Large-Capacity HDDs Applying SMR Technology for Data Centers

Large-capacity hard disk drives (HDDs) that allow data centers to store huge volumes of electronic information have become increasingly important as a key product for the development of information infrastructure in recent years. The two principal types of technical challenges to be addressed in order to increase the storage capacity of HDDs are (1) improvements in the performance of main parts including magnetic heads, magnetic disks, and the signal processing circuit in the case of conventional magnetic recording (CMR) HDDs; and (2) the overwriting of data on tracks, analogous to the shingling of a roof, in the case of shingled magnetic recording (SMR) HDDs.

Through the application of SMR technology and the optimization of magnetic heads and magnetic disks, Toshiba Electronic Devices and Storage Corporation is working toward realizing large-capacity HDDs for data centers. Experiments on a 3.5-inch CMR HDD product incorporating firmware appropriate for data writing in an SMR HDD have verified an approximately 25% increase in storage capacity.

1. Introduction

HDDs have evolved as secondary storage devices for computers. The applications of HDDs have expanded to include diverse fields ranging from TV recorders and other personal electronics to data center servers that store a huge volume of electronic information. Accompanying the progress of deep learning and other artificial intelligence (AI) technologies, the ever-growing amount of electronic information at data centers is now facilitating the secondary use of such information.

In light of the past 20 years and the trends likely to shape the next 20 years, high-capacity HDDs for data centers have evolved and will continue to evolve as they are essential for the development of information infrastructure.

SMR is suitable for increasing the capacity of HDDs, but does not deliver satisfactory performance. To resolve this issue, data center customers are also working on the development of host systems suited to SMR HDDs.

This report presents an overview of and discusses prospects for the SMR technology we intend to apply to the high-capacity HDDs we are developing for data centers[1].

2. Data write and read operations of CMR and SMR

The magnetic head of an HDD tracks a given track on the surface of a rotating disk. The magnetic head of an HDD has write and read elements. The write element records data by magnetizing bit cells on a circular track to exhibit south (S) and north (N) polarities. The write element records data by magnetizing bit cells on a track in the circumference direction to exhibit S and N polarities. The read element reads back the recorded data by sensing the polarities of the magnetized bit cells.

Generally, the per-platter capacity of CMR HDDs is increased by reducing the bit and track pitches and thereby increasing the magnetic recording density (**Figure 1**). To reduce the track pitch, CMR requires a reduction in the width of the write and read elements in the radial direction. Since smaller write and read elements have lower recording and read-back capabilities, it is also necessary to improve the characteristics of magnetic heads and disks.

In contrast, SMR HDDs make it possible for the write element and other components to accommodate a narrower track pitch than in the case of CMR HDDs because SMR HDDs write a new track, overwriting part of the previously written track in a manner analogous to the shingling of a roof (**Figure 2**).

SMR tracks on a magnetic disk consist of bands and guard bands. A band is defined as a unit consisting of shingled tracks with a narrow pitch that are sequentially written. Each band on an SMR HDD is separated from adjacent bands by gaps called guard bands so that a write access to the last track of a given band will not corrupt the data on the first track of the next band (**Figure 3**).

