

Жесткий диск



Накопитель на жёстких магнитных дисках, жёсткий диск или винчестер, (англ. Hard Disk Drive, англ. HDD) — энергонезависимое, перезаписываемое компьютерное запоминающее устройство. Является основным накопителем данных практически во всех современных компьютерах. В основе функционирования винчестера лежит принцип магнитной записи/считывания сигналов на вращающийся диск, покрытый магниточувствительным рабочим слоем - жёсткие (алюминиевые или стеклянные) пластины, покрытые слоем ферромагнитного материала. Каждая сторона диска, покрытая рабочим слоем, называется рабочей поверхностью.

При записи цифровые данные преобразуются в аналоговые электрические сигналы, создающие с помощью головки записи участки с различной намагниченностью, расположенные вдоль окружности по всей рабочей поверхности вращающегося диска (так называемые треки или дорожки). Размеры участков и расстояние между соседними дорожками определяют поверхностную плотность записи данных. При чтении участки диска движутся под магнитной головкой и индуцируют в ней электрические сигналы, которые преобразуются в цифровые данные.



Внешний жесткий диск

Устройство жесткого диска

Типичный современный накопитель на жестких дисках состоит из блока (пакета) дисков, шпиндельного двигателя привода вращения дисков, блока головок записи/чтения, предусилителя-коммутатора головок и контроллера (печатной платы с электронными схемами управления).

В нерабочем состоянии головка прижимается поводком к поверхности диска в специальной нерабочей зоне, называемой зоной парковки.

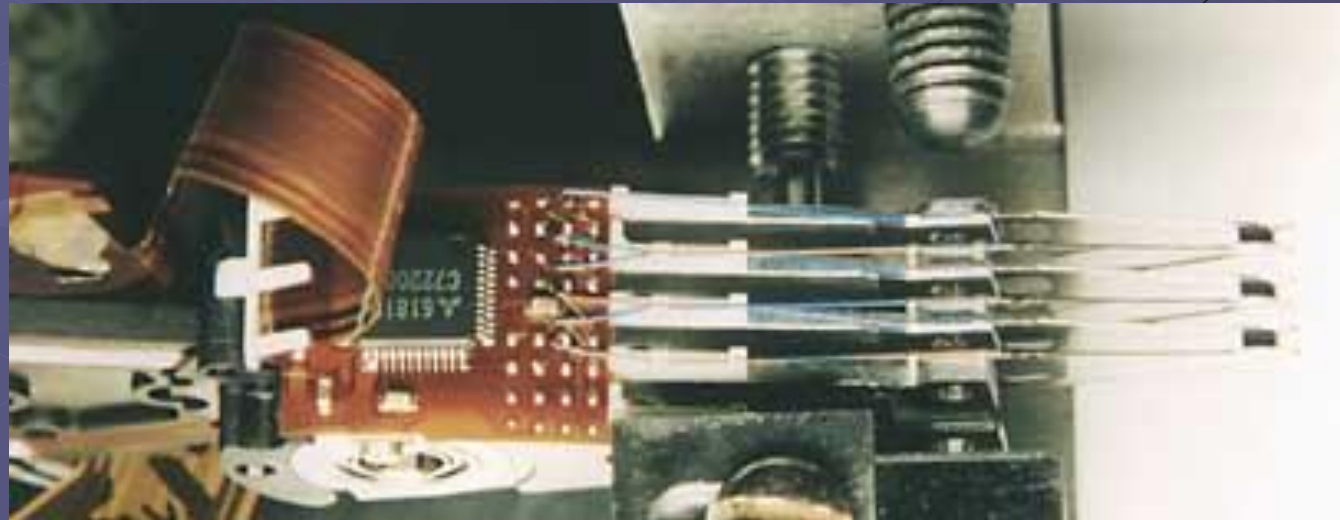
Первые модели винчестеров требовали выполнения специальной операции парковки головок, инициируемой программным обеспечением.

В современных винчестерах операция вывода головок в зону парковки выполняется автоматически при снижении скорости вращения двигателя ниже номинальной или при пропадании напряжения питания, а вывод головок в рабочую зону разрешается только после достижения номинальной скорости вращения дисков. Зазор между головкой и поверхностью диска в современных винчестерах составляет несколько сотых долей микрометра.



Головки жесткого диска

В большой степени максимальная плотность записи зависит от конструкции и характеристик головок записи/чтения. Раньше в винчестерах использовались магнитные головки, представляющие собой миниатюрные катушки индуктивности, намотанные на магнитный сердечник. Позднее стали использовать тонкопленочные магнитные головки, а в современных винчестерах используются высокочувствительные магниторезистивные головки чтения (принцип их работы основан на эффекте анизотропии некоторых полупроводниковых материалов в магнитном поле), конструктивно объединенные с тонкопленочными головками записи. Головки собираются в блок. В современных винчестерах используется система позиционирования блока головок с поворотной подвижной катушкой, помещенной в зазоре мощного постоянного магнита, которая и является исполнительным элементом системы позиционирования.



В основе этой системы лежит предварительная (произведенная при изготовлении винчестера) запись специальных цифровых последовательностей, которые называются сервометками, в специально отведенные для этого на каждой дорожке сектора. Во время работы контроллер винчестера ориентируется на эти сервометки, вырабатывая управляющие сигналы, подаваемые в подвижную катушку, и поворачивает головку таким образом, чтобы она установилась точно над дорожкой, а затем удерживает ее на этой дорожке до поступления команды о переводе головки в новое положение. Смещение головки определяется по разности амплитуды сигналов сервометок при помощи позиционера (head positioner). От того, насколько оперативно позиционер справляется со своими функциями, в немалой степени зависит общая скорость работы привода. Важнейший параметр - время позиционирования головок (seek time) - во многом зависит именно от этого модуля. Позиционер состоит из длинных тонких несущих и управляющего электромагнитного привода. Такую систему называют коромыслом. Обмотку привода окружает статор - неподвижный магнит. Когда по обмотке проходит ток - необходимой величины и полярности - коромысло совершает поворот в ту или иную сторону. Подобного рода устройства называют поворотными - головки в них перемещаются по дуге от центра диска к периферии.

Встречаются и линейные позиционеры, позволяющие перемещать головки не по дуге, а по радиусу диска. Несмотря на некоторые преимущества этой конструкции, из-за большой инерционности, низкой устойчивости к ударам и вибрациям линейные позиционеры не получили широкого распространения. Для увеличения плотности записи зазор между поверхностью диска и головкой необходимо уменьшить до минимума. В современных винчестерах эта задача решается с использованием аэродинамической подъемной силы, создаваемой потоком воздуха, который увлекает за собой вращающаяся рабочая поверхность диска. Для возникновения подъемной силы рабочим поверхностям головок придают специальную форму в виде крыла. Для того чтобы головка не «улетала» далеко от поверхности диска, она закрепляется на пружинящем поводке.

Поскольку величина подъемной силы определяется плотностью воздуха, которая зависит от атмосферного давления, то винчестеры общего применения имеют ограничения по максимальной высоте подъема над уровнем моря (приблизительно до 2000...3000 м). В современных накопителях скорость вращения пакета дисков может достигать 15 000 об/мин. Однако высокие скорости вращения порождают проблемы, связанные с его балансировкой, гироскопическим эффектом и аэродинамикой головок. Во время работы головки ни в коем случае не должны механически соприкоснуться с рабочими поверхностями – случайное касание поверхности практически всегда приводит к полному или частичному повреждению соответствующей дорожки рабочей поверхности и очень часто к обрыву самой головки.

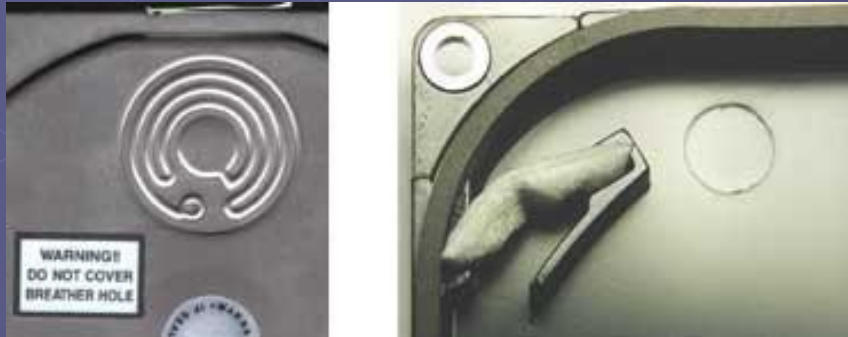
Гермоблок

Пакет дисков с двигателем и блок головок размещаются в специальном герметичном металлическом корпусе со съемной крышкой, который называется гермоблоком или камерой. Ее внутренний объем не изолирован от внешней среды – обязательно предусматривается возможность перетока воздуха снаружи в камеру и наоборот. Это необходимо для выравнивания давления внутри камеры с внешним давлением для предотвращения деформаций корпуса.

Достигается это с помощью так называемого барометрического фильтра, размещаемого внутри камеры. Фильтр способен задерживать частицы, размер которых превышает величину рабочего зазора между головкой и поверхностью диска. Так что под герметичностью камеры подразумевается только невозможность проникновения внутрь посторонних частиц, способных при попадании в рабочий зазор повредить рабочую поверхность и головку.

Кроме того, в камере винчестера обязательно размещается фильтр рециркуляции, предназначенный для улавливания частиц, которые могут возникать внутри самой камеры, например за счет осыпания поверхности дисков при «взлетах» и «посадках» головок в зоне парковки, или проникать внутрь камеры через барометрический фильтр. Место расположения фильтра рециркуляции выбирается с учетом движения воздушного потока и возможных траекторий движения частиц таким образом, чтобы обеспечить максимальную степень очистки воздушного потока внутри камеры.

Так как попадание посторонних частиц внутрь камеры абсолютно недопустимо, сборка винчестеров производится только в специальных «чистых помещениях», оборудованных дорогостоящими фильтровентиляционными установками тонкой очистки воздуха – в кубическом футе воздуха может быть не более 100 частиц (пылинок) размером более 0,5 мкм.



Барометрический и рециркуляционный фильтры

Электроника HDD

Часть электроники привода находится в блоке механики. Сигнал, снимаемый с магнитных головок очень слабый, и если проводники будут слишком длинными, он будет серьезно искажен. Прочитанный сигнал необходимо сразу же усилить - тогда проблема транспортировки исчезнет. С этой функцией успешно справляется предусилитель, расположенный в гермоблоке.

От предусилителя к позиционеру идет ленточный кабель или набор обычных одножильных проводов, а они довольно часто рвутся. Устранение подобной неисправности, увы, обходится в копейчку.

Остальная электроника винчестера менее уязвима и находится на отдельной плате за пределами гермоблока. По своей структурой она очень напоминает... отдельный компьютер! Действительно, среди основных компонент значатся: центральный процессор, ОЗУ (буфер диска), ПЗУ с программой управления (иногда часть ее записывают в служебную область самого диска), а также DSP (Digital Signal Processor), служащий для обработки считанных сигналов и подготовки записываемых.

На печатных платах многих жестких дисков встречается технологический интерфейсный разъем, с помощью которого их подключают к тестовому оборудованию. В ПЗУ находится специальная программа, позволяющая вести диалог, переназначать дефектные участки, производить ту же первичную разметку и пр. В современных моделях, изготавливаемых в рамках программы Energy Star, обязательно есть устройство для отключения винчестера при отсутствии запросов к нему и других функций энергосбережения.



Структура данных в HDD

Хранение данных

Жесткий диск, как и всякое другое блочное устройство, хранит информацию фиксированными порциями, которые называются блоками. Блок является наименьшей порцией данных, имеющей уникальный адрес на жестком диске. Для того чтобы прочесть или записать нужную информацию в нужное место, необходимо представить адрес блока в качестве параметра команды, выдаваемой контроллеру жесткого диска. Размер блока уже довольно с давних пор является стандартным для всех жестких дисков - 512 байт.

