

Технический Обзор
Решения с использованием SMR Дисков и Smart
Storage HBA и RAID контроллеров

Октябрь 2017

Содержание

1	Что такое SMR Диски и Для Чего Они Используются?.....	1
2	SMR Диски в Конфигурациях с HBA или RAID Контроллером	3
3	Как Это Работает в Реальном Мире?	4
4	Заключение	6

1 Что Такое SMR Диски и Для Чего Они Используются?

Черепичная магнитная запись (SMR) - это технология, которая увеличивает емкость жестких дисков (HDD) с помощью записи перекрывающихся дорожек. Для обычных жестких дисков используется традиционная магнитная запись (CMR), что и отличает их от SMR дисков, при этом минимальная ширина дорожек при записи для таких дисков определяется размером записывающей головки, которая сконструирована так, чтобы производить запись надежным образом. Традиционные диски не имеют перекрывающихся дорожек.

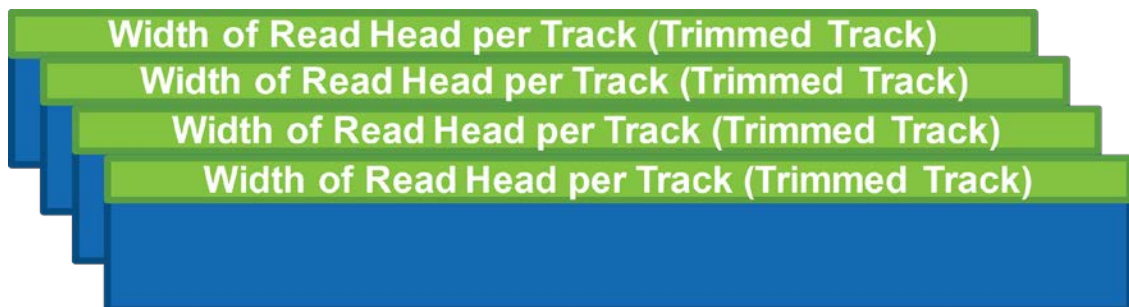
Рисунок 1. Традиционная запись (CMR диски и SMR зона традиционной записи [Conventional Zone]).



SMR диски получают преимущество от того факта, что головка для чтения HDD диска имеет гораздо меньшую ширину, чем головка для записи, чтобы читать данные надежно.

Этот рисунок показывает, как SMR диски осуществляют перекрытие дорожек во время записи, чтобы увеличить плотность размещения данных, что приводит к увеличению полезной емкости диска. Все работает хорошо, пока данные добавляются в конце записи. Если данные должны быть добавлены где-то в середине, для записи придется перезаписать много дорожек для выполнения этой операции.

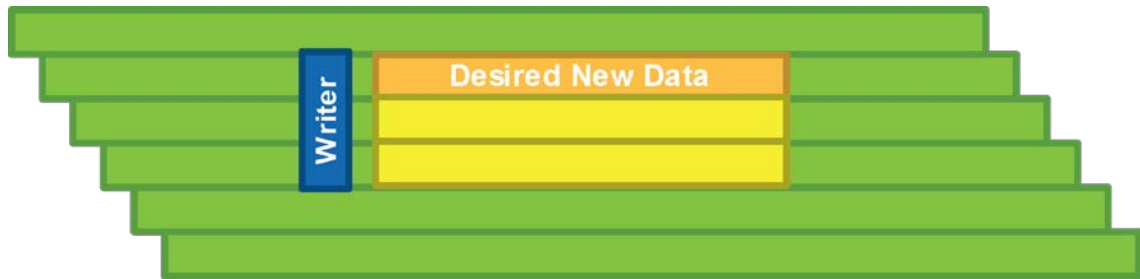
Рисунок 2. Перекрывающиеся SMR области записи.



Чтобы позволить случайные операции записи или для записи нескольких отдельных потоков данных, например, при записи отдельных видео потоков или при одновременном создании нескольких резервных копий данных, каждый SMR диск имеет несколько SMR областей (bands), иногда называемых зонами (zones), для ведения записи, что дает возможность добавлять данные в каждую такую область.

