

Лекция 06: Жесткие диски.

Основными носителями информации в современных компьютерах являются

- жёсткие диски (HDD)
- твердотельные диски (SSD)
- оптические диски (DVD и BD)
- флэш-накопители

Они отличаются по объему, скорости доступа, надёжности и цене. Так же принципиально у них отличается способ хранения информации.

Технически самый сложный в изготовлении и одновременно самый неустойчивый к механическим повреждениям во время работы это жесткий диск. Однако с точки зрения соотношения цена/объем информации жёсткий диск всё ещё вне конкуренции, так что начнем рассмотрение именно с него.

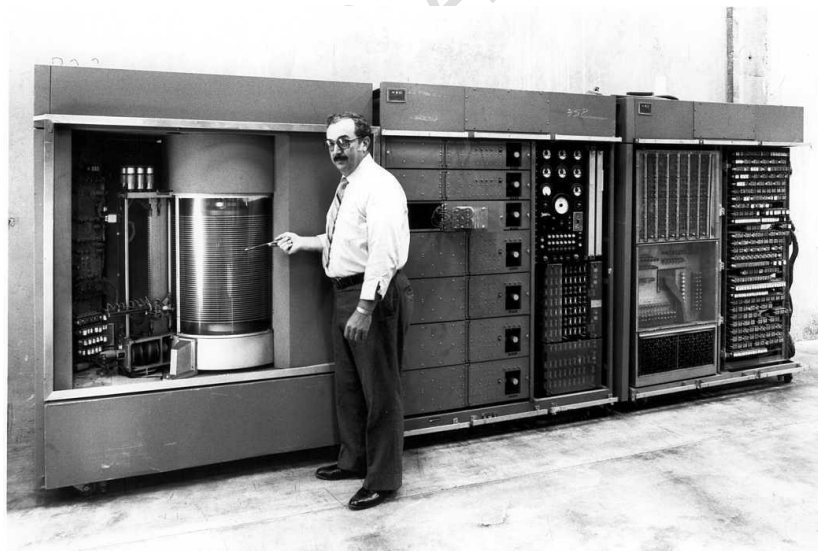
История создания и развития.

Первый жёсткий диск был выпущен компанией IBM 350 в 1956 году. Размеры были с холодильник, вес около тонны, а ёмкость всего 3,5 Мб. В качестве хранилища данных использовались 50 алюминиевых пластин диаметром 24», ну или более привычные нам 61 см.

На фото ниже блок пластин диска погружают на самолет для отправки потребителю.



Но кроме блока пластин была и управляющая электроника и интерфейс, так что в сборе это выглядело несколько масштабнее



Однако полноценным накопителем стоит считать устройство IBM 3340, выпущенное в 1973 году. Именно в нем был впервые применен отдельный чип для управления вращением дисков и перемещением считывающей головки, герметичный корпус для устранения пыли и так далее.



Технически это было вообще два жестких диска в одном корпусе, каждый по 30 Мб ёмкостью, причём одна из пластин была съёмной. По этой причине в маркировке этого устройства обычно писали 30-30, что у американцев вызывало ассоциации с винтовкой Winchester 30/30. Вскоре этот диск стали называть «винчестер», а в дальнейшем и все жесткие диски получили такое прозвище. Дальнейшее движение было в сторону увеличения ёмкости накопителя, размеры и вес производителей не заботили. Появление персональных компьютеров потребовало более компактных жёстких дисков.

- В 1980 году был выпущен первый накопитель формата 5,25» ST-506 компании Seagate объемом 5 Мб (для сравнения, в

«больших» компьютерах на этот момент объем был около 1 Гб).