К сожалению, достаточно часто происходит путаница между такими понятиями как "сектор", "кластер" и "блок". Фактически, между "блоком" и "сектором" разницы нет. Правда, одно понятие логическое, а второе топологическое. "Кластер" - это несколько секторов, рассматриваемых операционной системой как одно целое. Переход к кластерам произошел потому, что размер таблицы FAT был ограничен, а размер диска увеличивался. В случае FAT16 для диска объемом 512 Мб кластер будет составлять 8 Кб, до 1 Гб - 16 Кб, до 2 Гб - 32 Кб и так далее.

Для того чтобы однозначно адресовать блок данных, необходимо указать все три числа (номер цилиндра, номер сектора на дорожке, номер головки). Такой способ адресации диска был широко распространен и получил впоследствии обозначение аббревиатурой CHS (cylinder, head, sector). Именно этот способ был первоначально реализован в BIOS, поэтому впоследствии возникли ограничения, связанные с ним. Дело в том, что BIOS определил разрядную сетку адресов на 63 сектора, 1024 цилиндра и 255 головок. Однако развитие жестких дисков в то время ограничилось использованием лишь 16 головок в связи со сложностью изготовления. Отсюда появилось первое ограничение на максимально допустимую для адресации емкость жесткого диска:

$$1024 * 16 * 63 * 512 = 504 \text{ Мб.}$$

Со временем, производители стали делать HDD большего размера. Соответственно число цилиндров на них превысило 1024, максимально допустимое число цилиндров (с точки зрения старых BIOS). Однако, адресуемая часть диска продолжала равняться 504 Мбайтам, при условии, что обращение к диску велось средствами BIOS. Это ограничение со временем было снято введением так называемого механизма трансляции адресов.

Проблемы, возникшие с ограниченностью BIOS по части физической геометрии дисков, привели в конце концов к появлению нового способа адресации блоков на диске. Этот способ довольно прост. Блоки на диске описываются одним параметром - линейным адресом блока. Адресация диска линейно получила аббревиатуру LBA (logical block addressing). Линейный адрес блока однозначно связан с его CHS адресом:

$$lba = (cyl * HEADS + head) * SECTORS + (sector - 1);$$

Введение поддержки линейной адресации в контроллеры жестких дисков дало возможность BIOS'ам заняться трансляцией адресов. Суть этого метода состоит в том, что если в приведенной выше формуле увеличить параметр HEADS, то потребуется меньше цилиндров, чтобы адресовать то же самое количество блоков диска. Но зато потребуется больше головок. Однако головок-то как раз использовалось всего 16 из 255. Поэтому BIOS'ы стали переводить избыточные цилиндры в головки, уменьшая число одних и увеличивая число других. Это позволило им использовать разрядную сетку головок целиком. Это отодвинуло границу адресуемого BIOS'ом дискового пространства до 8Gb. Также был разработан режим Large Mode. Этот режим работы предназначен для работы жестких дисков объемом до 1 Гб. В Large Mode количество логических головок увеличивается до 32, а количество логических цилиндров уменьшается вдвое. При этом обращения к логическим головкам 0..F транслируются в четные физические цилиндры, а обращения к головкам 10..1F - в нечетные. Винчестер, размеченный в режиме LBA, несовместим с режимом Large, и наоборот.

Дальнейшее увеличение адресуемых объемов диска с использованием прежних сервисов BIOS стало принципиально невозможным. Действительно, все параметры задействованы по максимальной "планке" (63 сектора, 1024 цилиндра и 255 головок). Тогда был разработан новый расширенный интерфейс BIOS, учитывающий возможность очень больших адресов блоков. Однако этот интерфейс уже не совместим с прежним, вследствие чего старые операционные системы, такие как DOS, которые пользуются старыми интерфейсами BIOS, не смогли и не смогут переступить границы в 8GB. Практически все современные системы уже не пользуются BIOS'ом, а используют собственные драйвера для работы с дисками. Поэтому данное ограничение на них не распространяется. Но следует понимать, что прежде чем система сможет использовать собственный драйвер, она должна как минимум его загрузить. Поэтому на этапе начальной загрузки любая система вынуждена пользоваться BIOS'ом. Это и вызывает ограничения на размещение многих систем за пределами 8GB, они не могут оттуда загружаться, но могут читать и писать информацию (например, DOS который работает с диском через BIOS).

Разделы

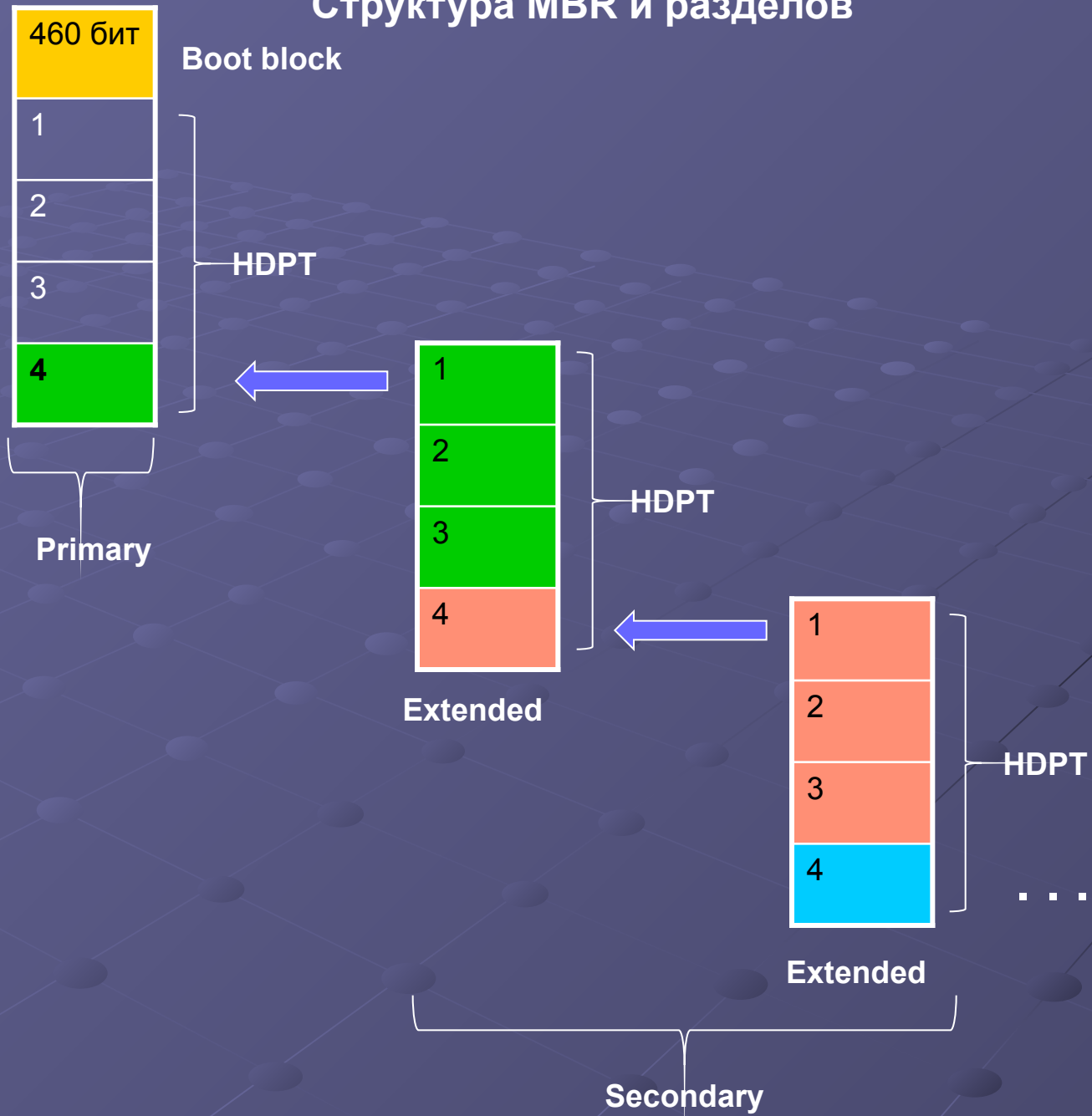
Для организации систем дисковое адресное пространство блоков разделяется на части, называемые разделами (partitions). Разделы полностью подобны целому диску в том, что они состоят из смежных блоков. Благодаря такой организации для описания раздела достаточно указания начала раздела и его длины в блоках. Жесткий диск может содержать четыре первичных раздела.

Во время загрузки компьютера, BIOS загружает первый сектор головного раздела (загрузочный сектор) по адресу 0000h:7C00h и передает ему управление. В начале этого сектора расположен загрузчик (загрузочный код), который прочитывает таблицу разделов и определяет загружаемый раздел (активный). А дальше все повторяется. То есть он загружает загрузочный сектор этого раздела на этот же адрес и снова передает ему управление.

Разделы являются контейнерами всего своего содержимого. Этим содержимым является, как правило, файловая система. Под файловой системой с точки зрения диска понимается система разметки блоков для хранения файлов. После того, как на разделе создана файловая система и в ней размещены файлы операционной системы, раздел может стать загружаемым. Загружаемый раздел имеет в своем первом блоке небольшую программу, которая производит загрузку операционной системы. Однако для загрузки определенной системы нужно явно запустить ее загрузочную программу из первого блока. Разделы с файловыми системами не должны пересекаться. Это связано с тем, что две разные файловые системы имеют каждая свое представление о размещении файлов, но когда это размещение приходится на одно и то же физическое место на диске, между файловыми системами возникает конфликт. Этот конфликт возникает не сразу, а лишь по мере того, как файлы начинают размещаться в том месте диска, где разделы пересекаются. Поэтому следует внимательно относиться к разделению диска на разделы. Само по себе пересечение разделов не опасно. Опасно именно размещение нескольких файловых систем на пересекающихся разделах. Разметка диска на разделы еще не означает создания файловых систем. Однако, уже сама попытка создания пустой файловой системы (то есть форматирование), на одном из пересекающихся разделов может привести к возникновению ошибок в файловой системе другого раздела. Все сказанное относится в одинаковой степени ко всем операционным системам, а не только самым популярным.

Диск разбивается на разделы программным путем. То есть, Вы можете создать произвольную конфигурацию разделов. Информация о разбиении диска хранится в самом первом блоке жесткого диска, называемым главной загрузочной записью (Master Boot Record (MBR)).

