ARCHIVE (SMR) HARD DRIVE DATA RECOVERY

These days Web 2.0, Cloud-based Services, and Big Data has created different demand when it comes to data storage. As much as a hard drive still holds its shape and size, much has changed within. The demand for scalable, cost-effective mass storage continues to challenge the technological capabilities of the storage industry while the data recovery industry is trying to keep up.

Conventional vs. Shingled

First, there was conventional perpendicular magnetic recording or conventional magnetic recording (CMR) hard disk drives which had a guard space between the tracks to prevent it from overwriting adjacent tracks. In contrast to CMR drives, SMR drives do not have guard space. Instead, they overwrite adjacent tracks and then rewrite them. Shingled magnetic recording (SMR) technology provided an additional 16% increase in areal density compared to same-generation drives using CMR technology. Physically, on these drives, data is written sequentially, then overlapped therefore “shingled,” with another track of data.

The write head designed for SMR drives is wider than required for a single track of data, and it produces a stronger magnetic field. 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 (usually 8GB) 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. Because of the overlapping, all data streams must be organized and written sequentially to the media. When data is modified because the modified data is potentially overlapped with another data set, direct modification is not permitted, the entire row of shingles above the modified track needs to be rewritten in the process.

New micro actuators for new drives

The second generation of Dual-stage actuators (DSA) are underway as Western Digital announced their Micro Actuator (WDMA). Small multilayer piezos are now attached to the flexure, so when differential voltage is applied to the Micro Actuator, one piezo element expand while the other contracts. This action causes a slight rotational motion of the read-write head. Since the WDMA’s moving portion is so small with lighter mass compared with Milli Actuator, first-generation DSA, the WDMA element’s vibrational resonance frequency is much higher than that of the previous generation. As a result, the WDMA can rapidly and accurately position the head element over the correct data track. The better WDMA mechanical dynamics enlarges the DSA servo bandwidth and also improves the loop-shaping capability, meaning these new heads should be less prone to vibration.

The hybrid era – Archive Drives

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 sitting offline on removable media is migrating now to data kept online and accessible. This is how manufacturers realized that their workloads are trending toward data that is written sequentially and rarely updated, then read randomly and frequently - giving rise to a new generation of true archive drives - SMR hard drives.

There are three types of SMR drives: Drive Managed SMR (DM-SMR) allows plug-and-play deployment; Host Managed SMR (HM-SMR) requires host-software modification but delivers maximum benefit and competitive advantage and Host Aware SMR offers the convenience and flexibility of DM-SMR, with the performance and control advantages of HM-SMR.

Most data recovery laboratories are dealing with DM-SMR simply because $/GB ratio (better capacity for less money) so many of us already purchased one. WD Passports 4TB in 2.5" form factor are a good example of DM-SMR drives, notorious for its "second translator." In their white paper from June 2018, "Shingled Magnetic Recording +

HelioSeal® Technology," Western Digital noted these drives for their "unpredictable performance." Nevertheless, DM-SMR drives are selling and they are selling well.

The major downside of DM-SMR HDDs is that these need to manage all random writes to sequential shingled writes by means of media caching and an indirection table. 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, DM-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, therefore are suitable for applications that have idle time for the drive to perform background tasks such as moving the data around. The entire process is hidden from the host and it is similar to Flash Translation Layer found on SSDs including support for TRIM command, designed to maintain the performance of SSD or HDD at an optimal level over the lifetime of the drive. The TRIM command enables an Operating System to inform a data storage device in which blocks (bands) of data are no longer considered to be in use and can be reclaimed again to ensure that later write operations perform at full speed. It maintains the performance at an optimal level and reduces the wear and tear of the drive as it reclaims the occupied garbage space in advance, avoiding unnecessary read-modify-write operations.

However, this is not always the case; there is a known "bug" on WD drives likely caused when user unplugged a drive which was in the process of internal migration of data from cache to sequentially shingled bands. In this case, all user data may be lost for good, likely because parts of the translator have not been "closed" properly. Seagate drives have also known issues related to Media Cache which can cause data loss when Media Scratch Pad, Media Cache Metadata Table becomes corrupted. Introduction of temporary cache in some implementations is allocated to static parts of the surface while in the other is dynamically allocated in different areas of the disk ro reduce wear and tear.

Although this is not new technology and most of the consumers are clearly not aware, these drives are different, so they are buying them based on $/GB, not specs. Many reported that external 8TB Seagate Archive drive actually caused Windows to return errors due to drive write delays. While at the beginning, everyone pointed fingers at Windows or the USB bus when the actual culprit was one of these SMR drives being used for an active workload. Toshiba's latest 8TB and 10TB MG model generation utilize Persistent Write Cache technology so we can assume their behavior will likely do differ.

Conclusion

Much work has been put to reverse engineer these drives as data recovery service providers are flooded with SMR drives. A possible solution for WD's "second translator bug", Seagate's media cache freeze or other issues that will arise over the next couple of years may be in separate software and appropriate "loader" (LDR) that will be used when hard drive internal firmware is no longer functioning or corrupted. Although this approach may lead to a recovery of raw files, issues will most likely be caused by the so-called Soft Error Rate (SER) – when the read-back signal deteriorates to the point where correction algorithm improvements and ECC extensions no longer good enough. Similar to SER another effect caused by further squeezing track density, a gradual deterioration of signal integrity of adjacent tracks when writing repeatedly at high magnetic fields, is known as Adjacent Track Interference (ATI). Also, Far Track Interference is increasing the write field also results in stray fields that impact neighborhood tracks in the vicinity of the present write location. For most of these known effects there is a part in drive's firmware that handles it but as usual our work begins where manufacturers stopped theirs.