- В 1983 году был выпущен первый диск размером 3,5».
- К 1990 году объем дисков достиг объема в 320 Мб.
- В 1991 году IBM выпускает первый диск формата 2,5», ёмкостью 63Мб. Появление этого формата дисков обусловлено появлением нового класса устройств – ноутбуков.
- Одним из способов ускорения работы диска было увеличение скорости вращения дисков. В 1992 году выпущены диски со скоростью 7200 об/мин. В 1996 году серверные диски Seagate Cheetah стали работать на скорости 10000 об/мин. В 2000 году появились серверные диски IBM и SeaGate со скоростью 15000 об/мин. На этом «гонка» скоростей закончилась – произвести диск с большей скоростью вращения пластин не удалось.
- В 1999 году выпущены диски microdrive для сверхкомпактных компьютеров. В силу сложности и хрупкости конструкции большого распространения они не получили.
- К 2005 году достигнута ёмкость в 500 Гб.
- В 2006 году появились «гибридные» диски – в жёсткий диск встроен небольшой флэш накопитель.
- В 2013 году выпускаются первые накопители с вновь герметичным корпусом, однако заполненные не воздухом, а гелием. Наличие менее плотной гелиевой атмосферы позволило уменьшить аэродинамические завихрения в корпусе. Что позволило в свою очередь сделать компоновку внутри накопителя более «плотной».

На текущий момент ситуация такова:

1. ёмкости настольных дисков достигли 20 Тб (первым диском

был WD Ultrastar DC HC650, представленный в 2019 году для использования в дата-центрах). На конец 2021 года диски 20 Тб представили и WD и Seagate. В основе жестких дисков 20 ТБ лежат привычные технологии:

1. заполнение гелием
2. пластины емкостью 2,2 ТБ
3. применение метода записи CMR.

Пластины обслуживаются блоком из 18 головок, которые поддерживают технологию двумерной магнитной записи (Two-dimensional magnetic recording, TDMR) для операций с очень плотно расположенными дорожками.

2. частоты вращения дисков остановились на 7200 об/мин (5400 об/мин для «тихих» моделей и 10000/15000 об/мин для серверных дисков)
3. из производителей остались только Seagate, Western Digital, Toshiba

Форм-факторов (габаритных типоразмеров) у жёстких дисков немного и различаются они по диаметру магнитных дисков

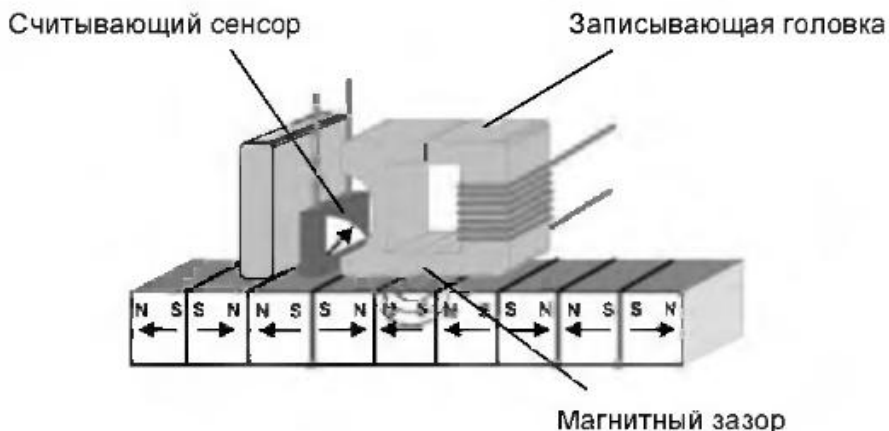
1. 3,5» используются в настольных компьютерах, серверах и дата центрах
2. 2,5» используются в ноутбуках и компактных серверах.

Остальные форм-факторы встречаются редко или уже не используются совсем.