Недостаток SMR дисков проявляется при необходимости перезаписи существующих данных, как показано на следующей картинке. Происходит это благодаря тому, что записывающая головка в силу своей ширины производит запись сразу в несколько перекрывающихся дорожек. Во избежание потери данных все данные из SMR области должны быть помещены в буфер, затем новые данные должны быть размещены в нужных позициях в буфере и, на заключительном этапе, вся область должна быть перезаписана (такой цикл называется “прочитать – модифицировать – записать”). Это означает, что при наличии случайных операций записи с маленьким размером блока требуемый размер данных для записи может вырасти до размеров всей SMR области (band или zone).

Рисунок 3. Перезапись существующих данных на SMR диске.



Кроме того, удаление данных в SMR области оставляет пустые места, поэтому время от времени должна быть запущена процедура очистки для данного диска, чтобы позволить новым данным быть добавленными к самому концу в SMR области. Этот процесс идентичен процессу во flash-based SSD дисках, хотя требуется он по другим причинам.

Перезапись данных и очистка диска - процессы на жестком диске, требующие времени. В силу своей механической природы, которая подразумевает перемещение головок, это занимает миллисекунды. Этот недостаток делает SMR диски даже еще менее подходящими для операций случайной записи, чем CMR диски, которые значительно менее производительные по сравнению с устройствами, использующими флэш память, такими как SSD диски.

SMR диски обычно обладают областью (conventional zone) с традиционной магнитной записью, где нет перекрытия дорожек и эта область может быть использована как буфер для операций случайной записи чтобы улучшить время выполнения или как область метаданных (служебная область), где хранятся записи об удаленных данных или указатели для позиции записи различных SMR областей.

Почему SMR диски используются, несмотря на дополнительные сложности, которыми они обладают? Ответ – они увеличивают емкость жесткого диска, не увеличивая его стоимость. По сравнению с таким же CMR диском, полезная емкость диска с SMR технологией на 15-20% выше.

С появлением новых технологий для систем хранения, таких как SSD, использующих технологию флэш памяти, или других видов энергонезависимой памяти (NVRAM), жесткие диски все реже используются для потоков данных от приложений, работающих с транзакциями. Однако, при предсказуемом росте данных остается очень большой спрос на устройства высокой емкости с оптимально низкой стоимостью, оптимизированных по критерию доллар за Гигабайт, нежели чем на величину случайной нагрузки.

Данные классифицируются как “горячие” или “холодные”. “Горячие” данные – это данные, к которым очень часто обращаются и которые также часто могут изменяться, и они размещаются на устройствах хранения, использующих флэш технологии или в оперативной памяти. К “теплым” данным обращаются время от времени, но их изменения маловероятны, а “холодные” данные заархивированы и крайне редко изменяются. Они обычно записываются один раз и время от времени прочитываются. Один из примеров “холодных” данных – архивы фотоальбомов в облачном хранилище. Новые фото обычно редактируются перед печатью, а старые альбомы остаются неизменными, они только просматриваются время от времени, чтобы освежить приятные воспоминания.

Внешние USB жесткие диски, так же как облачные хранилища для архивного или резервного хранения, являются идеальными случаями применения SMR дисков. На момент написания данной статьи, миллионы SMR дисков уже были реализованы как USB диски для резервного копирования, так как они предлагают больше емкости без увеличения цены. В будущем SMR диски будут использоваться для архивного и резервного хранения. Различные типы SMR дисков были разработаны для удовлетворения различных требований к облачным дискам и внешним USB дискам для резервного копирования, и эти варианты и различия между ними будут рассмотрены нами в следующей части.

2 SMR Диски в HBA или RAID Конфигурациях

Device-managed (DM) SMR диски могут использоваться с любыми стандартными HBA, которые поддерживают SMR диски. Эти диски обычно используются как внешние USB устройства для резервного копирования для десктопных компьютеров или ноутбуков.

