

Интерфейсы и периферийные устройства

Раздел 2. Устройства хранения данных

Лекция 6-1. Характеристики жестких дисков

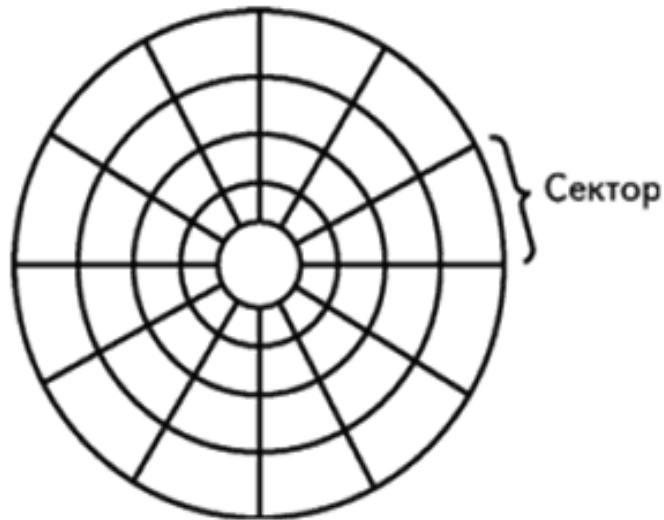
Логическая структура

Тракты чтения-записи

Методы повышения плотности записи

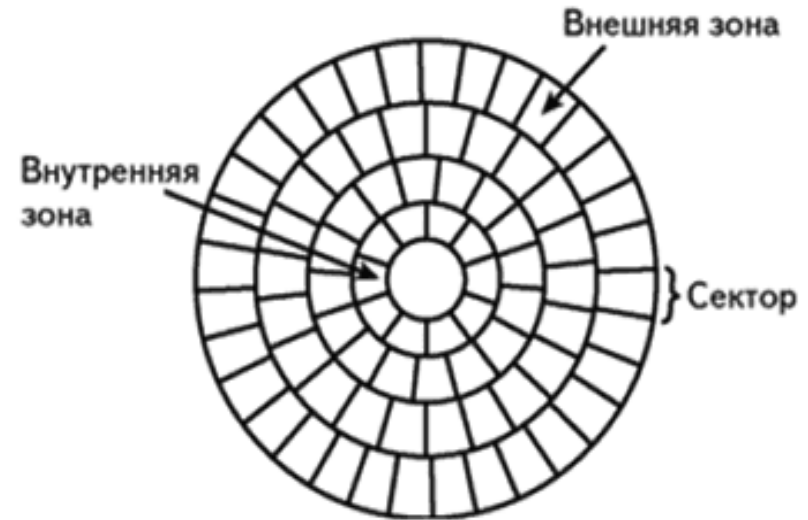
Классификация HDD

Логическая структура магнитного носителя



Стандартный способ форматирования

Фиксированное форматирование, когда каждая дорожка делилась на одинаковое количество секторов с одинаковыми угловыми размерами. Однако при этом линейные размеры секторов на разных дорожках были неодинаковыми (сокращались к центру диска).



Зональное форматирование

Поверхность разбивается на зоны (20-30), в рамках каждой дорожка делится на определенное количество секторов. Линейные размеры варьируются только для секторов одной зоны, ширина которой невелика.

Сектора

Номера секторов и их статус обычно хранятся на диске в специальной таблице в служебной зоне.

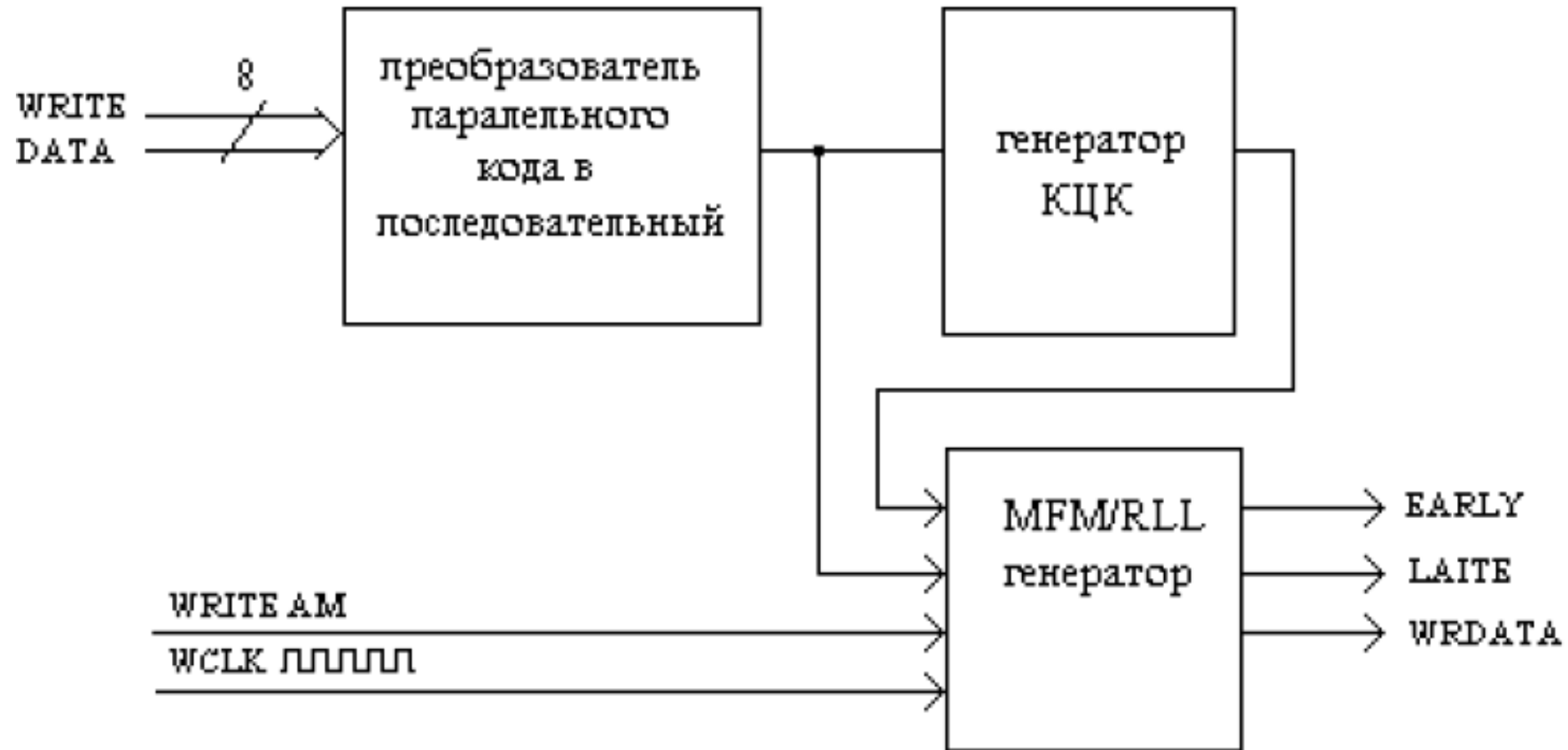
Сектор, помимо блока данных (типичный размер 512 байт), содержит коды проверки достоверности (ЕСС), позволяющие исправлять ошибки без повторного обращения к сектору. Обычно коды ЕСС перемешаны с данными (скрэмблированы).

Неисправный сектор помещается в дефект-лист диска, который хранится в служебной зоне. На замену ему выделяется новый сектор, которому присваивается номер вышедшего.

