



HGST Shingled Magnetic Recording + HelioSeal[®] Technology

Achieving Unprecedented Storage Capacity through Innovation

Introduction

The recent convergence of three macro forces—Web 2.0, Cloud-based Services, and Big Data—has created an infrastructural opportunity in which data that is properly stored and analyzed can be the basis for valuable business insights and new business opportunities. As data-in-flight transitions to data-at-rest, storage becomes one of the key infrastructural components. However, the demand for scalable, cost-effective mass storage at an attractive economic price point continues to challenge the technological capabilities of the storage industry. The upshot is that in cases where the cost of storage exceeds the value of the data at hand, data is left unstored—and untapped.

According to industry analysts, petabytes stored in the Capacity Enterprise Hard Drive Segment are increasing at a compound annual growth rate of 40%. Accompanying this hyper-growth, the storage industry is also undergoing a tectonic shift in which new market segments and storage tiers are emerging.

As Cold Data Warms, Enterprise Storage Categories Evolve

The emergence of Big Data analytics and bulk storage have changed the way data architects view and approach data. With increased processing power and data analytics technologies, data that was previously in “deep archive” (data sitting offline on removable media) is migrating to an “active archive” (data kept online and accessible), where continuous value can be extracted from the data set. The simplistic notion of archived data as data written to the cheapest media available and rarely accessed – namely tape drives – is rapidly changing.

In addition, many hyperscale and cloud storage customers are now beginning to realize that their workloads are trending toward data that is written sequentially and rarely updated, then read randomly and frequently. In these instances, customers are seeking storage solutions with the lowest total cost-per-terabyte and the maximum capacity. These active, sequentialized workloads have the potential to create a new paradigm opportunity for innovative storage solutions.

Challenges

The laws of physics are limiting how far storage vendors can continue to push existing perpendicular magnetic recording (PMR) technologies for significant capacity gains that meet the storage demands of business. Three factors that limit recording density are:

- **Soft Error Rate (SER)** - By further increasing bit density beyond present levels, the read-back signal deteriorates to the point where correction algorithm improvements and ECC extensions no longer yield net capacity gains at the required bit error rates of 1E-15 or better for enterprise class drives.

- **Adjacent Track Interference (ATI)** - By further squeezing track density, there is a gradual deterioration of signal integrity of adjacent tracks when writing repeatedly at high magnetic fields. This effect is called adjacent track interference. Today, drives can still manage that effect by monitoring write frequency and read-back error rates in order to initiate appropriate refresh cycles to meet data retention requirements. However, pushing track density further will prove increasingly difficult.
- **Far Track Interference (FTI)** - Increasing the write field also results in stray fields that impact neighborhood tracks in the vicinity of the present write location. This effect, known as far track interference, drives corrective actions very similar to ATI management.

These challenges, amongst others, will continue to slow PMR areal density growth for existing 3.5-inch form factor HDDs.

How Helium Technology Addresses the Capacity Challenge

Sealed, helium-sealed drives represent one of the most significant storage technology advancements in decades. In 2013, HGST introduced HelioSeal[®] technology, a foundational building block for high-capacity hard disk drives (HDDs). This technology innovation hermetically seals the HDD with helium, which is one-seventh the density of air. The less-dense atmosphere enables thinner disks and more of them for higher capacities in the same industry-standard form factor. Less air friction means less power required to spin the disks and less air turbulence for higher reliability. Helium drives expand the boundaries of conventional high-capacity HDDs, allowing for dramatic increases in efficiency, reliability, and value. HelioSeal technology delivers today’s lowest total cost of ownership (TCO) for hyperscale and data-centric applications.

Driving Down TCO: The Value of HelioSeal Technology

By reducing turbulence within the sealed HDD, HelioSeal technology allows unmatched storage density with industry-leading capacity in an eight-disk design and a 3.5-inch HDD footprint. Combining HelioSeal technology with HGST Host Managed SMR sets a new capacity standard, boosting 3.5-inch HDD capacity to 14TB, a 16% increase over 12TB drives. In addition to breakthrough drive capacity, HelioSeal technology provides the following benefits:



Greater power efficiency with 41% lower operating power than 8TB air drives. Our 14TB helium drives consume just 5.2 watts during idle operation, a 61% reduction in watts-per-TB compared to conventional 8TB air-based HDDs.