Идеальный HBA для host-aware (HA) SMR дисков - это такой HBA, который поддерживает ZBC/ZAC наборы команд. Поддержка этих наборов команд является обязательным для любого host-managed (HM) SMR диска. HA диски часто используются для приложений, которые получают прямой доступ к диску. На сегодняшний день, компании производители операционных и файловых систем разрабатывают решения для HM дисков. Требования для поддержки HM дисков отличаются от подобных требований для других типов дисков из-за того, что HM диски не поддерживают какие-либо шаблоны случайного трафика. Например, большинство современных операционных систем используют много потоков, исполняющихся на разных ядрах серверного процессора и много командных очередей, чтобы рационально использовать ресурсы оборудования и оптимизировать операции Ввода-Вывода для высокоуровневых систем хранения данных. При такой архитектуре не все I/O команды обрабатываются последовательно и этот процесс не поддерживается HM SMR дисками.

Пример выше иллюстрирует, почему SMR решения с большой вероятностью являются частными решениями и специфичными для некоторых приложений в гиперконвергентных центрах обработки данных. Такие центры обработки данных находятся среди первых, кто активно использует HM SMR технологию, поскольку они размещают огромное количество систем хранения, и требуют изолированного окружения, где они могут контролировать все программное обеспечение и выполняющиеся приложения.

Механизм Избыточного Массива Независимых Дисков (RAID) агрегирует множество устройств хранения в один логический диск, который виден как виртуальное устройство хранения в хост/операционной системе. Цель - увеличить доступность данных, добавляя избыточность и масштабируя производительность таких многодисковых систем. С этой целью производится зеркалирование (RAID1), чередование (RAID0) или добавление избыточной информации с целью увеличения надежности (RAID5, RAID6).

Специальные контроллеры для устройств хранения, которые называются RAID контроллерами, управляют RAID томами с энергонезависимой памятью и могут быть использованы для дальнейшего ускорения расчета функций RAID с помощью разгрузки расчетов избыточности со стороны хост системы.

RAID конфигурации используют специальные шаблоны, чтобы определить, как распределять данные по устройствам хранения в RAID томах, и по этой причине нет возможности использовать HM SMR диски, которые требуют, чтобы все записываемые данные были последовательные (нет поддержки для записи в определенном месте). RAID может использоваться с DM SMR дисками и HA дисками, которые поддерживают DM режим.

Несмотря на то, что RAID контроллеры совместимы и с SMR дисками и с CMR дисками, смешивание двух типов дисков в одном RAID томе не будет хорошей идеей, поскольку они имеют совершенно разные характеристики. Как гласит поговорка: "Прочность цепи определяется прочностью ее самого слабого звена", - таким же образом производительность RAID тома, где используются и SMR и CMR диски, будет похожа на RAID том, где используются только SMR диски. Из-за своей повышенной сложности SMR диски имеют ограничения в количестве IOPS, которые они могут обеспечить, и они обладают неустранимой задержкой при обработке запросов ввода-вывода при доминировании случайной записи. Объединение SMR дисков в RAID тома не отменяет эти недостатки.

В целом, SMR диски в RAID томах обладают теми же ограничениями, что и отдельные SMR диски. Однако, через механизмы RAID можно объединить производительность нескольких SMR дисков тем же самым образом, как и для CMR дисков. В результате, может быть достигнут общий, более высокий уровень производительности для рабочих нагрузок, при этом RAID обеспечит более высокую доступность данных.

Для поддержки SMR дисков в RAID массивах программное обеспечение RAID стека и энергонезависимый кэш могут помочь буферизировать данные. Однако, это не может изменить характеристики SMR диска. Это означает, что RAID контроллер может только улучшать ситуацию, ограничиваясь только буферизацией команд и обеспечивая более лучшие таймауты. SMR диски в таких RAID томах должны быть использованы только с приложениями и в средах, где требуется очень ограниченная производительность для случайных операций записи. Это правило также применимо

и для отдельных SMR дисков, подключенных к HBA. Иначе буфер диска и контроллера очень быстро переполнится, при этом приложение будет страдать от запросов ввода-вывода, которые не были обслужены за требуемое время.










Если использовать RAID тома таким неправильным образом, время, требующееся для обработки запросов ввода-вывода в RAID тома, состоящие из DM SMR HDD дисков, может значительно превысить таймауты операционной системы и таймауты приложений, вызывая зависания и даже случаи потери информации. Это же применимо и к отдельным SMR устройствам. Использование DM/HA SMR дисков в RAID томах таким образом разрешается только в таких условиях, где уровень операций ввода-вывода для случайной записи низкий и контролируемый.

