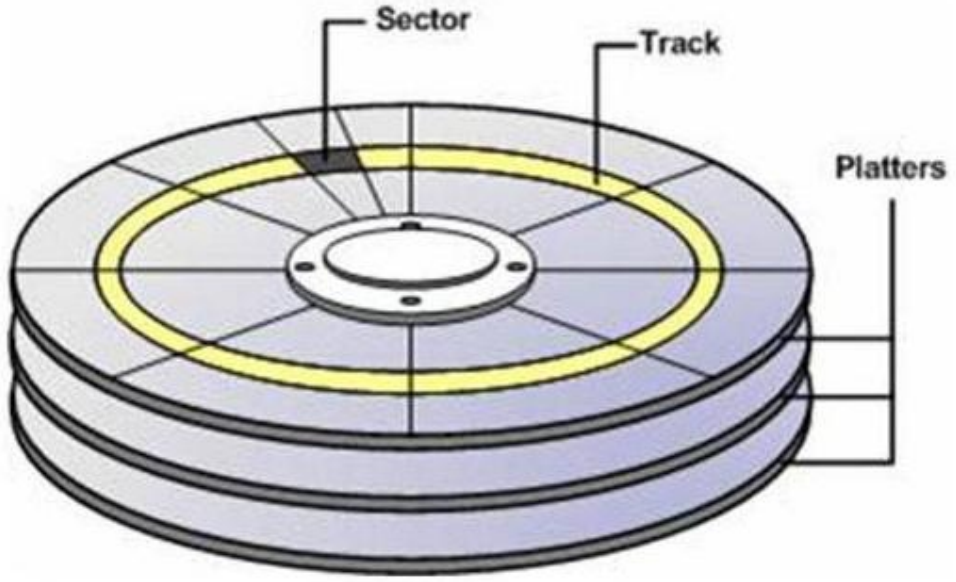
ХАРАКТЕРИСТИКИ УСТРОЙСТВ ХРАНЕНИЯ ДАННЫХ

Обзор дисковых интерфейсов и архитектур

HDD. Форм-фактор

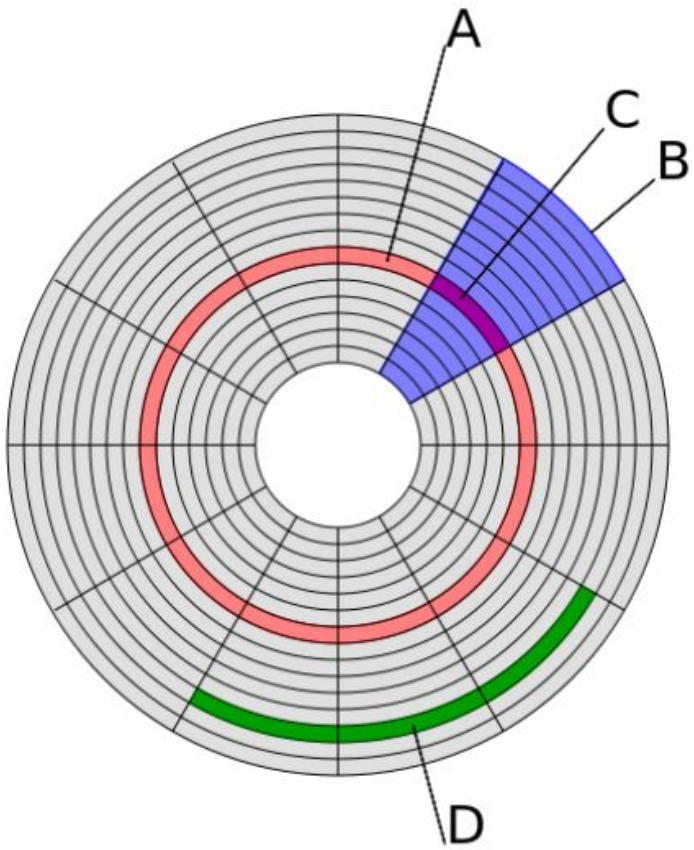
Форм фактор:

- 2,5 inch (толщина)
- 3,5 inch



Где длиннее дорожка, там выше скорость ☺️

Адресация HDD



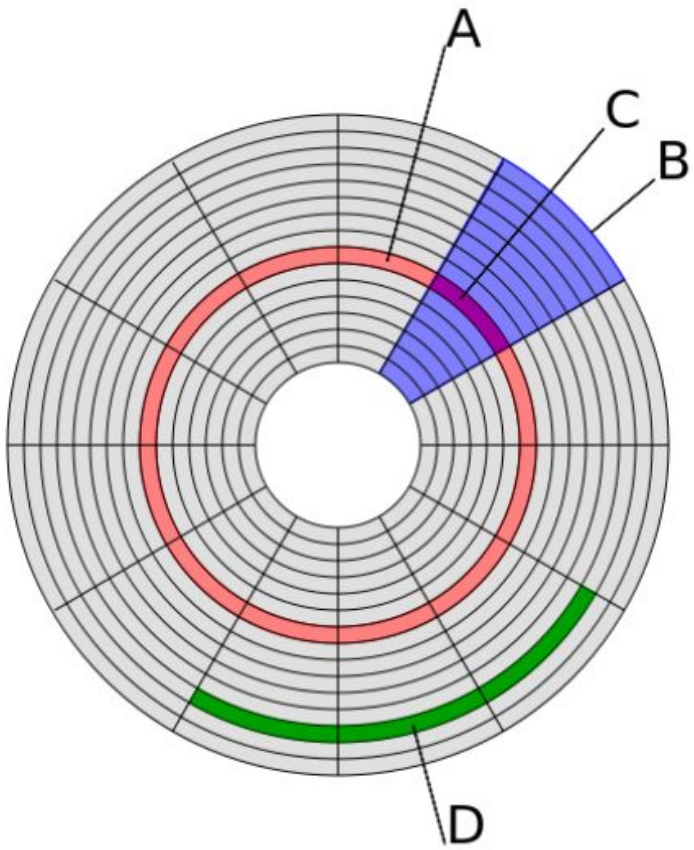
A - дорожка,

B - геометрический сектор,

C - сектор дорожки,

D - кластер

Адресация HDD



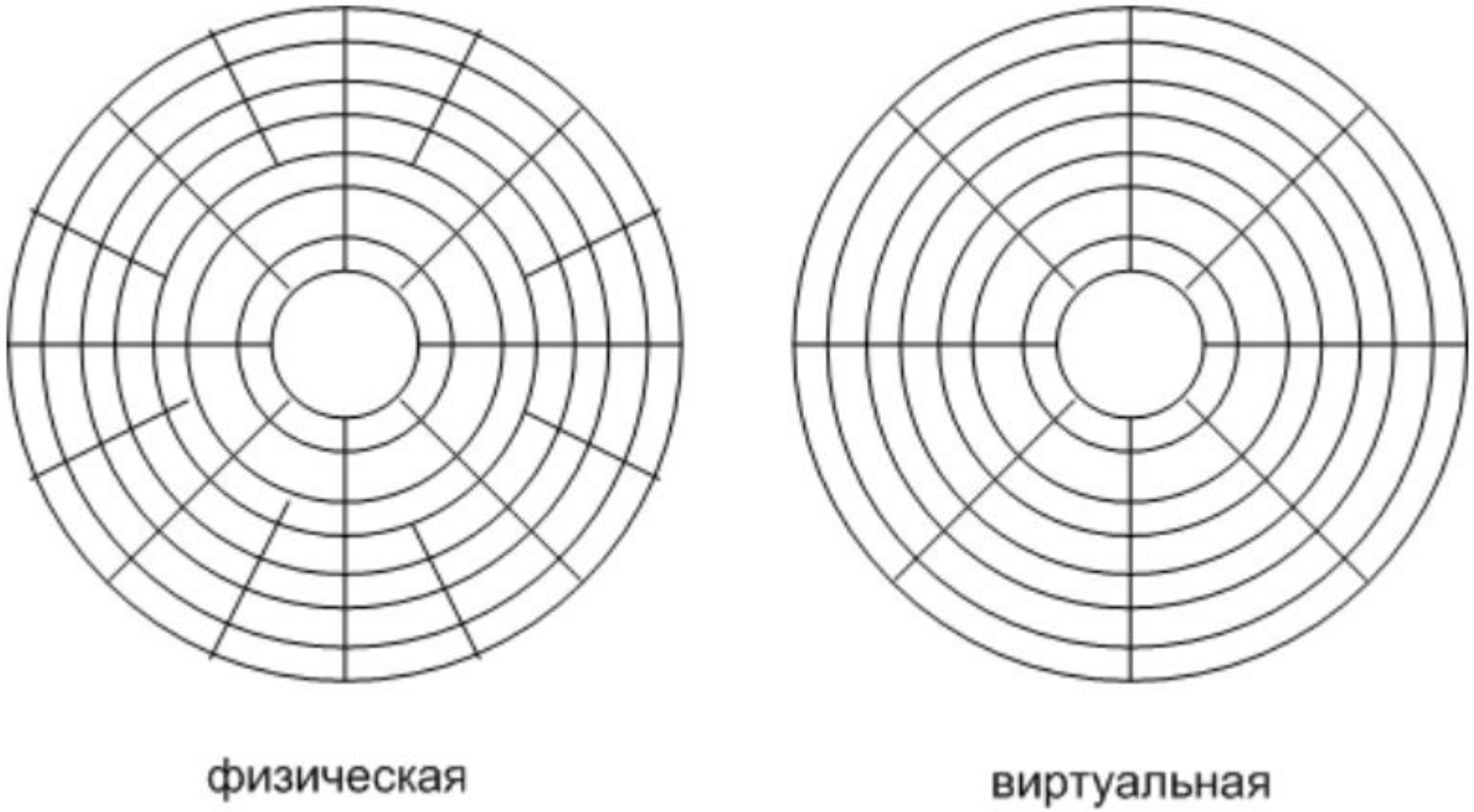
Сектор диска (Sector) – минимально адресуемая единица хранения информации на дисковых запоминающих устройствах.

Является частью дорожки диска.

HDD используют размер сектора 4096 байт (Advanced Format).