Lower cooling requirements with drives that typically run 4° to 5°C cooler to lower power and cooling costs. Cooler operation also results in better reliability and enables systems with higher storage densities.



Quieter operation and up to 50% lower weight-per-TB improve environmental conditions in high-density deployments and enable more storage capacity where building codes enforce floor loading limits.



More affordable than ever with the introduction of our high-volume, fourth-generation, 12TB PMR HelioSeal HDD -- higher volume leads to lower cost to market.

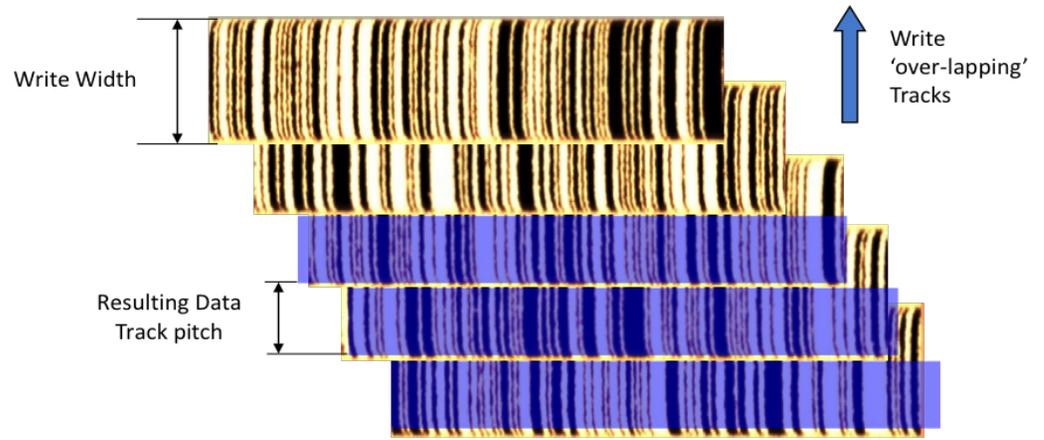


Figure 1: Shingled Magnetic Recording allows data tracks to be layered on top of each other to increase areal density up to 16%. However, to maintain predictable performance, this shingled approach requires all data streams to be organized and written sequentially to the media.

How Shingled Magnetic Recording Further Addresses the Capacity Challenge

Shingled magnetic recording (SMR) technology complements Helium technology by providing an additional 16% increase in areal density. Physically, this is done by writing the data sequentially, then overlapping (or “shingling”) it with another track of data. By repeating this process, more data tracks can be placed on each magnetic surface. See figure 1.(Perpendicular magnetic recording places gaps between recording tracks on HDDs to account for Track mis-registration (TMR) budget. These separators impact areal density, as portions of the platter are not being fully utilized. Shingled magnetic recording removes the gaps between tracks by sequentially writing tracks in an overlapping manner, forming a pattern similar to shingles on a roof.)

The write head designed for SMR drives is wider than required for a single track of data. It produces a stronger magnetic field suitable for magnetizing films of high coercivity. Once one track has been written, the recording head is advanced by only part of its width, so

the next track will partially overwrite the previous one, leaving only a narrow band for reading.

Overlapping tracks are grouped into bands (called zones) of fixed capacity for more effective data organization and partial update capability. Recording gaps between bands are laid to prevent data overwrite by the wide write head from one band to another.

Fundamental Implications of SMR

Because of the shingled format of SMR, all data streams must be organized and written sequentially to the media. While the methods of SMR implementation may differ (see SMR Implementations section below), the data nonetheless must be written to the media sequentially.

Consequently, should a particular track need to be modified or re-written, the entire “band” of tracks (zone) must be re-written. Because the modified data is potentially under another “shingle” of data, direct modification is not permitted, unlike traditional PMR drives. In the case of SMR, the entire row of shingles above the modified track needs to be rewritten in the process.

SMR HDDs still provide true random-read capability—allowing rapid data access like any traditional PMR drive. This makes SMR an excellent technology candidate for both active archive and higher-performance sequentialized workloads.

SMR Implementations

There are three shingled magnetic recording implementation options. It is important to understand their differences, as not all SMR options are appropriate for enterprise storage applications. The three options are:

- **Drive Managed SMR** allows plug-and-play deployment without modifying host software. It accommodates both sequential and random writing, but can result in highly unpredictable enterprise performance.
- **Host Managed SMR** requires host-software modification but delivers maximum benefit and competitive advantage. It accommodates only sequential write workloads, and delivers both predictable performance and control at the host level.
- **Host Aware SMR** offers the convenience and flexibility of Drive Managed SMR, with the performance and control advantages of Host Managed SMR. It accommodates both sequential writes and some random writes, but is the most complicated option to implement for HDDs. To obtain maximum benefit and predictability, host should do much of the same work as Host Managed SMR.