3 Как Это Работает В Реальном Мире?

И RAID контроллеры и HBA нуждаются в оптимизации для того, чтобы работать с подходящими SMR дисками. Поддержка SMR дисков для рекомендованных рабочих режимов, рассмотренных выше, доступна для следующих продуктов: Microsemi Adaptec® HBA 1100 Series, SmartHBA 2100 Series и SmartRAID 3100 Series.

С целью помочь пользователям идентифицировать DM и HA SMR диски, которые выглядят и ведут себя как обычные CMR диски, Microsemi Smart Storage контроллеры проверяют необходимые идентификаторы диска, которые отвечают за тип диска. Если диск определяется как SMR диск, контроллер покажет его с помощью отдельной графической иконки, которая предупреждает пользователя, что это не стандартный CMR HDD диск. Следующая таблица показывает эти графические иконки и дает описание их значений.

Рисунок 4. Графические элементы в Enterprise View области maxView Storage Manager.

Иконка	Описание
	Система с контроллером и напрямую подключенными дисками или бэкплейном
	Контроллер
	Бэкплейн
	Логический диск
	maxCache Диск (SSD кэш)
	Жесткий HDD диск
	SSD диск
	SMR диск
	Коннектор или другое физическое устройство

Эти графические значки были разработаны, чтобы показать типы устройств и предупредить случаи смешиваний CMR и SMR дисков в одном RAID или логическом томе. Контроллер покажет предупреждение (warning message), если пользователь попытается создать такой RAID том из смешанных устройств и контроллер будет препятствовать созданию такого RAID тома или логического устройства (LD). HBAs также использует такие значки, чтобы показать, что данный диск это HA SMR диск или HM SMR диск.

Обладая этой информацией, пользователь может создавать только SMR логические и RAID тома, чтобы увеличить емкость, производительность и доступность данных (в пределах ограничений, накладываемых SMR технологией). Те же самые уровни RAID, которые поддерживаются для обычных CMR дисков, доступны и для SMR дисков с теми же самыми ограничениями по количеству устройств для RAID или логического тома.

Современными SMR дисками поддерживается даже возможность загрузки операционной системы, поскольку этот процесс подразумевает большей частью чтение и крайне ограниченное количество операций случайной записи. Однако, существуют сценарии где SMR диски могут оказывать негативное влияние на производительность, и, как следствие, такие случаи использования не рекомендуются. Если SMR устройство нагрузить большим потоком операций случайной записи, необслуженные запросы, превысив размер стандартной зоны для записи, сделают доступ к диску практически нулевым. Такого сценария можно избежать, только если пользователь обеспечит работу в дружелюбном для SMR устройств окружении.



Power Matters.™

Использование NM SMR дисков и HA SMR дисков в режиме HA ограничено режимом raw устройств (HBA режим). Этот режим поддерживается всеми Microsemi Smart Storage контроллерами, включая HBA 1100, SmartHBA 2100 и SmartRAID 3100 продукты. Компания Microsemi уже обладает валидированными моделями различных HA SMR HDD дисков для данных контроллеров и работает с производителями операционных систем и приложений для реализации полной поддержки NM устройств, которые в настоящий момент не поддерживаются из-за их неспособности обслуживания любых операций случайной записи.

Если Вам нужно больше информации о совместимости SMR дисков и их поддержке, пожалуйста, или обратитесь к представителю Microsemi по продажам или посетите наш вебсайт www.microsemi.com.

4 Заключение

Черепичная Магнитная Запись увеличивает емкость жестких дисков с помощью записи перекрывающихся дорожек – это то, что обычный диск с традиционной магнитной записью делать не может. RAID конфигурации используют фиксированные шаблоны для распределения данных по устройствам хранения в RAID томах, что приводит к невозможности использования HM SMR дисков, требующих, чтобы все записываемые данные были последовательными. Однако, RAID тома могут быть использованы с DM SMR дисками и HA SMR дисками в drive-managed режиме. SMR диски в RAID томах имеют такие же ограничения, как и одиночные SMR диски. Несмотря на это, RAID конфигурации могут помочь увеличить производительность нескольких SMR дисков тем же способом, как это делается с обычными CMR дисками, таким образом, получая более высокие уровни производительности и доступности для рабочих нагрузок.

