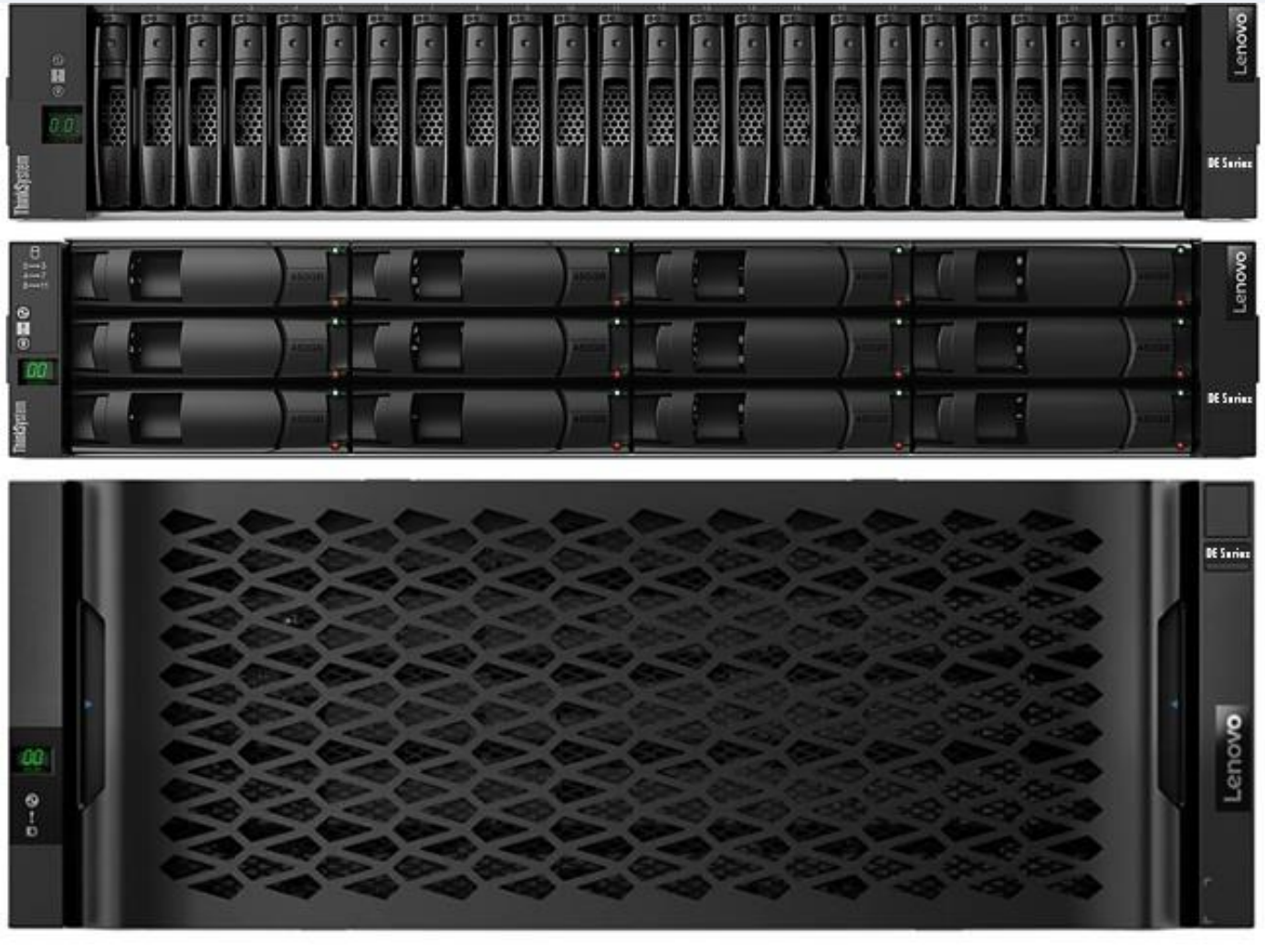
Nihol

СИСТЕМЫ ХРАНЕНИЯ ДАННЫХ ALL-FLASH

ОСОБЕННОСТИ КОНФИГУРИРОВАНИЯ, ОБЛАСТЬ ИСПОЛЬЗОВАНИЯ

ЧТО ТАКОЕ СИСТЕМА ХРАНЕНИЯ ДАННЫХ?

Система хранения данных (СХД) — комплекс аппаратных и программных средств, который предназначен для хранения и оперативной обработки информации большого объема. В качестве носителей информации используются жесткие диски, в основном SSD, а также гибридные решения, сочетающие SSD- и HDD-накопители в одной СХД.



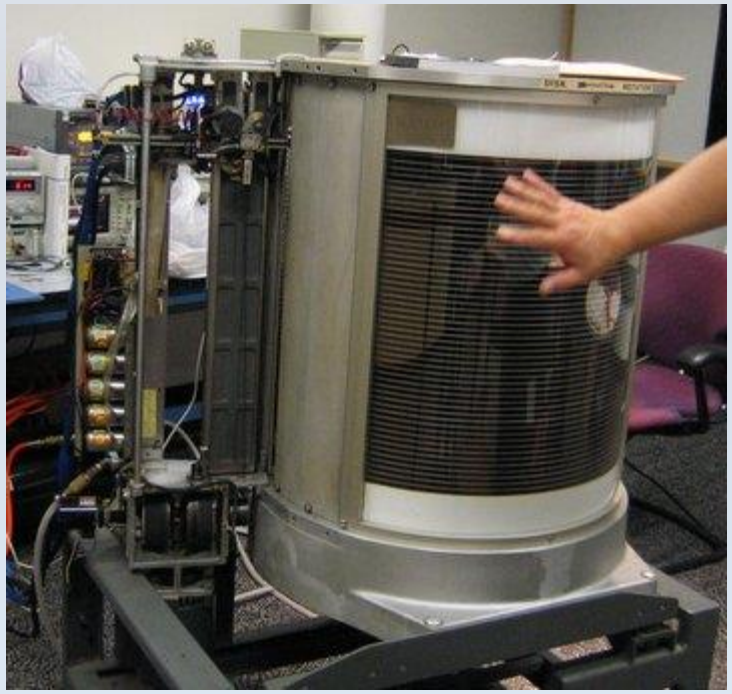
Системы хранения Lenovo ThinkSystem серии DE