Нумерация секторов начинается от маркера, который имеется на каждом диске

Адресация HDD



HDD. Объем

Seagate ST14000DM001

SSD 15,36 Тб компании Samsung

HDD. Объем

AWS Snowmobile

«жёсткий диск на колёсах» ёмкостью 100 петабайт
10 машин перевозять экзабайт примерно за полгода
(перекачка по 10 Gb\s каналу займет примерно 26 лет)

- Источник питания мощностью 350 киловатт (при чтении\записи),
- Защита от физического взлома,
- Шифрование,
- Система видеонаблюдения и GPS
- Вооруженная охрана

HDD. Скорость вращения шпинделя

- Жесткие диски для ноутбуков имеют скорость вращения 4200, 5400 и 7200 оборотов в минуту,
- для стационарных компьютеров 5400, 7200 и 10 000 об/мин.
- для серверов 7200, 10 000 и 15 000 об/мин.

HDD. Интерфейс подключения

- См. предыдущую лекцию

HDD. Объем буфера

- От 8 до 256 Мб.
- Больше не значит лучше.

HDD. Наработка на отказ (MTBF)

- Сервера ~ 1 мл часов.
  Актуально для серверов.

HDD. Среднее время ожидания

- параметр определяет то, сколько времени требует головке для перемещения на любой участок дорожки. Средняя производительность напрямую зависит от него.
- Головки серверных модели накопителей позиционируются за 2-14 мс. Этот параметр жестких дисков для домашней техники обычно равен 7-14 мс.

HDD. Количество операций ввода-вывода в сек.

- количество операций ввода-вывода в секунду. Обычно жесткий диск производит около 50 операций в секунду при произвольном доступе и около 100 при последовательном.

HDD. Уровень шума

- Чем медленней вращается шпиндель, тем меньше посторонних звуков возникает в процессе работы устройства. Показатель в 26 дБ считается довольно низким.

HDD. Ударостойкость

- сопротивляемость накопителя резким скачкам давления или ударам. Измеряется в единицах допустимой перегрузки (G) во включённом и выключенном состоянии.

HDD. Энергопотребление и тепловыделение

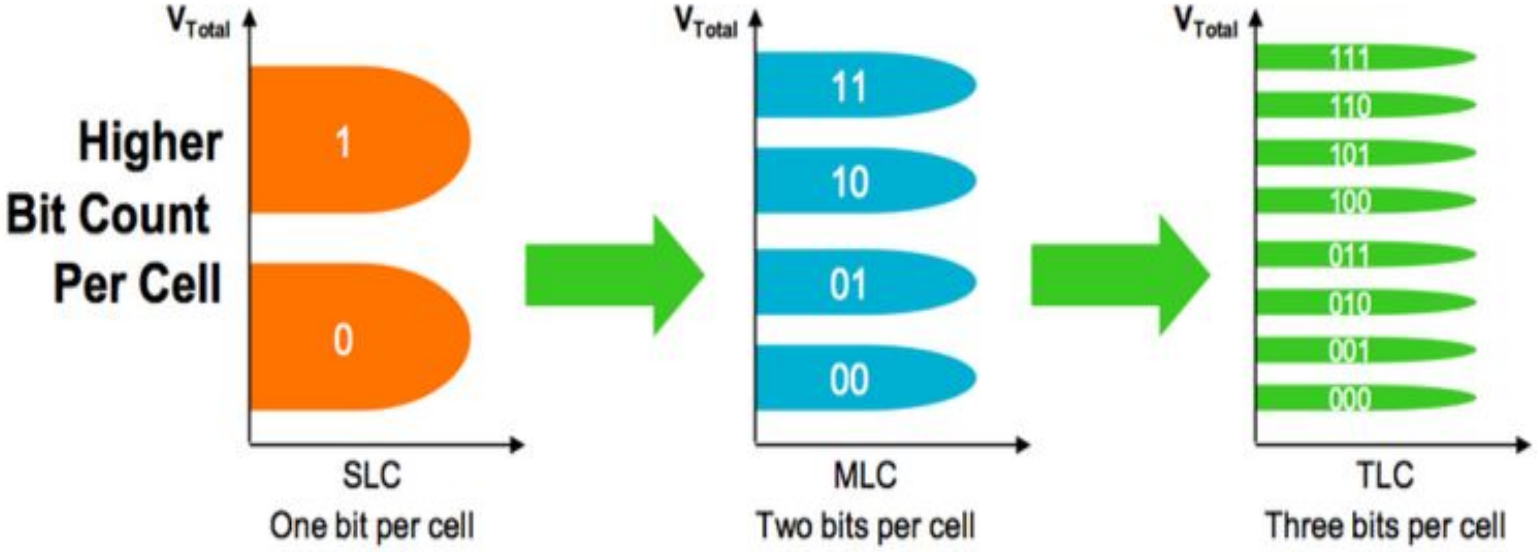
- Актуально для ЦОД

SDD

- Формфактор (2.5)
- Объем
- ничего не должно крутиться (работает тихо)
- Объем буфера (бывают без буфера)
- MTBF
- Кол-во операций в сек.
- Ударостойкость
- Энергопотребление \ тепловыделение
- Количество циклов перезаписи