W:write element   R:read element n:ordinal number
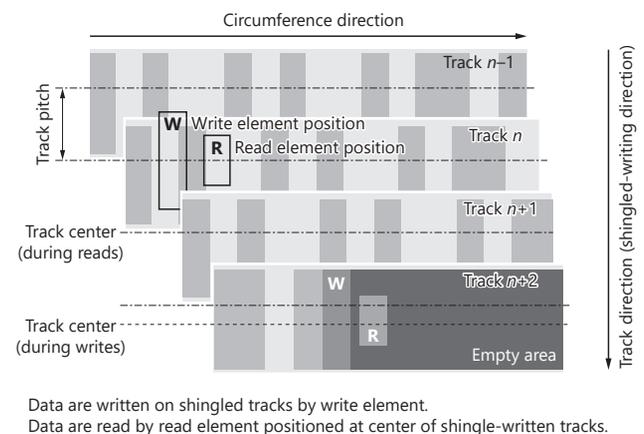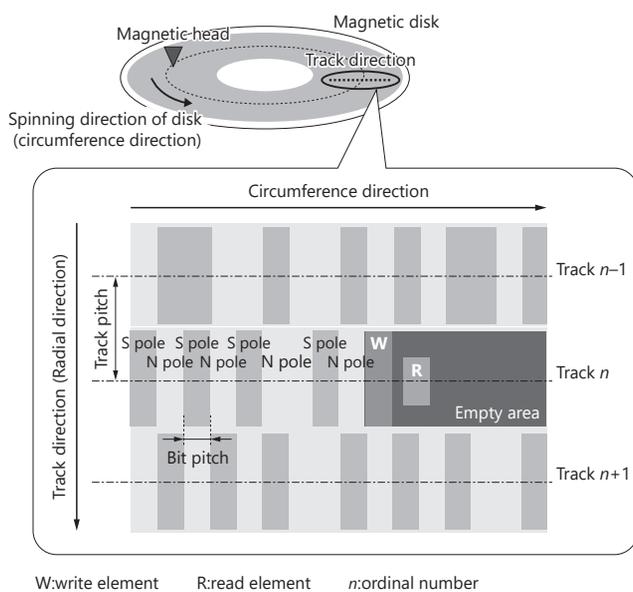
Figure 1. Outline of head and media of HDD
The write and read elements are positioned on the tracks of a spinning magnetic disk to write and read data.

Data are written on shingled tracks by write element.
Data are read by read element positioned at center of shingle-written tracks.

Figure 2. Schematic diagram of data tracks, write head, and read head of SMR HDD
The track pitch can be reduced by writing a new track, overwriting part of the previously written track analogous to the shingling of a roof.

N: ordinal number
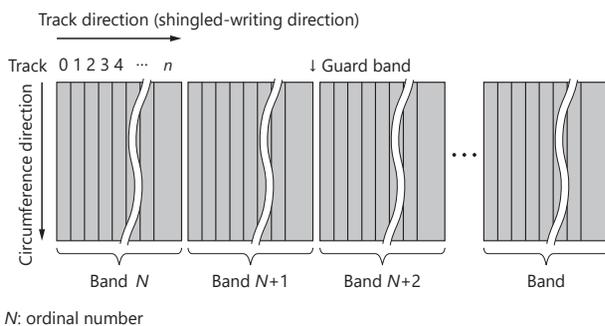
Figure 3. Configuration of shingled-written tracks and bands
The tracks that are sequentially shingle-written are stipulated in a specification or other document. Guard bands prevent the data on the first track of the adjacent band from being corrupted.

3. Specifications for the firmware of SMR HDDs for data centers

There are two types of firmware for SMR: drive-managed (DM) and host-managed (HM).

DM firmware is designed to execute the conventional HDD command set, making it possible to replace CMR HDDs with SMR HDDs. However, the random write performance of DM SMR HDDs varies depending on the host's access conditions. In contrast, HM firmware is designed to operate with HM commands. The host can optimize the write operations of an SMR HDD to achieve stable random write performance. It is necessary, however, to develop a filesystem etc. on the host side.

Data centers are considering the use of HM SMR HDDs. The command set for HM firmware has been standardized by the T10 and T13 Technical Committees of the International Committee for Information Technology (INCITS) that are responsible for the Serial Attached Small Computer System Interface (SAS) and the Serial Advanced Technology Attachment (SATA), respectively.

## 3.1 Random write operations and performance of DM SMR

Upon a write request from a host, a DM SMR HDD temporarily buffers the write data in a media cache (MC) provided on the magnetic disk. When the MC runs short of empty space, the buffered data are written to SMR bands. In the case of SMR, random write operations need to rewrite previously written data. Therefore, an SMR HDD reads all old data from an SMR band and replaces the buffered data in the MC with them, and sequentially writes them back to a new SMR band in a physically different location on a disk (**Figure 4**).

Since random write data are always buffered into the MC, DM SMR delivers high random write performance. However, it greatly depends on the access sequences from a host that are determined by data size and frequency. When random write requests occur frequently, the MC runs short of empty space even if a host writes data of small size, causing data to be moved from the MC to SMR bands frequently. This degrades random write performance (**Figure 5**).