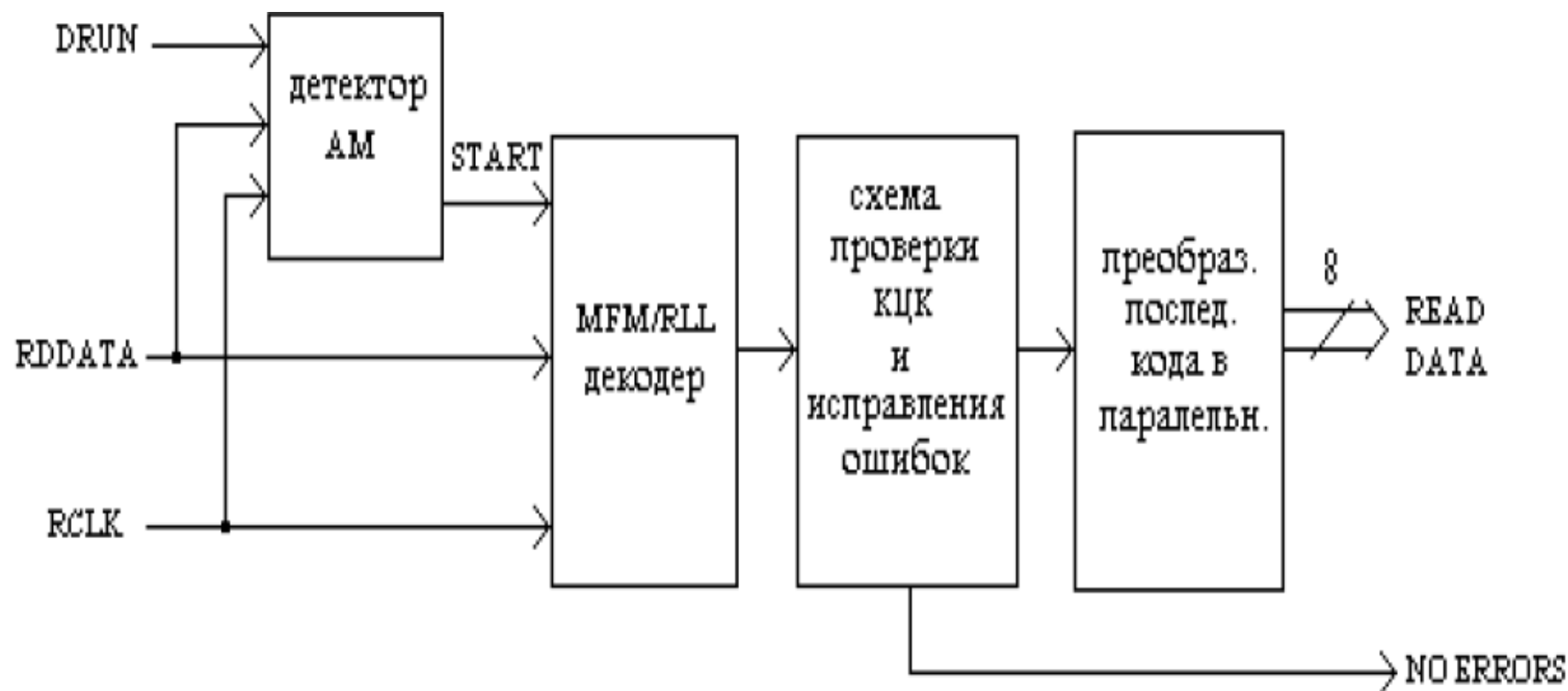
Проблема изменение размера сектора в большую сторону - потеря совместимости с ПО и контроллерами.

В 2006 году IDEMA объявила о переходе на размер сектора 4096 байт, который был завершен к 2010 году.

Тракт записи



Тракт чтения



Плотность записи

Плотность – количество элементов разметки или логических бит на единицу длины или площади. Именно этот показатель определяет потенциальную емкость жесткого диска.

Линейная плотность (Recording density) – количество бит на единицу длины дорожки. Считаются не только биты данных, но и служебные биты. Измеряется в BPI (Bits Per Inch - бит на дюйм).

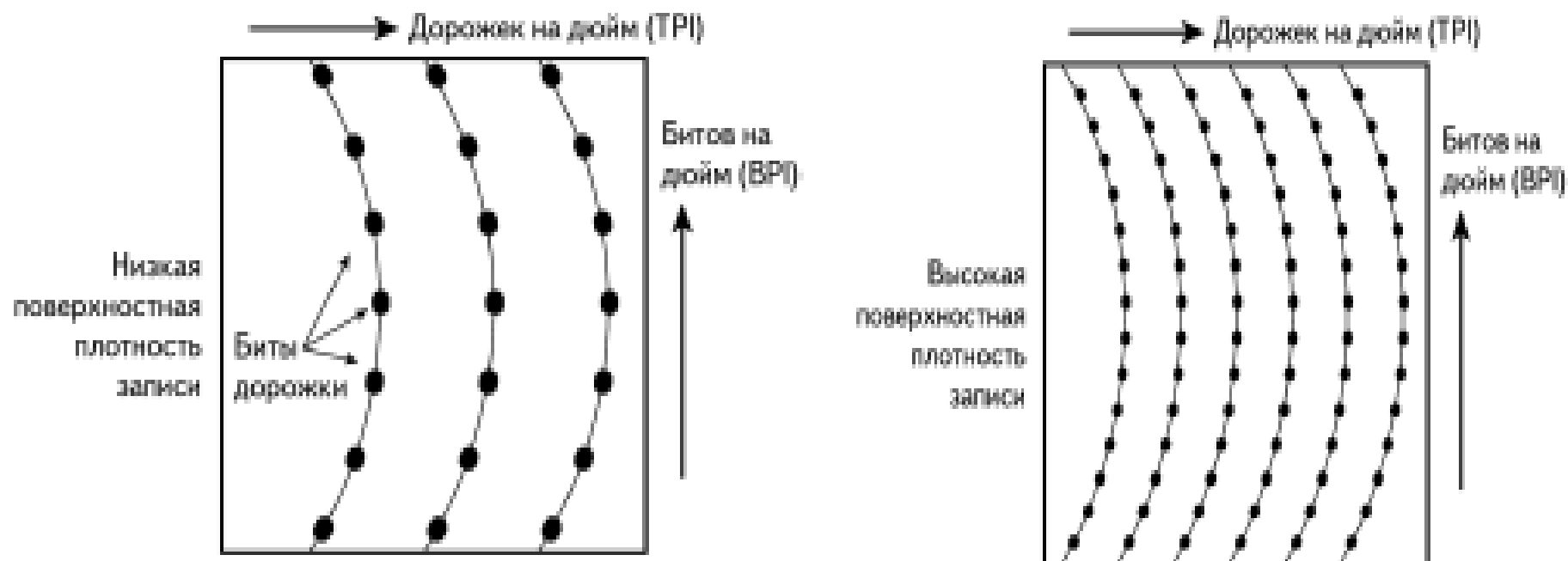
Плотность дорожек (Track density) – количество концентрических дорожек на единицу радиуса. Измеряется в TPI (Tracks Per Inch - треков на дюйм).

Поверхостная (площадная) плотность

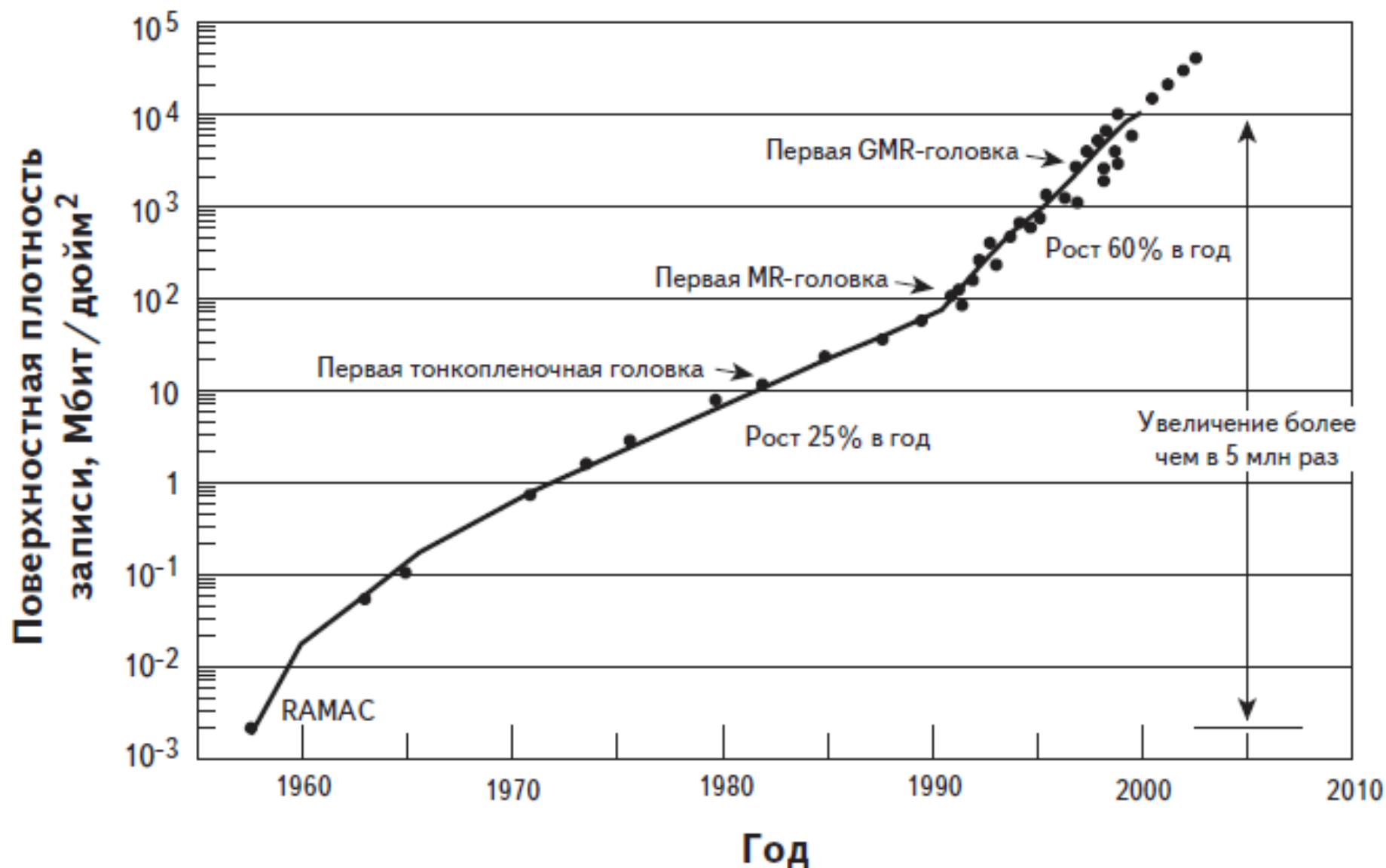
Площадная плотность (Areal density) – количество бит на единицу плотности. Измеряется в бит/кв. дюйм. Типичное значение на сегодня –300-500 Гбит/кв. д.