SDD. Конструктивные особенности

- Sigle-Layer Cell
- Multi-Level Cell
- Triple-Level Cell
- 3D NAND

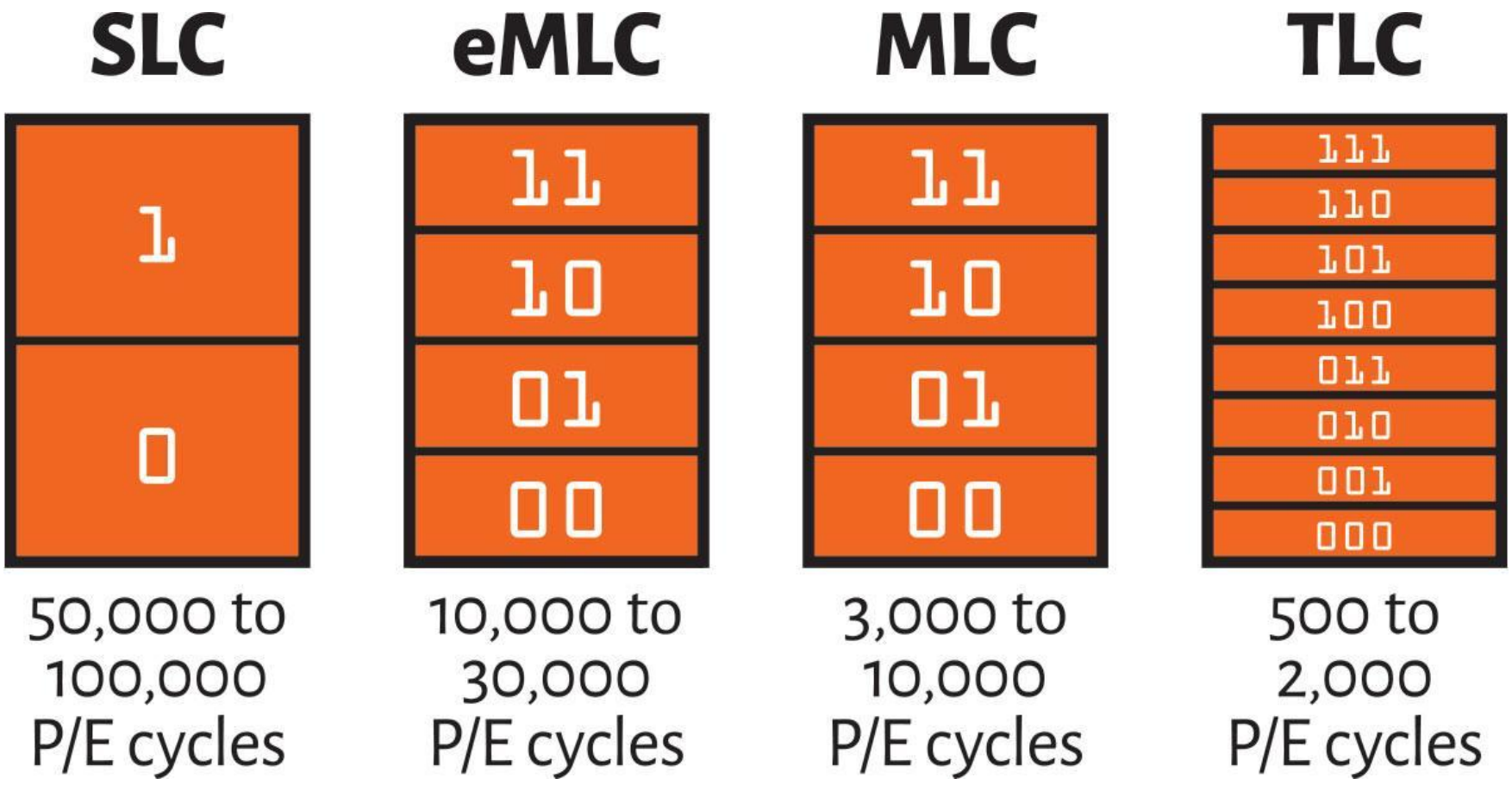


3D NAND – это скорее форма организации памяти, а не её новый тип. Существует как MLC, так и TLC 3D NAND.

По возрастанию стоимости/надёжности, условно:

MLC ≥ TLC ≥ 3D NAND.

Количество циклов перезаписи



Количество циклов перезаписи

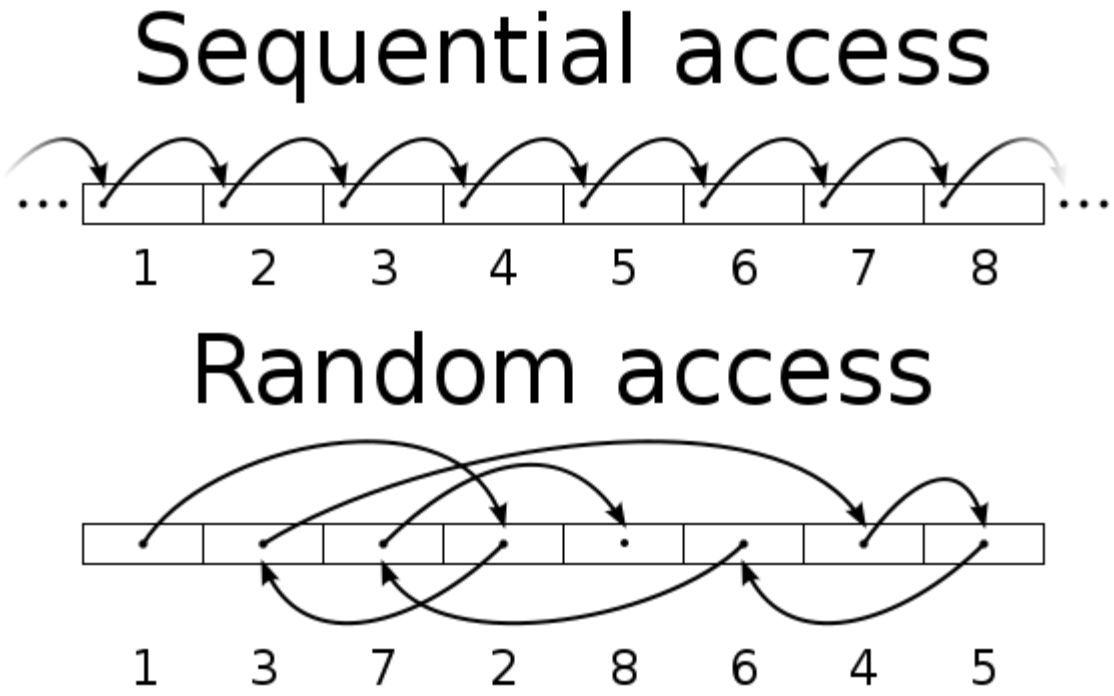
Ресурс твердотельного накопителя		
Количество циклов перезаписи	**3 000**	**5 000**
Емкость памяти	120 Гб	120 Гб
Дневной объем записи	12 Гб	12 Гб
Увеличение объема записи	10x	10x
Циклов перезаписи в день	1 (10x12Гб)	1 (10x12Гб)
Оценка ресурса диска	**8 лет**	**13.5 лет**