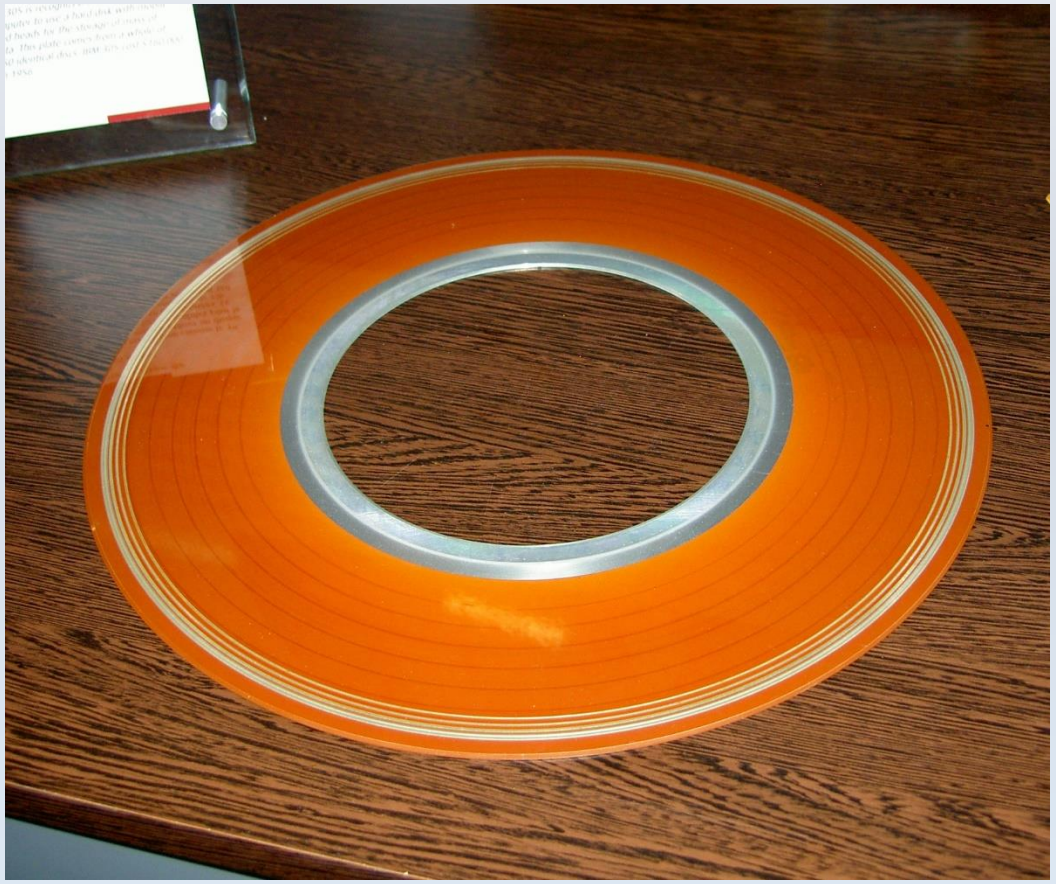
НЕМНОГО ИСТОРИИ

Дебютный жесткий диск HDD RAMAC 305 был выпущен корпорацией IBM в 1956 году. Он весил более тонны, его приходилось перемещать с помощью вилочных погрузчиков и доставлять большими грузовыми самолетами.
Объем такого «диска» составлял 3,75 Мб - это примерно ДО 2х фотографий на вашем современном смартфоне.

НЕМНОГО ИСТОРИИ

Состояла эта громадина из 50-ти дисков диаметром в 24 дюйма (610 мм).
Данную систему можно было арендовать за 3200 долларов США (что эквивалентно ~30 000$ США в 2020 году) в месяц.





СОВРЕМЕННЫЕ ДИСКИ

3.5 ДЮЙМОВЫЕ ЖЕСТКИЕ ДИСКИ

Основные характеристики 3.5 дюймовых дисков:

Интерфейс подключения
SATA, SAS
Ёмкость
Можно найти 18ТБ 7.200 SATA (7.500$)
Скорость вращения шпинделя
от 5.400 об/мин до 15.000 об/мин
Скорость передачи данных
До 200 МБ/с
IOPS - количество операций ввода-вывода в секунду
До 300

2.5 ДЮЙМОВЫЕ ЖЕСТКИЕ ДИСКИ

Основные характеристики 2.5 дюймовых дисков:

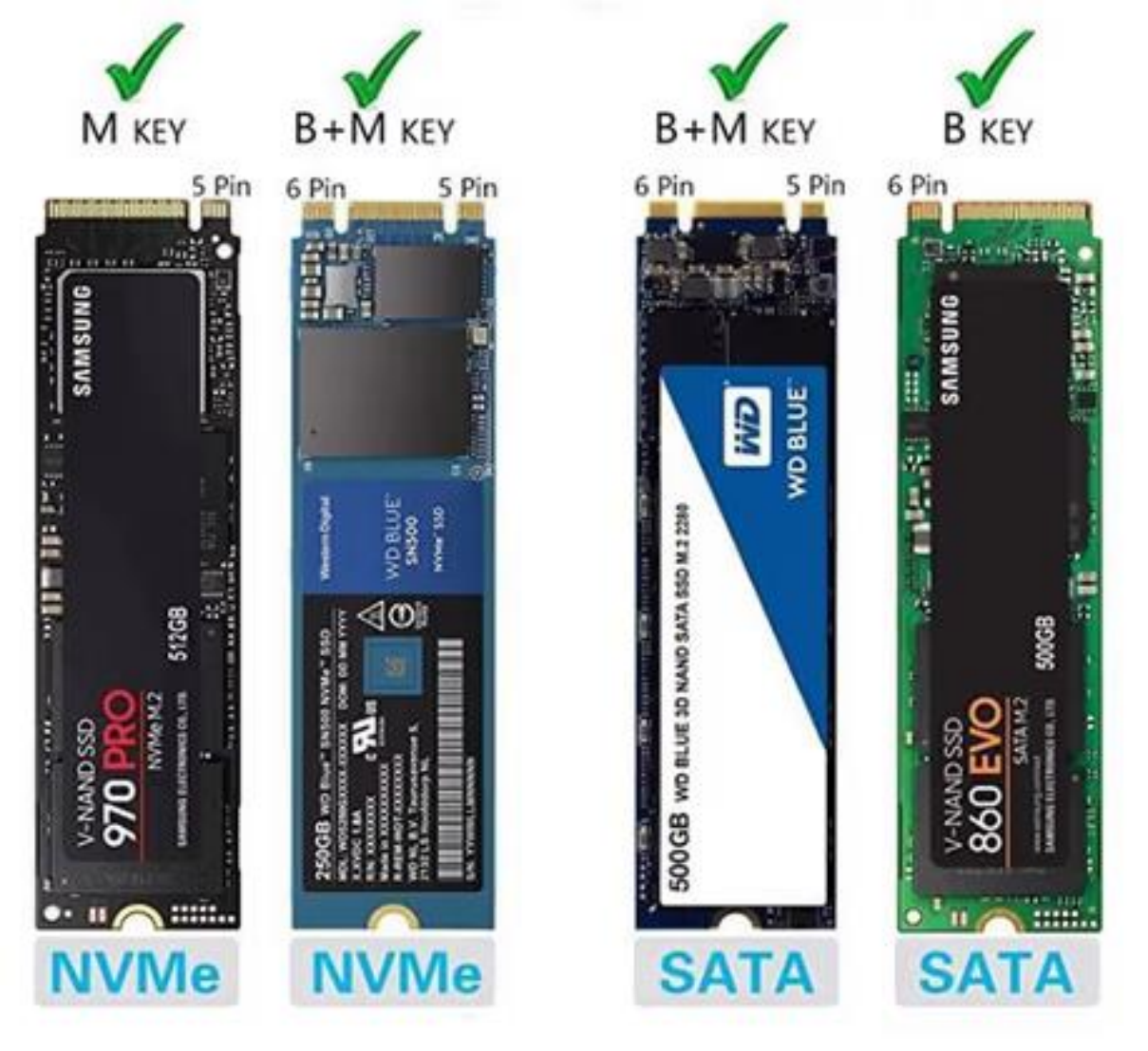
Интерфейс подключения
SATA, SAS
Ёмкость
HDD – 2.4ТБ 10.000 SAS (3.300$)
SSD – 15.36ТБ Entry SAS (12.000$)
Скорость вращения шпинделя
HDD – от 5.400 об/мин до 15.000 об/мин
Скорость передачи данных
HDD – До 200 МБ/с
SSD – До 1200 МБ/с
IOPS - количество операций ввода-вывода в секунду
HDD – До 300
SSD – До 240 000 / 200 000

HDD SSD

M.2 НАКОПИТЕЛИ

Основные характеристики M.2 накопителей:

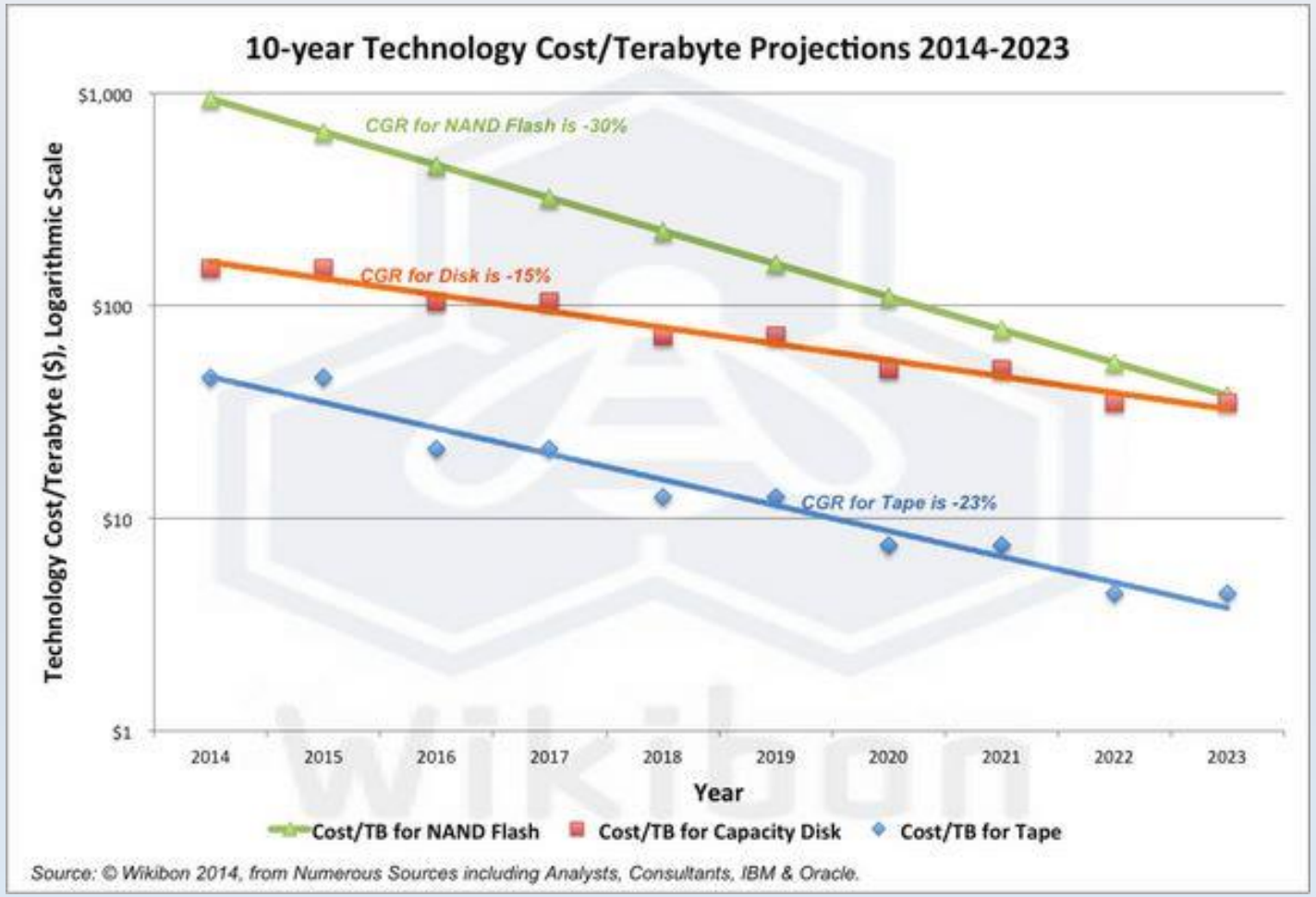
Интерфейс подключения
SATA, SAS, PCIe NVMe
Ёмкость
До 15.36ТБ
Скорость передачи данных
До 7200 МБ/с
IOPS - количество операций ввода-вывода в секунду – до 1 500 000
DWPD (Drive Write per Day) – Количество перезаписей всего объёма накопителя в день за определенный период времени – чем больше значение, тем лучше и дороже.
TBW (Total Bytes Written) – Суммарный объем данных, который гарантированно можно записать на накопитель. Указывается в терабайтах – чем больше значение, тем лучше и дороже.