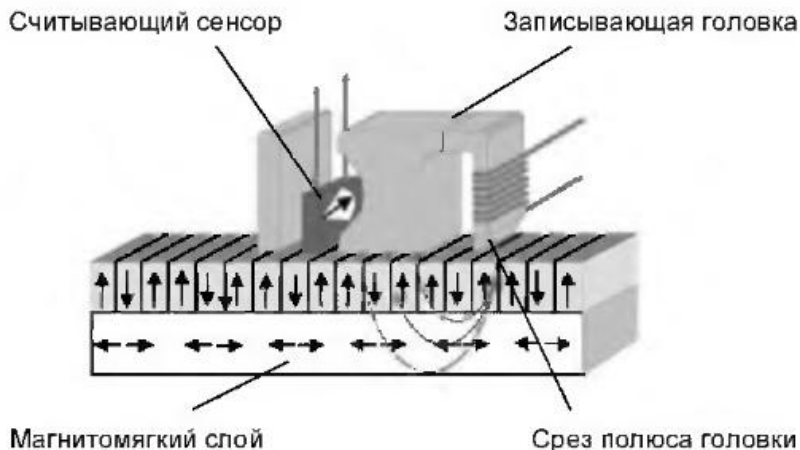
Принцип работы.

В работе используется принцип магнитной записи информации: под действием электромагнитного сигнала головки области поверхности диска (домены) меняют в определенном порядке свое магнитное поле. Домены могут иметь два направления полярности

магнитного поля, что и служит основой для записи информации. При считывании магнитные поля доменов приводят к изменению тока в катушке за счёт эффекта индукции, что позволяет считать записанную в доменах информацию. Первой физической технологией записи был **метод параллельной записи**.



Домены лежат вдоль поверхности диска и последовательно намагничиваются или считываются головкой. Физические размеры доменов получаются довольно большими — максимальная плотность записи по такой технологии не больше 23 Гбит на см^2 . Основным на данный момент способом хранения является **метод перпендикулярной записи**. Домены располагаются перпендикулярно поверхности диска, что позволяет существенно снизить площадь одного домена и повысить плотность записи информации (до 75 Гбит на см^2). Изменилась конструкция головки: теперь один полюс (рабочий) стал очень узким и напряженность магнитного поля на нём очень высока.



На втором полюсе, который намного шире, напряженность магнитного поля существенно меньше и её недостаточно для перемагничивания доменов. Второй слой на диске, магнитомягкий слой, служит для равномерного распределения магнитного поля, что так же позволяет не затрагивать магнитный момент соседних доменов.

Разработка этой технологии не окончена и в настоящее время имеется как минимум три способа улучшения технологии «холодной» записи

- EAMR – Energy Assisted Magnetic Recording (по сути это MAMR технология)
- TDMR – Two Dimensional Magnetic Recording: технология считывания несколькими головками. Применяется в дисках объемом от 14 Тб.
- SMR – Shingled Magnetic Recording: технология черепичной записи. При этом способе записи магнитные домены соседних дорожек частично перекрывают друг друга. Это приводит к уплотнению записи, но негативно сказывается на общей производительности устройства. Так же это может

вызывать сбои при работе таких накопителей в системах RAID. В мае 2020 года компания WD попала под иск за то, что не указывала на ряде своих накопителей факт использования SMR в них.

Все эти технологии легко могут быть объединены в одно устройство, как это сделал WD, выпустив в конце 2019 года серию дисков, объемом 20 Тб.

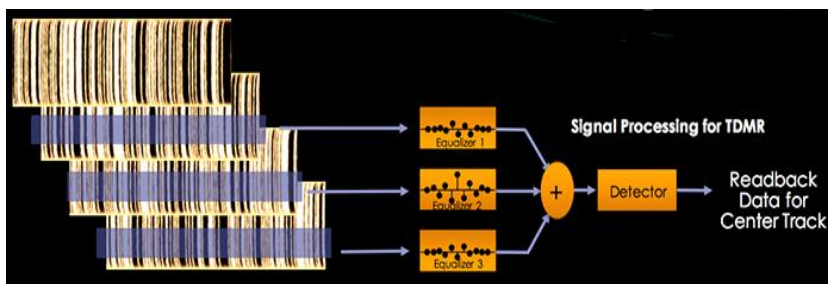
Перспективной разработкой является **метод тепловой магнитной записи**. Используется метод точечного нагрева магнитного слоя для изменения намагниченности, после остывания намагниченность сохраняется. Теоретически это позволит повысить плотность записи до 7,75 Тбит на см². Первый диск выпущен компанией SeaGate в декабре 2018 года. Вариант технологии от Seagate был назван HARM (Heat-Assisted Magnetic Recording) и первым диском оказался 16-ти Тб диск.