Data to be rewritten
New write data
Previously written data
Empty area
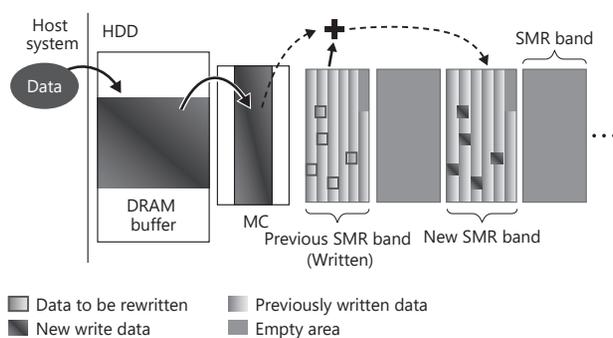
Figure 4. Data rewriting procedure in DM SMR HDD
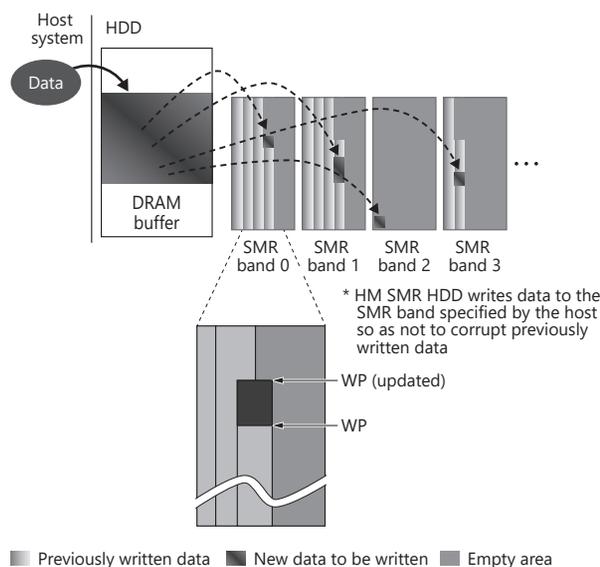The DM HDD uses an MC for random writes that need to rewrite the data in SMR bands.
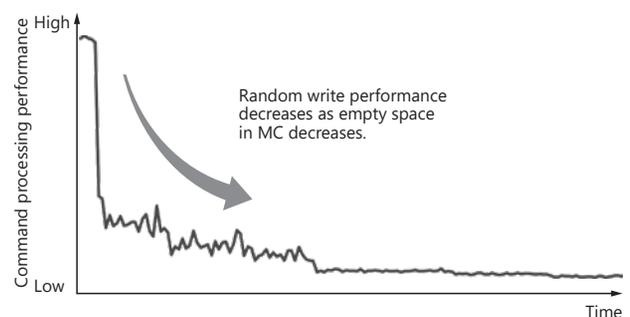
* HDDs are optimized according to their target applications to prevent sudden changes in performance so as to be able to maintain a certain performance level for a given period of time.

Figure 5. Example of random data writing performance of DM SMR HDD
The frequency of internal data movement from the MC to SMR bands affects random write performance.

3.2 Random write operations and performance of HM SMR

HM SMR directly writes data to SMR bands without using an MC. The write pointer (WP) points to the address within an SMR band at which the data can be written.

Each SMR band has a WP. HM SMR can write data only to sequential addresses in the spinning direction of a disk, starting with the address pointed to by a WP (**Figure 6**).

Since the SMR region of an HDD consists of a huge number of SMR bands, the SMR specification restricts the number of SMR bands that can be accessed at a time (maximum SMR band count) in order to simplify their management by the host. The host can use as many WPs as the maximum SMR band count to optimize HDD write operations so as to achieve the desired random write performance (**Figure 7**).

It is therefore necessary to develop a technique that allows the host to make the best use of HM SMR HDDs.

Figure 6. Data writing procedure in HM SMR HDD
The host writes data directly to SMR bands pointed to by WPs without using an MC.
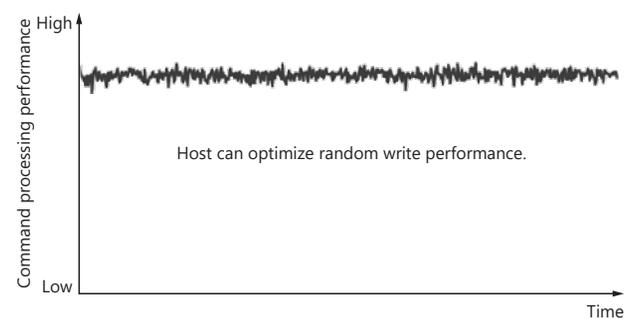
*The number of SMR bands that can be accessed at once (maximum SMR band count) is stipulated in a specification. (This figure shows an example of command processing performance when a write command is executed to write fixed-size data of 4,096 bytes to 128 SMR bands.)