ПОЧЕМУ НЕТ ALL-FLASH

У SSD два основных недостатка. Во-первых, пока еще высокая цена. Во-вторых, меньшее число циклов перезаписи по сравнению с HDD. Впрочем, ниже мы покажем, что это несколько надуманная проблема. Кроме того, есть некоторое ограничение по емкости, вызванное ограничениями плотности элементов памяти на единице площади кристалла кремния в микросхемах SSD. Однако прогресс технологий в ближайшем будущем должен минимизировать оба недостатка.

Прогнозы аналитиков IBM, Oracle и пр. говорят о том, что через 2-3 года цены в пересчете на терабайт как на SSD (NAND Flash), так и на HDD (Disk) практически сравняются.
Но у Covid-19 свои коррективы…



РАЗБЕРЕМ СХД

УРОВНИ ХРАНЕНИЯ

Блочное хранилище

СХД используется как обычный диск, который можно форматировать, устанавливать на него ОС, создавать логические диски. Данные хранятся не файлами, а блоками, что ускоряет операции ввода-вывода. Чаще используется в сетях типа SAN (Storage Attached Network).

Файловое хранилище

Данные хранятся в виде файлов, которые размещаются в каталогах. Такая СХД используется для хранения «холодной» информации, которая не требуется для операционных вычислений. На файловых хранилищах, как правило, строятся NAS (Network Attached Storage). При накоплении больших объемов данных усложняется иерархия папок, и скорость работы СХД постепенно снижается.

Объектное хранилище

Тип СХД, который ориентирован на работу с большими неструктурированными данными объемом до петабайтов. Информация хранится не в виде файлов, а в виде «объектов» с уникальными идентификатором и метаданными. Поэтому объектное хранилище похоже по структуре на БД. Используется в аналитике, big data, машинном обучении и т.д.

СЕТЕВОЙ ДОСТУП

NAS (network-attached storage)

Файловый сервер, который включен в локальную сеть. Доступ к дисковому хранилищу организован через протоколы NFS (в UNIX/Linux-системах) или CIFS (Windows). NAS используется для работы с данными файлового типа, к которым нужен коллективный одновременный доступ — например, к общим документами Word и Excel. NAS работает «поверх» существующей локальной сети, через общие коммутаторы/маршрутизаторы.

SAN (storage area network)

Сеть, которая годится для использования разнотипных хранилищ (диски, оптические приводы, ленточные массивы), но которые воспринимаются операционной системой как единое логическое хранилище данных, или как сетевой логический диск. Протоколы: iSCSI (IP-SAN) и FibreChannel (FC). Для подключения компьютеров используются адаптеры HBA (Host Bus Adapter). В SAN применяется в основном блочный тип хранения данных.

Разделение SAN/NAS уже не так строго, как было в начале 2000-х, поскольку с появлением протокола iSCSI производители стали выпускать гибридные решения.

ОТКАЗОУСТОЙЧИВОСТЬ

RPO (recovery point objective)

Период, за который будут потеряны данные, — между моментом аварии и временем, когда создалась последняя резервная копия. Если RPO равен 12 часам, при выходе из строя хранилища возможна потеря данных, накопившихся за последние 12 часов. RPO влияет на выбор технологии аварийного восстановления и зависит от стоимости потери конкретного объема данных.

RTO (recovery time objective)

