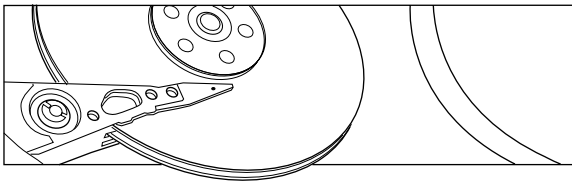

Time-Limited Error Recovery (TLER)

Error Recovery Fallout Prevention in WD Caviar RAID Edition Serial ATA Hard Drives

Meeting the demands of the enterprise environment means finding ways to improve performance, improve up time, improve compatibility, decrease down time, and reduce total cost of ownership. To this end, Western Digital has introduced Time-Limited Error Recovery (TLER) in its WD Caviar® RAID Edition (RE) hard drives to improve error coordination between WD hard drives and RAID controllers.

The Problem

Desktop hard drives are designed under the assumption that there is no RAID card. All desktop drives include error correction such as the ability to handle write errors and reallocate around bad blocks. During error correction, desktop hard drives do not issue error messages or respond to commands by adapters. Desktop hard drives are designed with the assumption they should do everything possible to complete error correction (the design assumes there is no RAID controller to help with error recovery). The difficulty comes when error correction takes longer than 8 seconds and RAID controllers assume that the non-responding disk has failed and the RAID controller drops the drive from the RAID volume.

ATA drives being "dropped" from a RAID volume was a commonly heard complaint regardless of manufacturer of the hard drive (that is, until WD delivered TLER). This error handling "mis-coordination" is encountered when drives are under a high I/O load such as a video surveillance server, a busy e-mail server, or a busy web server. And under high I/O load, the length of time needed to recover increases.

A drive under a continuous I/O load and performing its own error recovery can easily exceed 8 seconds of timeout, during which the normal desktop hard drive does not respond. Although there is no industry standard, RAID cards will typically wait 8 seconds for a drive to respond, and if the drive does not respond, RAID cards are programmed to take action. The "mis-coordination" of error handling occurs when desktop drives are programmed to take responsibility for all error recovery; while RAID cards are *also* programmed to take responsibility for error recovery.

The consequences of this mis-coordinated error handling are significant. After the drive has been dropped from the RAID volume, the RAID volume runs in degraded mode until a replacement drive is supplied. After a replacement drive is supplied, assuming it is configured as a RAID 5 volume, the RAID volume must be rebuilt from parity data.