Произведение линейной плотности записи вдоль дорожки (BPI), и количества дорожек на дюйм (TPI)

Kb KB Mb MB Gb GB Tb TB



Эволюция поверхностной плотности



Повышение плотности записи

Производители винчестеров ищут способы

- обойти суперпарамагнетический барьер;
- улучшить стабильность операций позиционирования, чтения и записи.

Модернизации подвергаются не только головки чтения/записи, но и магнитные пластины.

Традиционно изменялся тип сенсора чтения:

индуктивный

магниторезистивный

GMR

TMR.

Суперпарамагнетизм

Магнитный домен формируется из отдельных магнитных зерен с определенным направлением намагниченности. Для того, чтобы сформировать домен, необходимо упорядочить магнитную ориентацию отдельных магнитных зерен. При этом, естественно, всегда будет существовать некоторое количество зерен, магнитная ориентация которых не будет совпадать с направлением магнитной ориентации домена. И чем больше количество таких несориентированных зерен, тем выше уровень шума и соответственно ниже значение соотношения «сигнал/шум» (Signal-to-Noise Ratio, SNR).

$$SNR \sim \sqrt{n}$$

(n — число зерен, приходящееся на ширину дорожки)

Поэтому для того, чтобы сокращать размер магнитного домена, не меняя при этом соотношения «сигнал/шум», необходимо уменьшать размер самих магнитных зерен. Но делать это до бесконечности невозможно. Если зерно будет слишком маленьким, то его магнитная энергия станет настолько ничтожной, что сравняется с тепловой энергией. В результате за счет теплового движения неизбежно произойдет разупорядочивание отдельных магнитных зерен домена, а следовательно, его размагничивание. Данное явление получило название суперпарамагнетизма.

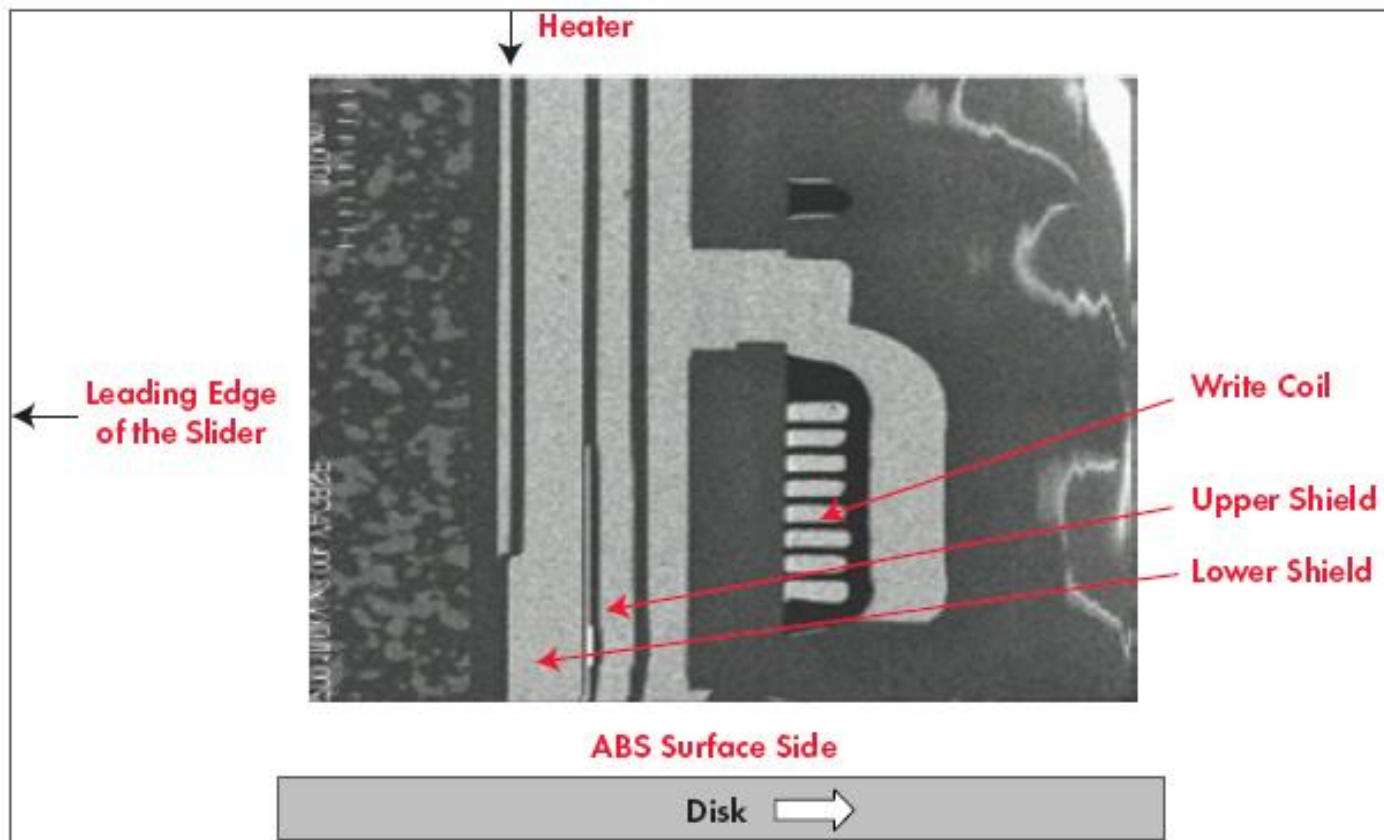
Методы повышения плотности записи

- 1) *Метод перпендикулярной записи* позволил улучшить все характеристики связки «поверхность-головки».
- 2) *Метод магнитной записи с подогревом*
Основан на технологии Thermal Flight Control
(тепловая магнитная запись (**HAMR**) (или термоассистируемая магнитная запись, Thermally-Assisted Recording, TAR)
- 3) *Создание неоднородной структуры поверхности носителя - .*
Структурированные носители данных), — перспективная технология хранения данных на магнитном носителе, в которой для записи используются массивы одинаковых магнитных ячеек, каждая из которых соответствует одному биту информации. Этим она отличается от представителей современных технологий магнитной записи, где бит информации записывается на нескольких магнитных доменах
SOMA, AFT, Patterned Media, Discrete Track Record
- 4) *Методы кодирования* PRML,

Технология Thermal Flight Control

Встраивание нагревателя в головку позволяет изменять ее габариты и форму для обеспечения стабильности и увеличения/уменьшения подъемной силы