Оценка производительности

- линейное чтение
- случайное чтение
- линейная запись
- случайная запись
- IOPS

Оценка производительности

- Под линейными операциям чтения/записи, при которых части файлов считываются последовательно, одна за другой, подразумевается передача больших файлов (более 128 К).
- При произвольных операциях данные читаются случайно из разных областей носителя, обычно они ассоциируются с размером блока 4 Кбайт.



IOPS

- IOPS (количество операций ввода/вывода – от англ. Input/Output Operations Per Second) – один из ключевых параметров при измерении производительности систем хранения данных, жестких дисков (НЖМД), твердотельных диски (SSD) и сетевых хранилища данных (SAN).
  По сути, *IOPS это количество блоков, которое успевает считаться или записаться на носитель*
- Для определения IOPS надо знать скорость и размер блока при операции чтения / записи (для RAID количество блоков и пенальти RAID).

IOPS

Параметр	Описание
Всего IOPS (Total IOPS)	Суммарное число операций ввода/вывода в секунду (при выполнении как чтения, так и записи)
IOPS произвольного чтения (Random Read)	Среднее число операций произвольного чтения в секунду
IOPS произвольной записи (Random Write)	Среднее число операций произвольной записи в секунду
IOPS последовательного чтения (Sequential Read)	Среднее число операций линейного чтения в секунду
IOPS последовательной записи (Sequential Write)	Среднее число операций линейной записи в секунду

IOPS

Величина IOPS зависит от многих параметров:

- конструкция и настройки оборудования (дисков и RAID);
- устройство и настройки драйвера;
- устройство и настройки драйвера файловой системы;
- устройство и настройки операционной системы;
- условия запуска программы, выполняющей тестирование производительности (бенчмарка):
- отношение количества операций чтения к количеству операций записи;
- размеры блоков для чтения и записи при последовательном и случайном доступе;
- количество потоков, выполняющих чтение и запись;
- размеры очередей и буферов;
- наличие фрагментации в файловой системе;
- наличие приложений, работающих в фоновом режиме;