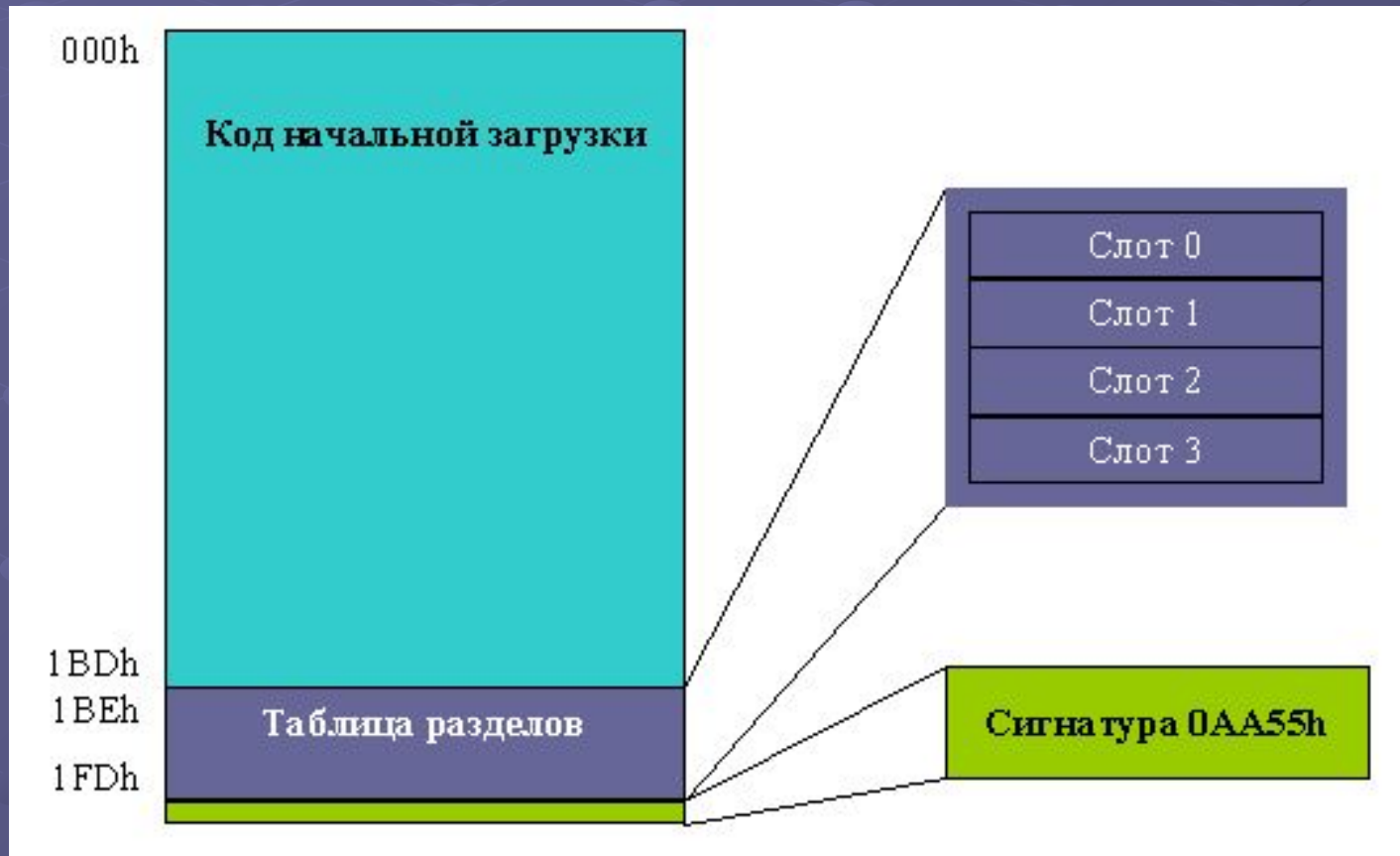
Структура MBR и разделов



MBR

MBR является основным средством загрузки с жесткого диска, поддерживаемым BIOS. Для наглядности представим содержимое загрузочной области в виде схемы.

Все то что находится по смещению 01BEh-01FDh называется таблицей разделов. Вы видите, что в ней четыре раздела. Только один из четырех разделов имеет право быть помеченным как активный, что будет означать, что программа загрузки должна загрузить в память первый сектор именно этого раздела и передать туда управление. Последние два байта MBR должны содержать число 0xAA55. По наличию этой сигнатуры BIOS проверяет, что первый блок был загружен успешно. Сигнатура эта выбрана не случайно. Ее успешная проверка позволяет установить, что все линии данных могут передавать и нули, и единицы. Программа загрузки просматривает таблицу разделов, выбирает из них активный, загружает первый блок этого раздела и передает туда управление.



Давайте посмотрим как устроен дескриптор раздела:

Смещение	Описание
0000h	маркер начальной загрузки
0001h	головка
0002h	сектор и цилиндр
0003h	цилиндр
0004h	системное описание
0005h	головка
0006h	сектор и цилиндр
0007h	цилиндр
0008h-000Bh	смещение секторов
000Ch-000Fh	количество секторов в разделе

* 0001h-0003h начало раздела
** 0005h-0007h конец раздела

С точки зрения разделов диска наиболее популярной до недавнего времени была и остается MS-DOS. Она забирает в свое пользование два из четырех разделов: Primary DOS partition, Extended DOS partition. Первый из них, (primary) это обычный досовый диск C:. Второй - это контейнер логических дисков. Они все находятся там в виде цепочки подразделов, которые так и именуются: D:, E:, ... Логические диски могут иметь и инородные файловые системы, отличные от файловой системы DOS. Однако, как правило, инородность файловой системы связана присутствием еще одной операционной системы, которую, вообще говоря, следовало бы поместить в свой собственный раздел (не extended DOS), но для таких целей часто оказывается слишком маленькой таблица разделов.

Отметим еще одно важное обстоятельство. Когда на чистый жесткий диск устанавливается DOS, то при загрузке нет никаких альтернатив в выборе операционных систем. Поэтому загрузчик выглядит весьма примитивно, ему не надо спрашивать у пользователя, какую систему тот хочет загрузить. С желанием иметь сразу несколько систем возникает необходимость заводить программу, позволяющую выбирать систему для загрузки.

Параметры HDD

Диаметр дисков (disk diameter)

параметр довольно свободный от каких-либо стандартов, ограничиваемый лишь форм-факторами корпусов системных блоков. Наиболее распространены накопители с диаметром дисков 2.2, 2.3, 3.14 и 5.25 дюймов. Диаметр дисков определяет плотность записи на дюйм магнитного покрытия. Накопители большего диаметра содержат большее число дорожек, и в них, как правило используются более простые технологии изготовления носителей, предназначенных для меньшей плотности записи. Они, как правило, медленнее своих меньших собратьев и имеют меньшее число дисков, но более надежны. Накопители с меньшим диаметром больших объемов имеют более высокотехнологичные поверхности и высокие плотности записи информации, а также, как правило, и большее число дисков.

Число поверхностей (sides number)

определяет количество физических дисков нанизанных на шпиндель. Выпускаются накопители с числом поверхностей от 1 до 8 и более. Однако, наиболее распространены устройства с числом поверхностей от 2 до 5. Принципиально, число поверхностей прямо определяет физический объем накопителя и скорость обработки операций на одном цилиндре.

Число цилиндров (cylinders number)

определяет сколько дорожек (треков) будет располагаться на одной поверхности. В настоящее время все накопители емкостью более 1 Гигабайта имеют число цилиндров более 1024, вследствие чего, для распространенных ОС применяются унифицированные режимы доступа с пересчетом и эмуляцией и виртуализацией числа головок, цилиндров и секторов (LBA и Large).

Число секторов (sectors count)

бщее число секторов на всех дорожках всех поверхностей накопителя. Определяет физический неформатированный объем устройства.

Число секторов на дорожке (sectors per track)

общее число секторов на одной дорожке. Часто, для современных накопителей показатель условный, т.к. они имеют неравное число секторов на внешних и внутренних дорожках, скрытое от системы и пользователя интерфейсом устройства.

Частота вращения шпинделя (rotational speed или spindle speed)

определяет, сколько времени будет затрачено на последовательное считывание одной дорожки или цилиндра. Частота вращения измеряется в оборотах в минуту (rpm). Для дисков емкостью до 1 гигабайта она обычно равна 5,400 оборотов в минуту, а у более вместительных достигает 7,200, 10000 rpm и более.

Время перехода от одной дорожки к другой (track-to-track seek time)

обычно составляет от 3.5 до 5 миллисекунд, а у самых быстрых моделей может быть от 0.6 до 1 миллисекунды. Этот показатель является одним из определяющих быстродействие накопителя, т.к. именно переход с дорожки на дорожку является самым длительным процессом в серии процессов произвольного чтения/записи на дисковом устройстве. Показатель используется для условной оценки производительности при сравнении накопителей разных моделей и производителей.

Время успокоения головок (head latency time)

время, проходящее с момента окончания позиционирования головок на требуемую дорожку до момента начала операции чтения/записи. Является внутренним техническим показателем, входящим в показатель - время перехода с дорожки на дорожку.

Время установки или время поиска (seek time)

время, затрачиваемое устройством на перемещение головок чтения/записи к нужному цилиндру из произвольного положения.

Среднее время установки или поиска (average seek time)

усредненный результат большого числа операций позиционирования на разные цилиндры, часто называют средним временем позиционирования. Среднее время поиска имеет тенденцию уменьшаться с увеличением емкости накопителя, т.к. повышается плотность записи и увеличивается число поверхностей. Среднее время поиска является одним из важнейших показателей оценки производительности накопителей, используемых при их сравнении.

Время ожидания (latency)

время, необходимое для прохода нужного сектора к головке, усредненный показатель – среднее время ожидания (average latency), получаемое как среднее от многочисленных тестовых проходов. После успокоения головок на требуемом цилиндре контроллер ищет нужный сектор. При этом, последовательно считываются адресные идентификаторы каждого проходящего под головкой сектора на дорожке. В идеальном, с точки зрения производительности случае, под головкой сразу окажется нужный сектор, в плохом - окажется, что этот сектор только что "прошел" под головкой, и, до окончания процесса успокоения необходимо будет ждать полный оборот диска для завершения операции чтения/записи.

Время доступа (access time)

суммарное время, затрачиваемое на установку головок и ожидание сектора. Причем, наиболее долгим является промежуток времени установки головок..

Среднее время доступа к данным (average access time)

время, проходящее с момента получения запроса на операцию чтения/записи от контроллера до физического осуществления операции - результат сложения среднего время поиска и среднего времени ожидания. Среднее время доступа зависит от того, как организовано хранение данных и насколько быстро позиционируются головки чтения записи на требуемую дорожку. Среднее время доступа – усредненный показатель от многочисленных тестовых проходов, и обычно, оно составляет от 10 до 18 миллисекунд

Скорость передачи данных (data transfer rate)

определяет скорость, с которой данные считываются или записываются на диск после того, как головки займут необходимое положение. Измеряется в мегабайтах в секунду (MBps) или мегабитах в секунду (Mbps) и является характеристикой контроллера и интерфейса. Различают две разновидности скорости передачи - внешняя и внутренняя. Скорость передачи данных, также является одним из основных показателей производительности накопителя и используется для ее оценки и сравнения накопителей различных моделей и производителей.

Размер кеш-буфера контроллера (internal cash size)

Встроенный в накопитель буфер выполняет функцию упреждающего кэширования и призван сгладить громадную разницу в быстродействии между дисковой и оперативной памятью компьютера. Выпускаются накопители с 128, 256 и 512 килобайтным буфером. Чем больше объем буфера, тем потенциально выше производительность при произвольном "длинном" чтении/записи.

Средняя потребляемая мощность (capacity).

При сборке мощных настольных компьютеров учитывается мощность, потребляемая всеми его устройствами. Современные накопители на ЖД потребляют от 5 до 15 Ватт, что является достаточно приемлемым, хотя, при всех остальных равных условиях, накопители с меньшей потребляемой мощностью выглядят более привлекательно

Уровень шума (noise level)

разумеется, является эргономическим показателем. Однако, он также, является и некоторым показателем сбалансированности механической конструкции, т.к. шум в виде треска - есть не что иное как звук ударов позиционера шагового или линейного механизма, а, даже микро- удары и вибрация так не желательны для накопителей и приводят к более быстрому их износу.

Среднее время наработки на отказ (MTBF)

определяет сколько времени способен проработать накопитель без сбоев

Сопrotивляемость ударам (G-shock rating)

определяет степень сопротивляемости накопителя ударам и резким изменениям давления, измеряется в единицах допустимой перегрузки g во включенном и выключенном состоянии. Является важным показателем для настольных и мобильных систем.

Физический и логический объем накопителей.

