# Dynamic Hybrid Shingled Magnetic Recording (DHSMR) Concept Description

**Date: 1/19/2018**

**Revision 0.1**

William Boyle

Western Digital Corporation
5601 Great Oaks Parkway
San Jose, CA 95119 USA

408.717.6000
William.B.Boyle@wdc.com

# Table Of Contents

# 1 Scope of Document

This document describes the Dynamic Hybrid Shingled Magnetic Recording (DHSMR) concept. This includes the high-level functional behavior of the hard disk drive (HDD) and the interface descriptions. The intention of this document is to provide the reader with a general understanding of this new HDD device type.

This document assumes the reader is familiar with existing ZAC and ZBC zone block device interfaces.

# 2 Glossary

The following terms are used in this document. This table can be referred to for explanations of acronyms used.

| | |
|---|---|
| 1. DHSMR | Dynamic Hybrid Shingled Magnetic Recording |
| 2. HDD | Hard Disk Drive |
| 3. CMR | Conventional Magnetic Recording |
| 4. SMR | Shingled Magnetic Recording |
| 5. ZAC | Zoned Device ATA Commands |
| 6. ZBC | Zoned Block Commands |
| 7. GPT | GUID Partition Table |
| 8. TCG | Trusted Computing Group |
| 9. LBA | Logical Block Address |
| 10. SWP | Sequential Write Pointer |
| 11. HM | Host Managed SMR system |
| 12. Seam | Boundary between a CMR Realm and an SMR Realm. |

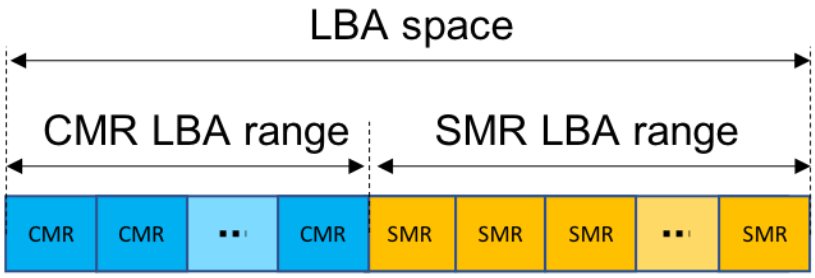
# 3 Goals for DHSMR

The goal of DHSMR is to enable host systems to dynamically configure portions of the HDD between CMR and SMR recording modes. This will allow these systems to gradually convert their access patterns to be compatible with SMR. CMR will generally store frequently written application data while SMR services colder data in a more cost-effective manner.

# 4 LBA Addressing System

The LBA system will enable separate addressing domains for CMR and SMR. The CMR domain will start at LBA 0 and the SMR domain will start after the CMR domain. It is preferable for the SMR domain to start immediately after the CMR domain to reduce the number of permanent 'Offline' zones that reside between the CMR and SMR domains. The ZAC/ZBC standard requires that all LBAs within the domain of the drive be represented in the zone table. The specific address ranges are described in the Realms Table shown later in this document.

The following diagram describes the LBA Addressing System. Each box represents a recording view (CMR or SMR) into a Realm which is described in more detail in the next section.