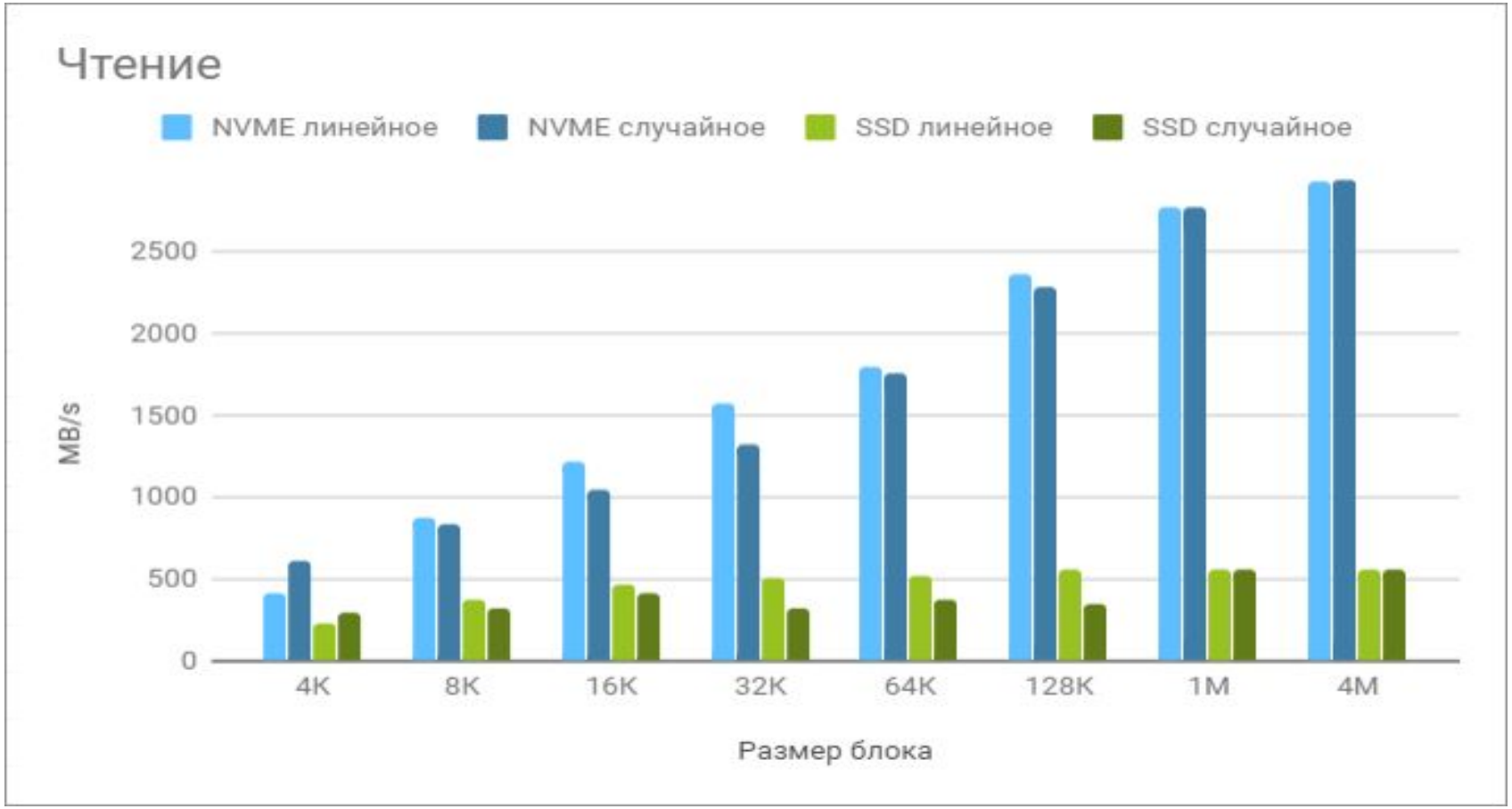
Пример оценки IOPS от компании firstVDS

- Операционная система: CentOS Linux 7 со штатным ядром - 3.10
- Планировщик ввода-вывода - none.
- Измерения утилитой fio при глубине очереди 16, блоками по 4K, 8K, 16K, 32K, 64K, 128K, 1M, 4M (4M не является обязательным вариантом). Условия теста исключительно синтетические, позволяют выжать из накопителей максимальные показатели.