Depending on the implementation type, the layout of the recording media and the performance characteristics of the resulting drive may differ.

Drive Managed SMR

The major selling point for Drive Managed SMR is that no change is required at the host side, resulting in a plug-and-play deployment with any existing system. To allow for this deployment ease, Drive Managed SMR HDDs need to manage all random writes to sequential shingled writes by means of media caching and an indirection table. No responsibility or action is required of the host. The media cache provides a distributed storage area to lay down data at the best-possible speed (regardless of the host-provided block address associated with that data). Buffered data will be migrated from the media cache to the final destination as part of the drive's background idle time function. In short, Drive Managed SMR requires stronger caching algorithms and more random write space to temporarily hold non-sequential data.

Because the HDD constantly works to optimize the caching algorithm and indirection table handling, performance is unpredictable at certain workloads such as large block random write with high duty cycles. Due to the wide range of performance variability and unpredictability, Drive Managed SMR is considered impractical and unacceptable for enterprise-class deployments.

Drive Managed SMR is suitable for applications that have idle time for the drive to perform background tasks such as moving the data around. Examples of appropriate applications include client PC use and external backup HDDs in the client space. In the enterprise space, parallel operations of backup tasks become random write operations within the HDD, which typically result in unpredictable performance or significant performance degradation.

Host Managed SMR

Host Managed SMR is emerging as the preferred option for implementing shingled magnetic recording. Unlike Drive Managed SMR, Host Managed SMR defers SMR management to the host to help manage the data streams, read/write operations and zone management. And as another key difference from Drive Managed SMR, Host Managed SMR does not allow any random write operations.

As the name implies, the host manages all write operations to be in sequence by following a write pointer. Once data is written to the zone, the write pointer increments to indicate the starting point of the next write operation in that zone. Any out-of-order writes or writes to areas not properly indexed for writes by associated counters will force the drive to abort the operation and flag an error. Recovery from such an error is the responsibility of the controlling host. This enforcement allows Host Managed SMR to deliver predictable, consistent performance.

With Host Managed SMR, data is organized in a number of SMR zones ranging from one to potentially many thousands. There are two types of SMR zones: a Sequential Write Required Zone and an optional Conventional Zone. The Conventional Zone, which typically occupies a very very small percentage of the overall drive capacity, can accept random writes but is typically used to store metadata. The Sequential Write Required zones occupy the majority of the overall drive capacity where the host enforces sequentiality of all write commands. (It should be noted that in Host Managed SMR, random read commands are supported and perform comparably to that of standard PMR drives.)

Unlike Drive Managed SMR, Host Managed SMR is not backwards-compatible with traditional drive technologies. However, Host Managed SMR allows

enterprises to maintain control and management of storage at the host level. And once developers gain SMR familiarity and optimize their applications for sequential writing, Host Managed SMR allows for a smooth, staged transition to Host Aware SMR in a future timeframe.

Host Aware SMR

Host Aware SMR is the superset of Host Managed and Drive Managed SMR. Unlike Host Managed SMR, Host Aware SMR is backwards-compatible with traditional PMR drives.

With Host Aware SMR, the host does not necessarily have to change the software or file system to make all writes sequential. However, unless host software is optimized for sequential writing, performance becomes unpredictable just like Drive Managed SMR. Therefore, if enterprise host applications require predictable and optimized performance, the host must take full responsibility for data stream and zone management, much in the same manner as Host Managed SMR. In enterprise applications where there may be multiple operations or multiple streams of sequential writes, without host intervention, performance will quickly become unpredictable if SMR management is deferred at the drive level.

Host Aware SMR also has two types of zones: The Sequential Write Preferred Zone and an optional Conventional Zone. By contrast to Host Managed SMR, the Sequential Write Required Zones are now Sequential Write Preferred Zones in Host Aware SMR, thereby allowing for random write commands. If the zone receives random writes that violate the write pointer, the zone becomes a non-sequential write zone.