В 2019 году каждый производитель выпустил свой вариант тепловой магнитной записи:

- MARM (MicroWave Assisted Magnetic Recording) у WD, выпустившей по этой технологии первые партии 18 Тб дисков в конце 2019 года.
- Toshiba продолжает опыты и с HARM и с MARM технологиями.

Считывающая головка выполняется во всех дисках одинаково и использует магниторезистивный эффект (TMR Head). Изменилось только количество считывающих головок. В этой сфере имеется два альтернативных варианта

- Несколько головок на одном актуаторе. На конец 2019 года эта технология улучшает считывание одной дорожки, что позволяет сделать её меньше.



Вычитание шумов крайних головок из центральной помогает лучше отфильтровывать помехи при чтении.

- Два набора головок, с собственными актуаторами.



Эта разработка является на конец 2019 года всего лишь перспективным прототипом, так как пока не понятно как её предоставлять интерфейсу: как два независимых диска, как RAID массив, или решать проблемы согласования чтения внутри накопителя. Теоретически эта разработка должна вдвое повысить количество операций чтения/записи, однако до коммерческого решения пока ещё дело не дошло.

Устройство диска.

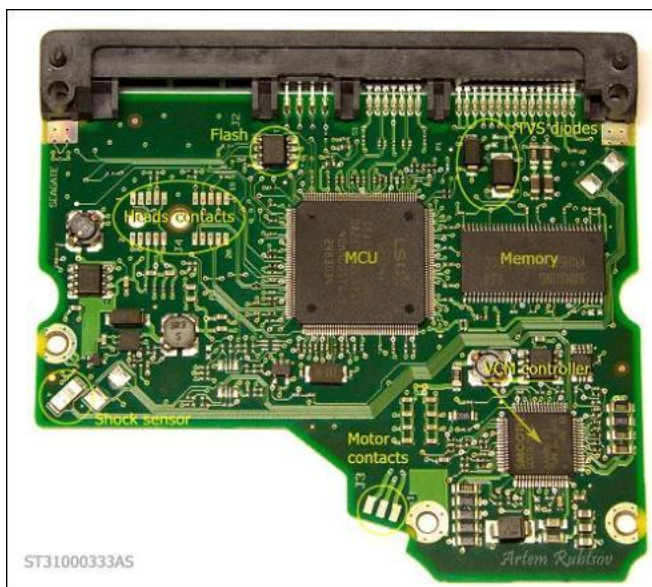
Рассмотрим устройство жёсткого диска на примере стандартного диска 3,5» от фирмы Seagate. Внешне современный диск представляет собой металлический корпус с платой на одной стороне.



В принципе – а не только внешне – жесткий диск и состоит из двух основных частей

- Металлический корпус, или **гермоблок**
- Блок электроники, или **плата управления**

Рассмотрим плату внимательнее.



На плате располагается собственно **чип контроллера**, отвечающий за передачу информации, управление остальными чипами на плате и обеспечивает работу интерфейса диска.

Чип памяти используется как кэш память для ускорения передачи информации. Так же этот чип используется для хранения прошивки в ходе работы диска. Копия прошивки хранится на отдельном чипе флэш памяти и подгружается на данный чип памяти для работы диска. Так что не весь объем чипа памяти используется под кэш при работе диска.

Следующий чип — **контроллер управления аппаратной частью**, или контроллер управления двигателем и головками чтения.

Дополнительно на плате есть разъемы питания и информационного интерфейса, датчики удара и два предохранителя.

Плата соединяется с **гермоблоком** (металлическим корпусом с механической частью диска) при помощи двух групп контактов

- группа контактов двигателя
- группа контактов информационных головок

Рассмотрим содержимое гермоблока.