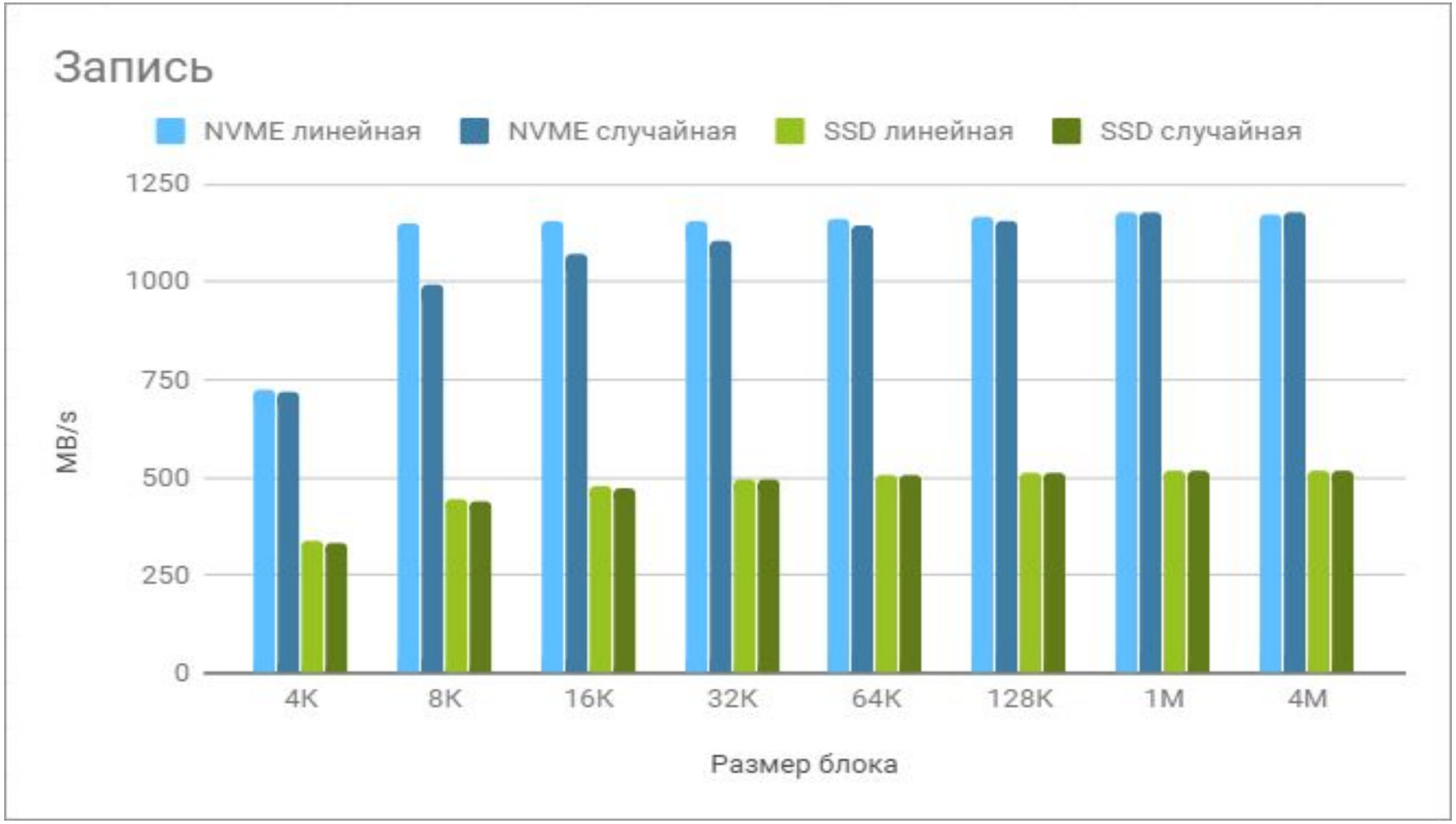
Пример оценки IOPS от компании firstVDS

	NVMe P4510	SSD D3-S4510
Дата выпуска	Q3'17	Q3'18
Тип литографии	64-Layer 3D TLC NAND	3D NAND TLC
Ёмкость	1 TB	1.92 TB
Последовательное чтение	2850 MB/s	560 MB/s
Последовательная запись	1100 MB/s	510 MB/s
Случайное чтение	465000 IOPS	97000 IOPS
Случайная запись	70000 IOPS	35500 IOPS
Задержка чтения	77 µs	36 µs
Задержка записи	18 µs	37 µs
Питание — активный режим	10W	3.2W

Пример оценки IOPS от компании firstVDS



Пример оценки IOPS от компании firstVDS



Пример оценки IOPS от компании firstVDS

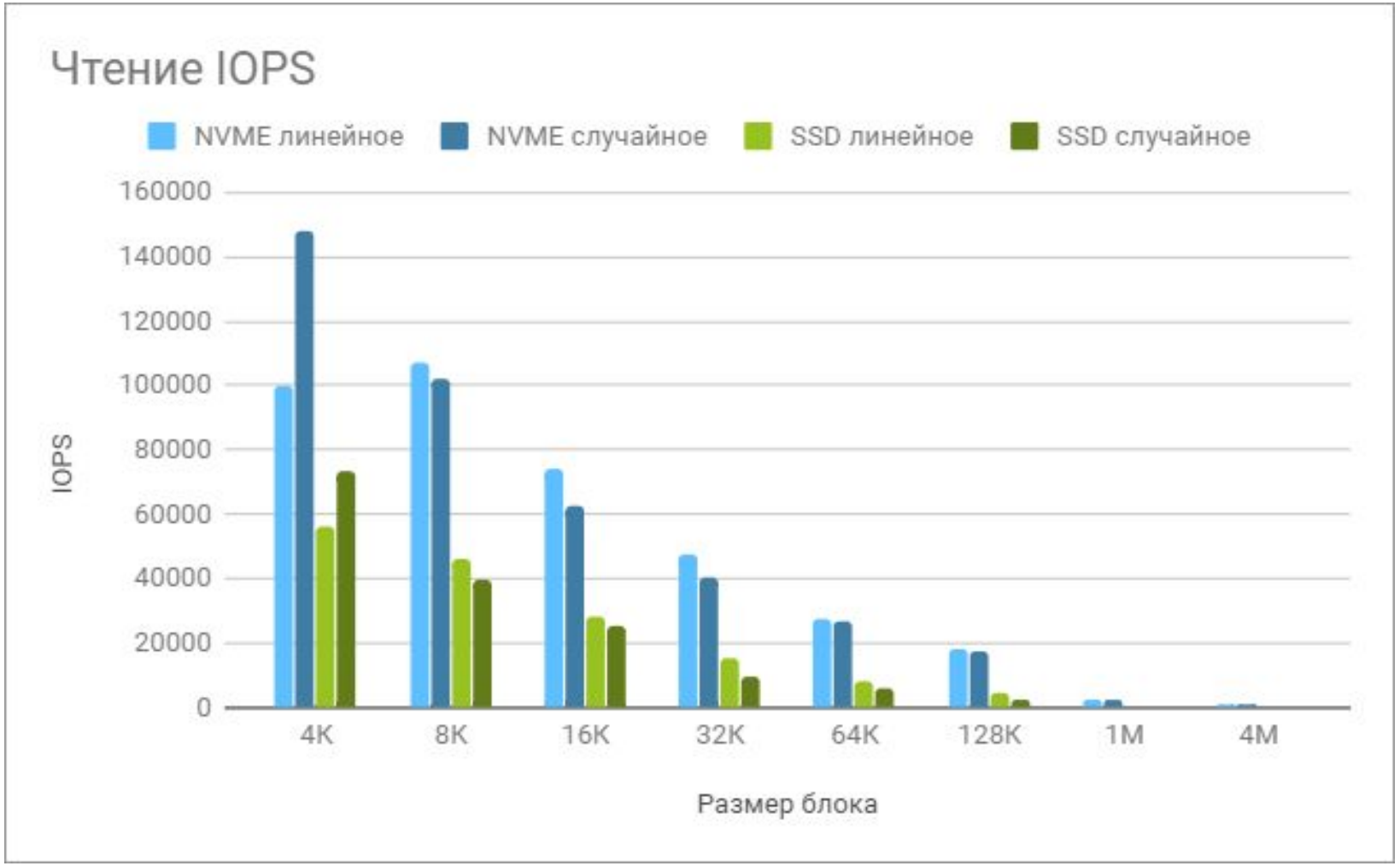
- NVMe выходит на пиковую скорость линейного чтения на блоках с 1M до 4M. А SSD выходит на свой «потолок» на блоках размером в 128K и зажимается пропускной способностью шины SATA.