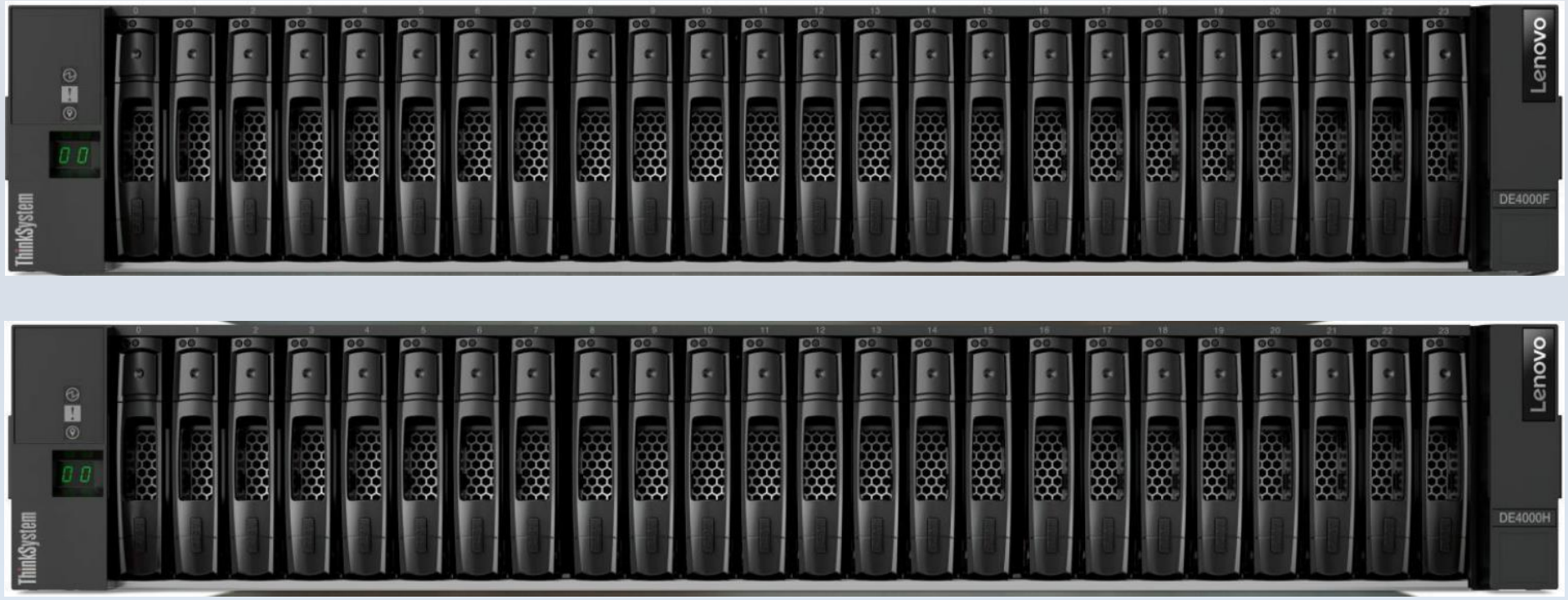
Время, за которое восстанавливается доступ к СХД. Значение RTO важно для оценки стоимости простоя системы.

РЕЗЕРВНОЕ КОПИРОВАНИЕ

Частота создания бэкапов выбирается исходя из конкретных задач и требуемого уровня защиты. То же касается и размещения: рабочие данные и их резервную копию можно хранить в географически распределённых СХД (например, в дата-центрах, расположенных в разных странах и даже континентах).

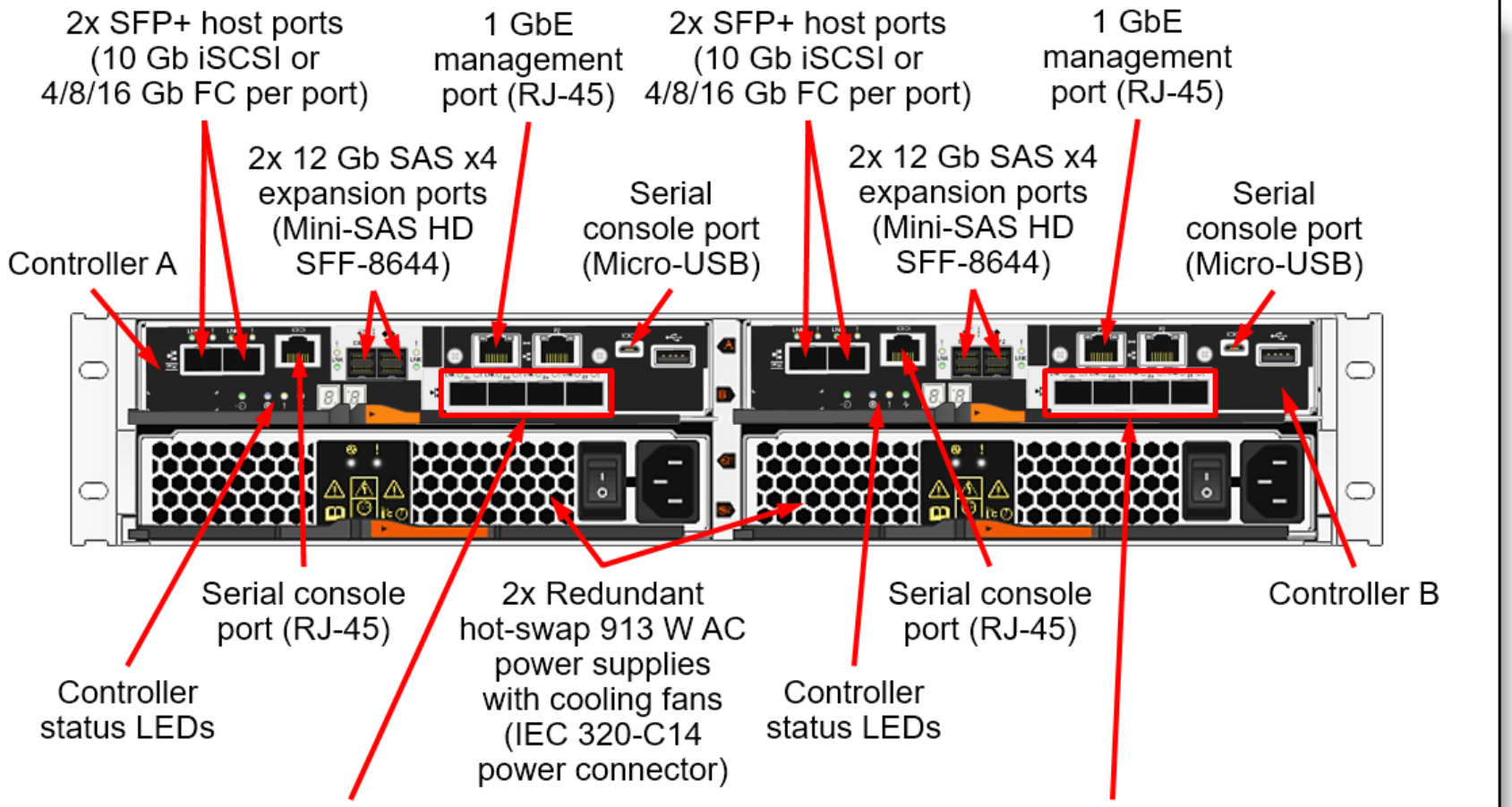
ЧТО ТАКОЕ ALL-FLASH СХД

All-Flash – это системы хранения данных в электронном виде. В отличие от СХД на жестких дисках (HDD), у СХД на твердотельных носителях (SSD) отсутствуют движущиеся детали точной механики, которая перемещает и позиционирует магнитные головки для считывания-записи по дискам в HDD.