Носители жестких дисков, в отличие от гибких, имеют постоянное число дорожек и секторов, изменить которое невозможно. Эти числа определяются типом модели и производителем устройства. Поэтому, физический объем жестких дисков определен изначально и состоит из объема, занятого служебной информацией (разметка диска на дорожки и сектора) и объема, доступного пользовательским данным. Физический объем жесткого диска, также, зависит от типа интерфейса, метода кодирования данных, используемого физического формата и др. Производители накопителей указывают объемы дисков в миллионах байт, предполагая исходя из десятичной системы исчисления, что в одном мегабайте 1000000 байт. Однако, ПО оперирует не десятичной, а двоичной системами, полагая, что в одном килобайте не 1000 байт, а 1024. Такие несложные разногласия в системах исчисления приводят к несоответствиям при оценке объема накопителей, данном в описании и - выдаваемом различными программными тестами.

Одним из возможных, но не желательных способов повышения физической емкости, для производителей, является увеличение емкости сектора. В настоящее время, стандартной емкостью сектора для IBM-совместимых компьютеров является 512 байт. Многие адаптеры позволяют, в процессе физического форматирования, программным путем, изменять емкость сектора, например, до 1024 байт. При этом, соотношение пользовательских данных и служебной информации для сектора улучшается, но снижается надежность хранения данных, т.к. тот же полином ECC будет использоваться для коррекции большего объема данных. Однако, выигрыш на физическом уровне еще не означает тот же результат на логическом, т.к. логическая структура диска может оказаться не эффективной, например, при использовании для работы с файлами малой длины (менее 1 К). Логический же объем зависит от того, как операционная система или программа записывает информацию в сектора. В случае использования программ и операционных систем с программной компрессией данных, можно повысить объем носителя на величину, зависящую от степени сжатия данных.

Для оптимального использования поверхности дисков применяется так называемая зонная запись (Zoned Bit Recording - ZBR), принцип которой состоит в том, что на внешних дорожках, имеющих большую длину (а следовательно - и потенциальную информационную емкость на единицу площади), информация записывается с большей плотностью, чем на внутренних. Таких зон с постоянной плотностью записи в пределах всей поверхности образуется до десятка и более; соответственно, скорость чтения и записи на внешних зонах выше, чем на внутренних.

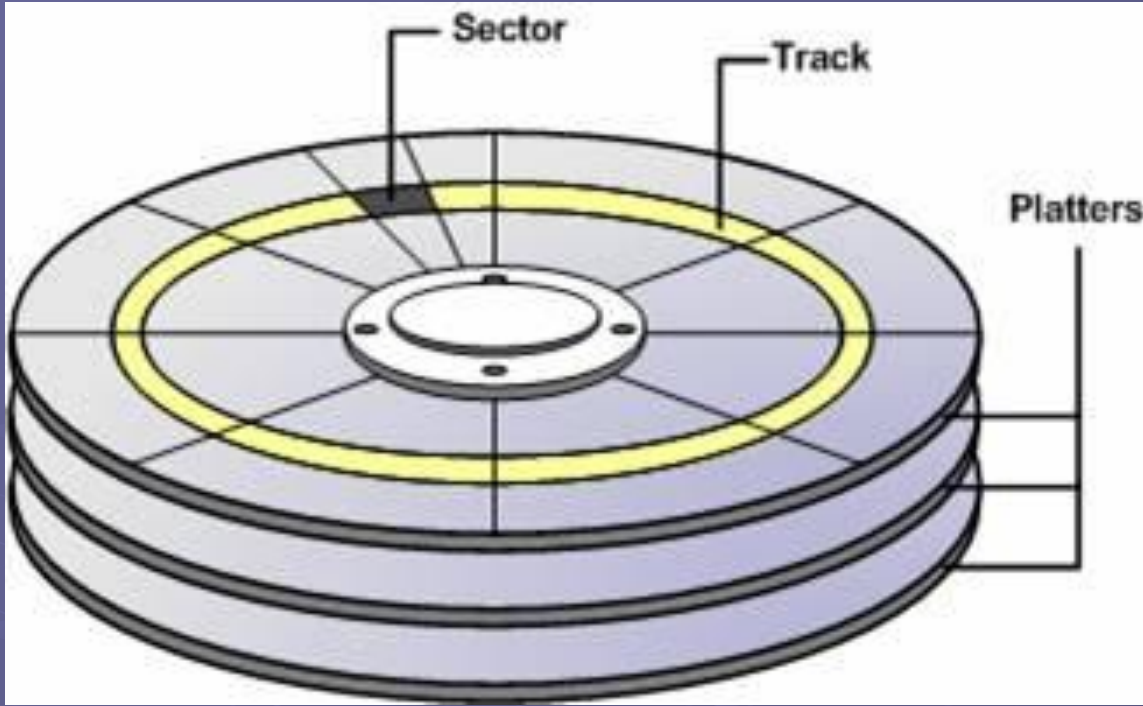
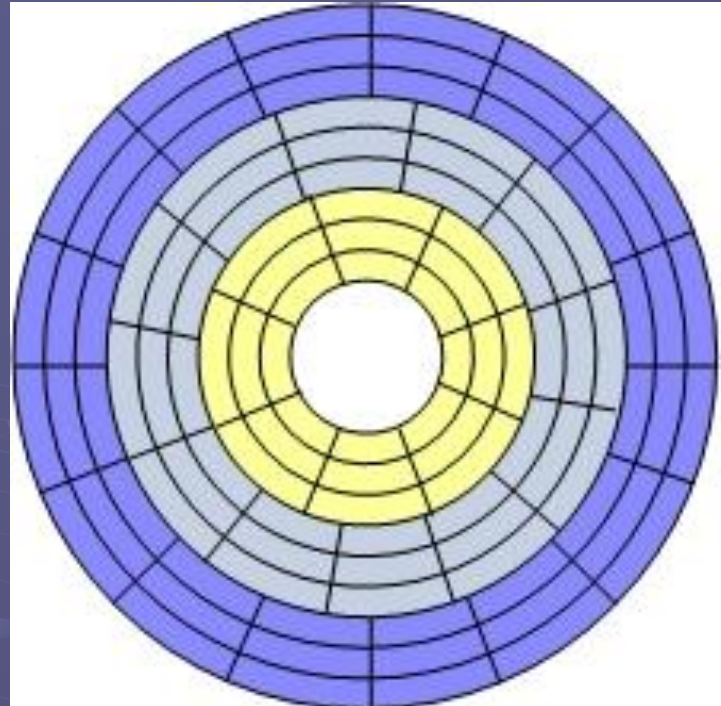


Схема строения жесткого диска



Зонная запись (ZBR)

Если вообразить себе поверхность пластины, учитывая то, что каждая дорожка разбита на одинаковое число секторов и то, что длина сектора растет с ростом радиуса дорожки, нетрудно заметить, что внешние сектора длиннее внутренних. Это означает, что внешние дорожки сильно недоиспользованы, потому как в теории они могут содержать значительно больше секторов с той же линейной плотностью записи. В целях увеличения емкости и уменьшения неиспользуемого пространства на современных моделях жестких дисков применяется технология зонной записи (ZBR). По этой технологии дорожки группируются в зоны в зависимости от их расстояния от центра диска. В каждой зоне свое количество секторов на дорожку. Если двигаться от внутренней части диска к его внешнему краю, то каждая следующая зона содержит больше секторов на дорожку, чем предыдущая. Это позволяет более эффективно использовать внешние дорожки диска. В настоящее время в дисках применяется много сложных внутренних структур, так что не существует простого пути выяснения реальной геометрии диска. На самом деле, число секторов на дорожках варьируется, благодаря зонной записи, так что не существует какого либо определенного числа секторов на дорожку.

Не менее 16% суммарной рабочей поверхности дисков отводится под служебную информацию, которая обеспечивает нормальную работу винчестера. В первую очередь это инженерная зона (секторы конфигурации, таблицы дефектов, рабочие программы винчестера).

Оставшееся дисковое пространство делится на зоны (для большинства винчестеров – от 8 до 20) с различным числом секторов в каждой зоне. Не все секторы используются в качестве рабочих. Часть секторов являются запасными. При первоначальной разметке дисков на заводе-изготовителе производится проверка поверхности диска и информация об обнаруженных дефектных участках записывается в таблицу дефектов, которая размещается в инженерной зоне.

В процессе функционирования винчестера эта таблица используется для переназначения (переадресации) обращения к дефектным участкам (секторам) на обращение к хорошим секторам, которые как раз и размещаются на запасных дорожках. Ввиду важности служебной информации инженерная зона различных моделей накопителей может содержать от 2 до 6 копий, а сервометки прописываются с запасом по количеству и более сильным магнитным полем.

На любом жестком диске есть заводской дефект-лист. В процессе производства жесткий диск проходит специальный цикл технологических тестов, суммарное время прохождения которых варьируется, в зависимости от модели и емкости диска и составляет от 2-х до 24-х часов. Цель некоторых тестов - выявить потенциальные ошибки поверхности, т.е., спрогнозировать нестабильные сектора и занести их в заводскую таблицу дефектов, чтобы не допустить попадания на такие сектора данных пользователя. Обычными тестами чтения записи такие сектора выявить невозможно, поэтому применяются методы повышения вероятности ошибки. Для этого накопитель искусственно ухудшает характеристики электронной схемы канала чтения-записи, причем используются всевозможные вариации, благо, что современные микросхемы позволяют программировать практически все свои параметры (подробнее об этом - здесь). Таким образом, диск, сходящий с конвейера, имеет некоторое количество записей в его заводской таблице дефектов, и это совершенно нормальное явление. Естественно, эти дефекты незаметны для пользователя, и, к тому же, имеют "мягкую" природу.

Интерфейсы HDD

Интерфейсом накопителей называется набор электроники, обеспечивающий обмен информацией между контроллером устройства (кеш-буфером) и компьютером. В настоящее время в настольных ПК IBM-PC, чаще других, используются три разновидности интерфейсов:

- Integrated Drive Electronics - IDE, Enhanced Integrated Drive Electronics - EIDE
- SCSI (Small Computers System Interface)
- SATA (Serial ATA)

Интерфейс IDE

Для подключения устройств IDE существует несколько разновидностей интерфейса:

- ATA (AT Attachment), он же AT-BUS - 16-битный интерфейс подключения к шине компьютера AT. В настоящее время это наиболее распространенный 40-проводной сигнальный и 4-проводной питающий интерфейс для подключения дисковых накопителей к компьютерам класса AT. Для миниатюрных (2,5" и меньших) накопителей используют 44-проводной кабель, по которому передается и питание.
- PC Card ATA - 16-битный интерфейс с 68-контактным разъемом PC Card (PCMCIA) для подключения к блокнотным ПК.
- XT IDE (8-бит), он же XT-BUS - 40-проводной интерфейс, похожий на ATA, но несовместимый с ним.
- MCA IDE (16-бит) - 72-проводный интерфейс, предназначенный специально для шины и накопителей PS/2. Как и компьютеры PS/2, по крайней мере в нашей стране устройства с этим интерфейсом встречаются редко.
- ATA-2 - расширенная спецификация ATA, включает 2 канала, 4 устройства, PIO Mode 3, multiword DMA mode 1, Block mode, объем диска до 8 Гбайт, поддержка LBA и CHS.
- Fast ATA-2 разрешает использовать Multiword DMA Mode 2 (13,3 Мбайт/с), PIO Mode 4.
- ATA-3 - расширение ATA-2. Включает средства парольной защиты, улучшенного управления питанием, самотестирования с предупреждением приближения отказа - SMART (Self Monitoring Analysis and Report Technology).
- ATA/ATAPI-ATAPI-4 - расширение ATA-3, включающее режим Ultra DMA со скоростью обмена до 33 Мбайт/с и пакетный интерфейс ATAPI.
- E-IDE (Enhanced IDE) - расширенный интерфейс, введенный фирмой Western Digital. Реализуется в адаптерах для шин PCI и VLB, позволяющий подключать до 4 устройств (к двум каналам), включая CD-ROM и стриммеры (ATAPI). Поддерживает PIO Mode 3, multiword DMA mode 1, объем диска до 8 Гбайт, LBA и CHS. С аппаратной точки зрения практически полностью соответствует спецификации ATA-2.