Тем самым достигается контроль расстояния от элемента чтения/записи до поверхности носителя



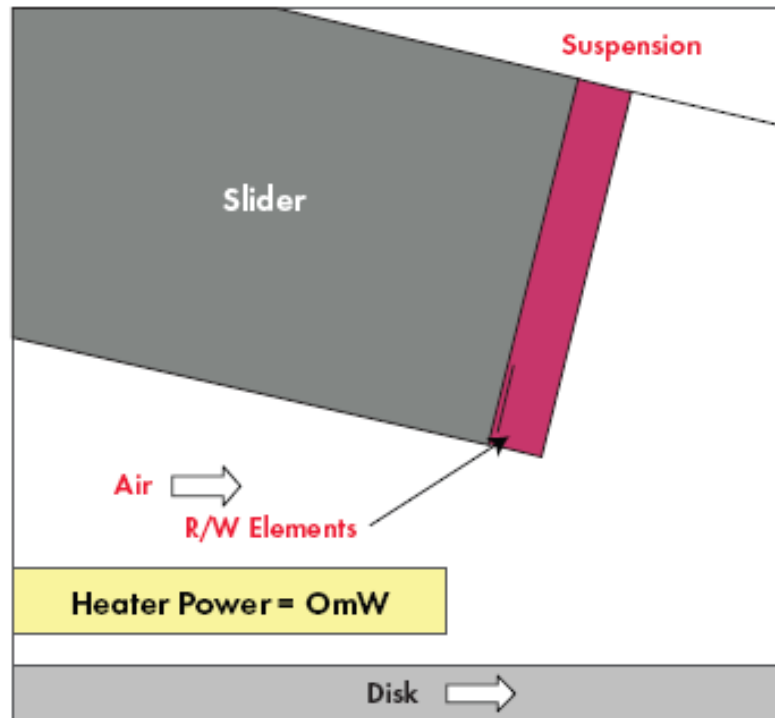


Figure 3: Nominal read/write element protrusion before TFC adjustment

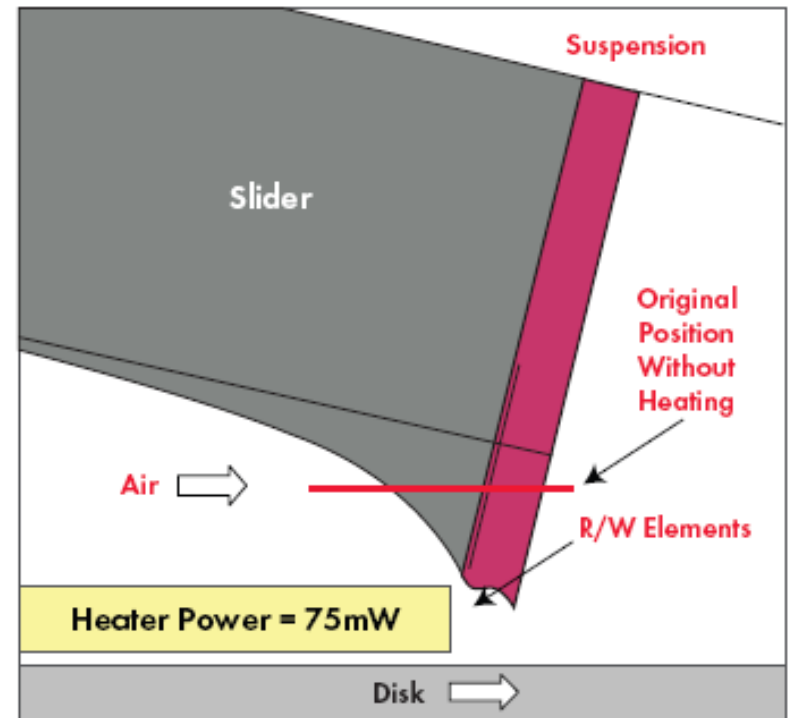


Figure 4: Read/write element protrusion with TFC adjustment

Технология TAMR (HAMR)

Thermal Assisted Magnetic Recording (TAMR) – *магнитной записи с подогревом* (тоже **Термоассистируемая магнитная запись**, *термомагнитная запись, тепловая магнитная запись* **HAMR** (*Heat-assisted magnetic recording*) — гибридная технология записи информации, комбинирующая магнитное чтение и магнитооптическую запись.

Размер бита определяется областью локализации магнитного поля.

При нагреве ферромагнитный слой становится более чувствительным к воздействию извне. Для записи требуется сравнительно небольшое магнитное поле.

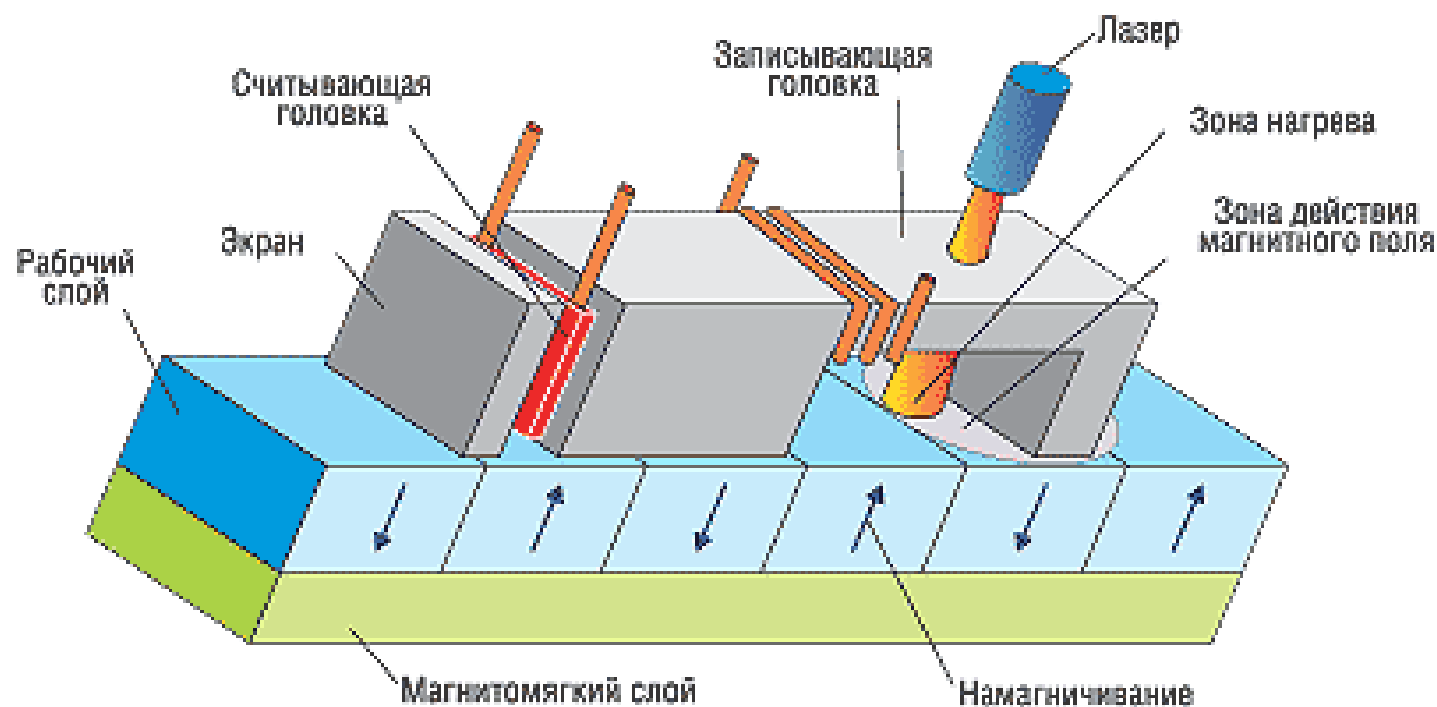
После охлаждения магнитные гранулы менее восприимчивы к воздействиям, в том числе со стороны своих соседей.

Это дает возможность повысить плотность размещения магнитных доменов, то есть информации на пластинах

Специальный тонокопленочный оптический элемент с лазером подогревает узкий участок диска, на который производится запись

Технология TAMR

Размер бита определяется областью локализации магнитного поля



Технология SOMA

SOMA (**S**elf-**O**rganized **M**agnetic **A**rray) - самоорганизующиеся магнитные решетки.

