

Технологии магнитной записи HDD: просто о сложном

Western Digital :: 21.10.2020



Первый в мире жесткий диск, IBM RAMAC 305, увидевший свет в 1956 году, вмещал лишь 5 МБ данных, а весил при этом 970 кг и по габаритам был сопоставим с промышленным рефрижератором. Современные корпоративные флагманы способны похвастаться емкостью уже в 20 ТБ. Только представьте себе: 64 года назад, для того чтобы записать такое количество информации, потребовалось бы свыше 4 миллионов RAMAC 305, а размеры ЦОДа, необходимого для их размещения, превысили бы 9 квадратных километров, тогда как сегодня для этого будет достаточно маленькой коробочки весом около 700 грамм! Во многом добиться столь невероятного повышения плотности хранения удалось благодаря совершенствованию методов магнитной записи.

В это сложно поверить, однако принципиально конструкция жестких дисков не меняется вот уже почти 40 лет, начиная с 1983 года: именно тогда свет увидел первый 3,5-дюймовый винчестер RO351, разработанный шотландской компанией Rodime. Этот малыш получил две магнитные пластины по 10 МБ каждая, то есть был способен вместить вдвое больше данных, чем обновленный ST-412 на 5,25 дюйма, выпущенный Seagate в том же году для персональных компьютеров IBM 5160.



Rodime RO351 — первый в мире 3,5-дюймовый винчестер

Несмотря на инновационность и компактные размеры, на момент выхода RO351 оказался практически никому не нужен, а все дальнейшие попытки Rodime закрепиться на рынке винчестеров потерпели фиаско, из-за чего в 1991 году компания была вынуждена прекратить свою деятельность, распродав практически все имеющиеся активы и сократив штат до минимума. Однако стать банкротом Rodime оказалось не суждено: в скором времени к ней начали обращаться крупнейшие производители винчестеров, желающие приобрести лицензию на использование запатентованного шотландцами форм-фактора. В настоящее время 3,5 дюйма является общепринятым стандартом производства как потребительских HDD, так и накопителей корпоративного класса.

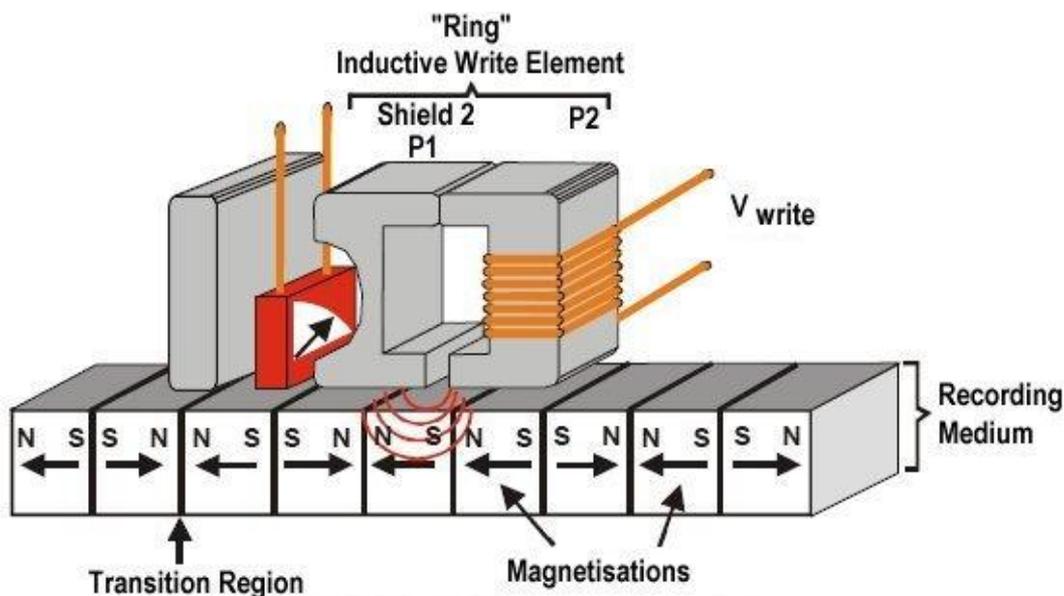
С появлением нейросетей, Deep Learning и интернета вещей (IoT) объем создаваемых человечеством данных стал лавинообразно расти. По оценкам аналитического агентства IDC, к 2025 году количество информации, генерируемой как самими людьми, так и окружающими нас девайсами, достигнет 175 зеттабайт (1 Збайт = 10^{21} байт), и это при том, что в 2019-м таковое составляло 45 Збайт, в 2016-м — 16 Збайт, а в далеком 2006-м общий объем данных, произведенных за всю обозримую историю, не

превышал 0,16 (!) Збайт. Справиться с информационным взрывом помогают современные технологии, среди которых не последнее место занимают усовершенствованные методы записи данных.