Внутри алюминиевого корпуса располагаются

- носители информации (диски)
- шпиндель и двигатель, вращающий диски в процессе работы
- блок головок, производящих чтение и запись
- плата интерфейса головок

Диски представляют собой пластины с нанесённым слоем магнитного материала. Сами диски изготавливаются из алюминия. Были попытки делать их из другого материала — керамики или стекла. Но диски получались избыточно хрупкими и очень часто выходили из строя даже не от внешних ударов, а от малейших перекосов и вибрации, возникающей при работе диска. Самый печальный пример — IBM серии DTLA. Из стоя очень быстро вышли все диски этой серии и это фактически привело к закрытию производства дисков под маркой IBM.

С 2017 года ситуация меняется и для производства дисков 2,5» используются всё чаще композитные материалы, содержащие

стекло (Glass Substrate). Такие материалы намного лучше подходят для использования в технологии HARM и почти не уступают алюминию по прочности.

В качестве рабочего слоя изначально использовалось полимерное покрытие из оксидов железа. Сейчас используется тонкоплёночный слой, создаваемый методом электролиза или напыления. Состав и технология являются одной из главных производственных тайн компании производителя.

В современных дисках обычно используется 2-3 пластины, диски с другим количеством пластин встречаются реже (до 6 штук). Чем меньше пластин, тем диск надёжнее, экономичнее и дешевле в производстве. Диск содержащий всего одну пластину, как правило, получает слишком мало ёмкости и становится не рентабельным.

Между дисками установлены специальные пластины — **демпферы**. Они стабилизируют вращение диска и равномернее распределяют воздушные потоки внутри корпуса. Изготавливают их из пластика и алюминия. В последнем случае они ещё играют роль радиаторов для отвода тепла.

Диски установлены на **шпиндель** — вертикальный вал, выходящий из двигателя. Вовремя работы диски вращаются со строго определенной скоростью. При стандартной скорости в 7200 об/мин головки «летят» над поверхностью диска со скоростью в 120-160 км/ч.

Вторым компонентом является **блок головок**.



Он состоит из блока рычагов на общей оси. На конце рычага находится головка чтения/записи (раньше делали головки отдельно, но теперь функции чтения и записи совмещены в одной головке). В основании блока рычагов находится катушка приводящего двигателя и выход для шлейфа с информационными проводниками головок.

Информационные головки на концах рычагов есть для каждого диска по две штуки, чтобы можно было читать данные с обеих сторон диска.

Для увеличения скорости работы дисков компания Seagate работает над архитектурой **Mach.2**, подразумевающей установку двух блоков головок. В теории это должно дать двукратный рост производительности. Опытные образцы представлены в марте 2018 года.

В работе головки не касаются поверхности диска, а парят над ним на небольшой высоте (5-10 нм), за счёт **воздушной подушки**, образующейся из-за быстрого вращения дисков. Парение головок необходимо для отсутствия трения между головкой и диском и — как следствие — отсутствие физического износа этих компонентов. Для правильной аэродинамики головки она имеет определённые выступы в корпусе и поддерживается небольшим рычажком-

слайдером.



Для парения головок так же необходим воздух, так что мнение что внутри диска вакуум в корне неверное. Воздух внутри диска есть и он необходим для его работы, но воздух должен быть сверхчистым. Толщина волоса человека в 2500 раз больше, чем воздушный зазор головки. Да и головка сама существенно меньше волоса.

Столкновение даже с мельчайшей пылинкой на тех скоростях, на которых парит головка является фатальным для неё. Поэтому сборку и ремонт жёстких дисков проводят только в «чистых» комнатах.

В дисках емкостью от 8Тб в качестве «наполнителя» вместо воздуха используют гелий. Его плотность ниже, соответственно ниже и аэродинамическое сопротивление, а значит механизм можно делать тоньше и плотность записи информации возрастает.

При завершении работы головки опускаются на диск, что может вызвать повреждение поверхности или головки. Поэтому в крайних положениях блока головок находится зона парковки (**L-Zone**), куда опускаются головки при выключении. Обычно используются крайние, внутренние части дисков, но могут и самые внешние.