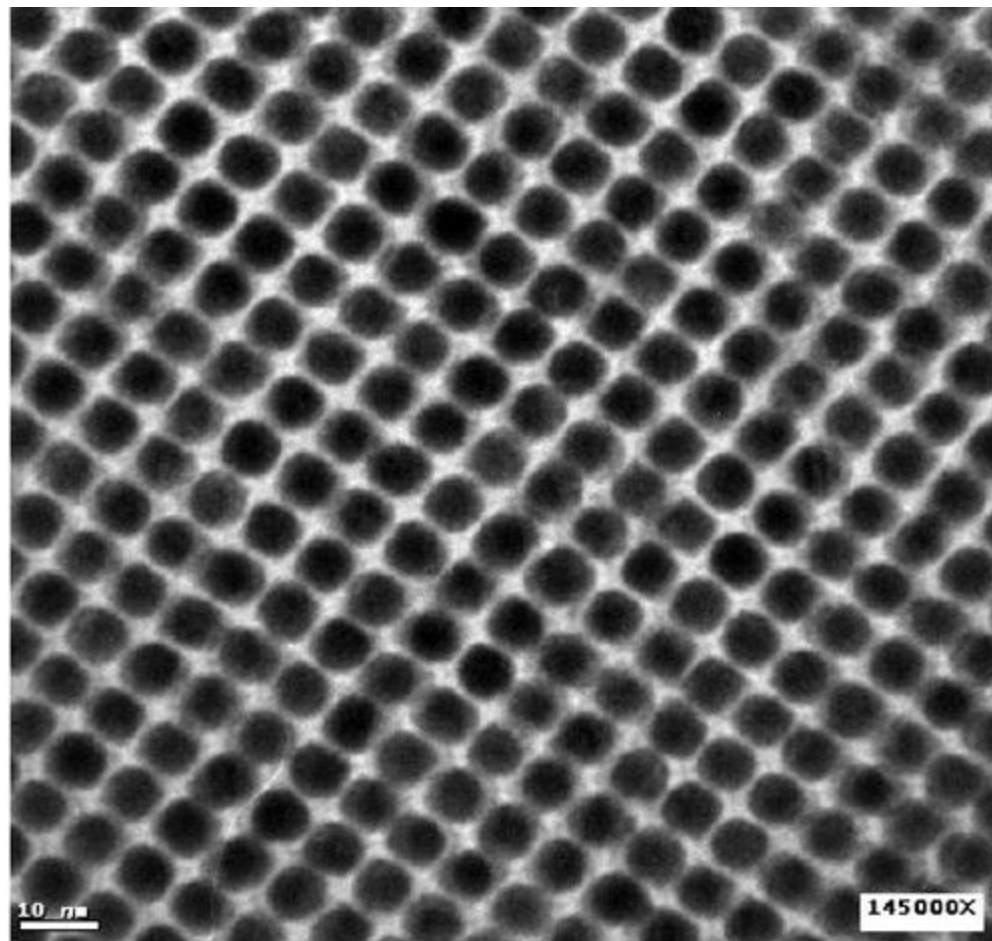
В основе технологии SOMA лежит идея создания небольших изолированных ячеек магнитного материала, организованных в регулярные массивы.

Данная технология предусматривает формирование на поверхности диска монодисперсного слоя "самоорганизующихся магнитных массивов" из мельчайших однородных железно-платиновых конгломератов размером около 3 нм (3 нм - это 10-15 атомов твердого вещества, выложенных в ряд).

Применение этой "нанотехнологии" позволит существенно снизить уровень нестабильности отдельных магнитных зерен и уменьшить размеры домена.

Предполагается, что сочетание технологий SOMA и HAMR позволит достичь плотности записи на магнитный диск 50 Тбит на 1 кв.дюйм.

Самоорганизующиеся магнитные решетки



Фотография массива с плотностью записи 9 Тбит/кв.дюйм, полученная при помощи электронного микроскопа с увеличением 145 тыс. раз, диаметр зерен магнитного материала менее 7 нм.

Технология AFC

AFC (**AntiFerromagnetically-Coupled**) - антиферромагнитная пара (магнитно-компенсированные пленки).

Суть идеи заключается в нанесении на диск винчестера трехслойного антиферромагнитного покрытия, в котором пара магнитных слоев разделена специальной изолирующей прослойкой из рутения. За счет того, что расположенные друг под другом магнитные домены имеют антипараллельную ориентацию магнитного поля, они образуют пару, которая оказывается более устойчивой к спонтанному перемагничиванию, чем одиночный "плоский" домен.

Patterned Media

Запись в ограниченные участки (Bit-Patterned Recording, BPR).

В идеале ячейки магнитного носителя не должны влиять друг на друга. Но поскольку носитель выполнен в виде однородной пленки с гранулами хаотичной формы, это влияние исключить нельзя.

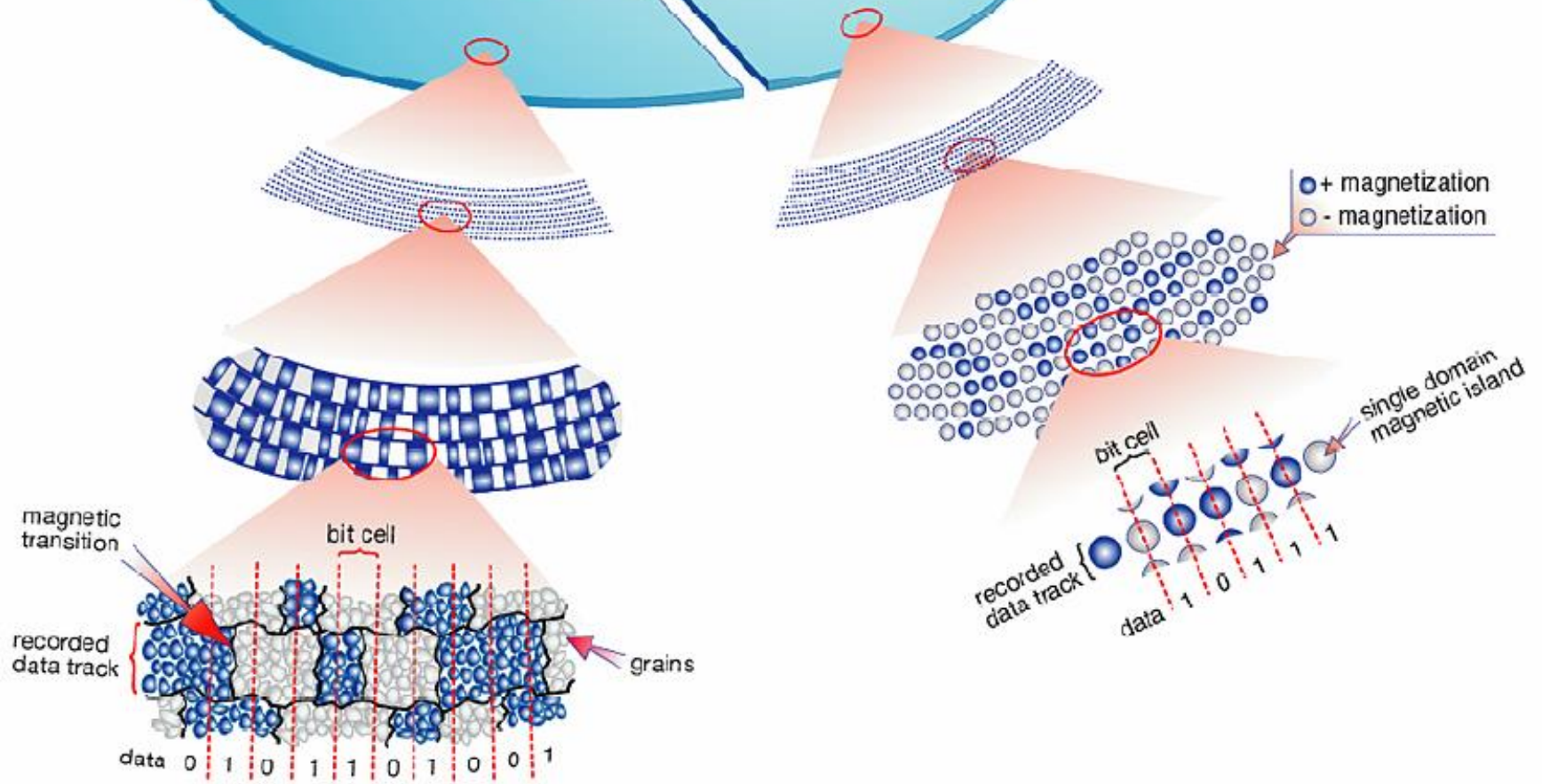
Выход – в отделении, или изоляции, ячеек друг от друга. Данная методика также позволит сделать ячейки одинаковыми по форме и размерам.