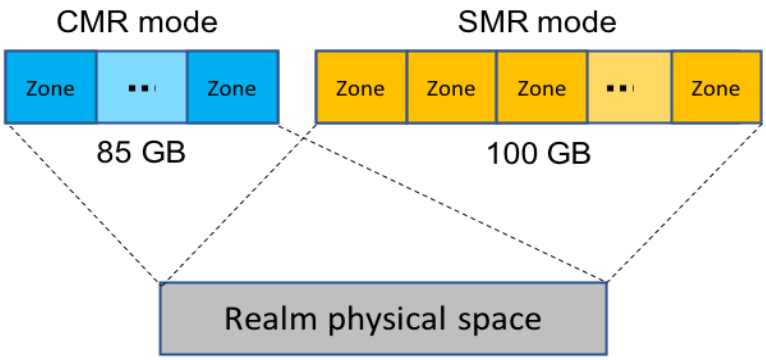



# 5 Realms

## 5.1 Definition

A Realm is the unit of converted between CMR and SMR recording modes. The concept provides a simple mechanism for conversions between CMR and SMR recording modes. The associations between the CMR and SMR LBA ranges in Realms are defined by the drive vendor and are fixed, providing clear boundaries for the host system to convert between recording modes. Conversion of a Realm does not affect the data contained within other Realms. The conversion process invalidates all data contained within a Realm. Realms may be locked to avoid accidental conversion that would result in data loss. Realms are locked through a vendor factory configuration process and reflected in the configurable 'Convertible' flag bit shown in the Realms Table described later in this document. Convertible Realms can currently be switched between SMR and CMR recording modes. This proposal does not preclude additional recording modes for Realms in the future. Non-Convertible Realms will be set in the factory and will return an error condition if a conversion is attempted. Typically, the Non-Convertible Realms are in CMR mode and enable backward compatibility of the HDD in legacy systems. The typical case would have the first Realm in CMR mode and be configured as Non-Convertible.

Realms in SMR mode can have varying capacities but the typical configuration will have a common SMR capacity across the Realms. Each Realm in CMR mode will likely have different capacities. The following diagram shows an example of a Realm in the two different recording modes.




## 5.2 Sizes

### 5.2.1 Realms

The Realm capacity in SMR mode must be an integer multiple of an SMR Zone (see below). Realm size is set in the factory and is not changeable by the customer. A Realm capacity in CMR mode will likely vary because the conversion ratio varies across the stroke and from drive to drive. It will also be an integer multiple of a CMR Zone.

### 5.2.2 SMR Zones

SMR zones are as described in the ZAC/ZBC specifications and are only relevant when the Realm is in SMR mode. In industry, a typical SMR zone size is 256MiB.

## 5.3 Read Capacity

The drive shall report its capacity based on a configuration when all Realms are in CMR mode. The reported capacity shall not change when Realms are converted to SMR mode. The actual dynamic capacity (combination of Realms currently configured in CMR mode plus Realms currently configured in SMR mode) can be derived by reading the Report Realms Table data (see below) or the report zones data.

## 5.4 CMR Zones

CMR Realms are sub-divided into CMR Zones. The CMR zone size will typically match that of the SMR zone size. All zone sizes being equal and 256 MiB is required to be backward compatible with existing Linux systems.

## 5.5 CMR SWP

Since the recording mode conversion process is destructive, reading of newly activated CMR LBAs would not result in readable sectors. SMR zones have an inherent mechanism to protect the host from trying to read sectors that have not been initialized through a host write. A Sequential Write Pointer (SWP) is used to validate read commands for initialized sectors. A similar mechanism is used here for newly converted CMR zones.

Each CMR Zone shall have an SWP associated with it. The drive has the option of using this pointer to validate write commands on a newly converted zone for sequential initialization. The SWP can also be used to validate read commands for access to previously written data only.

### 5.5.1 CMR Realm Write Commands Validation

By default, the command checking will be disabled in order to allow random write capability within the CMR zone and preserve backward compatibility with the ZBC/ZAC defined behavior of conventional zones. The host system can enable the checking of write commands relative to the CMR SWP in the zone where the LBA resides. When enabled, only write commands that have their start LBA at or before the SWP are accepted for execution. Write commands that have their start LBA after the SWP shall return an aborted status to the host system. Accepted commands that have the range of LBAs in the command inclusive of the SWP will advance the SWP to the end of the LBA range upon successful completion of the command. Valid Write commands that span across CMR Zones shall be accepted and upon command completion the SWP of the first Zone shall no longer be needed and the SWP of the second Zone shall advance to the end of the command.

### 5.5.2 CMR/SMR Read Commands Validation

By default, the read command checking will be disabled in order to allow read capability within the conventional zones and preserve backward compatibility with the ZBC/ZAC defined behavior of conventional zones. The host system can enable the checking of read commands relative to the SWP for the CMR Zone where the LBAs of the command reside. If this checking is enabled, then only read commands that have the complete range of LBAs in the command addressing sectors prior to the SWP will be accepted for execution. Read commands that attempt to address LBAs at or past the SWP shall return an aborted status to the host system. Read commands can straddle CMR Zones but all the LBAs of the command must have been written prior to the read command as determined by the SWP of each of the Zones servicing the command. SMR zones will continue to use the URSWRZ bit to determine the availability of reading from unwritten sectors.