Figure 7. Example of random data writing performance of HM SMR HDD
The host can optimize HDD write operations according to the HM firmware specification to obtain a desired random write performance.

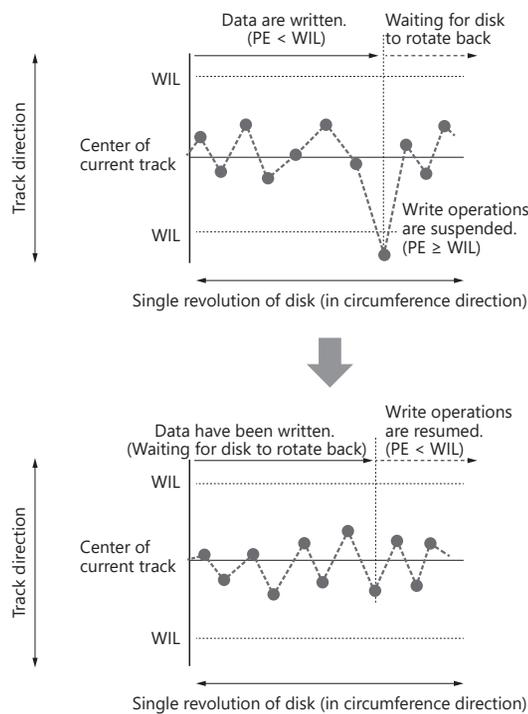## 4. Technology to ensure the data quality, taking advantage of the SMR characteristics

In the event that an HDD is exposed to rack and other external vibrations while the write head element is being positioned at the center of a track during a write operation, the magnetic head could move sideways, increasing its positioning error (PE). If the PE exceeds the write inhibit level (WIL), the magnetic head could go off track. In this case, the CMR HDD temporarily suspends the write operation to prevent data on an adjacent track from being overwritten and corrupted. The CMR HDD lets the disk keep spinning until the WIL and then resumes the write operation. In this way, the CMR HDD maintains the quality of the data recorded on adjacent tracks (**Figure 8**).

If the WIL is set too small, the CMR HDD will frequently suspend write operations, reducing write performance. In order to reduce the track pitch, the WIL should be set to a level that helps improve the write performance while ensuring the quality of the written data.

The CMR HDD has a fixed WIL whereas the SMR HDD, which

is characterized by sequential writes, uses dynamic write inhibit levels (DWILs), i.e., different WILs dynamically for the written and unwritten tracks. **Figure 9** illustrates DWILs.

During a sequential data write operation, the SMR HDD saves the PE of the write head on track Trk_{n-1} (PE_{n-1}). Then, when writing data to track Trk_n, the SMR HDD sets a dynamic WIL based on PE_{n-1} so that the read width limit for the data on Trk_{n-1} (i.e., the minimum width that allows a correct read-back of the written data) will be satisfied. Unlike the fixed WIL, the dynamic WIL becomes smaller when the previous track Trk_{n-1} comes close to the current track Trk_n and larger when Trk_{n-1} moves away from Trk.

Next, let's see when the SMR HDD suspends a write operation.

The write operation to the current track Trk_n is suspended when the positioning error PE_n increases toward the unwritten track Trk_{n+1} and exceeds the WIL. Conversely, if the PE increases toward the previously written track Trk_{n-1}, the write operation is suspended only when the PE exceeds the dynamic WIL.

This approach helps to reduce the suspending of write operations without affecting the quality of the data written on the previous track even when the track pitch is reduced[2].