Lenovo DE4000F и DE4000H – найди отличия

THINKSYSTEM DE6000F 2U SFF, ВИД СЗАДИ



1x Host interface card:
- 4x 12 Gb SAS x4 ports (SFF-8644)
- 4x 10/25 Gb iSCSI ports (SFP28)
- 4x 8/16/32 Gb FC or NVMe/FC ports (SFP+)
- 2x 25/40/100 Gb NVMe/RoCE ports (QSFP28)

1x Host interface card:
- 4x 12 Gb SAS ports (SFF-8644)
- 4x 10/25 Gb iSCSI ports (SFP28)
- 4x 8/16/32 Gb FC or NVMe/FC ports (SFP+)
- 2x 25/40/100 Gb NVMe/RoCE ports (QSFP28)

Nihol

	DE6000H	DE4000H	DE2000H	DE6000F	DE4000F
Form Factor	4U, 60 drives 2U, 24 drives	2U, 24 drives 2U, 12 drives 4U, 60 drives	2U, 24 drives 2U, 12 drives	2U, 24 drives	2U, 24 drives
Maximum Raw Capacity	Supports up to 8.64PB	Supports up to 3.456PB	Supports up to 1.728PB	Supports up to 1.84PB	Supports up to 1.47PB
Maximum Drives	Supports up to 480 HDDs / 120 SSDs	Supports up to 192 HDDs / 120 SSDs	Supports up to 96 HDDs/SSDs	Supports up to 120	Supports up to 96
Maximum Expansion	2U/12, 2U/24: up to 7 expansions 4U/60: up to 3 expansions	2U/12, 2U/24: up to 7 expansions 4U/60: up to 3 expansions	2U/12, 2U/24: up to 3 expansions	2U/24: Up to 4 DE240S expansions	2U/24: Up to 3 DE240S expansions
System Memory	32GB/128GB	16GB/64GB	16GB	128GB	64GB
Base I/O Ports (per system)	4x 10Gb iSCSI (optical) 4x 16Gb FC	4x 10Gb iSCSI (optical) 4x 16Gb FC	4x 10Gb iSCSI (optical) 4x 16Gb FC	4x 10Gb iSCSI (optical) 4x 16Gb FC	4x 10Gb iSCSI (optical) 4x 16Gb FC
Optional I/O Ports (per system)	8x 16GB/32GB FC 8x 10GB/25GB iSCSI optical 8x 12GB SAS	8x 16GB/32GB FC 8x 10GB/25GB iSCSI optical 8x 12GB SAS	4x 10GB iSCSI Base-T 4x 12GB SAS	8x 16GB/32GB FC 8x 10GB/25GB iSCSI optical 8x 12GB SAS	8x 16GB/32GB FC 8x 10GB/25GB iSCSI optical 8x 12GB SAS
Optional Software Feature	Drive Encryption, Snapshot upgrade, asynchronous mirroring, synchronous mirroring (DE6000H and DE4000H only)				
System Maximums	Hosts/partitions: 512 Volumes:2,048 Snapshot copies: 2,048 Mirrors:128	Hosts: 256 Volumes: 512 Snapshot copies: 512 Mirrors: 32	Hosts: 256 Volumes: 512 Snapshot copies: 512 Mirrors: 32	Hosts/partitions: 512 Volumes: 2,048 Snapshot copies: 2,048 Mirrors: 128	Hosts/partitions: 256 Volumes: 512 Snapshot copies: 512 Mirrors: 32

RAID

RAID (англ. Redundant Array of Independent Disks — избыточный массив независимых дисков) — технология виртуализации данных, которая объединяет несколько дисков в логический элемент для повышения производительности.

RAID 0 — striping (чередование). Массив при котором информация разбивается на одинаковые по длине блоки, а затем записывается поочерёдно на каждый диск в структуре.

RAID 1 – зеркальный дисковый массив.

RAID 5 – n-1 - дисковый массив с чередованием, в том числе данных чётности (нет диска, выделенного для хранения чётности — блоки чётности чередуются с блоками данных на каждом диске).

RAID 6 – n-2 - дисковый массив с чередованием, использующий две контрольные суммы, вычисляемые двумя независимыми способами

RAID 10, 50, 60 – объединение 1 и 0, 5 и 0 и т.д.