LMR, PMR, CMR и TDMR: в чем разница?

Принцип работы жестких дисков достаточно прост. Тонкие металлические пластины, покрытые слоем ферромагнитного материала (кристаллического вещества, способного сохранять намагниченность даже при отсутствии воздействия на него внешнего магнитного поля при температуре ниже точки Кюри) движутся относительно блока пишущих головок на большой скорости (5400 оборотов в минуту или более). При подаче электрического тока на пишущую головку возникает переменное магнитное поле, которое изменяет направление вектора намагниченности доменов (дискретных областей вещества) ферромагнетика. Считывание данных происходит либо за счет явления электромагнитной индукции (перемещение доменов относительно сенсора вызывает в последнем возникновение переменного электрического тока), либо за счет гигантского магниторезистивного эффекта (под действием магнитного поля изменяется электрическое сопротивление датчика), как это реализовано в современных накопителях. Каждый домен кодирует один бит информации, принимая логическое значение «0» или «1» в зависимости от направления вектора намагниченности.

Долгое время жесткие диски использовали метод продольной магнитной записи (Longitudinal Magnetic Recording, LMR), при котором вектор намагниченности доменов лежал в плоскости магнитной пластины. Несмотря на относительную простоту реализации, данная технология имела существенный недостаток: для того чтобы побороть коэрцитивность (переход магнитных частиц в однодоменное состояние), между треками приходилось оставлять внушительную буферную зону (так называемое guard space — защитное пространство). Вследствие этого максимальная плотность записи, которой удалось добиться на закате данной технологии, составляла всего 150 Гбит/дюйм².



В 2010 году LMR была практически полностью вытеснена PMR (Perpendicular Magnetic Recording — перпендикулярная магнитная запись). Главное отличие данной технологии от продольной магнитной записи состоит в том, что вектор магнитной направленности каждого домена располагается под углом

90° к поверхности магнитной пластины, что позволило существенно сократить промежуток между треками.

За счет этого плотность записи данных удалось заметно увеличить (до 1 Тбит/дюйм² в современных устройствах), при этом не жертвуя скоростными характеристиками и надежностью винчестеров. В настоящее время перпендикулярная магнитная запись является доминирующей на рынке, в связи с чем ее также часто называют CMR (Conventional Magnetic Recording — обычная магнитная запись). При этом надо понимать, что между PMR и CMR нет ровным счетом никакой разницы — это всего лишь другой вариант названия.

Изучая технические характеристики современных жестких дисков, вы также можете наткнуться на загадочную аббревиатуру TDMR. В частности, данную технологию используют накопители корпоративного класса [Western Digital Ultrastar 500-й серии](#). С точки зрения физики TDMR (что расшифровывается как Two Dimensional Magnetic Recording — двумерная магнитная запись) ничем не отличается от привычной нам PMR: как и прежде, мы имеем дело с непересекающимися треками, домены в которых ориентированы перпендикулярно плоскости магнитных пластин. Разница между технологиями заключается в подходе к считыванию информации.

В блоке магнитных головок винчестеров, созданных по технологии TDMR, на каждую пишущую головку приходится по два считывающих сенсора, осуществляющих одновременное чтение данных с каждого пройденного трека. Такая избыточность дает возможность контроллеру HDD эффективно фильтровать электромагнитные шумы, появление которых обусловлено межтрековой интерференцией (Intertrack Interference, ITI).