- В операциях же со случайным чтением SSD выходит на максимальную скорость на блоках в 128K, в то время как у NVMe картина не меняется.

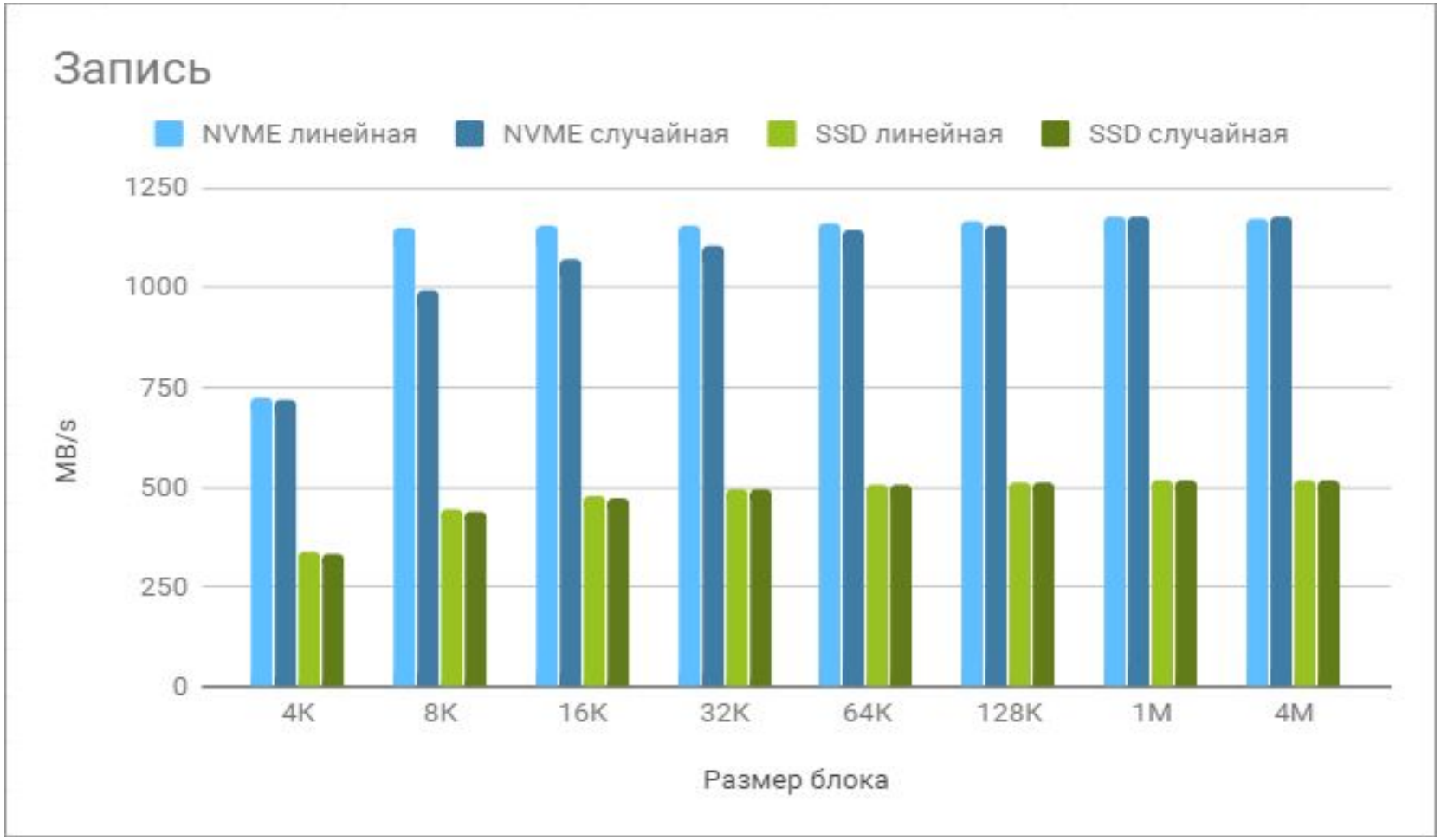
- В операциях линейной записи NVMe выходит на максимальную скорость уже на блоках в 8K, SSD — с 16к-32к.

- В операциях случайной записи картина схожая.

Пример оценки IOPS от компании firstVDS



Пример оценки IOPS от компании firstVDS



Задача

1) С помощью утилит выяснить все доступные формальные характеристики дисков на ваших компьютерах.
(Windows (https://www.hdtune.com/files/hdtune_255.exe) , Linux (hddtemp или smartmontools)).
Дать (по возможности) качественную оценку значениям параметров.

2) Ознакомиться с параметрами SMART. Подробно описать смысл 6 параметров.

3) ознакомиться с утилитами:

- iostat
- iotop
- smatctl
- blkid
- blockdev
- sar -d
- pidstat -d

Подготовить демонстрацию их использования.