Как только продукты Microsemi Adaptec HBA 1100, SmartHBA 2100 и SmartRAID 3100 Series станут доступными, контроллеры для систем хранения данных от компании Microsemi будут поддерживать нужные командные расширения для HM SMR или HA SMR дисков: Zone Block наборы команд или Zoned-Device ATA наборы команд (ZBC/ZAC). HM SMR диски и HA SMR диски, работающие в HA режиме должны быть использованы только как raw устройства и обладать поддержкой со стороны операционной системы, файловой системы и приложения для того, чтобы правильно работать.

Если Вам нужно больше информации по продуктам компании Microsemi, посетите www.microsemi.com.



Штаб-квартира Корпорации Microsemi
 One Enterprise, Aliso Viejo, CA 92656 USA
 Для звонков внутри США: +1 (800) 713-4113
 Для звонков из других стран: +1 (949) 380-6100
 Для клиентов в России +1 (949) 215-4996
 email: sales.support@microsemi.com
 storage.microsemi.com/ru-ru/

©2017 Microsemi Corporation. Все права защищены. Microsemi и логотип Microsemi являются зарегистрированными торговыми марками корпорации Microsemi. Все другие упомянутые торговые марки и сервисы являются собственностью соответствующих владельцев.

Корпорация Microsemi (Nasdaq: MSCC) предлагает значительный набор полупроводниковых и системных решений в области коммуникаций, безопасности и обороны, а также для воздушно-космических систем и для различных производственных областей. Линейка продуктов включает в себя высокопроизводительные и защищенные от радиации аналоговые и аналого-цифровые интегральные микросхемы, FPGAs, SoCs и ASICs, системы управления питанием, компоненты систем сверхточного времени и синхронизации, устанавливающих международные стандарты точности, устройства обработки голосовой информации, системы для работы с высокочастотными и дискретными сигналами, системы хранения и обмена информацией корпоративного класса, технологии защиты информации и продукты для предотвращения несанкционированного доступа к информации, Ethernet решения, микросхемы и продукты Power-over-Ethernet, а также сервисы и возможности для создания продуктов и систем по заказам клиентов. Штаб-квартира корпорации Microsemi расположена в Алисо-Вьехо в Калифорнии. Штаб-квартира компании, примерно, 4800 человек по всему миру. Более подробную информацию можно получить на Web сайте www.microsemi.com.

Компания Microsemi не принимает никаких рекламаций и не предоставляет никаких гарантийных или иных обязательств относительно информации, содержащейся в данном документе, или соответствия данных продуктов и сервисов для каких-либо практических целей. Также, компания Microsemi не несет никакой ответственности за ситуации, возникающие при использовании данных приложений, продуктов или их подсистем. Любые упомянутые продукты или какие-либо другие продукты, проданные компанией Microsemi, подвергаются ограниченному тестированию и в силу этого не могут быть использованы совместно с критически важными приложениями или критически важным оборудованием. Любые предоставленные спецификации по производительности считаются надежными, но они не подвергаются широкой и детальной проверке. Покупатели должны проводить тестирование производительности и проверять другие характеристики продуктов изолированно и совместно с другими конечными продуктами или методом установки в другие конечные продукты. Покупатели не должны основываться на данных и спецификациях или параметрах, предоставляемых компанией Microsemi. Покупатель несет ответственность за независимую оценку применимости продуктов, а также за тестирование и верификацию этого. Информация, предоставляемая компанией Microsemi в данном документе, предоставляется по принципу "как есть", и все возможные неисправности и риски, связанные использованием данной информации, целиком оцениваются покупателем и целиком в его ответственности. Компания Microsemi не предоставляет, явно или не явно, никому никаких прав на использование патентов, лицензий или каких-либо других прав на интеллектуальную собственность, независимо по отношению к данной информации или к чему-либо, связанному с данной информацией. Информация, предоставляемая в данном документе, является собственностью компании Microsemi, компания Microsemi оставляет за собой право вносить какие-либо изменения в данный документ или в какие-либо продукты и сервисы в любое время без предварительного уведомления.

ESC-2171613