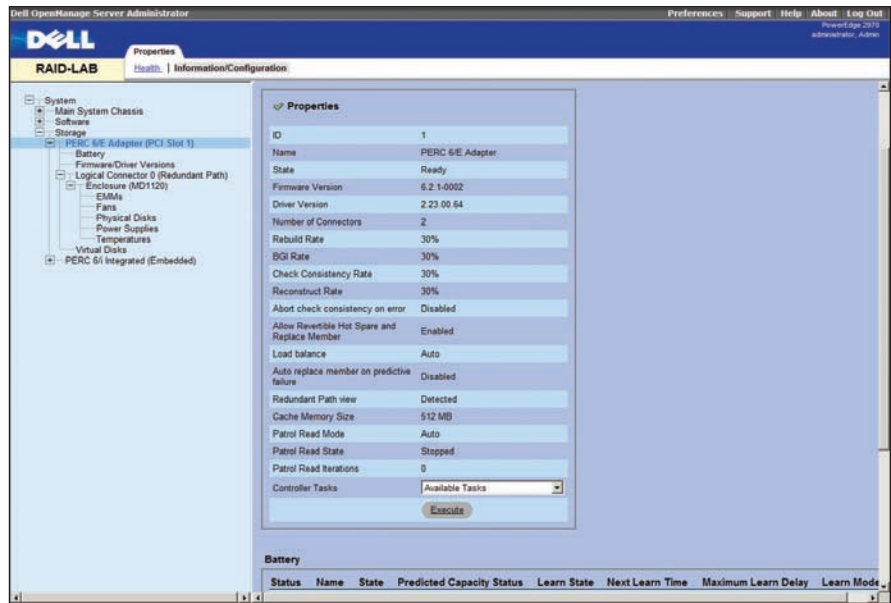


Figure 5. Merged Dell PERC 6 controller ports in OMSA Storage Management

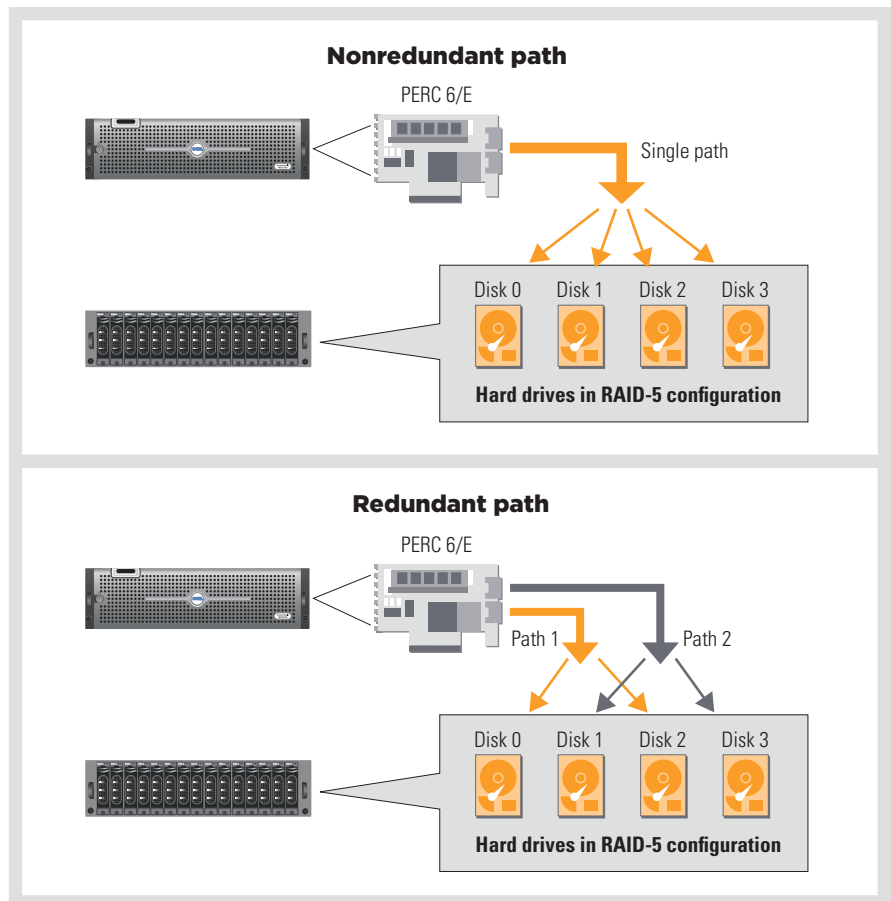


Figure 6. Nonredundant and redundant path configurations for the Dell PERC 6

²Based on Iometer tests performed in August 2008 by Dell engineers using a Dell PowerEdge 2970 server configured with two dual-core AMD Opteron™ 2210 processors at 1.79 GHz, 8 GB of RAM, a PERC 6/E with firmware version 6.1.1, and the 64-bit Microsoft Windows Server™ 2003 Release 2 (R2) OS, connected to a Dell PowerVault MD1120 storage enclosure configured with twenty-four 10,000 rpm hard drives. All data was generated by configuring Iometer to ramp up I/O for 10 seconds and then capture performance data by averaging each run over a 45-second interval.

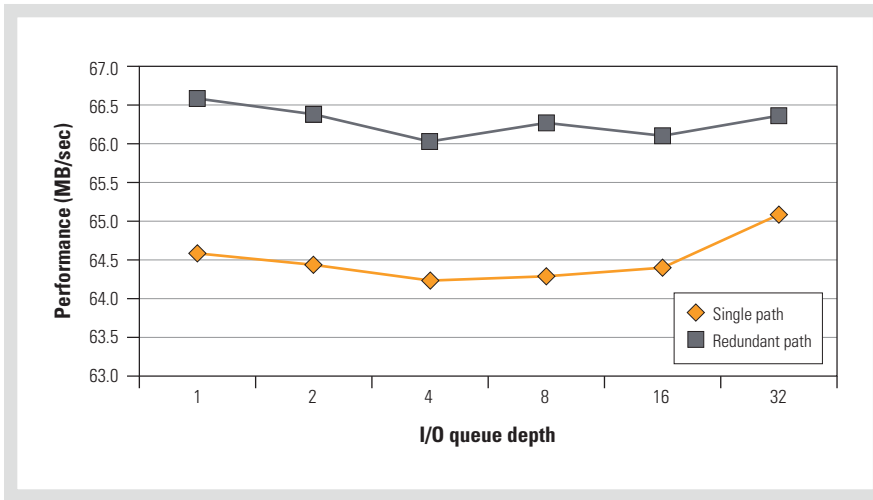


Figure 7. Dell PERC 6 performance results for 128 KB random write operations using a single path and a redundant path

drive manually. After this rebuild, the originally assigned hot spare becomes part of the RAID volume. If administrators need to assign another hot spare, they must manually reassign the replaced drive in that slot location within the array as a hot spare.

This method of hot spare allocation is still available for administrators who prefer it. However, the PERC 6.1 firmware now also supports revertible hot spares. When a drive fails, the hot spare rebuilds into the array, and administrators still replace the failed drive manually. At this point, the former hot spare can automatically begin replicating data back to the replaced drive. When this replication completes, the original hot spare drive reverts back to a hot spare drive automatically, without requiring administrator intervention.

Revertible hot spares also allow automatic drive replacement through Self-Monitoring, Analysis, and Reporting Technology (SMART), which can help detect predictable physical disk failures. If a disk detects such a failure, the controller can automatically copy that disk's contents to an available hot spare. On completion, the disk is marked as failed, and the hot spare becomes part of the virtual disk. The virtual disk does not become degraded during this process.

This feature is disabled by default. Administrators can enable it using the GUI in OMSA Storage Management or the CLI.

