The Private Lives of Disk Drives

How NetApp Protects Against Five Dirty Secrets

By Rajesh Sundaram

NetApp builds resiliency into its storage systems at every level to ensure that critical data is always protected. If you've been involved with NetApp for a while, you've probably heard a lot about technologies such as SnapMirror®, SnapVault®, and Snapshot™ that protect you from events ranging from sitewide disasters to user and application errors. NetApp also offers a unique degree of resiliency against problems that occur within disk drives themselves, but you've probably heard much less about these technologies.

You may be surprised by some of the "secret" problems that still lurk inside disk drives despite their remarkable dependability. Below are five of the most troublesome disk problems and the resiliency technologies that NetApp Engineering has developed to protect against them.

Secret 1: Drives fail suddenly!

NetApp solution: unique RAID-DP™ technology that provides extra protection against failure.

Okay, this isn't really a secret. Despite their dependability, we all know that disk drives still occasionally fail. When you consider the relatively short production lifecycles of disk drives (most models are only manufactured for a year or two) combined with the huge production volumes for popular enterprise disks (tens of millions of units annually), it's obvious that some problems are going to occur. Occasionally, a component change, manufacturing facility change, generational drive transition, or some other perturbation will result in the production of a less reliable lot. NetApp uses stringent drive screening criteria that meet or exceed industry norms, but some failure modes are extremely time-dependent. This not only causes drives to fail after a significant amount of time has passed but also increases the likelihood of two drives failing near the same time.

One common failure mode is for a disk to suddenly cease functioning. One moment it's working fine, and the next it's gone with no warning. Everyone knows that the way to protect against this sort of failure is RAID, but what if two drives fail in the same RAID group or an uncorrectable media error occurs during RAID reconstruction? Given the rapid market adoption of large-capacity SATA drives, many customers don't realize the odds of "double failure" are increasingly stacked against them.

NetApp protects you from these types of failures with RAID-DP (RAID Double Parity). RAID-DP adds a second parity stripe to drastically increase data availability without sacrificing performance or capacity utilization. Aggregates and volumes using RAID-DP can withstand up to two failed disks in a RAID group-or the increasingly common event of a single disk failure followed by an uncorrectable bit read error from a second disk during reconstruct. (Disk capacities continue to increase, while the media error rate stays about the same. A disk reconstruction must read many more bits of data now than in the past, significantly increasing the risk of a bit error.)

RAID-DP offers the data reliability of mirroring (RAID1) at the price of RAID4. Before the release of RAID-DP, storage administrators typically limited the size of RAID groups to protect against these types of failure. With RAID-DP, NetApp customers can feel confident using larger RAID groups and aggregates.

Secret 2: Drives slowly degrade.

NetApp solution: unresponsive drive protection takes the drive offline and regenerates data from parity.

Another common disk failure mode is for a drive to slowly degrade away, resulting in a steady performance decline over time. This can happen for any number of reasons. If you've seen this problem, you know that you can often read all the data stored on the drive before it fails completely.

You may not be aware of the impact such a drive can have on storage performance. A drive with multiple media errors—or a drive with a servo problem—may take several minutes retrying a read until it succeeds. In server environments, the resulting long I/O response times can lead to unwanted connection terminations and noticeable delays on clients.

NetApp engineers specifically designed the Data ONTAP® operating system to anticipate and circumvent potential performance issues. In the event an unresponsive or semi-responsive disk emerges within the system, Data ONTAP ceases all I/O operations to the affected disk, marks it as offline, and serves reads from parity while queuing writes until the disk recovers. If the disk fails to recover, it is marked as failed and reconstructed to a spare.

The innovative disk offline feature (which is available only from NetApp) ensures high performance consistency that is critical to applications that demand consistent quality of service in from the storage subsystem.

Secret 3: A bad drive can lock up an entire FC loop.

NetApp solution: dual pathing, ESH2, and local SyncMirror prevent data lockout.

Sometimes firmware bugs or disk failures can result in a single disk locking up an entire Fibre Channel loop, blocking access to up to 84 drives. In these scenarios, the remaining drives are in perfect working condition but temporarily inaccessible until the communication path is unblocked. An errant drive may generate a LIP (loop initialization primitive) storm; the drive continuously issues LIP requests that interrupt ongoing data transmissions.