# 6 Conversion of Realms

## 6.1 Conditions for Conversion

### 6.1.1 SMR

A Realm in SMR mode is allowed to be converted to CMR mode when it is configured as a Convertible Realm.

There is an optional check by the drive for all SMR zones within the SMR Realm to have their Sequential Write Pointers (SWP) reset. This condition is consistent with the SMR Zones being in the 'Empty' state. This check by the drive can be enabled by the host. If enabled and any zone within the Realm is not 'Empty', then the conversion process will abort with an error.

### 6.1.2 CMR

A Realm in CMR mode is allowed to be converted to SMR mode if it is configured as Convertible.

## 6.2 Conversion Process

New ZAC and ZBC commands will be developed to enable the conversion of Realms between SMR and CMR recording modes. The command will return successful status if the conditions for the conversion have been met. If any condition has not been met, then the conversion process will abort with error status. The maximum time to convert a Realm shall be reasonable, in the range of 100mS – 200mS. A conversion command that describes multiple Realms to convert shall have a maximum time equal to the sum of the conversion times of each of the Realms.

## 6.3 State of Realm After Conversion

### 6.3.1 SMR

The state of the Realm after is has been converted to SMR recording mode will have the SWP for all SMR Zones reset and will be in the 'Empty' state. None of the data that was valid prior to conversion in the Realm is accessible.

### 6.3.2 CMR

The state of the Realm after it has been converted to CMR recording mode has indeterminate sector data. The CMR SWP will be reset back to the beginning of the Realm's LBA range and the condition of the zones within the Realm will be in "CMR Write Pointer" state. The write pointer field in the zone table will describe the current condition of the CMR write pointer.

# 7 Realm Behavioral Requirements

## 7.1 SMR Recording Mode

SMR Zones within an SMR Realm shall conform to the HM ZAC/ZBC standard. The host can enable the drive to ensure all of a read command's LBAs occur prior to the zone's SWP. The drive can be configured to abort commands with LBAs that address sectors at or past the SWP.

The drive will abort any media access commands to LBAs in Realms that are not currently active. This applies to both inactive CMR and SMR LBA ranges.

## 7.2 CMR Recording Mode

CMR Realms shall behave as legacy CMR HDD drives after the sectors have been initialized through host writes. Random writes are allowed under the restrictions described in section 5.6.1 above.

### 7.2.1 Initialization After Conversion

The state of the data in sectors of a Realm that have been converted to CMR mode is indeterminate. Read commands to those sectors will not be able to recover data. Therefore, initialization of the CMR Realm must be performed. The primary method of sector initialization is through write commands from the host system. The drive may augment some of the initialization through the second mode shown below. There are two host-selectable initialization modes.

1. Command Checking
    a. Drive will abort any write command that does not start before or at the SWP. This mode facilitates checking of host software for bugs if the intention is for the host to sequentially initialize the Realm.
    b. Drive will abort any read command that contains an LBA that addresses sectors at or past the SWP. This facilitates checking of host software for bugs if the intention is that no sector will be read prior to being written by the host system.
2. No Command Checking
    a. Drive will accept write commands anywhere within the Realm. The drive may have to pad some portion of the media with a predefined data pattern for writes that do not occur at or before the SWP. This mode may incur extra command execution time for writing more data sectors on the media than was contained in the write command. These additional data sectors are part of capacity of the CMR Zone and are initialized with a pre-defined data pattern.
    b. Drive will return a predefined data pattern for any sectors that have not been initialized with a previous write command. This mode does not inform the host system when reads occur at LBAs that have not been previously written by the host.