Проблема только в том, что современные технологии фотолитографии не позволяют сделать ячейки достаточно миниатюрными для обеспечения высокой плотности записи.

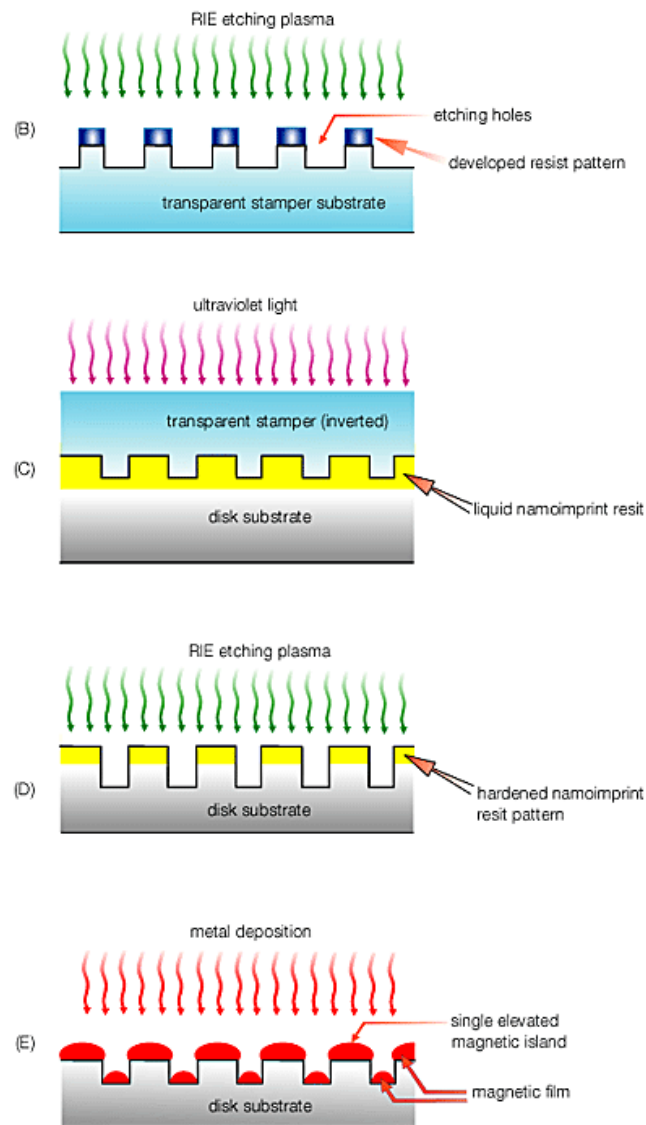
В качестве альтернативы традиционной фотолитографии могут применяться методики электронно-лучевой литографии и нанопечати. Технология записи в ограниченные участки (BPR) предусматривает создание при помощи литографии изолированных друг от друга "магнитных островов" на поверхности диска. Каждая единица информации как бы помещается внутрь высокого "забора", который нельзя преодолеть как снаружи, так и изнутри.

Conventional
Multigrain Media

Patterned Magnetic Media



Технология нанопечати



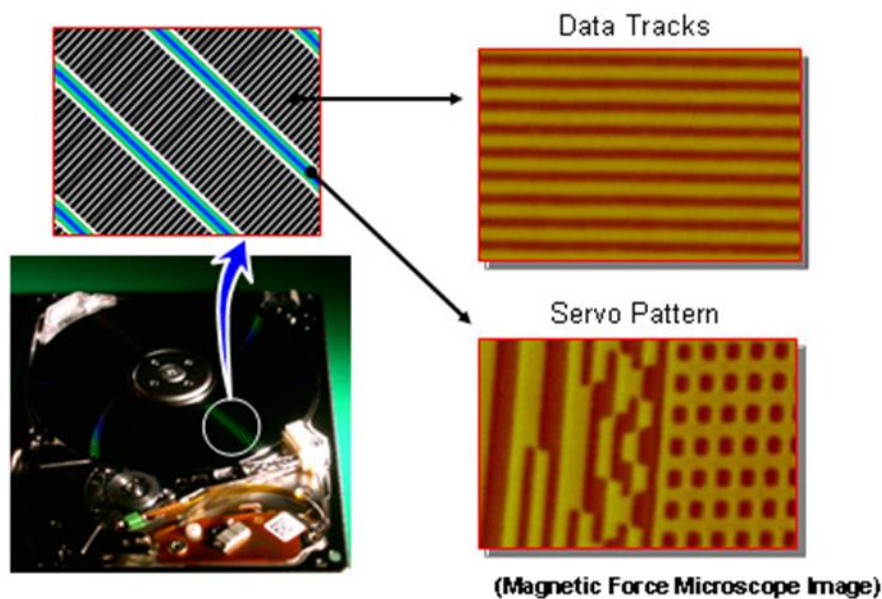
Технология Discrete Track Record

Более дешевой и простой альтернативой Patterned Media является технология Discrete Track Record.

Она предполагает изоляцию не каждой ячейки, а только дорожек и сервометок.

Данная технология может быть внедрена без существенной переделки остальных узлов жесткого диска, однако достигаемый при этом результат невысок – примерно 50% прироста плотности записи.

DTR Pattern Details



Метод кодирования PRML

PRML (Partial Response Maximum Likelihood) - максимальное правдоподобие при неполном отклике.

Это алгоритм преобразования аналогового сигнала, записанного на магнитный диск, основанный на ряде положений теории распознавания образов. В методе PRML для декодирования применяется набор образцов, с которыми сравнивается считанный сигнал, и за результат принимается наиболее похожий.

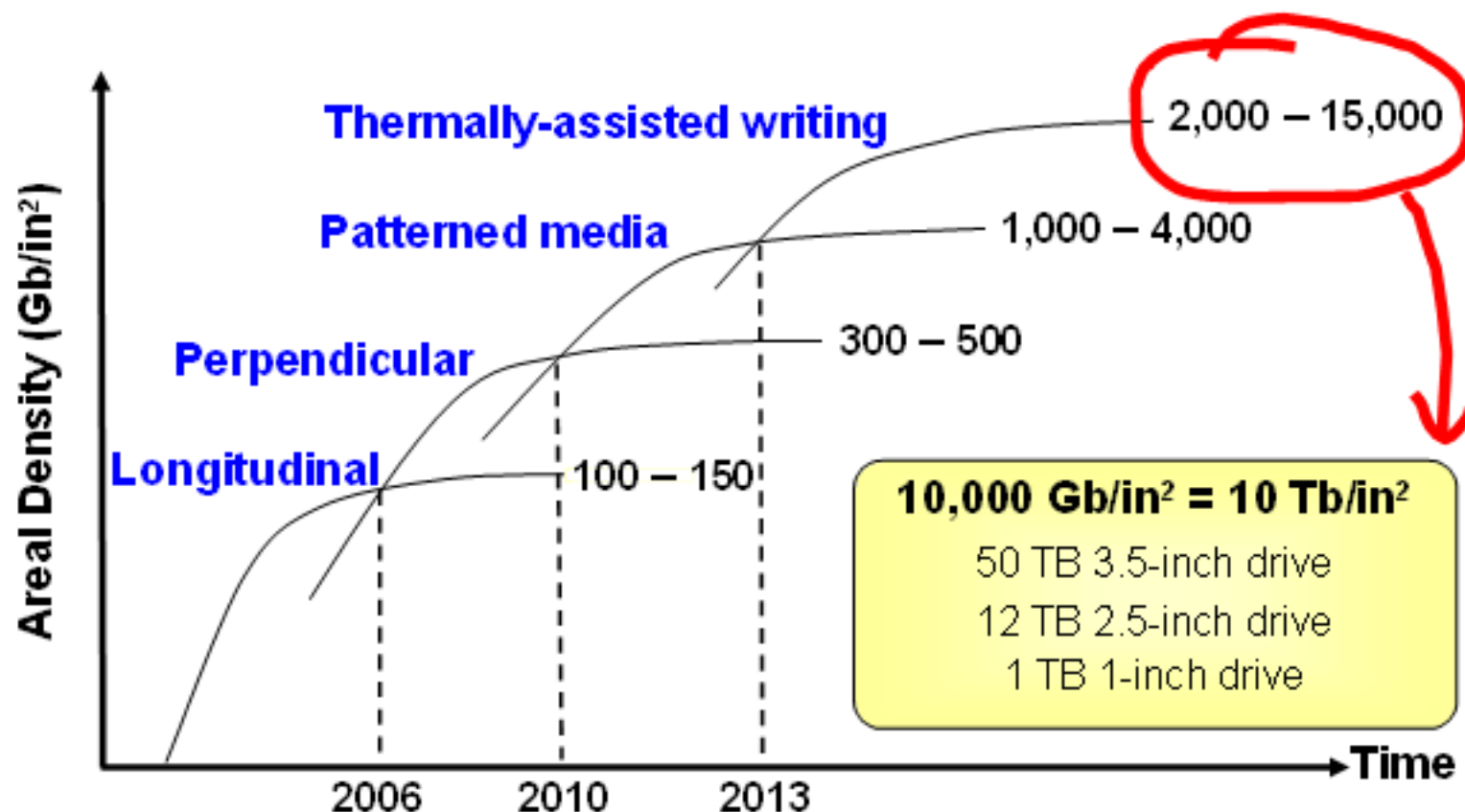
Состоит из двух частей - подсистема Partial Response (частичный отклик) переводит сигнал из аналоговой формы в цифровую, минимизируя шумы, а подсистема Maximum Likelihood (максимальное правдоподобие) производит цифровую обработку сигнала для восстановления наиболее правдоподобной его формы. Данный алгоритм и его развитие EPRML применяется практически во всех современных жестких дисках.