Устройства ATA IDE, E-IDE, ATA-2, Fast ATA-2, ATA-3 и ATA/ATAPI-4 электрически совместимы, степень логической совместимости достаточно высока (все базовые возможности ATA доступны). Однако для полного использования возможностей всех расширений необходимо соответствие спецификаций устройств, хост-адаптера и его программного обеспечения. В настоящее время наиболее широко распространен и четко стандартизован интерфейс, официально называемый ATA-2.

Интерфейс IDE разрабатывался как недорогая и производительная альтернатива высокоскоростным интерфейсам ESDI и SCSI. Интерфейс, предназначен для подключения двух дисковых устройств. Отличительной особенностью дисковых устройств, работающих с интерфейсом IDE состоит в том, что собственно контроллер дискового накопителя располагается на плате самого накопителя вместе со встроенным внутренним кэш-буфером. Такая конструкция существенно упрощает устройство самой интерфейсной карты и дает возможность размещать ее не только на отдельной плате адаптера, вставляемой в разъем системной шины, но и интегрировать непосредственно на материнской плате компьютера. Интерфейс характеризуется чрезвычайной простотой, высоким быстродействием, малыми размерами и относительной дешевизной.

На смену интерфейсу IDE пришло детище фирмы Western Digital - Enhanced IDE, или сокращенно EIDE. Долгое время это был лучший вариант для подавляющего большинства настольных систем. Жесткие диски EIDE заметно дешевле аналогичных по емкости SCSI-дисков и в однопользовательских системах не уступают им по производительности, а большинство материнских плат имеют интегрированный двухканальный контроллер для подключения четырех устройств. Что же появилось нового в Enhanced IDE по сравнению с IDE ?

Во-первых, это большая емкость дисков. Если IDE не поддерживал диски свыше 528 мегабайт, то EIDE поддерживает объемы до 8.4 гигабайта на каждый канал контроллера.

Во-вторых, к нему подключается больше устройств - четыре вместо двух. Раньше имелся только один канал контроллера, к которому можно было подключить два IDE устройства. Теперь таких каналов два. Основной канал, который обычно стоит на высокоскоростной локальной шине и вспомогательный.

В-третьих, появилась спецификация ATAPI (AT Attachment Packet Interface) дающая возможность подключения к этому интерфейсу не только жестких дисков, но и других устройств - стриммеров и дисководов CD-ROM.

В-четвертых - повысилась производительность. Накопители с интерфейсом IDE характеризовались максимальной скоростью передачи данных на уровне 3 мегабайт в секунду. Жесткие диски EIDE поддерживают несколько новых режимов обмена данными. В их число входит режим программируемого ввода-вывода PIO (Programmed Input/Output) Mode 3 и 4, которые обеспечивают скорость передачи данных 11.1 и 16.6 мегабайт в секунду соответственно. Программируемый ввод-вывод - это способ передачи данных между контроллером периферийного устройства и оперативной памятью компьютера посредством команд пересылки данных и портов ввода/вывода центрального процессора.

В пятых - поддерживается режим прямого доступа к памяти - Multiword Mode 1 DMA (Direct Memory Access) или Multiword Mode 2 DMA и Ultra DMA, которые поддерживают обмен данными в монопольном режиме (то есть когда канал ввода-вывода в течение некоторого времени обслуживает только одно устройство). Периферийные устройства обслуживает специальный контроллер DMA. Скорость при этом достигает 13.3 и 16.6 мегабайта в секунду, а при использовании Ultra DMA и соответствующего драйвера шины - 33 мегабайт в секунду. EIDE-контроллеры используют механизм PIO точно так же, как это делают и некоторые SCSI-адаптеры, но скоростные адаптеры SCSI работают только по методу DMA.

В шестых - расширена система команд управления устройством, передачи данных и диагностики, увеличен кеш-буфер обмена данными и существенно доработана механика.

Фирмы Seagate и Quantum вместо спецификации EIDE используют спецификацию Fast ATA для накопителей, поддерживающих режимы PIO Mode 3 и DMA Mode 1, а работающие в режимах PIO Mode 4 и DMA Mode 2 обозначают как Fast ATA-2.

В спецификации ATA фигурируют следующие компоненты:

Хост-адаптер - средства сопряжения интерфейса ATA с системной шиной (в простейшем случае - набор буферных схем между шинами ISA и ATA). Хостом будем называть компьютер с хост-адаптером интерфейса ATA.

Кабель-шлейф с двумя или тремя 40-контактными IDC-разъемами. В стандартном кабеле одноименные контакты всех разъемов соединяются вместе.

Ведущее устройство (Master) - периферийное устройство, в спецификации ATA официально называемое Device-0 (устройство-0).

Ведомое устройство (Slave) - периферийное устройство, в спецификации официально называемое Device-1 (устройство-1).

Если к шине ATA подключено одно устройство, оно должно быть ведущим. Если подключены два устройства, одно должно быть ведущим, другое - ведомым. Все иные варианты назначения устройств неработоспособны (если "хитрый" хост-адаптер не возьмет на себя некоторые функции ведущего устройства). О своей роли (ведущее или ведомое) устройства обычно "узнают" с помощью предварительно установленных конфигурационных джамперов. В редких случаях, когда применяется "кабельная выборка", о которой будет сказано ниже, роль устройства определяется его положением на специальном кабеле-шлейфе.

Оба подключенных устройства воспринимают команды от хост-адаптера одновременно. Однако исполнять команду будет лишь выбранное устройство. Если бит DEV=0, выбрано устройство-0, если бит DEV=1 - устройство-1. Выходные сигналы на шину ATA имеет право выводить только выбранное устройство. Такая система выбора устройства подразумевает, что начав операцию обмена с одним из устройств, хост-адаптер не может переключиться на обслуживание другого устройства той же шины ATA до завершения начатой операции обмена. Параллельно могут работать только устройства IDE, подключаемые к разным шинам (каналам) ATA. Спецификация ATA-4 определяет возможности обхода этого ограничения.

Выполняемая операция и направление обмена данными между устройством и хост-адаптером определяется предварительно записанной командой. Непременным компонентом устройства является буферная память. Ее наличие позволяет выполнять обмен данными в темпе, предлагаемом хост-адаптером (конечно, в пределах возможности устройства), без оглядки на внутреннюю скорость передачи данных между носителем и буферной памятью периферийного устройства.

Все информационные сигналы интерфейса передаются через 40-контактный разъем, у которого ключом является отсутствующий на вилке и закрытый на розетке контакт №20. Для соединения устройств спецификация требует использования плоского многожильного кабеля, длина которого не должна превышать 0,46 м (18"), допустимая емкость проводников не более 35 пФ. Специальные терминаторы стандартом не предусматриваются (они имеются в каждом устройстве и хост-адаптере), но если кабель с тремя разъемами (розетками) используют для подключения одного устройства, то и его и хост-адаптер рекомендуется подключать к противоположным концам кабеля. Для устойчивости работы в режиме UltraDMA при высокой скорости обмена рекомендуется применение сборки 40-контактных разъемов и 80-проводных кабелей, обеспечивающих чередование сигнальных цепей и проводов схемной земли.

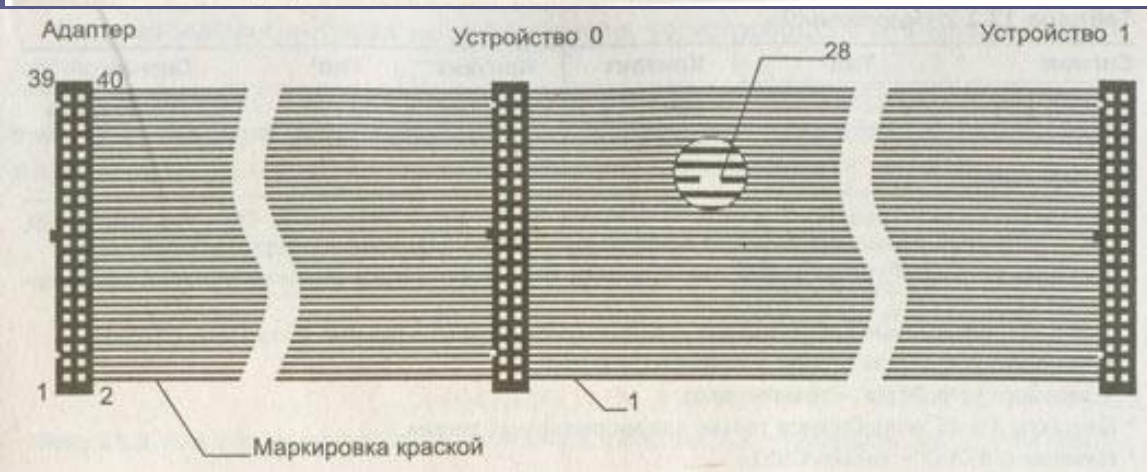


Сигнальный разъем

Сигнальный разъем IDE/EIDE и разъем питания



Интерфейсный кабель



Интерфейсный кабель с кабельной выборкой

Максимально возможная скорость передачи данных для интерфейса IDE (он же ATA)

single word DMA 0	2.1 MByte/s
PIO mode 0	3.3 Mbyte/s
single word DMA 1, multi word DMA 0	4.2 MByte/s
PIO mode 1	5.2 MByte/s
PIO mode 2, single word DMA 2	8.3 MByte/s

Максимально возможная скорость передачи данных для интерфейса EIDE (он же ATA-2)

PIO mode 3	11.1 MB/s
multi word DMA 1	13.3 MB/s
PIO mode 4, multi word DMA 2	16.6 MB/s

Максимально возможная скорость передачи данных для интерфейса Ultra-ATA (он же Ultra DMA/33)

multi word DMA 3	33.3 MB/s
------------------	-----------

Система команд АТА

Стандарт АТА задает систему команд, ориентированную, опять-таки, на накопители на магнитных дисках. Для операций, связанных с обменами данных, имеются команды, использующие обмен данных в режиме PIO или по каналу DMA. Режимы PIO и DMA, включая и Ultra DMA, программируются для устройств специальными командами.

Свое основное назначение устройства АТА реализуют с помощью команд чтения и записи данных, минимальной адресуемой единицей которых является 512-байтный сектор.

Команды чтения и записи буфера служат для обмена информации в режиме PIO с 512-байтной буферной памятью устройства (но не сектором носителя).

Команда форматирования трека по входным параметрам специфична для каждого устройства, и ее использование в целевой системе не рекомендуется. Многие устройства ее отвергают как недопустимую.

По команде поиска устройство устанавливает головки на заданный цилиндр, трек и считывает идентификатор сектора. Команда рекалибровки заставляет устройство найти нулевой цилиндр. Эту команду обычно применяют при обработке ошибок: часто после такого "встряхивания" ошибка не повторяется.

Для накопителей со сменными носителями в АТА-2 были предназначены команды загрузки и выгрузки, подтверждения смены носителя, блокировки и разблокировки дверец, их реализация специфична для каждой модели устройства. В АТА-4 набор этих команд сокращен, здесь может работать механизм уведомления о смене носителя Removable Media Status Notification.

Для запоминающих устройств на флэш-памяти в АТА-4 ввели специальную группу команд Команда установки свойств имеет ряд подкоманд, позволяющих управлять режимом обмена, кэшированием, параметрами режимов энергосбережения и т.п.

Команда диагностики, в отличие от остальных всегда адресуясь к нулевому устройству, выполняется одновременно обоими. О ее результате устройство-1 сообщает не хост-контроллеру, а устройству-0. При этом состояние обоих устройств определяется по диагностическому коду, который потом считывается из регистра ошибок нулевого устройства.