### 7.3 Unexpected Power Loss Handling

The drive shall power up in an operational state after an unexpected power loss while the drive is performing any operation. This includes the conversion of Realms. The drive shall report the state of all Realms through the Report Realms ext command (see below) after a power up. If a drive was in the process of converting a Realm when a power loss occurred, then the state of the Realm after a subsequent power up shall be in a valid, stable, and static condition. This may require the HDD to finish the conversion process prior to posting Ready status. This same process may be required when a Reset command is issued to the drive.

# 8 Dependencies on Other Standards

## 8.1 Security

### 8.1.1 TCG Implications

Only Realms in the global range can be converted between SMR and CMR. If any user range is configured in a Realm, then any conversion attempt of that Realm will be rejected. The range must be released back into the global range before the Realm can be converted.

## 8.2 ZAC/ZBC

This proposal requires minimal modifications to the ZAC/ZBC standard if command checking is disabled and/or an initialization process for Realms in CMR mode is introduced into the conversion process on the host. Backward compatibility to existing standards is an important consideration for quicker adoption of this new device type in the industry.

### 8.2.1 Realm Conversion Command

There shall be a new ZAC and ZBC command defined that allows the host system to perform conversion operations on Realms to switch between CMR and SMR recording modes. This conversion command shall conform to the restrictions described above for SMR and CMR Conditions for Conversion. Only Realms with the Convertible indication in the Realms Table Log (described below) shall be convertible with this command.

### 8.2.2 Report Zones Ext Command

The CMR Zones in the table will use the SWP field to show the sequential write point for each CMR Zone while in “CMR Write Pointer” condition. All LBAs shall be accounted for in the zone map from the beginning of the CMR domain through the end of the SMR domain. If there are any zones not represented in any of the Realms then those zones will be marked as in an “Offline” condition. Realms that are currently configured in a CMR mode shall have the SMR zones in that Realm marked as in an inactive state in the zone condition field. This new inactive state will be a new condition for zones defined to support this proposal. This applies equivalently for the CMR zones when the Realm is configured in SMR mode.

CMR zones that have not been completely initialized with all the sectors written shall be in a “CMR Write Pointer” Condition in the Zone Descriptor. Once all the sectors have been initialized in a zone through writing to the last sector of the zone then the Zone Condition shall report the normal “NOT WRITE POINTER” condition for a conventional zone.

The “Reporting Options” field shall add the option to report the “CMR Write Pointer” condition to accommodate systems to retrieve a list of CMR zones needing initialization.

### 8.2.3 Report Realms Ext Command

There shall be a new command that returns the current state and restrictions of all the Realms. The following table lists the minimum set of items that must reside in this table:

1. Header information
2. Per Realm Information
    a. Current mode (SMR or CMR)
    b. CMR mode starting LBA
    c. CMR mode ending LBA
    d. SMR mode starting LBA
    e. SMR mode ending LBA
    f. Convertible flag(s)

# 9 Disk Enumeration Requirements

The HDD shall be configured to work in legacy systems that are designed to expect a CMR drive. This requires the identification of the peripheral device type to be a standard HDD (standard legacy signature). A new zone block device type will be defined to identify this HDD as a DHSMR device.

# 10 Special Case Realms

There are some special configurations of Realms that may be configured by the drive vendor to enable host system needs. Here are some examples

## 10.1 Redundant GPT

Some BIOSs use the GUID Partition Table (GPT) partitioning system. This requires the last 32KB of the capacity of the drive in CMR mode to support random writes. This will be accomplished by providing the last Realm to be a Non-Convertible CMR Realm. Since all Realms must contain an integer number of zones, then this last Realm will contain a single CMR zone.

## 10.2 First Realm

To make the drive backward compatible with legacy BIOS systems, the first Realm (starting at LBA 0) shall be in CMR mode and marked as Non-Convertible. This ensures the legacy BIOSs will be able to boot the system prior to any upper level device drivers being loaded.

# 11 **Factory Configuration Examples**

## 11.1 Backward Compatibility with Existing Legacy Host Systems

To support existing legacy host systems that do not have SMR aware software, the drive vendor can configure all the Realms in CMR mode as the default factory configuration. All the Realms shall have all the sectors initialized to allow for reads to any LBA consistent with previous CMR products.

# 12 **Recording Modes Configuration Examples**