Будущее и перспективы магнитной записи


Hard Drive Technology Roadmap

HITACHI
Inspire the Next

Technology changes coming in the next 10 years will be as radical as the changes that have occurred over the past 50 years.



Сочетание SOMA и HAMR -> плотность ~ 10-15 Тбит/кв.дюйм



Учёные разрабатывают самые необычные способы хранения данных, например, **голографическую или молекулярную** запись на полимеры, но все они требуют принципиального изменения технологии и существенно удорожают производство.

В сравнении с такими предложениями относительной простотой внедрения обладают *методы тепловой магнитной записи (или термоассистированной магнитной записи, Thermally-Assisted Recording, TAR)* и *записи в ограниченные участки (Bit-Patterned Recording, BPR)*.

Эти два способа по-разному борются с главным препятствием дальнейшему "уплотнению" записываемых данных – так называемым явлением суперпарамагнетизма. Дело в том, что при последовательном уменьшении намагничиваемых частиц снижается энергетический барьер, нужный для смены их заряда. В результате частицы становятся очень нестабильными и достаточно минимальной энергии, чтобы получить массив хаотически заряженных частиц, а это означает полное уничтожение записанных ранее данных.

GST технология

Учёные из Hitachi Global Storage Technologies придумали способ объединить преимущества этих двух технологий и минимизировать недостатки. Благодаря "магнитным островам" исчезла необходимость в дорогостоящем мелкозернистом покрытии, а благодаря технологии нагрева четко определённой площади пропала нужда в прецизионных размерах пишущей головки.

В построенном учёными прототипе используется луч лазера, направляемый по волноводу к плазмонной антенне, в которой свет преобразуется в заряд, осуществляющий саму запись. Антенна выполнена в форме буквы "Е", у которой верхняя и нижняя перекладины осуществляют функции заземления, а средняя – функцию разрядника, концентрирующего заряд на небольшой площади поверхности носителя. При длине разрядника 20-25 нм и расстоянии между дорожками записи 24 нм диаметр пятна записи составляет всего 15 нм. При передаче через волновод на антенну пришло около 40% начальной энергии, но частота ошибок оказалась несущественной, а экспериментальная скорость записи составила 250 Мбит/с.

В ходе лабораторных испытаний была достигнута плотность записи до 1,5 Пбит/см²(*). при этом ученые предполагают, что их метод позволит увеличить плотность магнитной записи данных еще в десять раз.

- Если разработанная в Hitachi GST технология будет использована в производстве, то современные винчестеры через некоторое время будут казаться такими же малоемкими, какими сейчас представляются 10-мегабайтные винты, ёмкость которых ещё четверть века назад казалась гигантской. Возникает вопрос: смогут ли тогда догнать их по ёмкости твёрдотельные SSD-накопители?

() 1 Пбайт = 1024 Тбайт = 104876 Гбайт = 128 Тбайт = 131072 Гбайт*

Технические характеристики и показатели HDD

- **Интерфейс.**
- **Емкость.**
- **Форм-фактор.**
- **Скорость вращения шпинделя.**
- **Время произвольного доступа.** Диапазон данного параметра составляет от 2,5 до 16 мс. Обычно, минимальным временем обладают серверные диски, а самым большим — диски для портативных устройств. К примеру, для сравнения, у SSD-накопителей данный параметр меньше 1 мс.
- **Надежность.** Это среднее время наработки на отказ (*MTBF*). Большинство современных дисков поддерживают технологию *S.M.A.R.T.*
- **Кол-во операций ввода-вывода в секунду.** Современные диски производят порядка 50 операций в секунду при произвольном доступе к накопителю, и около 100 операций в секунду при последовательном доступе.

Технические характеристики и показатели HDD (продолжение)

- **Потребление энергии.** Является весьма важным показателем особенно для мобильных устройств.
- **Сопrotивляемость ударам.** Сопrotивляемость резким перепадам давления и другим воздействиям измеряется в единицах допустимой перегрузки во включенном и выключенном состоянии.
- **Скорость передачи данных при последовательном доступе:**
 - внутренняя зона диска: от 44,2 до 74,5 Мб/с;
 - внешняя зона диска: от 60,0 до 111,4 Мб/с.
- **Объем буфера.** Буфер - это промежуточная память, задачей которой является сглаживание различий скорости чтения/записи и передачи по всему интерфейсу. Он обычно варьируется от 8 до 64 Мб.
- **Уровень шума.** Производимый механикой шум накопителя во время его работы. Рассчитывается в децибелах. Тихими накопителями считаются устройства с уровнем шума примерно 26 дБ и ниже. Шум устройства заключается в двух источниках шума: от вращения шпинделя и от позиционирования.

Классификация жестких дисков

Основной критерий – область применения:

- Настольные (**Desktop**) – ПК, рабочие станции
- Мобильные (**Mobile**) – ноутбуки, Tablet PC
- Для бизнеса (**Enterprise**) – серверы, рабочие станции, сетевые системы (устройства) хранения (NAS - Network-Attached Storage) и сети хранения данных (SAN -Storage Area Network)
- Потребительские (CE - Consumer Electronics) – бытовые рекордеры, плееры, КПК
- Промышленные (**Automotive**) – встраиваемые системы, авто, космос и т.д.

Второй критерий – форм-фактор, то есть диаметр пластин

- 3.5 дюйма
- 2.5 дюйма
- 1.8 дюйма
- 1 дюйм, 0.85 дюйма (практически не применяются, вытеснены флэш-памятью)

Классификация жестких дисков

Третий критерий – тип применяемого интерфейса

- Parallel ATA (**PATA**),, (ATA= **AT Attachment**, EIDE= **Enhanced IDE = Integrated Drive Electronics** — параллельный интерфейс подключения накопителей к компьютеру
- Serial ATA = *Serial ATA* — последовательный интерфейс обмена данными с накопителями информации. SATA является развитием параллельного интерфейса ATA (IDE), который после появления SATA был переименован в PATA (Parallel ATA).
- eSATA
- SCSI- *Small Computer System Interface*
- SAS - Serial Attached SCSI - компьютерный интерфейс, разработанный для обмена данными с жесткими дисками и ленточными накопителями

Классификация жестких дисков

Третий критерий – тип применяемого интерфейса – (продолжение)

- FC-AL - *Fibre Channel Arbitrated Loop* — кольцо волоконно-оптического канала с арбитражем, управляемая петля — топология высокоскоростного интерфейса передачи данных, используемое для взаимодействия рабочих станций, мейнфреймов, суперкомпьютеров и систем хранения данных, построенное на базе волоконного канала. Данная технология может использоваться для создания NAS на базе решений устройств хранения данных с применением волоконного канала^[1].
- CE-ATA/ZIF **Consumer Electronics ATA (CE-ATA)** - для соединения устройств хранения и хостов в CE таких как мобильные и портативные. Одна из основных целей – стандартизация соединений дисков с малым форм-фактором, как 1"- микропривод
- ZIF Zero Insertion Force самофиксирующийся разъём
- LIF (Low Insertion Force) с малым усилием сочленения

Классификация жестких дисков

Третий критерий – тип применяемого интерфейса

- *PCMCIA*

CompactFlash - разработан компанией SanDisk Corporation в 1994 году для одного из первых форматов карт флеш-памяти (эти карты всё ещё популярны в фототехнике благодаря рекордным показателям скорости и емкости).