NetApp offers multiple levels of protection for this problem. Every NetApp drive uses dual pathing in which two independent loops are connected to each drive. If one loop is down, the other provides continued access. If a rare drive failure blocks both loops, dual redundant shelf I/O modules containing second-generation electronically switched hubs (ESH2) detect and bypass disk drives that can disrupt FC operations. In fact, the ESH2 module with firmware revision 15 (FW15) and higher is designed to specifically protect against LIP storms. By electrically isolating these drives from the loop via intelligent point-to-point switching, the ESH2 provides a safety net in addition to dual pathing.

For maximum data availability, customers can deploy NetApp SyncMirror to achieve a level of resiliency that no other storage vendor offers. SyncMirror is local RAID mirroring between two separate volumes on the same storage system. While it also provides improved read performance (similar to RAID1+0) and is an instrumental part of the NetApp MetroCluster disaster recovery solution, SyncMirror stands on its own for customers demanding the ultimate level of local storage resiliency. By ensuring two mirrors are stored on separate failure domains, SyncMirror protects your data against a wide range of rare and unpredictable failures, including dual cable breaks, power strip failures, dual loop failures, disk shelf backplane failures, HBA failures, and even up to five concurrent disk failures on mirrored RAID groups if also using RAID-DP.

For unparalleled local storage resiliency, NetApp recommends SyncMirror for business and mission-critical applications requiring the highest level of data availability.

Secret 4: Firmware bugs can cause silent data corruption.

NetApp solution: checksums and RAID scrubs ensure that correct data is always returned.

It's a well-known fact in the storage world that firmware bugs (and sometimes hardware and data path problems) can cause silent data corruption; the data that ends up on disk is not the data that was sent down the pipe. To protect against this, when Data ONTAP writes data to disk, it creates a checksum for each 4kB block that is stored as part of the block's metadata. When data is later read from disk, the checksum is recalculated and compared to the stored checksum. If they are different, the requested data is recreated from parity. In addition, the data from parity is rewritten to the original 4kB block, then read back to verify its accuracy.

To ensure the accuracy of archive data that may remain on disk for long periods without being read, NetApp offers the configurable RAID scrub feature. A scrub can be configured to run when the system is idle and reads every 4kB block on disk, triggering the checksum mechanism to identify and correct hidden corruption or media errors that may occur over time. This proactive diagnostic software promotes self-healing and general drive maintenance.

To NetApp, rule number 1 is to protect our customer data at all costs. Protection against firmware-induced silent data corruption is an example of NetApp's continuing focus on developing innovative storage resiliency features to ensure the highest level of data integrity.

Secret 5: Committed writes can get dropped!

NetApp solution: lost write protection—the only solution in the industry to protect against this threat.

Brace yourself, because we saved the most insidious disk problem for last. With extreme rarity, a disk malfunction occurs in which a write operation fails but the disk is unable to detect the write failure and signals a successful write status. This event is called a "lost write," and it causes silent data corruption if no detection and correction mechanism is in place. You might think that checksums and RAID will protect you against this type of failure, but that isn't the case. Checksums are written in the block metadata—co-resident with the block—during the

same I/O. In this failure mode, neither the block nor the checksum gets written, so what you see on disk is the previous data that was written to that block location with a valid checksum.

Only NetApp, with its innovative WAFL (Write Anywhere File Layout) storage virtualization technology closely integrated with RAID, identifies this failure. WAFL never rewrites a block to the same location. If a block is changed, it is written to a new location, and the old block is freed. The identity of a block changes each time it is written. WAFL stores the identity of each block in the block's metadata and cross checks the identity on each read to ensure that the block being read belongs to the file and has the correct offset. If not, the data is recreated using RAID. The check doesn't have any performance impact.

NetApp always uses WAFL at the lowest level of disk organization, so even block-oriented, SAN installations have this protection.

Conclusion

Since every storage vendor uses more or less the same disk drives, no one is immune to these problems. Not all vendors, however, can offer equal protection against them. Innovative NetApp technologies, including RAID-DP, unresponsive drive protection, SyncMirror, ESH2, RAID scrubs, and lost write protection, offer a level of security against disk malfunctions that other vendors don't.

Are *you* protected?

Rajesh Sundaram
Storage Resiliency Architect, NetApp

During his 8½ years in NetApp Engineering, Rajesh Sundaram has worked on many of the most significant resiliency projects in the company. In addition to being an early member of the team that worked on the WAFL® file system, Rajesh helped lead the rearchitecture of the RAID subsystem and the development of SyncMirror®. Rajesh is currently focused on designing unique new resiliency technologies and improved on-site drive diagnostics.