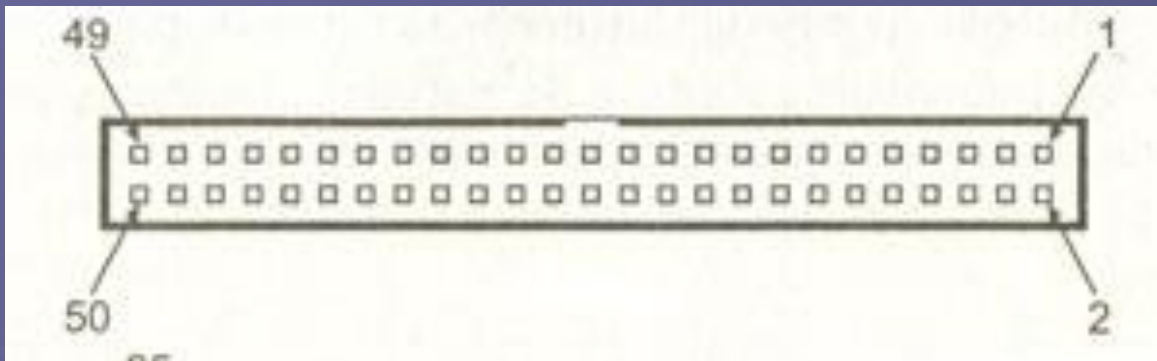
Команда загрузки микрокода позволяет модифицировать firmware - встроенное программное обеспечение устройства. В зависимости от кода в регистре свойств загруженный микрокод будет действовать временно, то есть до выключения питания, или постоянно. Эта команда, как и флэш-BIOS, является палкой о двух концах: возможность модификации кода может обернуться выводом из строя устройства загрузкой некорректного микрокода.

Средства управления энергопотреблением - Power Management не являются обязательными. С точки зрения потребления различают следующие состояния, перечисленные в порядке возрастания потребления: Sleep - "заснувшее" устройство потребляет минимум энергии, "разбудить" его может только сброс. Время "пробуждения" может доходить до 30 с. В состоянии Standby Mode (дежурный режим) устройство способно принимать команду по интерфейсу, но для доступа к носителю может потребоваться такое же большое время. И наконец, в активном режиме Active Mode устройство все запросы обслуживает за кратчайшее время. Устройства могут поддерживать и расширенное управление энергопотреблением APM (Advanced Power Management). При этом задается уровень APM Level, определяющий степень активности: 01h - минимальное потребление, FEh - максимальная производительность. Уровень выше 80h не позволяет устройству останавливать шпиндельный двигатель. APM управляется подкомандами Set Features. Устройства с пакетным интерфейсом для управления энергопотреблением могут использовать и команды пакетного протокола. Начиная с ATA-3 в стандарт введена группа команд защиты - Security. Защищенное устройство по включению питания или аппаратному сбросу будет находиться в заблокированном состоянии, при котором любой доступ к информации носителя запрещается. Система защиты поддерживает два пароля - главный (Master Password) и пользовательский (User Password). Разблокировать можно только специальной командой, в которой необходимо указать пароль пользователя. Если пароль утерян, то можно использовать главный пароль, но доступ к данным будет получен только если была выбрана высокая степень защиты (High). Если была выбрана максимальная степень защиты, то разблокировать устройство по главному паролю можно только командой защитного стирания всей информации с носителя. Для осложнения подбора пароля (его длина составляет 32 байт) служит счетчик неудачных попыток разблокировки, по срабатывании которого команды разблокировки будут отвергаться до выключения питания или аппаратного сброса.

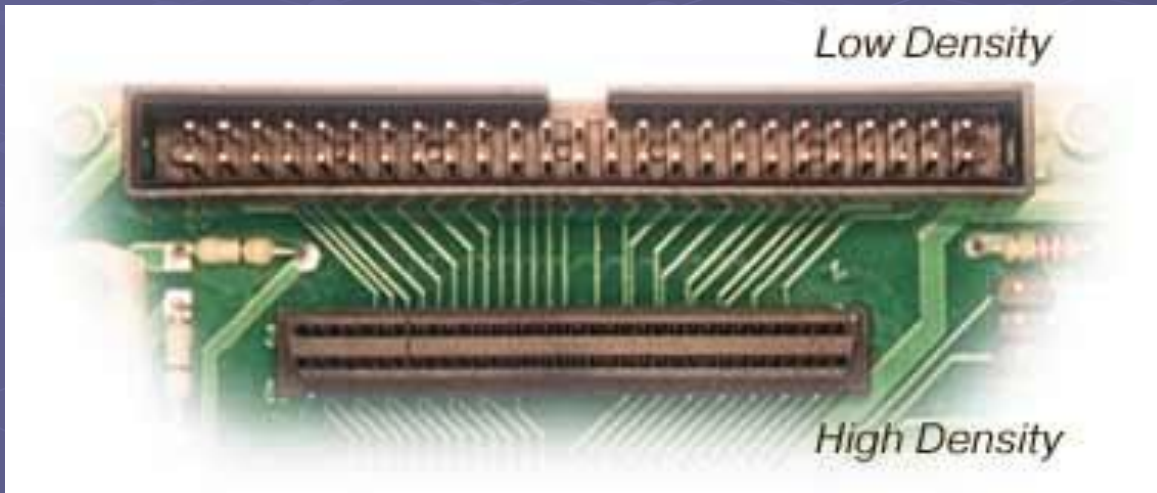
Интеллектуальный многофункциональный интерфейс SCSI

SCSI был разработан еще в конце 70-х годов в качестве устройства сопряжения компьютера и интеллектуального контроллера дискового накопителя. Интерфейс SCSI является универсальным и определяет шину данных между центральным процессором и несколькими внешними устройствами, имеющими свой контроллер. Помимо электрических и физических параметров, определяются также команды, при помощи которых, устройства, подключенные к шине осуществляют связь между собой. Интерфейс SCSI не определяет детально процессы на обеих сторонах шины и является интерфейсом в чистом виде.

Применяются в основном два стандарта - SCSI-2 и Ultra SCSI. В режиме Fast SCSI-2 скорость передачи данных доходит до 10 мегабайт в секунду при использовании 8-разрядной шины и до 20 мегабайт при 16-разрядной шине Fast Wide SCSI-2. Появившийся позднее стандарт Ultra SCSI отличается еще большей производительностью - 20 мегабайт в секунду для 8-разрядной шины и 40 мегабайт для 16-разрядной. В новейшем SCSI-3 увеличен набор команд, но быстродействие осталось на том же уровне. Все применяющиеся сегодня стандарты совместимы с предыдущими версиями "сверху - вниз", то есть к адаптерам SCSI-2 и Ultra SCSI можно подключить старые SCSI-устройства. Интерфейс SCSI-Wide, SCSI-2, SCSI-3 - стандарты модификации интерфейса SCSI, разработаны комитетом ANSI. Общая концепция усовершенствований направлена на увеличение ширины шины до 32-х, с увеличением длины соединительного кабеля и максимальной скорости передачи данных с сохранением совместимости с SCSI. Это наиболее гибкий и стандартизованный тип интерфейсов, применяющийся для подключения 7 и более периферийных устройств, снабженных контроллером интерфейса SCSI. Интерфейс SCSI остается достаточно дорогим и самым высокопроизводительным из семейства интерфейсов периферийных устройств персональных компьютеров, а для подключения накопителя с интерфейсом SCSI необходимо дополнительно устанавливать адаптер, т.к. немногие материнские платы имеют интегрированный адаптер SCSI.



Для подключения внутренних приборов применяются разъемы: для подключения 8-битных устройств – IDC-50 (сейчас его, как правило, называют Low Density); для подключения 16-битных устройств, в том числе и приборов LVD-68 – контактный разъем High Density. Low Density предназначен для подключения «узких» – Narrow (8-битных) устройств. High Density предназначен для подключения «широких» – Wide (16 битных) устройств,



High density 80 pin

В последнее время очень широкое распространение получили RAID-системы, базирующиеся на SCSI-интерфейсе. Такие системы подразумевают наличие возможности «горячей» замены отказавших дисков. Для этого разработан разъем, через который подключаются как сигнальные цепи, так и цепи питания жесткого диска. Это разъем "High density 80 pin".

Существует также Ultra2 SCSI (LVD) Low Voltage Differential Parallel SCSI Interface, т.е. низковольтный дифференциальный параллельный SCSI интерфейс. Этот вариант SCSI существенно отличается от всех своих предшественников по двум параметрам:

Скорость передачи увеличена до 80 MB/s

Максимальная длина соединительного кабеля может достигать 12 метров

Кроме этого, к одному шлейфу можно подключить до 15 устройств. Обратная совместимость, как это принято для SCSI устройств, также выдерживается и устройство с Ultra2 SCSI LVD можно подключить к обычному контроллеру SCSI. С этим интерфейсом выпускаются только жесткие диски в вариантах с 68-контактным разъемом и SCA.

Максимально возможная скорость передачи данных для интерфейса SCSI и максимально допустимые длины кабелей

Частота шины SCSI	8 бит (50-ти контактный разъем)		16 бит (68-и контактный разъем, Wide SCSI)	
	Максимальная скорость обмена	Максимальная длина кабеля	Максимальная скорость обмена	Максимальная длина кабеля
5 MHz (SCSI 1)	5 MB/s	6 m	N/A*	-
10 MHz (Fast SCSI, SCSI II)	10 MB/s	3 m	20 MB/s	3 m
20 MHz (Fast-20, Ultra SCSI)	20 MB/s	1.5 m	40 MB/s	1.5 m
40 MHz (Ultra-2 SCSI) LVD	N/A	N/A	80 MB/s	12 m

Интерфейс SATA

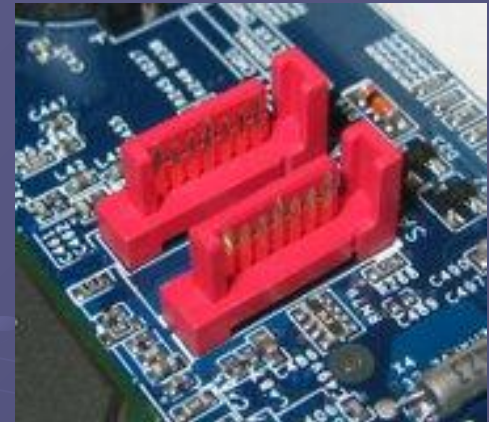
SATA (англ. Serial ATA) — последовательный интерфейс обмена данными с накопителями информации (как правило, с жёсткими дисками). SATA является развитием интерфейса ATA (IDE), который после появления SATA был переименован в PATA (Parallel ATA).

Первоначально стандарт SATA предусматривал работу шины на частоте 1,5 ГГц, обеспечивающей пропускную способность приблизительно в 1,2 Гбит/с (150 МБ/с). (20%-я потеря производительности объясняется использованием системы кодирования 8B/10B, при которой на каждые 8 бит полезной информации приходится 2 служебных бита). Пропускная способность SATA/150 незначительно выше пропускной способности шины Ultra ATA (UDMA/133). Главным преимуществом SATA перед PATA является использование последовательной шины вместо параллельной.

SATA/300

Стандарт SATA/300 работает на частоте 3 ГГц, обеспечивает пропускную способность до 2,4 Гбит/с (300 МБ/с). Впервые был реализован в контроллере чипсета nForce 4 фирмы Nvidia. Весьма часто стандарт SATA/300 называют SATA II или SATA 3.0. [1] Теоретически SATA/150 и SATA/300 устройства должны быть совместимы (как SATA/300 контроллер и SATA/150 устройство, так и SATA/150 контроллер и SATA/300 устройство) за счёт поддержки согласования скоростей (в меньшую сторону), однако для некоторых устройств и контроллеров требуется ручное выставление режима работы (например, на НЖМД фирмы Seagate, поддерживающих SATA/300 для принудительного включения режима SATA/150 предусмотрен специальный джампер).

Стандарт SATA предусматривает возможность увеличения скорости работы до 600 МБ/с (6 ГГц).