Современные диски проводят парковку самостоятельно, даже при внезапном пропадании питания на диске; для совсем старых

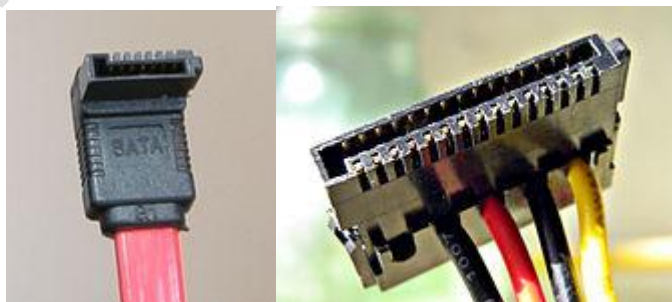
дисков парковку приходилось проводить при помощи специальных утилит перед выключением компьютера.



Если снять диски и блок головок, то станет видно вращающий диски **двигатель** и фильтр очистки, расположенный на входном отверстии. Перекрывать на диске это отверстие не стоит, это может привести к перегреву диска.

Интерфейсы.

Стандартом на данный момент является интерфейс **SATA**: интерфейс последовательной передачи данных. Для передачи информации используется 7 контактный разъём, а для питания специализированный 15 контактный разъём.



Стандарт в первой версии (sata 1.0) был опубликован в 2003 году. Частота работы была заявлена до 1,5 ГГц, что с учётом служебной информации позволяло передавать 150Мбайт/сек. Это было немногим больше существовавшего тогда стандарта UDMA133, но стандарт sata выигрывал за счёт более простого и надёжного кабеля и более простого подключения устройств.

Дальнейшие стандарты sata 2.0 и sata 3.0 (в 2008 году) направлены на увеличение скорости передачи информации до 300 Мбайт/сек и 600 Мбайт/с соответственно. Для устройств сохранена обратная совместимость — устройство по стандарту sata 1.0 можно подключить к контроллеру sata 2.0 и 3.0, правда работать оно конечно будет на скоростях версии 1.0 .

В производительных серверных компьютерах используется интерфейс **SAS**. Принципиально это тоже последовательный интерфейс для подключения жёстких дисков, но он имеет ряд улучшений:

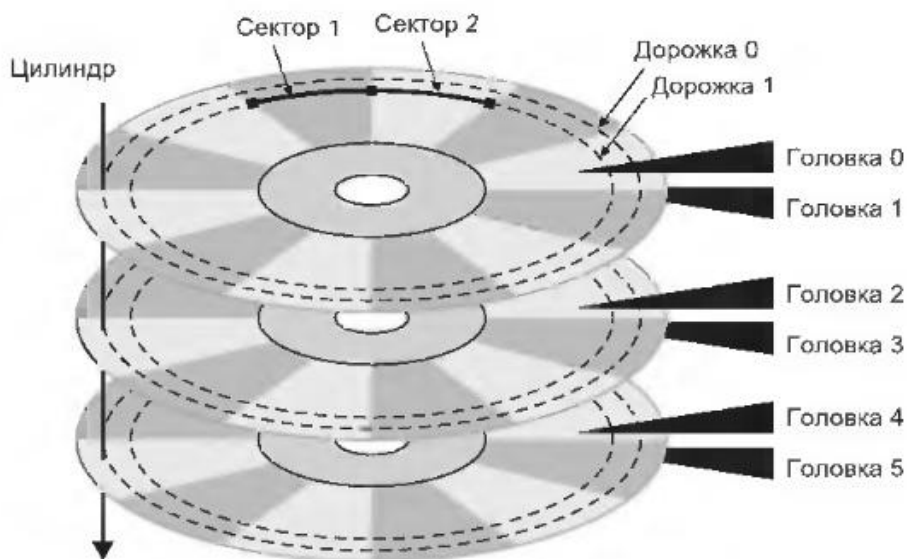
- поддержка подключения во время работы (горячее подключение) позволяет менять вышедшие из строя диски, не выключая весь компьютер. В сочетании с RAID технологией позволяет организовать быструю замену вышедшего из строя диска незаметно для пользователей
- наиболее совершенные диски имеют два порта подключения, позволяющие использовать их с двух разных компьютеров
- сквозная проверка целостности данных
- ускоренная работа за счёт увеличения буферов ввода/вывода и переупорядочивания команд.
- Диски SAS существенно дороже дисков SATA.