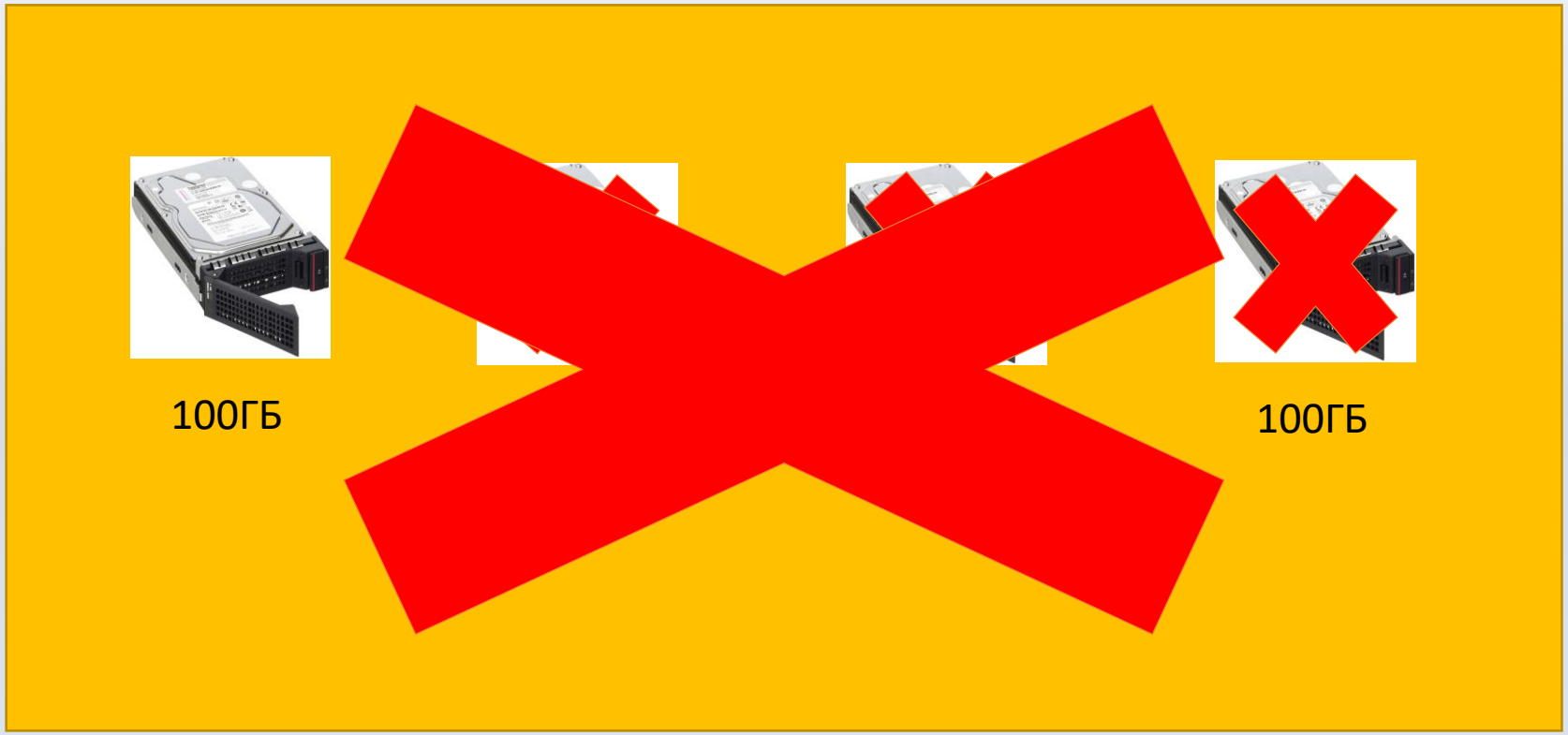
RAID 10 – это не одно и то же, что и RAID 01 !