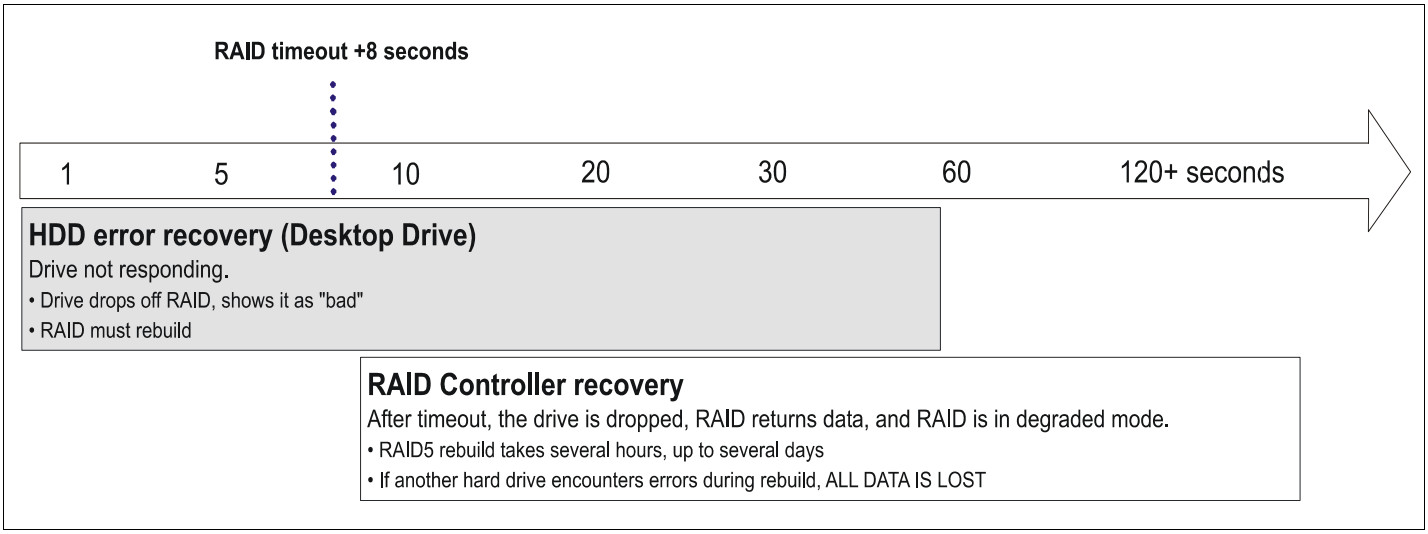

Figure 1. Desktop Drive Without TLER

Western Digital

While the RAID volume is running in degraded (parity recovery mode), the disks work harder to process the normal I/O load *and* process the parity I/O. This further increases the likelihood that an error recovery will exceed 8 seconds. Moreover, once the dropped drive is replaced, the RAID volume must be rebuilt. For large volumes (1 to 10 terabytes), this rebuild process can take hours to days. Like driving a car without a spare tire, the consequence of the second failure is far worse than the first failure. Thus, if another drive fails during recovery, all data on the volume is lost. The probability of this happening is increased when all drives work harder to handle both the normal I/O load and the parity rebuild I/O load.

The Solution

Hard drives for servers are designed with the assumption that there is a RAID controller present and some coordination of error management must occur. WD has delivered that coordinated error management in the form of Time Limited Error Recovery (TLER).

TLER-capable hard drives will perform the normal error recovery, and after 7 seconds, issue an error message to the RAID controller and defer the error recovery task until a later time.

With coordinated error handling, the hard drive is not dropped from the RAID array, thereby avoiding the entire RAID recovery, replacement, rebuild, and return experience.

The error handling is further coordinated between the TLER-capable hard drive and the RAID card. The TLER-capable drive will respond without waiting on the error to be resolved. RAID cards are very capable of handling this with a combination of parity protection and journaling. The RAID card flags the error in the error log and proceeds to deliver data using parity protection until the drive retries its own error recovery and corrects the error. This is quite similar to error management proven in SCSI-RAID for many years. It is important to realize TLER-capable hard drives should not be used in non-RAID environments.

Conclusion

Through coordinated error handling, TLER prevents hard drive error recovery fallout by limiting the time the drive spends in error recovery, providing increased performance, improved availability, and lower total cost of ownership in RAID arrays.

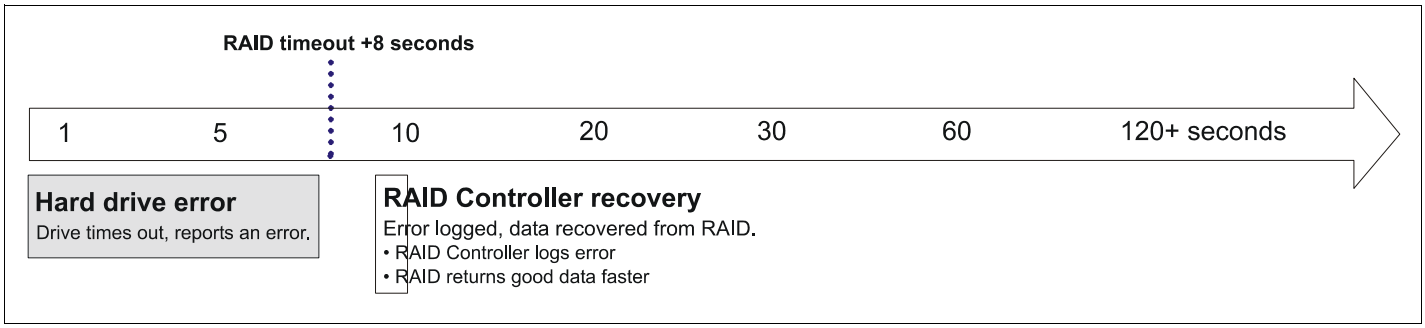

Figure 2. WD Caviar RAID Edition Drive With TLER

Western Digital
20511 Lake Forest Drive
Lake Forest, California 92630 U.S.A.

2579-001098-A00 10/04

For service and literature:

800.ASK.4WDC North America
949.672.7199 Spanish
+800.6008.6008 Asia Pacific
+31.20.4467651 EMEA
support.wdc.com
www.westerndigital.com