Устаревшими на данный момент являются интерфейсы

- SCSI
- ATA
- ESDI
- ST-412

Логическая структура диска.

При неподвижном положении головки вращение диска обеспечивает формирование окружности на поверхности диска из намагниченных доменов. Эта окружность называется **дорожкой**. При каждом положении головки относительно диска получается по одной дорожке информации на его поверхности.



Головки диска расположены на одном блоке и двигаются по поверхности дисков синхронно. Совокупность всех дорожек в заданном положении называется **цилиндром**. Цилиндры нумеруются с внешнего края диска и нумерация начинается с 0.

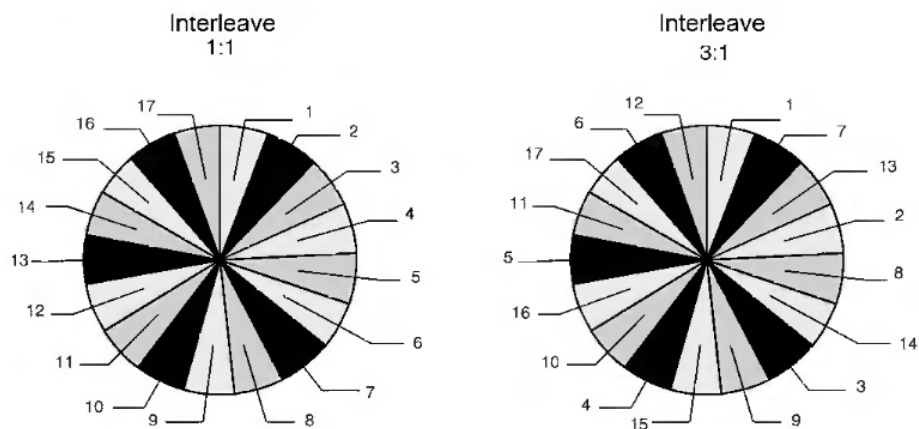
Нумерация дорожек в цилиндре так же начинается с 0. Дорожка разбивается на **секторы** – сегменты, содержащие одинаковое количество информации. Нумерация секторов начинается с 1.

Количество секторов зависит от накопителя. Каждый сектор содержит данные и служебную информацию. Для 512 байт информации сектор чаще всего имеет размер 571 байт. Начинается сектор с заголовка, а заканчивается контрольной суммой, которая служит для проверки целостности данных. Подобная схема организации дискового пространства называется низкоуровневым форматированием и выполняется на заводе перед отгрузкой диска потребителю. Проводить самостоятельно форматирование повторно крайне не рекомендуется, так как шанс получить хотя бы как то работоспособный диск после самостоятельного низкоуровневого форматирования крайне низок.

Для ускорения работы диска в низкоуровневом форматировании применены несколько ухищрений:

- перекос секторов
- перекос дорожек

Перекос секторов на дорожке выглядит так



Сектора на дорожке изначально нумеровались последовательно. Информация имеет тенденцию к последовательному расположению на диске, т. е. чаще всего требуется информация из нескольких последовательных секторов. Чтение сектора разбивается на этапы чтения и передачи информации от блока головок к кэш памяти. За время передачи диск успевает повернуться и под головками оказывается не следующий сектор, так что для получения информации следующего сектора требуется подождать один оборот диска. Перекос секторов призван исправить эту ситуацию: при низкоуровневом форматировании сектора нумеруются не подряд, а с пропусками. Допустим, что при после сектора 0 при передаче данных диск успел повернуться на 4 сектора. Под головкой теперь сектор 5. Требуется подождать больше половины круга. Теперь при форматировании с перекосом для дорожки с 11 секторами нумерация из 0,1,2,3,4,5,6,7,8,9,10 при перекосе в 5 сектора превратится в 0,3,5,7,9,1,4,6,8,10,2.