RAID 0

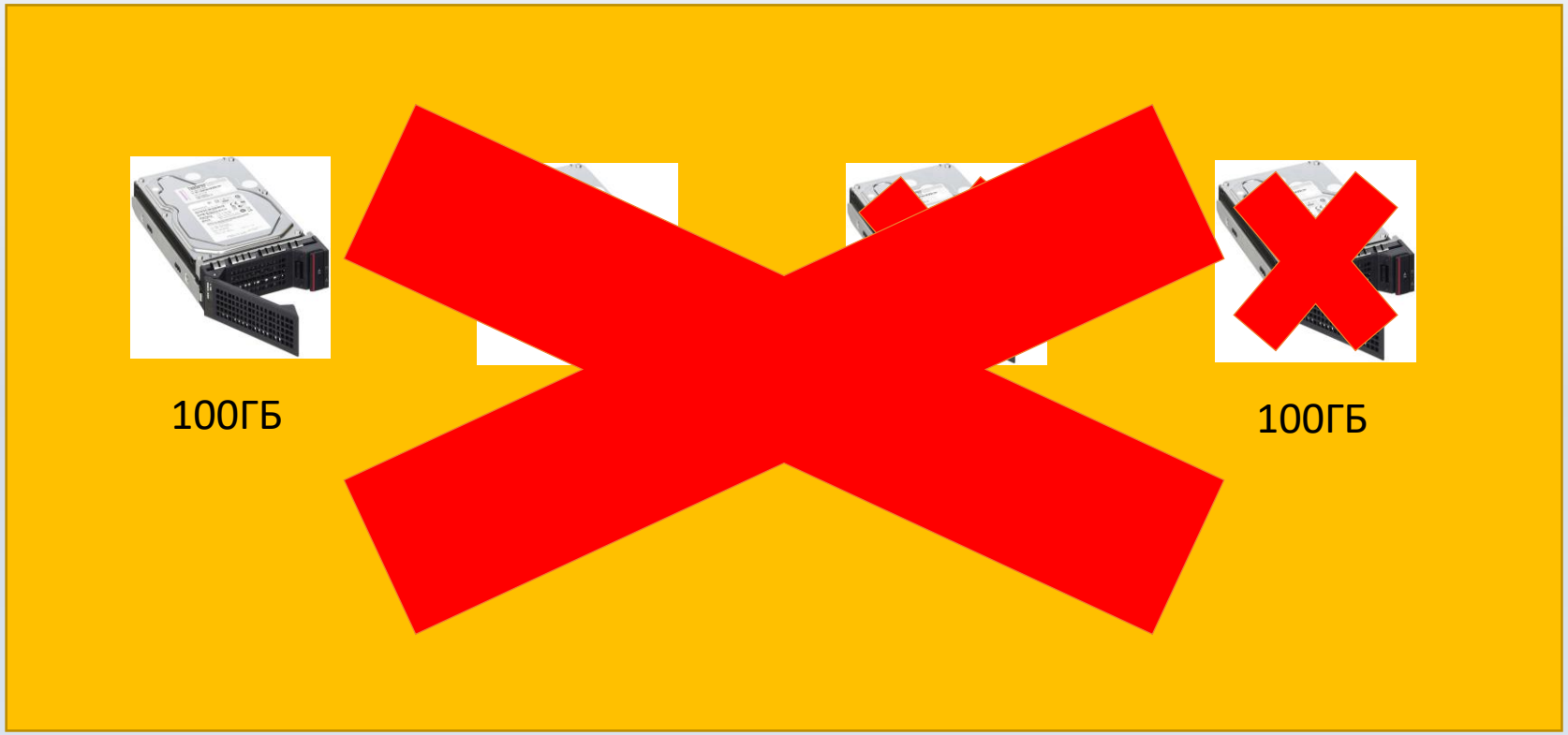
100ГБ

100ГБ

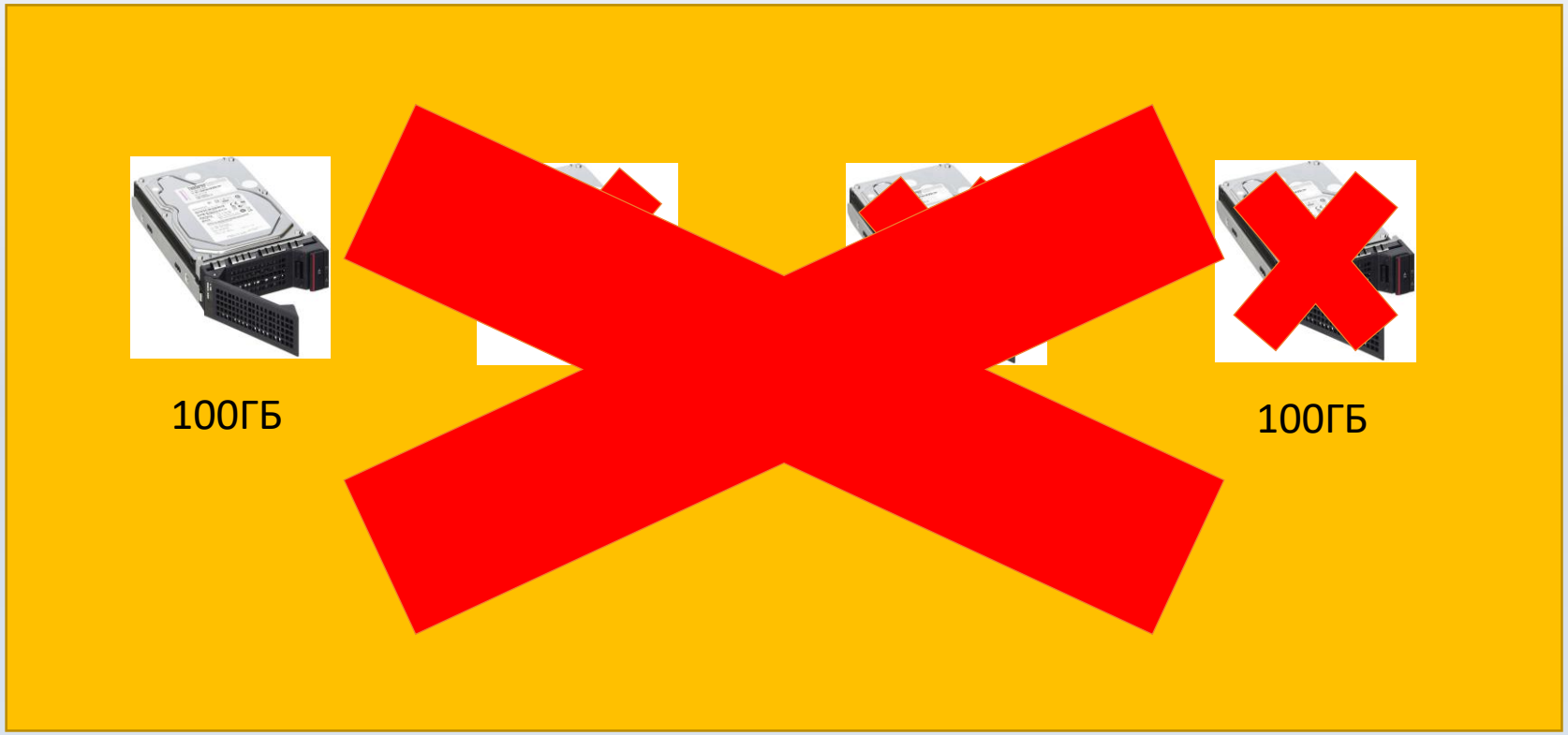
RAID 1



RAID 5



RAID 6



RAID 10

100ГБ

100ГБ

ОСОБОЕ ВНИМАНИЕ

FlashCore Modules

Модули IBM FlashCore (FCM) — это семейство высокопроизводительных флэш-накопителей в стандартном форм-факторе 2,5", 15 мм. В конструкции модуля FlashCore используется протокол NVMe, интерфейс PCIe Gen3 U.2 и высокоскоростная память NAND для обеспечивают высокую пропускную способность и количество операций ввода-вывода в секунду с постоянной и предсказуемой задержкой. Каждый FCM содержит одну FPGA со шлюзом NVMe и многоядерными процессорами ARM.

Модули FlashCore доступны в следующих размерах:

4,8 Тбайт (полезные терабайты) / 21,99 Тбайт (эффективные терабайты), 9,6 Тбайт / 21,99 Тбайт, 19,2 Тбайт / 43,98 Тбайт, 38,4 Тбайт / 87,96 Тбайт.

Поддерживается аппаратное сжатие данных и самошифрование.

ИНТЕРЕСНО ПОЧИТАТЬ

https://itelon.ru/blog/all-flash-skhd-perspektivy-i-problemy/

СПАСИБО!

Если Вы заинтересованы во внедрении в своей организации, то можете обратиться к нам:

Узбекистан, г. Ташкент, 100187, ул. Интизор, 26, Группа компаний NIHOL;

наши телефоны: (998-71) 208-58-44, 208-58-45, 208-58-47, 208-58-48, 266-58-46;

e-mail: info@nihol.uz