Series 8 RAID Adapter Family

White Paper

Optimizing SMR Drives in RAID Configurations

Microsemi
Power Matters.™

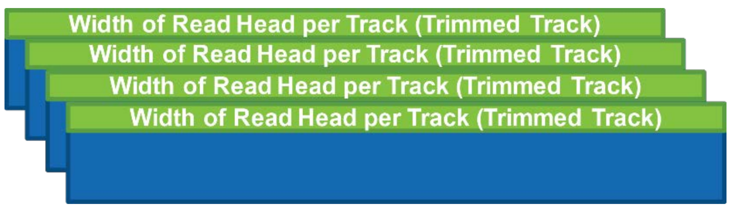
What are SMR Drives?

Shingled Magnetic Recording (SMR) increases hard disk drive capacity by writing overlapping tracks. In Conventional Magnetic Recording (CMR), the minimal width of write tracks for HDDs is limited by the size of the write head, which is constructed to write data reliably. Conventional HDDs do not overlap tracks.

Figure 1 Conventional Writes (CMR Drives and SMR Conventional Zone)

SMR drives take advantage of the fact that width of the read head does not need the full width of the written track to read data reliably.

Figure 2 Overlapped SMR Writes



SMR drives do overlap tracks during writes to increase data density, which results in higher usable capacity.

To allow random access or to store data in multiple sequential streams, each SMR drive has multiple bands with a write pointer, which allows for data to be appended to each band.

The pitfall of SMR drives is overwriting existing data, as shown in the following illustration. This is due to the fact that the write head's width touches multiple overlapping tracks. To avoid data loss, the entire band's data has to be read into a buffer, then the new data must be placed at the appropriate position in the buffer, and finally the whole band must be rewritten (read-modify-write cycle).

Figure 3 Overwriting Existing Data on an SMR Drive

Deleting data from a band leaves holes, therefore garbage collection has to take place from time to time to clean up bands in order to allow new data to be appended at the end of the band. This process is similar to flash-based SSDs, although it occurs for a different reason.

Overwriting data and garbage collection are very time-consuming processes on hard disk drives. Due to their mechanical nature, they have existing time-consuming output operations on random workloads that could increase the time required to perform cleanup operations.

SMR drives usually have a conventional zone similar to CMR drives that do not overlap tracks and can be used either as a cache or for metadata storage supporting random writes.

Why SMR Drives?

Why are SMR drives deployed despite the additional complexity that they introduce? The answer is that they increase the capacity of HDDs without increasing cost.

With the introduction of new storage tiers, such as flash-based SSDs or other non-volatile memory technologies, hard disk drives are being replaced by new technologies that offer higher performance in workflows. However, with the data growth predicted, there is still a large demand for cost-efficient, high-capacity storage devices.

Many usage models do not require frequent changes to older data. In fact, much of the data is written once and then only read from time to time if it gets older. Photos archived on an external USB hard disk drive or cloud storage are one such example. Newer photos are edited once before being printed, but older albums are left untouched, occasionally being looked through for revisiting fond memories.

External USB HDDs, as well as cloud storage for archival or backup purposes, are ideal use cases for SMR drives. Today, millions of SMR drives are used as USB backup drives, as they offer more capacity without increasing the price.

SMR Drive Variants

Different types of SMR drives are available for different use cases. Drive-managed (DM) SMR drives are currently the most popular type. Such drives use their conventional zone as a buffer and have an algorithm to internally manage their SMR nature. To the host, they function just like a normal hard disk drive as long as they are not exposed to many random writes over a long period of time. Too many random writes can eventually overload the drive’s conventional zone with data, slowing its response time to new requests until the buffer is cleaned up.

DM SMR drives can be used by any existing adapters, file systems, and applications that only require limited amounts of random write operations.

Host-managed (HM) SMR drives do not manage the bands themselves. Instead, they expect that all writes are sequential, and that the file system, adapter, and applications that use such drives support their specific nature through an extended command set available for both SATA and SAS as an extension to the existing command set for CMR drives and SSDs called ZAC/ZBC.