Решение проблемы с ITI обеспечивает два чрезвычайно важных преимущества:

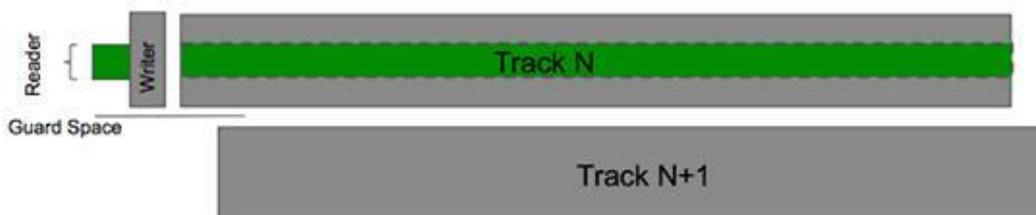
1. снижение коэффициента помех позволяет повысить плотность записи за счет уменьшения расстояния между треками, обеспечивая выигрыш по общей емкости вплоть до 10% по сравнению с обычной PMR;
2. в сочетании с технологией RVS и трехпозиционным микроактуатором, TDMR позволяет эффективно противостоять ротационной вибрации, вызванной работой винчестеров, что помогает добиться стабильного уровня производительности даже в наиболее сложных условиях эксплуатации.

Что такое SMR и с чем его едят?

Размеры пишущей головки примерно в 1,7 раза больше по сравнению с размерами считывающего сенсора. Столь внушительная разница объясняется достаточно просто: если записывающий модуль сделать еще более миниатюрным, силы магнитного поля, которое он сможет генерировать, окажется недостаточно для намагничивания доменов ферромагнитного слоя, а значит, данные попросту не будут сохраняться. В случае со считывающим сенсором такой проблемы не возникает. Более того: его миниатюризация позволяет дополнительно снизить влияние упомянутой выше ITI на процесс считывания информации.

Данный факт лег в основу черепичной магнитной записи (Shingled Magnetic Recording, SMR). Давайте разбираться, как это работает. При использовании традиционного PMR пишущая головка смещается относительно каждого предыдущего трека на расстояние, равное ее ширине + ширина защитного пространства (guard space).

Conventional Writes



При использовании черепичного метода магнитной записи пишущая головка смещается вперед лишь на часть своей ширины, поэтому каждый предыдущий трек оказывается частично перезаписан последующим: магнитные дорожки накладываются друг на друга подобно кровельной черепице. Такой подход позволяет дополнительно повысить плотность записи, обеспечивая выигрыш по емкости до 10%, при этом не отражаясь на процессе чтения. В качестве примера можно привести [Western Digital Ultrastar DC HC 650](#) — первые в мире 3.5-дюймовые накопители объемом 20 ТБ с интерфейсом SATA/SAS, появление которых стало возможным именно благодаря новой технологии магнитной записи. Таким образом, переход на SMR-диски позволяет повысить плотность хранения данных в тех же стойках при минимальных затратах на модернизацию IT-инфраструктуры.

Несмотря на столь значительное преимущество, SMR имеет и очевидный недостаток. Поскольку магнитные дорожки накладываются друг на друга, при обновлении данных потребуются перезапись не только требуемого фрагмента, но и всех последующих треков в пределах магнитной пластины, объем которой может превышать 2 терабайта, что чревато серьезным падением производительности.

Решить данную проблему помогает объединение определенного количества треков в обособленные группы, называемые зонами. Хотя такой подход к организации хранения данных несколько снижает общую емкость HDD (поскольку между зонами необходимо сохранять достаточные промежутки, препятствующие перезаписи треков из соседних групп), это позволяет существенно ускорить процесс обновления данных, так как теперь в нем участвует лишь ограниченное количество дорожек.

Черепичная магнитная запись предполагает несколько вариантов реализации:

- **Drive Managed SMR (SMR, управляемая диском)**

Основным ее преимуществом является отсутствие необходимости в модификации программного и/или аппаратного обеспечения хоста, поскольку управление процедурой записи данных берет на себя контроллер HDD. Такие диски могут быть подключены к любой системе, в которой присутствует необходимый интерфейс (SATA или SAS), после чего накопитель будет сразу готов к работе.

Недостаток этого подхода заключается в изменчивости уровня производительности, в связи с чем Drive Managed SMR оказывается неподходящей для корпоративных приложений, в которых постоянство быстродействия системы является критически важным параметром. Тем не менее такие диски хорошо показывают себя в сценариях, предоставляющих достаточное время для выполнения фоновой дефрагментации данных. Так, например, DMSMR-накопители [WD Red](#), оптимизированные для использования в составе малых NAS на 8 отсеков, станут отличным выбором для системы архивирования или резервного копирования, предполагающей долговременное хранение бэкапов.



