Hybrid-SMR Product Requirements Proposal for OCP

Theodore Ts'o

November 13, 2017

Introduction

This document describes proposed requirements for a Hybrid SMR drive where portions of the storage media can be dynamically converted between Conventional Magnetic Recording (CMR) and Shingled Magnetic Recording (SMR).

Must be usable as 100% CMR drive by Legacy Software

In the factory-default state, the HDD must be 100% backwards compatible with traditional CMR HDD's. For this reason, the number of accessible sectors from IDENTIFY DEVICE shall return the lesser size between: 100% CMR disk size or the size from SET ACCESSIBLE MAX ADDRESS EXT. In addition, GET NATIVE MAX ADDRESS EXT should return the 100% CMR disk size. The CMR capacity should conform to the industry specification SFF-8447 on HDD sizes.

Short-stroking compatibility

As with current CMR HDD's available on the market today, larger LBA sectors should roughly correspond with locations closer to the inner diameter (ID). This should be true for both LBA sector numbers in the CMR and SMR space; the smallest CMR or SMR sector numbers should correspond to physical sectors at the outer diameter, while the largest CMR or SMR sector numbers should correspond to physical sectors at the inner diameter. Replacement sectors in response to grown defects are an exception to this rule.

The reason for this requirement is so that the "hottest" CMR data can be located at the OD. Since SMR data tends to be cold, the SMR region will be located at the ID. (See below for more details about the CMR->SMR conversion.)

Support for other advanced HDD features

The Hybrid SMR drive must support other advanced HDD features, including Head Depop, TCG/Opal Storage Specification, ATA Security, GPL, NCQ, SMART, Sense Data Reporting, Write-Read-Verify, etc. Interactions with the SMR zones will be based on the ZAC specification, potentially with modifications as required to support the Hybrid SMR feature.

CMR->SMR conversion

The hybrid SMR drive must be able to support CMR->SMR conversion where the size of the SMR region created from the CMR region is a multiple of 100 GiB (400 SMR zones); conversions at a smaller granularity may be supported. That is, the hybrid SMR drive must be able to support the creation of 100 GiB volumes, starting at the ID of the disk (with room for the GPT partition table in CMR space). Each successive 100 GiB SMR volume will be adjacent to the previous volume, located closer to the OD of the disk, and further away from the ID.

The hybrid SMR drive must have an interface by which the Host software can determine what range of CMR sectors must be taken off-line in order to bring each new 100 GiB SMR region on-line. This is to allow the Host software to copy out any data which must be preserved before asking the drive to trigger the CMR->SMR conversion process. No other data stored on the HDD, other than the CMR sectors being taken off-line, should be affected by this process.

There must be a way that the Host OS can determine which SMR sectors can not be brought on-line in order to protect a range of CMR sectors from being brought off-line. One use of this functionality would be to assure that the GPT partition table or a small file system located at the OD of the disk will not be brought off-line.

As each 100 GiB SMR volume is brought on-line, starting at the ID and growing “down” in the LBA number space, this shall result in a contiguous range of SMR sectors being brought on-line and a contiguous region of CMR sectors being brought off-line. The region of on-line CMR sectors after the conversion shall remain contiguous.

The CMR<->SMR conversion command must be available as a non-queueable command.

SMR->CMR conversion

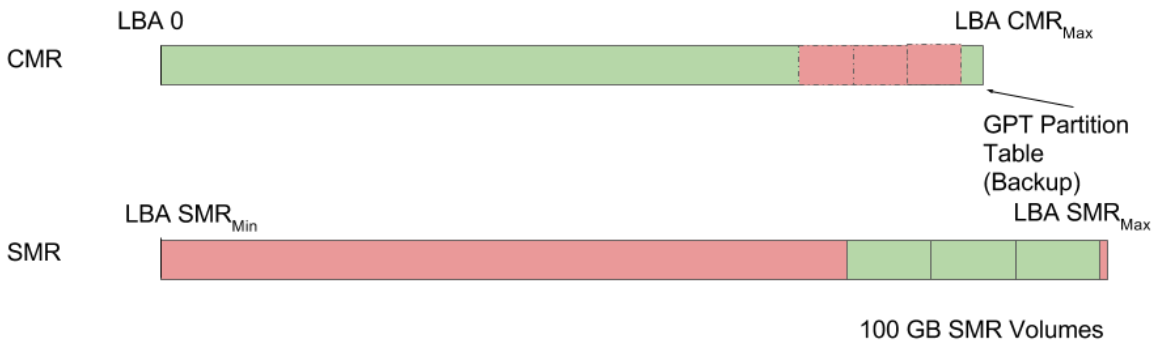
The hybrid SMR drive must be able to support converting a 100 GiB SMR volume back to CMR. It is sufficient that the drive only supports conversion of the 100 GiB SMR volume closest to the OD of the drive before other SMR volumes can be converted back to CMR space. In other words, SMR volumes will be created (by conversion from CMR space) from the ID to the OD, and they will be released and converted back to CMR space from the OD to the ID. There shall be no restrictions on the number of times that a portion of the recording media can be converted between CMR and SMR and vice-versa.

CMR / SMR sector addressing

The hybrid SMR drive should export the CMR and SMR storage as a set of CMR-capable zones and SMR zones. As described above, the CMR-capable zones must be addressed using low-numbered sectors. The SMR zones will start at a starting LBA number, SMR_{MIN}, where $SMR_{MIN} > CMR_{MAX}$. In both the CMR and SMR LBA number space, lower number sectors shall represent sectors closer to the outer diameter (OD) while higher number sectors shall represent sectors closer to the inner diameter (ID) of the disk platters.

When disk space is converted from CMR to SMR, some number of CMR-capable zones shall be taken off-line, and some number of SMR zones will be enabled. Similarly, when disk space is

converted from SMR back to CMR, this will be done by disabling SMR zones and enabling CMR zones.



Performance Requirements

The time to create a 100 GiB SMR volume should be no more than two hundred milliseconds (ms); with typical conversion time under 50 ms. The time to convert the space containing the 100 GiB SMR volume back to CMR should be similarly quick. The time to query the hybrid SMR drive to determine which CMR sectors should be vacated in order to make the 100 GiB SMR volume come on line should be no more than 50 ms.

The performance of the drive in CMR mode should be similar to that of the same-generation CMR-only HDD's.

Conversion Atomicity

A CMR->SMR or SMR->CMR conversion operation must be atomic. That is, if the conversion operation is interrupted for any reason (power failure, bus reset, etc.) the state of the hybrid SMR drive must either be unchanged, or the conversion operation must be fully completed by the time the drive is ready to accept commands from the host.

Grown Defect handling

When a defect is discovered in a portion of the recording media while writing to a portion of the disk in one mode (CMR or SMR) and it is remapped to spare portion of the disk, that fact that a particular region of the recording media has a defect must be remapped when it is converted to the other mode (e.g., SMR or CMR). The host software should not need to discover the media defect again after the recording mode conversion. In addition, the host software should have visibility to all of the sectors that were reallocated (in both CMR and SMR modes).