PCMCIA - спецификация на модули расширения, разработана ассоциацией PCMCIA (*Personal Computer Memory Card International Association*). Широко используются в ноутбуках, модули расширения, изготовленные в соответствии с этой спецификацией обычно называются «PC-карты» (*PC Card*).

People Can't Memorize Computer Industry Acronyms (люди не в состоянии запоминать сокращения компьютерной индустрии)

Personal Computer Manufacturers Can't Invent Acronyms (изготовители персональных компьютеров не умеют придумывать сокращения).

Классификация жестких дисков

Четвертый критерий – обороты шпинделя.

Сегодня шпиндельные двигатели винчестеров обеспечивают обороты:

- 3600 RPM
- 4200 RPM
- 5400 RPM
- 7200 RPM
- 10000 RPM
- 15000 RPM

Ряд новых моделей способен снижать обороты для экономии энергии, при этом головки обязательно выводятся из рабочей зоны.

Особенности винчестеров Desktop

Баланс между производительностью, надежностью, емкостью и себестоимостью:

- Форм-фактор: 3.5". Переход к 2.5" возможен в недалеком будущем.
- Тип интерфейса: Serial ATA.
- Обороты шпинделя – 7200 RPM. Выпускаются специальные «зеленые» серии с оборотами 5400 или 6000 RPM.
- Объем буфера – 16 или 32 Мб, реже встречаются модели с 8 Мб.
- Диапазон емкостей – от 80 Гб до 2000 Гб.
- Количество пластин – от 1 до 5.
- Количество головок чтения/записи – от 1 до 10
- Плотность записи – от 250 до 500 Гб на пластину
- Производители начали вводить деление на серии: стандартную, экономичную (с пониженными оборотами, но высокой емкостью), дешевую (за счет чего – непонятно), производительную (более высокая скорость позиционирования).

Особенности винчестеров Mobile

На первом месте – экономичность и надежность:

- Форм-фактор – 2.5". Существуют модели 1.8", но они мало распространены.
- Тип интерфейса – Serial ATA, для 1.8" обычно PATA, планируется применение microSATA.
- Обороты шпинделя – 5400, реже 7200 RPM. У 1.8" обычно 4200 или 5400 RPM, у старых моделей может быть 3600 RPM.
- Объем буфера – 8, реже 16 Мб. Встречаются модели с буфером 32 Мб.
- Диапазон емкостей – от 80 до 1000 Гб
- Количество пластин – 1 или 2, головок – от 1 до 4. Реже встречаются модели увеличенной толщины с 3 пластинами
- Обязательно датчики ускорения для предотвращения выхода из строя головок при падении ноутбука

Форм-фактор 1.8"

На данный момент освоен **Toshiba**, Seagate, Hitachi GST и Samsung.

Винчестеры имеют обороты 5400 или 4200 RPM, реже 3600 RPM.

Рассчитаны на применение в ноутбуках, КПК, планшетных ноутбуках, UMPC = Ultra mobile personal computing (тип мобильных компьютеров, нечто среднее между планшетным и КПК), камкодерах (адаптерах маленькой кассеты к стандартной VHS, видеокамера которая режет изображение сразу на мини DVD-диск), плеерах и т.д.

Емкость винчестеров составляет до 120 Гб при одной и до 240 Гб при двух пластинах.

Скорее всего, от них откажутся в пользу SSD - **Твердотельный накопитель** (*solid-state drive*) — ЗУ на основе микросхем памяти..

Особенности винчестеров Enterprise

Делятся (условно) на винчестеры для серверов, сетевых массивов (NAS и SAN), для файл-серверов, рабочих станций, медиа-приложений (видео, звук, телевидение, видеонаблюдение) и т.п. Используются модифицированные модели из других классов – Desktop и Mobile.

Основные требования для этого класса:

- Надежность (MTBF= Mean Time Between Failures) – среднее количество часов наработки до появления отказа
- Высокая готовность – малое время отклика, большой буфер
- Способность работать в непрерывном режиме (24/7)
- Устойчивость к влиянию вибрации, передаваемой от соседних винчестеров в корпусе
- Система предсказания отказов и мониторинга
- Оптимизация под RAID-массивы (*redundant array of independent disks* — избыточный массив независимых жёстких дисков) — массив из нескольких дисков, управляемых контроллером, взаимосвязанных скоростными каналами и воспринимаемых внешней системой как единое целое.)

Винчестеры с оборотами 15K RPM

Предназначены для применения в серверах, рабочих станциях, сетевых массивах. Основные задачи:

- Транзакции и базы данных
- Системы принятия решений
- Интернет
- E-mail
- NAS, SAN
- Научные и инженерные расчеты

Главное требование – высокая скорость выполнения операций ввода/вывода (IO per second, IOps)

Скорость доступа – максимальна, объем и плотность записи – минимальные

Применяются пластины уменьшенного диаметра и мощные актуаторы (VCM)

Винчестеры с оборотами 10K RPM

Применяются в файл-серверах, серверах среднего и малого класса, рабочих станциях

Основные задачи:

- Ведение архивов деловой информации
- Работа с видео, фото и другими мультимедиа-данными
- NAS/SAN
- RAID-массивы большой емкости
- Стоечные массивы хранения данных
- Остальные задачи, требующие высокой плотности и хорошей скорости

Баланс между скоростью доступа и объемом

Форм-фактор – 3.5”, обычные пластины

Развитие этого направления замедляется ввиду перехода к 15000 RPM

Винчестеры 2.5" с оборотами 10K RPM

Выпускаются Seagate, Toshiba и Hitachi GST.

Оптимизированы для применения в Blade-серверах и системах хранения данных высокой плотности.

Имеют схожую с мобильными винчестерами конструкцию, например, тоже содержат не более двух пластин, но отличаются меньшей плотностью и более высокой скоростью доступа.

В одну корзину для 3.5" винчестера помещается два винчестера 2.5"

По времени наработки на отказ и условиям гарантии не отличаются от других винчестеров Enterprise.

Особенности винчестеров класса СЕ

Единого класса не существует, к этой категории могут относиться различные винчестеры – от 0.85” до 3.5”

Можно разделить на следующие подклассы:

- Внешние накопители
- Винчестеры для бытовой электроники
- Винчестеры для плееров, КПК, видеокамер
- Винчестеры в форм-факторе карточек памяти CompactFlash и карточек PC Card

Внутренние винчестеры такого типа в широкую продажу не поступают.

Внешние винчестеры

Основной интерфейс – USB, реже применяют **IEEE 1394 (FireWire, i-Link)** и eSATA (external SATA)

Основное требование – максимальный объем при минимальной массе

Мостовая схема обеспечивает буферизацию и согласование двух интерфейсов (обычно – SATA и USB (FireWire)). Для eSATA мост получается проще и обеспечивает только преобразование уровней

Основное назначение больших винчестеров – резервное копирование и перенос файлов

Особенности винчестеров Automotive

Основное отличие – способность работать при повышенной температуре, влажности, высоте над уровнем моря

Обычно имеют форм-фактор 2.5” (Hitachi GST) или 1.8” (Toshiba)

Емкость небольшая – в пределах 30-60 Гб

Новая серия Hitachi способна работать в диапазоне температур (-30, +85), обычные серии рассчитаны на (-10,+70), стандартные винчестеры – (0,+60)

Интерфейс – только ATA

Предназначены для:

- Автомобильных систем
- Систем GPS
- Промышленных роботов и оборудования

Скорее всего, будут вытеснены устройствами типа SSD.