HM SMR drives are designed for applications that take control of the SMR nature of the drives in order to optimize certain workloads and applications. At the moment, applications and filesystems that can manage SMR drives are rare and typically proprietary. However, it is expected that more host-managed systems will be developed over time for use cases where the cost per gigabyte advantage of the SMR drives justifies the added complexity and additional development effort.

Host-aware (HA) SMR drives are a variant that supports both the SMR-specific ZAC/ZBC command sets and randomness in the write input and output operations. They are not as highly-optimized for applications that are not aware of the SMR nature of the drives as DM drives. They use some of their internal resources, such as cache memory and compute power, to allow both drive-managed and host-managed modes of operation.

These variants of SMR drives enable them to function in as many environments as possible, giving users a choice on how much development effort they are willing to spend in order to support their application.

SMR Drives in RAID Configurations

Redundant array of independent disks (RAID) aggregates multiple storage devices into one logical unit seen as a storage device by the host/operating system. The goal is to increase the data availability by adding redundancy and to scale the performance of those multiple storage devices. For this purpose, that is mirrored (RAID1) or striped (RAID0) and redundancy information is added for increased availability (RAID5, RAID6).

There are dedicated storage adapters, called RAID adapters, that manage RAID arrays with nonvolatile memory, and they can be used to further accelerate RAID by caching and offloading redundancy calculation from the host.

RAID configurations use a fixed pattern to determine how to distribute data among storage devices in a RAID array, and are therefore not suitable for HM SMR drives that require all write data to be sequential. RAID can be used with DM SMR drives and HA drives that support drive-managed mode.

Despite RAID being compatible with both SMR drives and CMR drives, mixing the two drive types within the same RAID array is not a good idea, as they have very different performance characteristics. As the saying goes, “The chain is only as strong as its weakest link;” likewise, the performance of a RAID array that mixes SMR and CMR drives would be similar to an SMR-only RAID array. Due to their additional complexity, SMR drives have limits in the number of IOPS they can provide and their consistent latency in response time to I/O requests when performing random write workloads. Incorporating SMR drives into RAID arrays does not change this fact.

In summary, SMR drives in RAID arrays have the same limitations as individual SMR drives. However, the RAID can help aggregate the performance of multiple SMR drives in the same way that it does for CMR drives. As a result, an overall higher level of performance can be achieved in workloads, while the RAID provides higher data availability.

In order to support SMR drives in RAID arrays, the RAID FW has to be tuned and optimized to adjust to the timeout behavior that comes with that device type. This is a task that goes along with a lot of testing and needs to stay within the boundaries that are provided by the host operating systems to avoid error messages and the host operating system to stall or even crash.

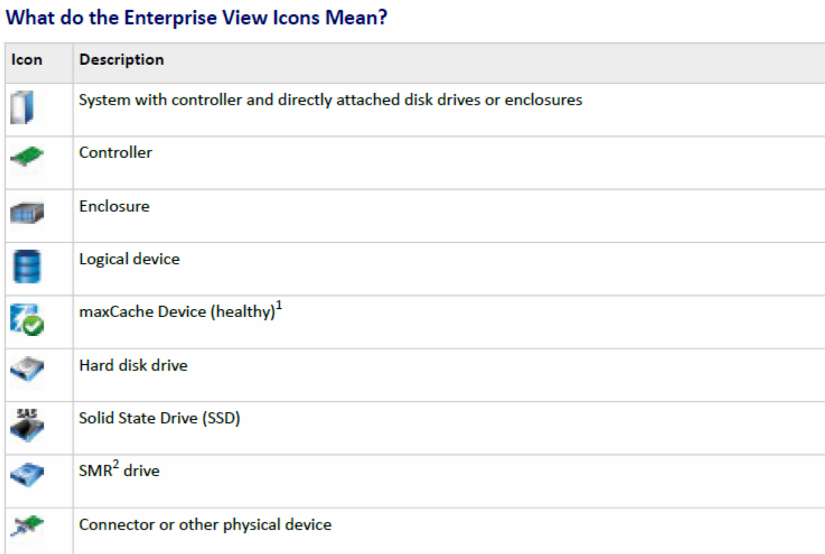
Of course, the RAID adapter, with its nonvolatile cache, can only do so much in terms of buffering commands and providing more relaxed timeouts. SMR drives in such RAID arrays should only be used with applications and in environments that require very limited random write performance. Otherwise, the disk and adapter’s buffers will quickly run over, starving the application on I/O requests that are not getting processed in a timely manner.