A Host Aware SMR implementation provides internal measures to recover from out-of-order or non-sequential writes. To manage out-of-order writes and background defragmentation, the drive holds an indirection table, which must be permanently maintained and protected against power disruptions. Out-of-order data is recorded into the caching areas. From there, the drive can append the data to the proper SMR band once all blocks that are required for band completion have been received.

Host Aware SMR is ideal for sequential read/write scenarios where data is not overwritten frequently and latency is not a primary consideration. Although it requires the optimization of data sets and data streams.

Enterprise-Grade SMR: Performance Without Compromise

Some storage vendors achieve higher capacities and lower cost per terabyte by eliminating features from their drives. For example, reducing reliability specifications, product warranties, and performance specifications is one way to drive down costs. But the end result is desktop-class HDDs.

Data is the lifeblood of any business. That's why HGST believes cost savings cannot come at the expense of availability and reliability. The HGST Host Managed SMR solution maintains the industry's highest levels of reliability, quality, and performance. Rather than eliminating features to achieve higher densities and cost savings, the HGST solution actually adds features, such as write-verify capabilities, design optimizations, and patented HelioSeal technology to reduce both cost per terabyte of storage and TCO.

How to Implement Host Managed SMR: High-level Considerations

Host Managed SMR will be the first SMR implementation to gain enterprise market adoption, even though it requires customers to modify their storage software stack. HGST believes that the initial software investment required for Host Managed SMR will be fully utilized in future Host Aware SMR implementations.

ZBC/ZAC Command Sets

First and foremost, a new specification of commands have been defined for Host Managed SMR. These new command protocols are all standards-based and developed by the T10/T13 committee. (There is no industry standard for Drive Managed SMR because it is purely transparent to hosts.) Please refer to the latest draft of the ZBC/ZAC when designing to Host Managed SMR.

Optimizing Write

Data should be written in 4KB portions and must be delivered in sequence without gaps within a zone. This pointer will be incremented at the completion of every write command. Consecutive writes must start from the last pointer address—otherwise the drive will flag an error and abort the operation.

Multiple SMR bands may be opened for writes at any time, where the exact number is a function of available RAM resources. The maximum number is documented in each drive's identify data.

Read is as Good as PMR

Reading data from an SMR drive is no different from conventional PMR drives. However, the reduced track width may impact settling times and re-read counts, which depend on the actual track width and the drive's tracking capability.

Performance Expectations

To ensure write operations do not corrupt previously written tracks, some Host Managed SMR HDDs perform read verification of the previous track under certain conditions. This read-verification process ensures delivery of maximum capacity at maximum bit integrity, while delivering write throughput near but slightly lower than PMR HDDs.

Ecosystem Considerations

Due to the nature of Host Managed SMR design, it is not a plug-and-play implementation with existing systems. There are three ways to design to Host Managed SMR depending on the customer's system structure and ability to modify their application layer.

1. Customers that own software applications and do not depend on the operating system and file system to control the HDD can start with new ZBC/ZAC command specifications that are newly defined for Host Managed and Host Aware SMR. These new command sets are based on T10/T13 committee standards. The applications will need to be re-written with the new command sets as well as all data streams sequentialized.
2. Customers that run on open-source Linux and relies on its own kernel to control the HDD will need to modify the file system, block layer, and driver. (The alternative option is to await for Linux and file system support). In either case, a thorough understanding of ZBC/ZAC is required to modify the kernel layer.
3. If a customer lacks the software control at the host and kernel levels and relies on a host-bus adaptor (HBA) to help manage the SMR responsibilities, no software modification is required. However, this customer needs to work directly with the HBA vendor for SMR support.

Conclusion

With data and storage continuing to grow at unprecedented rates, the capacity enterprise market is starting to see an emergence of new storage segments that are predominately sequential, highly accessible, and at unprecedented capacity points. To effectively service these new segments requires a purpose built solution that leverages innovation such as Host Managed SMR. To capitalize on the capacity advantages, customers do need to make certain changes to accommodate for Host Managed SMR – ie. the host software (or end application) does need to be modified, all data streams need to be sequentialized, and new command sets are required. Because host-based type of SMR will become more pervasive and ubiquitous in the future, the investment is fully leveraged for subsequent generations of SMR products and even recording technologies. By building today's highest-capacity SMR storage solutions on the industry's most reliable helium-sealed HDD platform, HGST is addressing the pressing need for true enterprise-grade mass storage at a compelling Total Cost of Ownership.