Background Patrol Read

for Dell PowerEdge RAID Controllers

Background Patrol Read, a new feature in Dell™ PowerEdge™ RAID Controllers (PERCs), is designed to help prevent data loss in a redundant array. This article describes how Background Patrol Read works and how it interoperates with Consistency Check and SMART alerts within the PERC Fault Management Suite.

BY DREW HABAS AND JOHN SIEBER

Related Categories:

Dell PowerEdge RAID Controller (PERC)

Dell PowerEdge servers

Fault tolerance

RAID

Storage

Systems management

Visit www.dell.com/powersolutions for the complete category index.

Over time, sectors on a hard drive can become damaged and unusable. To help address this problem, hard drives keep track of unusable sectors in a grown-defect list. Once a sector is added to this list, future attempted accesses to it are remapped to a good sector—a back-up sector in the hard drive designated for this purpose.

Typically, bad sectors are discovered through normal read or write accesses to the drive. Using a redundant RAID level such as RAID-1 or RAID-5, a drive array can survive this data integrity threat by reconstructing the data from the good sectors of the other drives and remapping the data to an unused backup sector on the affected drive. Dell PowerEdge RAID Controller (PERC) products are designed to manage this process seamlessly and independently from the rest of the server without administrative intervention. As long as the array remains redundant, the server can remain online and running with data intact.

However, when a drive fails and there is an unknown bad sector on one of the nonfailed drives, problems may occur. As an example, consider a RAID-5 array that has experienced a drive failure and—unknown to the RAID controller—one of the nonfailed drives contains a hard-drive media defect. The administrator replaces the failed drive with a new one, and the array starts to rebuild the data onto the new drive based on the parity and peer data sectors from the remaining good drives. While data is being rebuilt on the new drive, the drive with the bad sector is read, but because of the media defect the data cannot be read. Without all of the available peer data and the parity, the new drive's sector cannot be regenerated. At this point, the array has lost data.

In the past when drive capacities were smaller, this type of problem was less likely to occur because small arrays typically contain fewer media defects than large arrays. Large hard drives are more prone to media defects. Media defects are specified as *x* number of defects per *y* number of bits. Therefore, larger drives are prone to more defects because they contain a greater number of bits. Today, hard-drive capacities have increased remarkably, and the likelihood has grown that one or more media defects will occur over the lifespan of the drive. In addition, large arrays take longer to rebuild than small arrays, thus increasing the amount of time the array is not redundant. Today's RAID systems need a proactive tool—such as the Background Patrol Read feature of Dell PERCs—to help avert such data problems by fixing the bad sectors when all of the drive array members are online and redundant.

Understanding how Background Patrol Read works

Background Patrol Read is designed to proactively detect hard-drive media defects while the array is online and redundant, and then proceeds to recover data. This tool provides three functions: data protection, variable run modes, and dynamic performance management.

Data protection

This function concerns data reconstruction and remapping. Background Patrol Read issues commands to each drive in the array to test all sectors. When a bad sector is found, the PERC instructs the hard drive to reassign the bad sector, and then reconstructs the data using the other drives. The affected hard drive then writes data to the newly assigned good sector. These operations continue so that all sectors of each configured drive are checked, including hot spares. As a result, bad sectors can be remapped before data loss occurs.[1]

Variable run modes

Background Patrol Read includes two run modes to help enhance flexibility and data protection:

- **Auto mode:** In this mode, the tool is always on. Once Background Patrol Read has checked all sectors of the array, it repeats this check indefinitely. This mode is recommended so that the PERC can maintain optimal array health.
- **Manual mode:** This mode is used to perform a single, quick check of the array. After Background Patrol Read has checked all sectors of the array, it stops and will not start again until started manually by the administrator. In Manual mode, Background Patrol Read commands are given a higher priority than in Auto mode so that the check completes significantly faster.

Manual mode is recommended during periods of low drive activity or during system maintenance. Background Patrol Read can be controlled by running the MegaPR utility within a Microsoft Windows® or Linux® OS.

Note: As its name indicates, Background Patrol Read is always a background or secondary process. Data I/O remains the highest priority for the RAID subsystem, whereas Background Patrol Read uses spare bandwidth.

Dynamic performance management

Background Patrol Read is designed to provide a balance between data protection and high performance. As a result, it uses an intelligent algorithm to adjust how much bandwidth it consumes based on the current data workload. During periods of light workload, the volume of Background Patrol Read commands increase. Conversely, during periods of heavy workload, the volume of Background Patrol Read commands decrease.