*: PE at time of sampling regarding positioning control

*Firmware lets the disk keep spinning until PE drops below WIL. Write performance decreases if this occurs frequently.
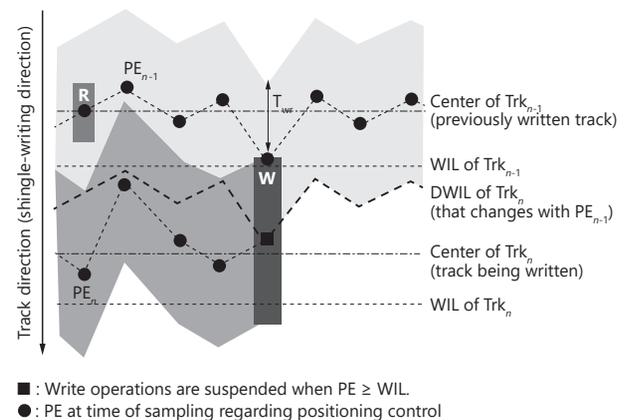
Figure 8. Relationship between head PEs and data writing operation
The firmware monitors the PE of the write head and suspends write operations when it exceeds the design threshold. The write operation restarts when the PE drops below the design threshold. This prevents the data on an adjacent track from getting corrupted.

■ : Write operations are suspended when PE ≥ WIL.
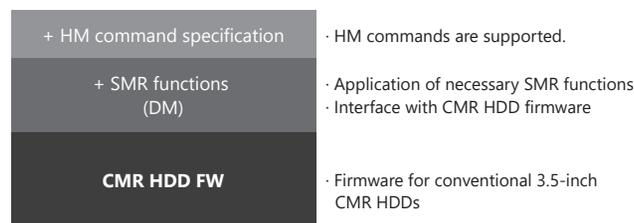● : PE at time of sampling regarding positioning control

Figure 9. Data writing of SMR HDD applying DWIL
The SMR HDD improves the write performance without affecting the quality of the data on adjacent tracks by dynamically changing the WIL.

5. Application of SMR technology to HDDs for data centers

To meet the requirements of data centers for HDDs, we have applied the SMR technology cultivated through the development of 2.5-inch DM SMR HDDs to the existing 3.5-inch CMR HDDs. We are also developing high-capacity HDDs incorporating HM firmware as well as the optimized magnetic heads and disks (**Figure 10**).

At present, we are working on the development of higher-capacity SMR HDDs for data centers. The eight- and nine-platter SMR HDDs will achieve a capacity of 15 TB (TB; 1 TB = 10^{12} bytes) and 17 TB, respectively, whereas the current eight- and nine-platter CMR HDDs provide a capacity of 12 TB and 14 TB, respectively.

+ HM command specification	· HM commands are supported.
+ SMR functions (DM)	· Application of necessary SMR functions · Interface with CMR HDD firmware
CMR HDD FW	· Firmware for conventional 3.5-inch CMR HDDs

Figure 10. Structure of firmware for SMR HDDs for data centers
To meet the requirements of data centers for HDDs, this SMR HDD incorporates SMR technology and HM firmware.

6. Future outlook for SMR HDDs

To combine the CMR advantage of excellent random access performance with the SMR advantage of high capacity, the HDD industry is now considering a new type of HDD that uses CMR to handle frequently accessed data and SMR to handle infrequently accessed data.

INCITS is working on the standardization of the command for the changeover between SMR and CMR areas. This command allows the host to change the ratio of the SMR and CMR areas on a magnetic disk.

We will apply this technology to the next-generation SMR HDDs.

7. Conclusion

This report outlined the SMR technology that is being employed to develop high-capacity SMR HDDs for data centers.

Nearly 10 years has passed since the concept of SMR appeared. SMR has become an important technology for increasing HDD capacity without relying only on the improvement of the major components of magnetic heads and disks.

The standardization of the commands for HM SMR HDDs is facilitating the innovations necessary for a host to make the best use of SMR HDDs. These innovations are expected to further increase the demand for SMR HDDs, mainly for data centers.

The next step is to develop magnetic heads, magnetic disks, signal processing circuits, and other components tailored to SMR in order to further increase the HDD capacity. Host systems that are highly compatible with SMR HDDs will also be developed to overcome the performance issue.

We will employ the ever-evolving SMR technology to increase the capacity of HDDs for data centers in order to contribute to the transformation of society that is being accelerated by cyber-physical systems (CPS).

References

(1) Shimomura, K. 2015. "Shingled Magnetic Recording Technologies for Large-Capacity Hard Disk Drives. " Toshiba Review: 70(8): 29–32.

(2) Toshiba Corporation. Kawabe, K.2012. "Disk recording device and disk control device and method. " Japanese Patent No. 5100861. December 19, 2012.