Introduction
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 enterprise hard drives to improve error handling compatibility between WD hard drives and RAID controllers.

The Problem
All drives include error correction such as the ability to handle read errors. During error correction, SATA hard drives do not issue error messages or respond to commands by adapters. Desktop SATA 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 issue with using normal desktop SATA drives in a RAID environment occurs when error correction takes longer than 8 seconds. When this happens, the RAID controllers assume that the non-responding disk has failed and therefore drop the drive from the RAID volume.

ATA drives being “dropped” from a RAID volume was a commonly heard complaint regardless of manufacturer of the drive (that is, until WD delivered TLER). This error handling incompatibility 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 incompatibility 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 incompatibility in 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.

<table>
<thead>
<tr>
<th>RAID timeout +8 seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

**HDD error recovery (Desktop Drive)**
- Drive not responding.
  - Drive drops off RAID, shows it as “bad”
  - RAID must rebuild

**RAID Controller recovery**
- After timeout, the drive is dropped, RAID returns data, and RAID is in degraded mode.
  - RAID5 rebuild takes several hours, up to several days
  - If another hard drive encounters errors during rebuild, ALL DATA IS LOST

Figure 1. Desktop Drive Without TLER
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 degraded or recovery mode, 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

WD has delivered 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. Though TLER is designed for RAID environments, it is fully compatible with, and will not be detrimental when 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 system performance, improved data availability, and lower total cost of ownership in RAID arrays.

<table>
<thead>
<tr>
<th>RAID timeout +8 seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

**Hard drive error**

Drive times out, reports an error.

**RAID Controller recovery**

Error logged, data recovered from RAID.

- RAID Controller logs error
- RAID returns good data faster

Figure 2. WD Enterprise Drive With TLER