Western Digital

TECH BRIEF

ArmorCache™ Write Cache Data Protection

An OptiNAND™ Technology Feature for HDDs

The Performance of Write Cache Enable Combined with Write Cache Disable Data Protection

Introduction

To achieve optimal performance, hard disk drives (HDDs) are designed to cache write data in DRAM before recording the data to the physical media. As shown in Figure 1, once the data is stored in DRAM, a write complete signal can be sent back to the host to accept the next command. After caching data in DRAM, the drive can then stage writes and reorder when necessary, which optimizes the drive's write performance. However, DRAM is volatile, and its contents are lost when power is unexpectedly shut off, so data in DRAM which has not yet been written to media is lost. Because of this, HDDs have long had the option to enable or disable this write cache.

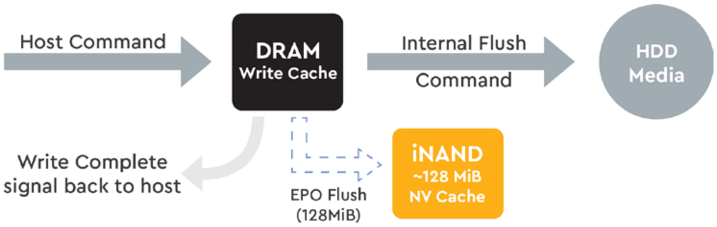


Figure 1: Emergency Power Off (EPO) Flush

The option to enable or disable write cache forces system designers into a tough decision. With Write Cache enabled (WCE) mode, the drive will achieve optimal write performance by accepting commands, signaling command completion to the system after caching data in DRAM, and writing them to media as soon as possible. However, if power is unexpectedly lost between the command being reported as complete and the data being written, the data will be lost. In Write Cache disabled (WCD) mode, the drive does not signal command completion for any writes that it cannot complete. This ensures that data is never lost, but it also means that the drive cannot cache any commands that it cannot guarantee will be written, which significantly limits write performance.

With the ArmorCache feature, enabled by OptiNAND™ technology, this choice is no longer necessary. When operating in WCE mode, drives can ensure that the DRAM cache will be safely written upon loss of power, and no data is lost. When operating in WCD mode, a larger cache size guarantees more data will be safely written upon power loss, and the drive operates with WCE-equivalent performance. Whichever the system chooses, ArmorCache technology combines the performance of WCE mode with the data protection of WCD mode, offering the user the best of both worlds.

ArmorCache Technology

In an emergency power off (EPO) scenario, a drive has very limited time to perform operations to ensure a proper shutdown. Upon recognition of power failure, a drive must first retract the heads from the media and then perform various "housekeeping" tasks to commit drive metadata in volatile DRAM to a non-volatile memory (NVM). While enterprise solid-state drives (SSDs) require large banks of capacitors to ensure enough remaining power to properly shut down, HDDs have a built-in energy source that they can use to provide this power: spinning disks.

Upon EPO, the HDD uses the rotational energy of the spinning disks to generate electricity. Akin to that created by a turbine generator, or by regenerative braking on an electrical vehicle, this energy gives the HDD time to perform its necessary housekeeping and write critical data to NVM. To date, however, the amount and speed of NVM used has been insufficient to write a significant amount of cached user data upon EPO. OptiNAND adds a high-capacity and high-speed NVM device to the HDD. With the addition of OptiNAND, the capacity and speed of NVM is now sufficient to write all cached user data in DRAM to NVM.

Upon EPO, the HDD can flush the entire DRAM cache to the onboard iNAND® device. Crucially, it is able to do this in both WCE and WCD modes. This guarantees in WCE mode that any cached data will be successfully written on EPO, and no user data will be lost. This also provides significantly larger caching capability in WCD mode, providing full performance while still guaranteeing all writes will be committed upon EPO.

Write Cache Enabled Data Protection

When operating in WCE mode, HDDs without ArmorCache require that the host system gracefully shut down before power is removed. This means that the host will typically send an ATA "Flush Cache" command or a SCSI "Synchronize Cache" command to inform the drive that any cached data needs to immediately be committed to media. The host would then not send any new data for a short amount of time in order to allow the drive to complete the writes, and then would finally power off.

In a real-world scenario, power may be lost unexpectedly and without warning to a host system, in which a graceful shutdown is not always possible. This can be mitigated with expensive solutions such as battery backup, or with complex system solutions to manage the data such that it is retained and rewritten upon subsequent power-up. However, many systems simply accept the risk of data loss, hoping that the lost data isn't critical enough to cause problems with system operation.

One of the most common methods to mitigate WCE data risk is via issuing flush cache commands after writing critical data. This tells the drive to immediately commit any cached data to media and not accept any new commands until the cache flush is complete. While this method does not provide perfect protection—an EPO could occur in the middle of a flush cache operation, leaving data unprotected—it reduces the risk. However, a forced flush cache command diminishes performance, as the HDD is designed to optimize the order in which it services commands. A forced cache flush impedes this optimization and reduces performance. With ArmorCache technology, host flush cache commands are no longer necessary to protect data, and the drive's maximum performance in WCE mode can be unleashed.

This data protection feature benefits users in three ways:

1. **No data loss during EPO** | For those who have a system that simply accepts the risk of data loss, it ensures that there will no longer be data loss during an EPO. If the lost data would consist of critical files or critical metadata, this will improve the overall resiliency of the system to corruption or failure.
2. **Reduced software complexity to manage WCE** | Systems which manage the data integrity via software require checking the media for data coherence upon power-up after an EPO, and then must manually rewrite any data which was lost. This increases the time before they can resume normal write operations while they are performing these checks. If the drive guarantees data is written, it will reduce this time as there are no rewrites needed, or eliminate it entirely. Finally, systems which rely on frequent flush cache commands can eliminate these, and thus will enjoy higher overall performance.
3. **Reduced system cost and complexity** | Systems which mitigate this cache with battery backup can reduce cost and complexity of their system by no longer needing battery backup to ensure volatile cached data is successfully written on EPO.