- **Host Managed SMR (SMR, управляемая хостом)**

Host Managed SMR — наиболее предпочтительный вариант реализации черепичной записи для использования в корпоративной среде. В данном случае за управление потоками данных и операциями чтения/записи отвечает сама хост-система, задействующая для этих целей расширения интерфейсов ATA (Zoned Device ATA Command Set, ZAC) и SCSI (Zoned Block Commands, ZBC), разработанные комитетами INCITS T10 и T13.

При использовании HMSMR весь доступный объем накопителя разделяется на зоны двух типов: Conventional Zones (обычные зоны), которые используются для хранения метаданных и произвольной записи (по сути, играют роль кэша), и Sequential Write Required Zones (зоны последовательной записи), занимающие большую часть общей емкости жесткого диска, в которых данные записываются строго последовательно. Неупорядоченные данные сохраняются в области кэширования, откуда затем могут быть перенесены в соответствующую зону последовательной записи. Благодаря этому все физические сектора записываются последовательно в радиальном направлении и перезаписываются только после циклического переноса, что позволяет добиться стабильной и предсказуемой производительности системы. При этом HMSMR-диски поддерживают команды произвольного чтения аналогично накопителям, использующим стандартный PMR.

Host Managed SMR реализована в жестких дисках enterprise-класса [Western Digital Ultrastar HC DC 600-й серии](#).



Линейка включает в себя SATA- и SAS-накопители высокой емкости, ориентированные на использование в составе гипермасштабных центров обработки данных. Поддержка Host Managed SMR существенно расширяет сферу применения таких винчестеров: помимо систем резервного копирования, они прекрасно подойдут для облачных хранилищ, CDN или стриминговых платформ. Высокая емкость жестких дисков позволяет существенно повысить плотность хранения (в тех же стойках) при минимальных затратах на апгрейд, а низкое энергопотребление (не более 0,29 Ватта на каждый терабайт сохраненной информации) и тепловыделение (в среднем на 5 °С ниже, чем у аналогов) — дополнительно сократить операционные расходы на обслуживание ЦОДа.

Единственным недостатком HMSMR является сравнительная сложность имплементации. Все дело в том, что на сегодняшний день ни одна операционная система или приложение не умеют работать с подобными накопителями «из коробки», в силу чего для адаптации ИТ-инфраструктуры требуются серьезные изменения стека программного обеспечения. В первую очередь это касается, конечно же, самой ОС, что в условиях современных ЦОД, использующих многоядерные и многосокетные сервера, является достаточно нетривиальной задачей. Узнать подробнее о вариантах реализации поддержки

Host Managed SMR можно на специализированном ресурсе [ZonedStorage.io](https://zonedstorage.io), посвященном вопросам зонального хранения данных. Собранные здесь сведения помогут предварительно оценить степень готовности вашей IT-инфраструктуры для перевода на зональные системы хранения.

- **Host Aware SMR (SMR, поддерживаемая хостом)**

Устройства с поддержкой Host Aware SMR сочетают в себе удобство и гибкость Drive Managed SMR и высокую скорость записи Host Managed SMR. Такие накопители обратно совместимы с устаревшими системами хранения и могут функционировать без непосредственного контроля со стороны хоста, однако в этом случае, как и при работе с DMSMR-дисками, их производительность становится непредсказуемой.

Подобно Host Managed SMR, Host Aware SMR использует два типа зон: Conventional Zones для произвольной записи и Sequential Write Preferred Zones (зоны, предпочтительные для последовательной записи). Последние, в отличие от упомянутых выше Sequential Write Required Zones, автоматически переводятся в разряд обычных в том случае, если в них начинает вестись неупорядоченная запись данных.

Реализация SMR с поддержкой хоста предусматривает внутренние механизмы восстановления после непоследовательной записи. Неупорядоченные данные записываются в области кэширования, откуда диск может переносить информацию в зону последовательной записи, после того как будут получены все необходимые блоки. Для управления неупорядоченной записью и фоновой дефрагментацией диск использует таблицу косвенного обращения. Однако, если корпоративным приложениям требуется предсказуемая и оптимизированная производительность, достичь этого по-прежнему можно лишь в случае, когда хост берет на себя полное управление всеми потоками данных и зонами записи.