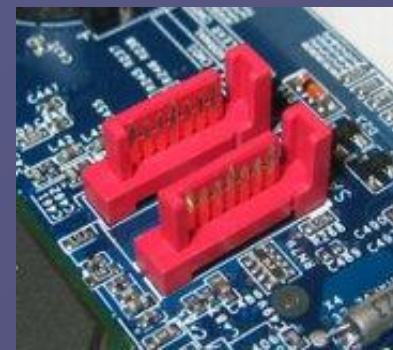
SATA использует 7-контактный разъём вместо 40-контактного разъёма у PATA. SATA-кабель имеет меньшую площадь, за счёт чего уменьшается сопротивление воздуху, обдуваемому комплектующие компьютера; улучшается охлаждение системы.

SATA-кабель за счёт своей формы более устойчив к многократному подключению. Питающий шнур SATA так же разработан с учётом многократных подключений. Разъём питания SATA подаёт 3 напряжения питания: +12 В, +5 В и +3,3 В; однако современные устройства могут работать без напряжения +3,3 В, что даёт возможность использовать пассивный переходник с стандартного разъёма питания IDE на SATA. Ряд SATA устройств поставляется с двумя разъёмами питания: SATA и Molex.

Стандарт SATA отказался от традиционного для PATA подключения по два устройства на шлейф; каждому устройству полагается отдельный кабель, что снижает задержки при одновременной работе двух устройств на одном кабеле, уменьшает возможные проблемы при сборке (проблема конфликта Slave/Master устройств для SATA отсутствует), устраняет возможность ошибок при использовании нетерминированных PATA-шлейфов.

Стандарт SATA предусматривает горячую замену устройств и функцию очереди команд (NCQ).

SATA устройства используют два разъёма — один, семи контактный, для подключения шины данных и второй, 15-ти контактный, для подключения питания. Стандарт SATA предусматривает возможность использовать вместо 15-ти контактного разъёма питания стандартный 4-х контактный разъём Molex. Использование одновременно обоих типов силовых разъёмов может привести к повреждению устройства.



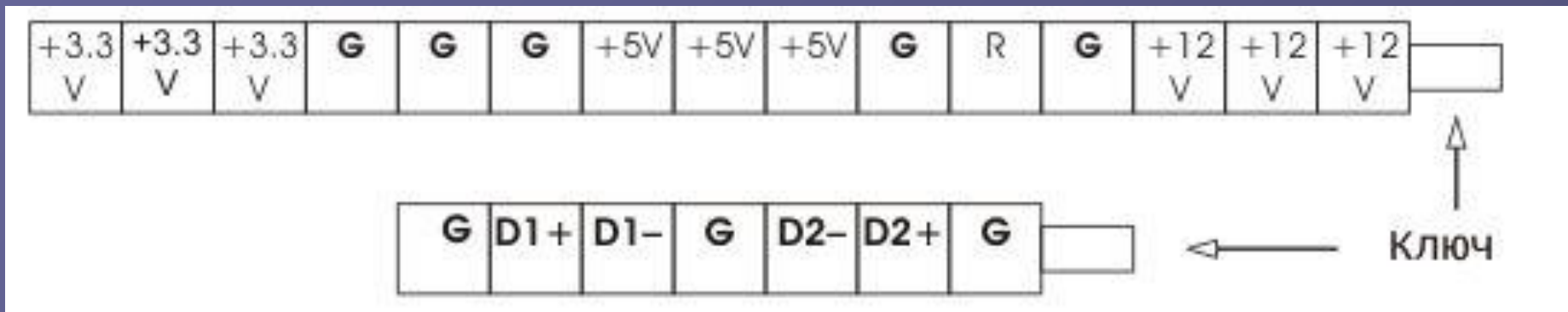
Сигнальный разъём



Сигнальный кабель



Разъём питания



G — заземление (англ. Ground)

R — зарезервировано

D1+, D1-, D2+, D2- — два канала передачи данных (от контроллера к устройству и от устройства к контроллеру соответственно). Для передачи сигнала используется технология LVDS, провода каждой пары (D1+, D1- и D2+, D2-) являются экранированными витыми парами.

eSATA (External SATA) — интерфейс подключения внешних устройств, поддерживающий режим «горячей замены» (англ. Hot-plug). Был создан несколько позже SATA (в середине 2004).

Основные особенности eSATA:

Требует для подключения два провода: шину данных и силовой кабель.

Ограничен по длине кабеля данных (около 2 м).

Средняя практическая скорость передачи данных выше, чем у USB или IEEE 1394.

Существенно меньше нагружается центральный процессор.

Интерфейс SAS (англ. Serial Attached SCSI) обеспечивает подключение по физическому интерфейсу, аналогичному SATA, устройств, управляемых набором команд SCSI. Обладая обратной совместимостью с SATA, он даёт возможность подключать по этому интерфейсу любые устройства, управляемые набором команд SCSI — не только НЖМД, но и сканеры, принтеры и др. По сравнению с SATA, SAS обеспечивает более развитую топологию, позволяя осуществлять параллельное подключение одного устройства по двум или более каналам. Так же поддерживаются расширители шины, позволяющие подключить несколько SAS устройств к одному порту.

Интерфейс SATA II

Введен с 2004 г. Увеличилась пропускная способность (со 150 до 300 MBps). Учитывая, что скорость чтения с одиночного диска на данный момент приближается к 70 MBps, пропускной способности первой версии стандарта скорее всего с головой хватит на ближайшие несколько лет.

Во-вторых, поддержка Native Command Queuing (NCQ), или технологии маршрутизации команд, стала фактически неотъемлемой частью стандарта SATA II, до этого же NCQ являлась необязательным дополнением SATA 1.0. NCQ позволяет переупорядочивать до 32 команд чтения/записи жесткого диска с целью достижения оптимальной производительности и снижения износа его механизмов. Ее работу можно проиллюстрировать следующим примером: предположим, диску поступают подряд несколько команд на чтение секторов с номерами 3000, 2000, 7000, 5000. Диск без NCQ считал бы сектора именно в этом порядке, в то время как диск с NCQ изменит последовательность чтения на 2000, 3000, 5000, 7000, совершив при том меньшее количество перемещений головок. Кроме того, результаты тестов показали, что вследствие такого упорядочивания в части задач (к примеру, дефрагментации) может наблюдаться ощутимый прирост производительности. Для использования данной технологии необходимы три условия: контроллер SATA II или SATA 1.0 с поддержкой NCQ (на плате или отдельный), драйверы для операционной системы с поддержкой команд NCQ и жесткий диск с NCQ. С первым и вторым на данный момент проблем нет – контроллеры SATA II имеются на всех материнских платах с чипсетами i915/925 и выше и nForce4, соответствующие драйверы для них уже есть. Однако в связи с тем, что часть работы с NCQ выполняет драйвер, при включении данной технологии может наблюдаться некоторое повышение загрузки процессора.

И в-третьих, добавлена функция горячего подключения, ранее являвшаяся опциональной. После установки соответствующих драйверов жесткий диск стандарта SATA II в системе определяется как съемное устройство и может быть в любой момент безопасно отключен. Что еще более удобно – в комплекте с некоторыми материнскими платами поставляется специальная планка на заднюю стенку с двумя SATA-разъемами и разъемом питания, благодаря чему можно подключить SATA II-винчестер, не вскрывая системный блок и не используя дополнительные приспособления вроде USB- или FireWire-карманов.

В целом, вторая версия SATA является скорее стандартизацией тех возможностей SATA, которые ранее были отданы производителям жестких дисков/контроллеров, а не чем-то радикально новым. Тем не менее на данный момент уже имеет смысл обращать внимание на тип и версию интерфейса винчестера, а при покупке отдавать предпочтение более новой версии, не забывая, конечно, и о цене.

RAID

RAID (англ. redundant array of independent/inexpensive disks) — дисковый массив независимых дисков. Служат для повышения надёжности хранения данных и/или для повышения скорости чтения/записи информации (RAID 0)

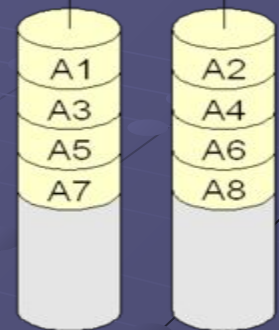
Уровни RAID

RAID 0 («Striping») — дисковый массив с отсутствием избыточности. Информация записывается сквозь все диски последовательно, разбиваясь на блоки данных (A_i). За счёт этого существенно повышается производительность, но страдает надёжность всего массива. (При выходе из строя любого из входящих в RAID 0 винчестеров полностью и безвозвратно пропадает вся информация). В соответствии с теорией вероятности, надёжность массива RAID 0 равна произведению надёжностей составляющих его дисков, каждая из которых меньше единицы, т. о. совокупная надёжность заведомо ниже надёжности любого из дисков.

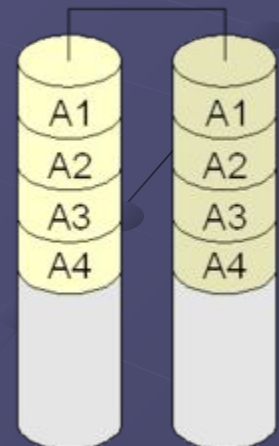
RAID 0 может быть реализован как программно так и аппаратно

RAID 1 (Mirroring — «зеркало»). отказоустойчивый массив из пары дисков. При записи данные с первого диска дублируются на втором (так называемое зеркалирование). Это эффективное и сравнительно простое в реализации решение обладает существенным недостатком — объем дискового пространства массива равен емкости наименьшего диска (например, если объединить в RAID1 диски на 30 и 40 Гбайт, то доступно будет только 30 Гбайт). Имеет защиту от выхода из строя половины имеющихся аппаратных средств (в частном случае — одного из двух жёстких дисков), обеспечивает приемлемую скорость записи и выигрыш по скорости чтения за счёт распараллеливания запросов. Зеркало на многих дисках — RAID 1+0. При использовании такого уровня зеркальные пары дисков выстраиваются в «цепочку», поэтому объём полученного тома может превосходить ёмкость одного жёсткого диска. Достоинства и недостатки такие же, как и у уровня RAID 1. Как и в других случаях, рекомендуется включать в массив диски горячего резерва HotSpare из расчёта один резервный на пять рабочих.

RAID 0



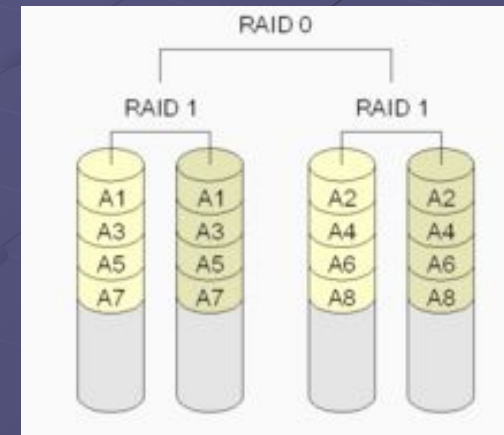
RAID 1



RAID 5



RAID 10



RAID 5 использует независимый доступ к дискам, так что запросы к разным дискам могут выполняться параллельно. Избыточность достигается путем размещения блоков четности циклически по всем дискам массива. При сбое какого-либо диска информацию можно восстановить, используя данные на оставшихся дисках. Емкость такого массива равна $(N-1) \times$ (емкость наименьшего диска). Минимальное число дисков для реализации массива равно трем. Причем для вычисления четности (а это нужно при каждой операции записи) необходимо вычислить функцию XOR для записываемых битов. Большинство контроллеров используют для этого специализированные процессоры, но некоторые выполняют операцию программно, силами центрального процессора.

Итак, независимый доступ дает возможность массивам RAID5 достигать высокой производительности в серверных приложениях с высокой частотой транзакций, при этом достаточно экономно расходуя дисковое пространство для обеспечения избыточности. RAID5 — очень распространенное и разумное решение для самых разных серверов.