SIMPLIFIED STORAGE ADMINISTRATION

Administrators can upgrade the enclosure and hard drive firmware online without rebooting, helping keep critical storage systems available. In addition, administrators can add Dell PowerVault MD1000 and PowerVault MD1120 storage enclosures quickly and efficiently: the PERC 6.1 firmware can automatically detect the new enclosures and drives and enables administrators to configure them through OMSA

Storage Management. Administrators can also select whether the write cache is enabled for SAS and SATA drives; by default, the write cache is disabled for SAS drives and enabled for SATA drives.

The PERC 6.1 firmware also introduces a feature within OMSA Storage Management that lets administrators preview foreign configurations so they can verify and validate them before importing them into the environment. Additional information is presented to help identify potential problems with newly added configurations: for example, if a virtual disk is in a foreign state without all members, or offline, OMSA Storage Management flags this condition and notes what is missing from the disk group. This feature helps administrators identify actions required to complete the array and helps avoid unnecessary rebuilds or other additional functions.

INCREASED BATTERY SHELF LIFE

Significant enhancements have been made to help increase the shelf life of PERC 6 batteries and help improve battery recovery following an error. Updated PERC 6 batteries are designed to provide a shelf life of up to two years before the battery charge is fully depleted, when stored under appropriate conditions. In addition, the PERC 6.1 firmware enables a battery that was offline because of excessive drain, over-temperature conditions,

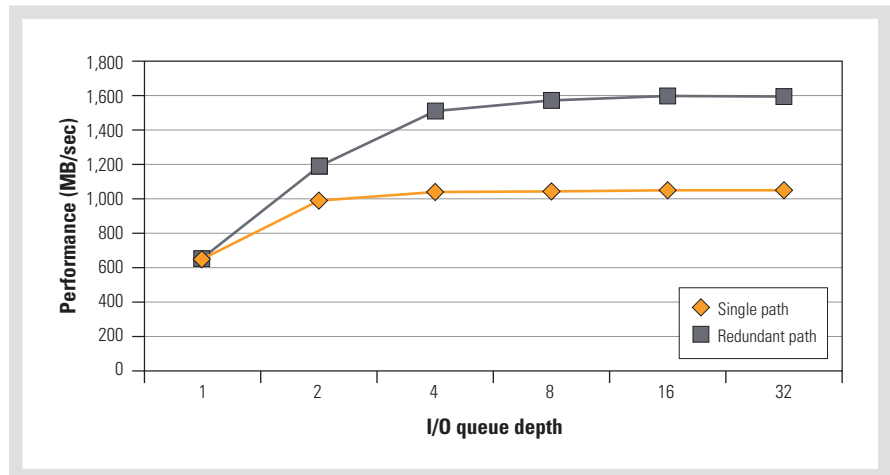


Figure 8. Dell PERC 6 performance results for 128 KB sequential read operations using a single path and a redundant path

or another failure to be brought back online, helping prevent system reboots and helping virtual disks to safely use write-back caching. When the battery is discovered, OMSA Storage Management sends a message that the battery is present, and the system can automatically adjust its cache policies to help increase performance.

ENHANCED RAID STORAGE MANAGEMENT

The PERC 6.1 firmware update is designed to enhance reliability, performance, and ease of use in enterprise IT environments. To help organizations continue meeting the challenges of ongoing data growth, Dell plans to continue enhancing the PERC 6 family of RAID controllers, including anticipated future support for solid-state drives, additional performance upgrades, and the upcoming enhancements offered by 6 Gbps SAS ecosystems. [u](#)

André Dumouchelle is an engineer in the Dell Server Storage Development Group. He has a B.S. in Computer Engineering from North Carolina State University.

Joe H. Trickey III is a storage product marketing consultant in the Dell Product Group. He has a B.A. in Political Science from the University of Texas at Austin.

John Seward is an engineer in the Dell Server Storage Development Group. He has a B.S. in Electrical and Computer Engineering from Baylor University.