Most HDDs ship into the market with a default of WCE. Unless a host system manually disables this to go into WCD mode, the drives will remain in WCE mode. ArmorCache protects those drives from the risk of data loss due to EPO.

For system designers preferring or requiring a traditional WCE experience, the write cache data protection feature protecting DRAM cache can be disabled via ATA Set Features or SCSI Mode Select command.

Write Cache Disabled Performance Gains

Many systems are intolerant of data loss on power failure, and those systems put the drive into WCD mode. HDDs without ArmorCache technology suffer a significant performance penalty as a result. The drive in WCD must be able to guarantee that all data for which the drive has issued a command completion is safe. As a result, the performance of those drives is limited to only caching the quantity of data that it knows can be safely saved during an EPO.

Recent generations of Capacity Enterprise HDDs from Western Digital and competitors have had small amounts of NVM available to assist with EPO flush. This ensures that for small block sizes, host commands may be cached, queued, and completed as soon as possible. However, the amount of NVM necessary to successfully flush data during an EPO escalates quickly as command transfer length increases. With 128MiB of NVM allocated for EPO flush, more internally queued commands and ones of greater length can be safely written than on drives without OptiNAND.

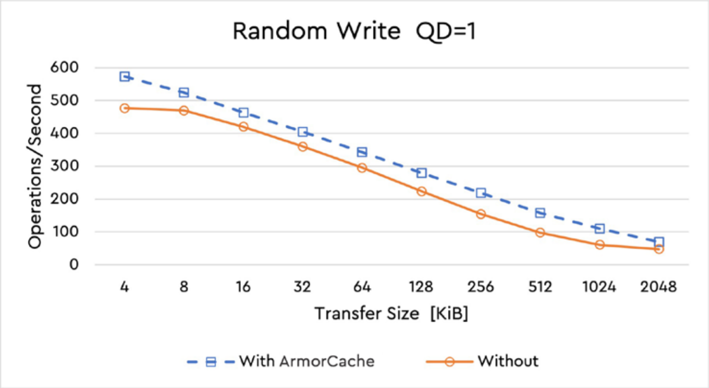
Many HDD specifications quote random write performance using the 4KB transfer size, and with very limited onboard NVM, these HDDs specify equal performance for these 4KB transfers whether the drive is in WCD or WCE mode. However, Table 1 shows that for larger transfer lengths, even at low queue depths, the amount of NVM necessary to flush DRAM grows quickly.

Commands Queued	Command Transfer Length	NVM Needed for EPO
128	4KB	512KB
64	32KB	2MB
32	256KB	8MB
24	512KB	12MB
16	1MB	16MB

Table 1: NVM capacity needed for successful EPO flush

While 4KB is the transfer length with the highest random write I/O per second (IOPS) value, 4KB is becoming increasingly irrelevant for HDD-based storage. Applications such as video recorders, object storage, network-attached storage (NAS) systems, and IOT sensor data use significantly larger transfer sizes. These large and rapidly growing applications for HDD storage receive the highest performance benefit from ArmorCache technology.

Figure 2 charts increased performance with transfer sizes of 4KiB up to 2048Kib as a result of the ArmorCache feature. At more common HDD application transfer lengths of 256KB or greater, IOPS improve by more than 40% relative relative to a drive without ArmorCache, with a peak improvement of over 80% at 1MB transfer length. OptiNAND's performance advantage is greatest in the most difficult and aggressive write workloads.



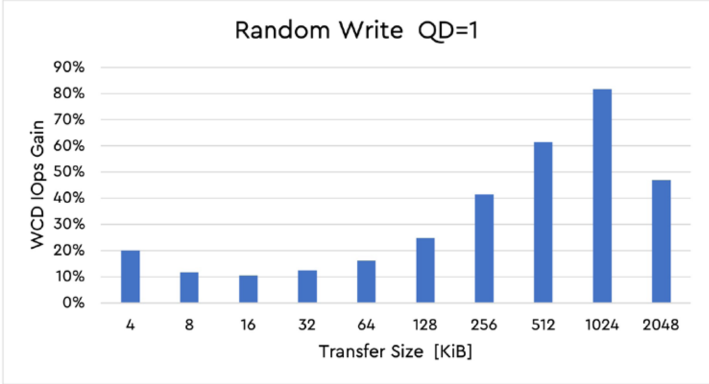


Figure 2: Performance Increase Graphed vs Transfer Length

While the data is shown with a queue depth of 1, this is the host queue depth. Internally, commands are still being staged and the internal command queue may achieve higher levels, as the 6Gbps SATA interface or 12Gbps SAS interface can send commands to the HDD faster than the HDD can write the data.

As a result, the performance gains at higher transfer sizes are directly attributable to the capacity of NVM devoted to cache flush being larger. A larger NVM allows more commands to be staged in WCD mode and thus provide much higher performance as they can all be guaranteed to be completed during EPO. Furthermore, because the NVM devoted to cache flush is equal to the size of the DRAM cache, the WCD performance is now equal to WCE performance.

Conclusion

ArmorCache eliminates the need for system designers to differentiate between the capabilities of WCE and WCD modes in HDDs. Guarantee of WCD data protection during EPO while the drive is in WCE mode is now available. The performance advantage of WCE is now fully available in WCD mode. Western Digital HDDs with ArmorCache offer the benefits of both modes regardless of the method chosen by the host.

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
www.westerndigital.com