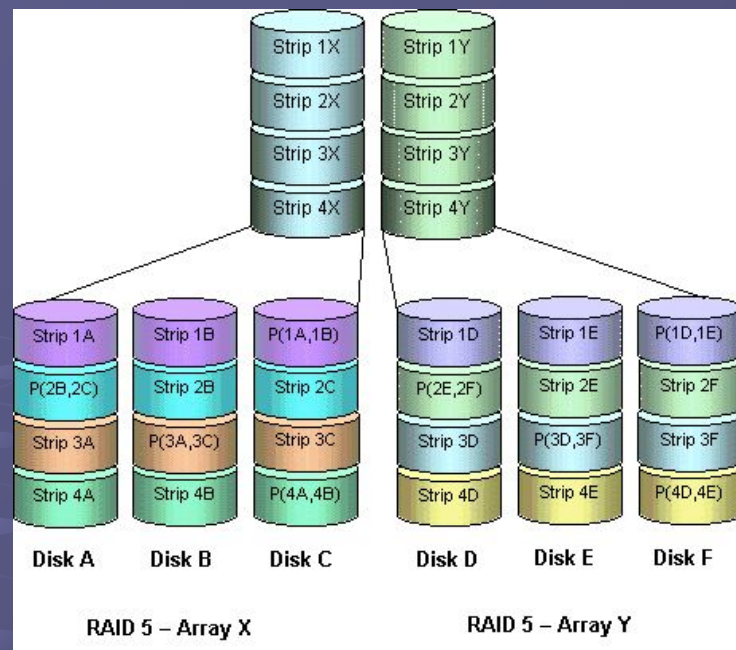
RAID10 — массив RAID0, элементами которого являются массивы RAID1.

Цель — объединить высокую производительность RAID0 с отказоустойчивостью RAID1. В последнее время часто встречается в недорогих контроллерах.

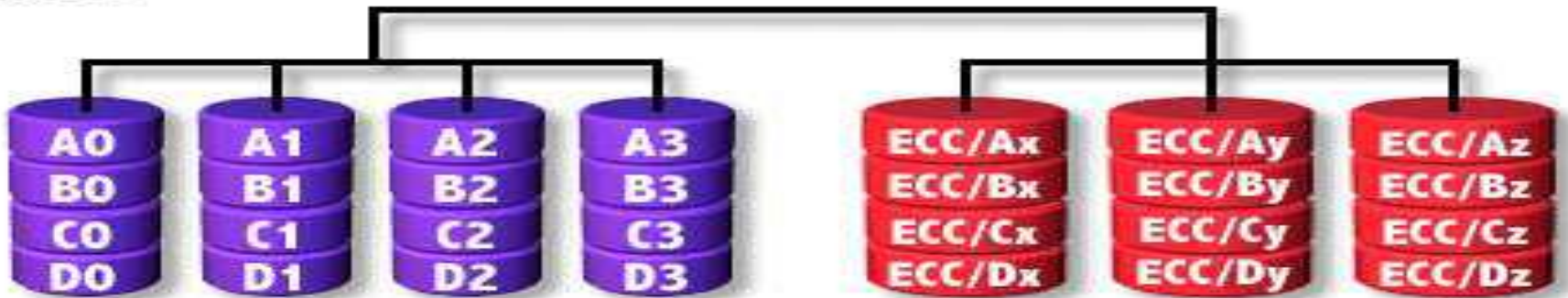
RAID 2 зарезервирован для массивов, которые применяют код Хемминга.

RAID 3, 4, 5, 6 используют четность для защиты данных от одиночных неисправностей.

RAID50 объединяет отказоустойчивость и высокую скорость обработки транзакций RAID5 с высокой скоростью потокового чтения/записи RAID0. RAID50 представляет собой RAID0-массив, элементами которого являются массивы RAID5. Это хорошее серверное решение, но для его реализации требуется как минимум шесть дисков.



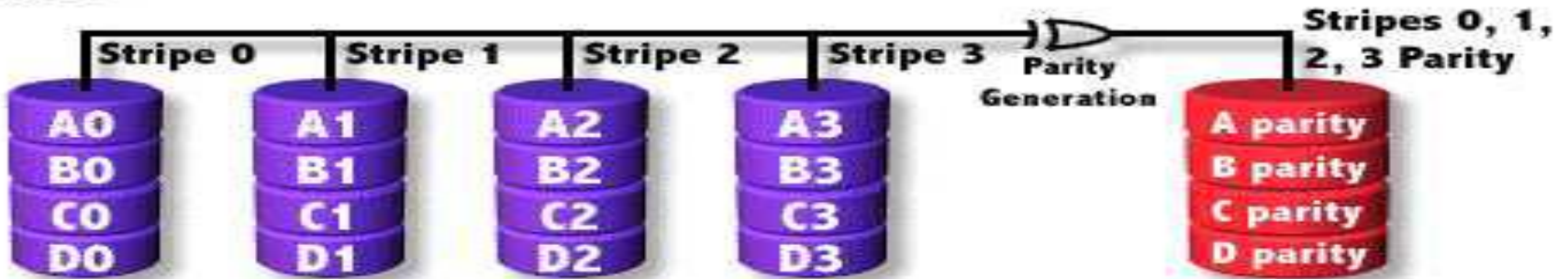
RAID 2



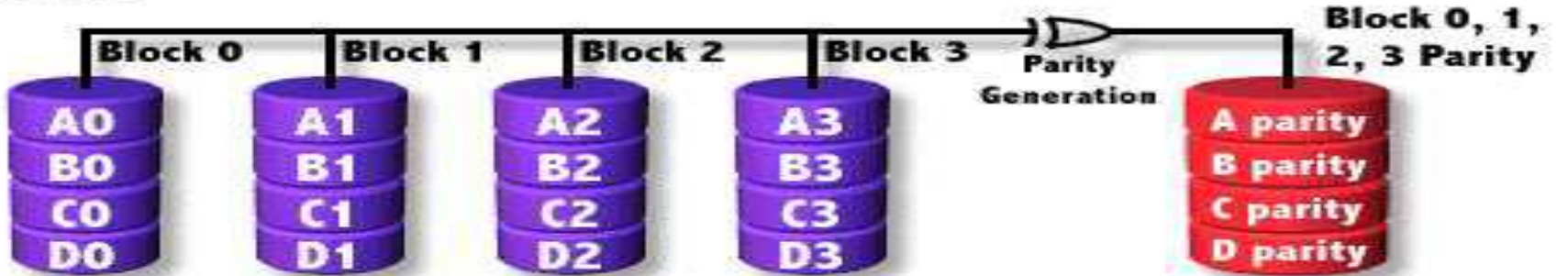
A0 to A3=Word A; B0 to B3 = Word B;
C0 to C3=Word C; D0 to D3 = Word D

ECC/Ax to Az=Word A ECC; ECC/Bx to Bz = Word B ECC;
ECC/Cx to Cz=Word C ECC; ECC/Dx to Dz = Word D ECC

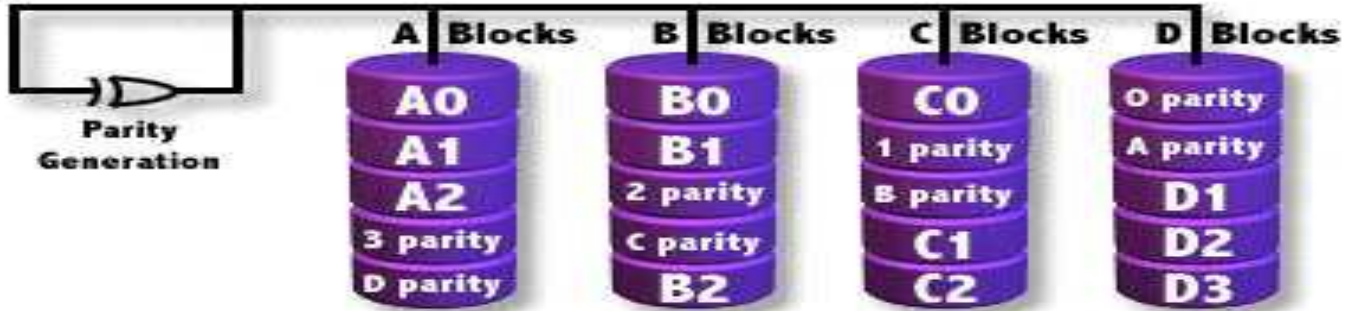
RAID 3



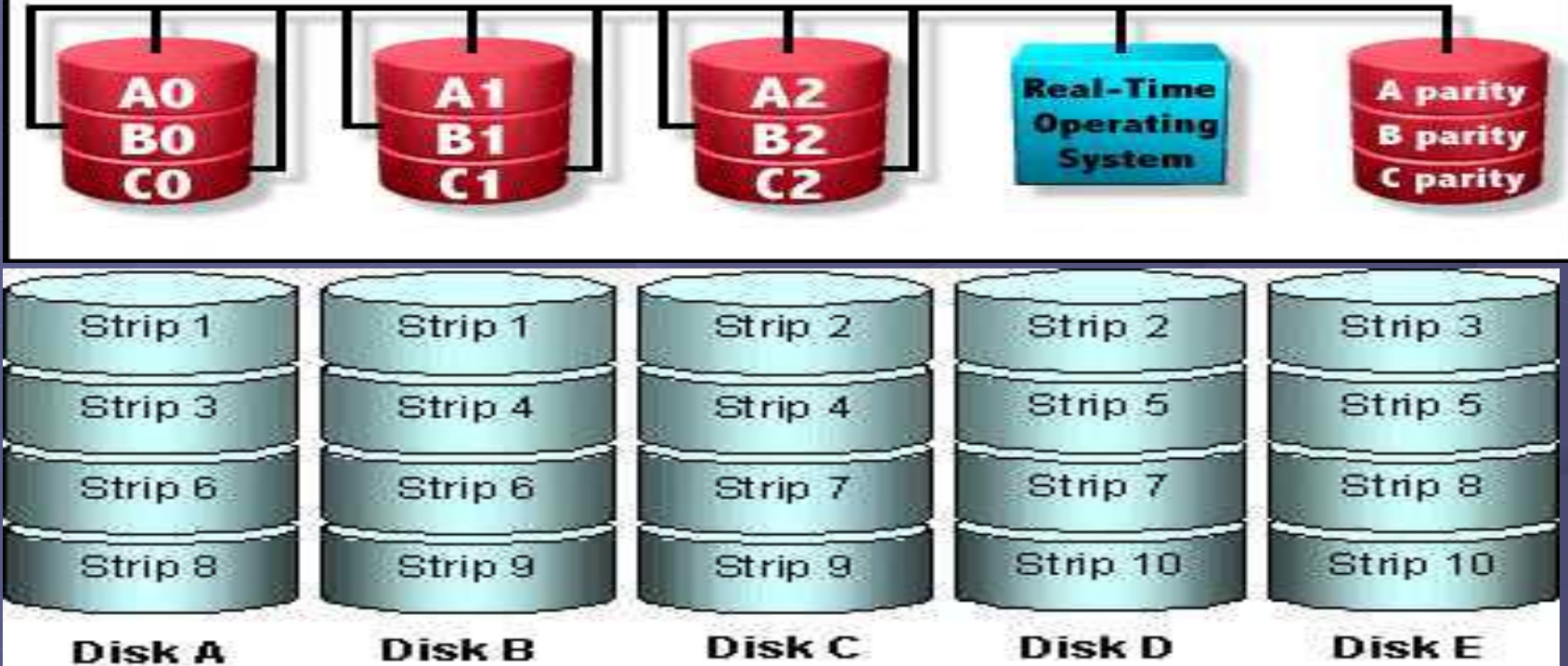
RAID 4



RAID 6



RAID 7®



RAID 1E