На этом же основан **перекос дорожек**. Нумерация нулевого сектора каждой дорожки сдвигается на величину подобранную так, чтобы при переходе с последнего сектора предыдущей дорожки мы оказывались на нулевом секторе.

Низкоуровневое форматирование диска так же исправляет следующую ситуацию: дорожки на краю диска имеют намного большую длину, чем в центре диска. Поэтому плотность записи информации в центре диска много выше, чем с краю. Для этого применяется метод зонно-секционной записи. Дорожки делятся на несколько зон и количество секторов в зонах, при движении к центру диска уменьшается. На внешней зоне число секторов на дорожке может в два раза превышать число секторов дорожки в самой последней зоне. Зон чаще всего 8 и более.

Основные характеристики диска.

Основными характеристика диска являются

- ёмкость
- быстродействие
- время безотказной работы

Рассмотрим их.

Под **ёмкостью** обычно понимается максимальный объём полезной информации, которую может хранить диск. Зачастую именно по ёмкости покупатель и выбирает диск. Производители часто вводят в заблуждение и указывают ёмкость не в Гб, в миллиардах байт. Итоговая ёмкость от этого может меняться существенно, например полезная ёмкость диска объемом 1 гб окажется меньше 900 Мб. Часть информации конечно была использована для служебных целей, но и «потери» от «хитростей» производителей очень существенны.

Быстродействие диска складывается из нескольких факторов.

Среднее время доступа к информации — время которое требуется диску чтобы начать обмен данными после получения команды от контроллера. Напрямую оно зависит от скорости перемещения головок внутри диска и складывается из среднего времени поиска и среднего времени ожидания. Среднее значение для дисков сейчас составляет 7-10 мс.

Среднее время поиска это время за которое блок головок перемещается от предыдущего цилиндра к требуемому. Чаще всего производитель указывает результаты в виде среднего времени по итогам своих тестов. Для современных это значение составляет менее 4-7 мс.

Среднее время ожидания это время за которое будет получен нужный цилиндр с данными. Сейчас это обычно 3 мс.

И ещё на быстродействие влияет скорость передачи данных. Диск не сможет работать быстрее своего интерфейса, однако его быстродействие ещё определяется возможностями механической части

$$MDTR = SRT * 512 * RPM/60 * \text{количество головок чтения}$$

где MDRT это максимальная скорость передачи данных, SRT количество секторов на дорожке, а RPM скорость вращения диска. Например, для диска со скоростью вращения 7200 об/мин и 250 секторами на дорожке и четырьмя головками получим максимальную скорость 60 Мб/с. Сравним со скоростями интерфейсов и станет понятно, что даже sata 1.0 является избыточным для такого диска.

Скорость увеличивается с увеличением числа секторов на дорожке. Так как число секторов неравномерно на каждой дорожке, то и скорость передачи данных зависит от места информации на диске.

Время безотказной работы. В документации часто указывают среднестатистическое время между сбоями. Эта характеристика призвана свидетельствовать о надёжности накопителя. Для нормального диска оно равно 250000-500000 часов. При круглосуточной работе это 57 лет. Понятно, что это значение не соответствует истине. Получено оно так: возьмем 100 дисков и начнём их тестировать. Если через год из строя выйдет один диск, то значит вероятность выхода за год диска из строя равно 1%. Исходя из таких данных по выбранной производителем формуле и вычисляется время безотказной работы.

Увеличению производительности диска способствует **кэш память**. Обычно в диске её от 2 до 32 Мб. Она используется для хранения данных, получаемых из диска по заранее запрограммированному в контроллере алгоритму и с определенной вероятностью нужная информация уже окажется в кэш памяти и соответственно будет считана быстрее.