How Does This Work in the Real World?

It is clear from the previous examples that RAID adapters need some optimizations in order to work with suitable SMR drives (DM or HA in drive-managed mode). This optimization is available in the Microsemi Adaptec Series 8 RAID Adapter family of products.

In order to help customers to identify SMR drives that look and behave like CMR drives, the Microsemi Adaptec Series 8 RAID adapter checks the necessary identifiers of the device that report drive type. If it is identified as an SMR drive, the adapter will display it with a separate icon in order to alert the user that it is not a standard CMR hard disk drive. The following illustration shows the icons and describes their meanings.

Figure 4 Enterprise View Icons



What do the Enterprise View Icons Mean?

Icon	Description
	System with controller and directly attached disk drives or enclosures
	Controller
	Enclosure
	Logical device
	maxCache Device (healthy)[1]
	Hard disk drive
	Solid State Drive (SSD)
	SMR[2] drive
	Connector or other physical device

These notifications are designed to prevent accidental mixing of CMR and SMR drives within the same RAID or logical volume. The adapter will display a warning message if a user tries to create such RAID array of mixed devices, but does not actively prevent the use of such an array in case a customer is willing to compromise performance for cost. Microsemi does not recommend this practice.

Equipped with this information, the user can create SMR-only logical volumes and RAID arrays to aggregate capacity and increase performance, capacity, and data availability (within the established limits of SMR technology). The same RAID levels that are supported with CMR drives are open to SMR devices with similar device counts per logical volumes and RAID arrays.

Even OS boot is supported by modern SMR drives, as it mostly involves only reads and limited random writes. However, the user is responsible for verifying that the SMR drive will work in his or her specific setup. If the SMR drive is exposed to heavy random write access, dirty data will build up on the drive’s random write zone and eventually become unresponsive. This scenario can only be avoided by the user ensuring that the drive is only used in an SMR-friendly environment.

Conclusion

Shingled magnetic recording increases the capacity of hard disk drives by writing on overlapping tracks, something that conventional HDDs cannot do. RAID configurations use a fixed pattern to distribute data amongst those storage devices in a RAID array, making them unsuitable for HM SMR drives that require all write data to be sequential. However, RAID can be used with DM SMR drives and HA SMR drives that support drive-managed mode. SMR drives in RAID arrays have the same limits as SMR drives used as individual drives. Despite this, RAID can help aggregate the performance of multiple SMR drives in the same way that it does with CMR drives, therefore achieving higher overall performance levels and data availability in workloads.

For maximum performance and reliability, RAID adapters must be optimized to work with suitable SMR drives (DM or HA in drive-managed mode). Microsemi’s Adaptec Series 8 RAID adapter family of products can provide this optimization.

Microsemi Corporate Headquarters
One Enterprise, Aliso Viejo,
CA 92656 USA
Within the USA: +1 (800) 713-4113
Outside the USA: +1 (949) 380-6100
Fax: +1 (949) 215-4996
Email: **sales.support@microsemi.com**
www.microsemi.com

About Microsemi
Microsemi Corporation (Nasdaq: MSCC) offers a comprehensive portfolio of semiconductor and system solutions for aerospace & defense, communications, data center and industrial markets. Products include high-performance and radiation-hardened analog mixed-signal integrated circuits, FPGAs, SoCs and ASICs; power management products; timing and synchronization devices and precise time solutions, setting the world's standard for time; voice processing devices; RF solutions; discrete components; enterprise storage and communication solutions; security technologies and scalable anti-tamper products; Ethernet solutions; Power-over-Ethernet ICs and midspans; as well as custom design capabilities and services. Microsemi is headquartered in Aliso Viejo, California, and has approximately 4,800 employees globally. Learn more at **www.microsemi.com**.

ESC-2170006