To gauge the intensity of the present workload, Background Patrol Read senses the volume of outstanding I/O operations pending for each hard drive. With this information, Background Patrol Read then adjusts both the frequency and size of the commands it sends to each drive.

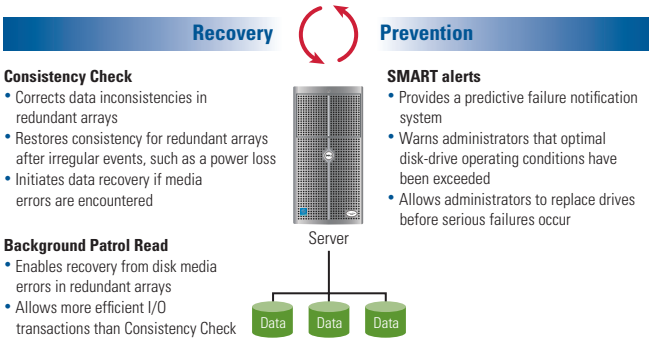

Figure 1. PERC Fault Management Suite

Exploring other components of the PERC Fault Management Suite

Background Patrol Read is just one of several PERC features that constitute the PERC Fault Management Suite. These features work in concert to help provide robust data protection. Other components of this suite are Consistency Check and SMART alerts. Figure 1 summarizes the components of the PERC Fault Management Suite and their functions.

Consistency Check

Consistency Check is designed to correct data inconsistencies in redundant arrays. A RAID-5 array is inconsistent when the data and parity do not match. Likewise, a RAID-1 array is inconsistent when the data and mirror do not match. Data inconsistencies can arise when all writes to an array are not completed because of catastrophic events such as a power loss.

Performance problems can arise when an array is inconsistent. RAID-5 arrays employ consistency checks to help improve write performance. For example, if the RAID controller must write new data to only a portion of a data stripe, it does not need to access every drive in the array if the array is consistent. When the array is consistent, the controller can read the new data from the host, read the old data from the affected drive or drives, and read the parity information. The RAID controller can then calculate the new parity and write the information to only the affected drives.[2]

An added bonus of running Consistency Check is that drive media are checked for errors and data recovery is initiated, similar to Background Patrol Read. However, Background Patrol Read is more efficient than Consistency Check in addressing this concern because Consistency Check is a data-level check and requires more controller resources to read and compare data. Also, because of the additional resources required, Consistency Check is not designed to

[1] Hard drives that are not part of a RAID array or are not assigned as hot spares are not scanned for media defects because these drives do not yet contain data.

[2] For more information about this read-modify-write operation, see "Understanding RAID-5 and I/O Processors" by Paul Luse in *Dell Power Solutions,* May 2003, www1.us.dell.com/content/topics/global.aspx/power/en/ps2q03_luse?c=us&l=en&s=corp.

run continuously. Rather, it should be scheduled to run at a regular interval, preferably during periods of low drive activity.

SMART alerts

Self-Monitoring Analysis and Reporting Technology (SMART) gauges hard-drive health. The PERC and hard drives work together to monitor various aspects of drive performance. They determine whether the drives are behaving normally and provide status information. Administrators can then choose to replace a drive before a failure occurs. SMART focuses on predictive errors, which are errors that occur over a long period and provide early warning signs that the situation is deteriorating.[3]

SMART alerts work in concert with Background Patrol Read and Consistency Check. During the life of a hard drive, Background Patrol Read and Consistency Check help maintain drive health and protect data. As the drive ages, more sectors are added to the grown-defect list and early warning conditions arise—thus triggering SMART alerts. SMART alerts allow administrators to assess a drive's health and consider replacement before failures occur.

Providing robust data protection

The PERC Fault Management Suite emphasizes proactive error detection, data recovery, and error prevention to help keep data safe and minimize downtime. In particular, Background Patrol Read blends proactive data protection with dynamic performance management and run-mode flexibility. This tool is designed to both protect data and provide optimal end-user experience by automatically maintaining performance and allowing administrators to make runtime adjustments to fit their data environments.

Drew Habas is a RAID product marketing manager at Dell. He is responsible for planning and managing Dell's RAID products. He has a B.S. in electrical engineering from the University of Illinois at Urbana–Champaign and an M.S. in engineering management from The University of Texas at Austin.

John Sieber is the lead engineer for Dell's SCSI RAID development group. He has extensive experience with RAID, storage area network, and network attached storage technologies. John has a B.S. in computer engineering from Texas A&M University.

[3] For more information about this technology, see www.pcguide.com/ref/hdd/perf/qual/featuresSMART-c.html.