

# УСТРОЙСТВА ХРАНЕНИЯ ДАННЫХ

Обзор дисковых интерфейсов и архитектур

# Немного истории



Сотрудники IBM грузят жесткий диск объемом 5 МБ, 1956 год.

Железо предназначалось для первого суперкомпьютера с жестким диском 305 RAMAC.

Весила система около тонны, — получается по 0,2 грамма за байт (или 5 килобайт в 1 кг) и состояла из 50-ти дисков диаметром в 24 дюйма (610 мм).

# Иерархия подсистем хранения

- Непосредственно подключенные устройства
- СХД DAS
- СХД NAS
- СХД SAN
- SDS (Software-Defined Storage)
- Облачное хранение?  
"Нет облака, есть чужой компьютер"

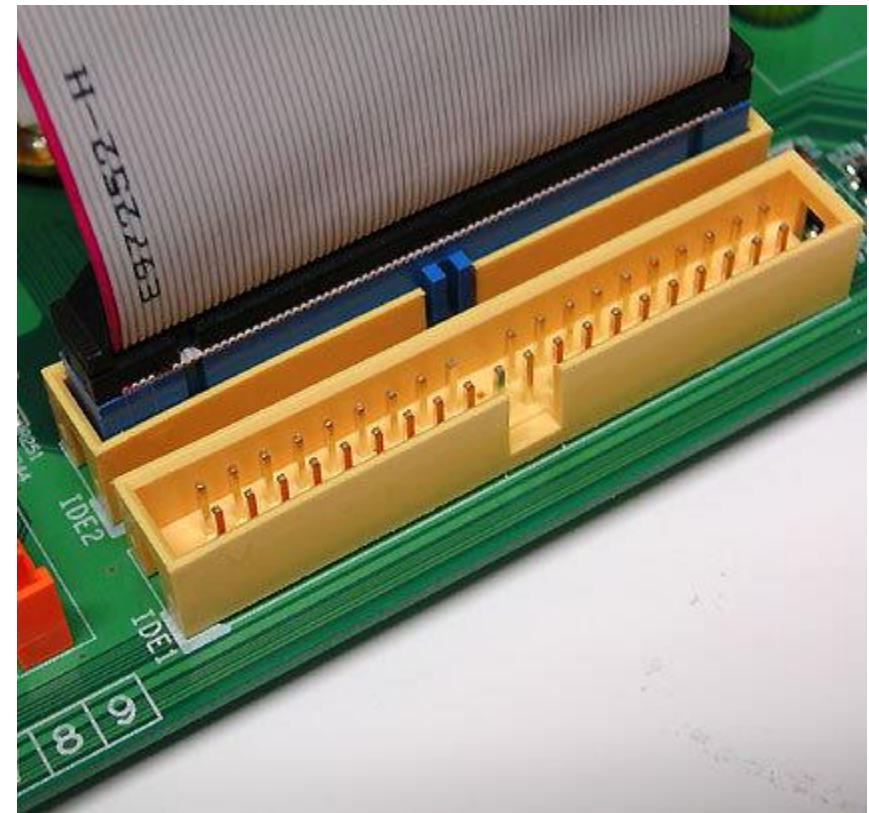
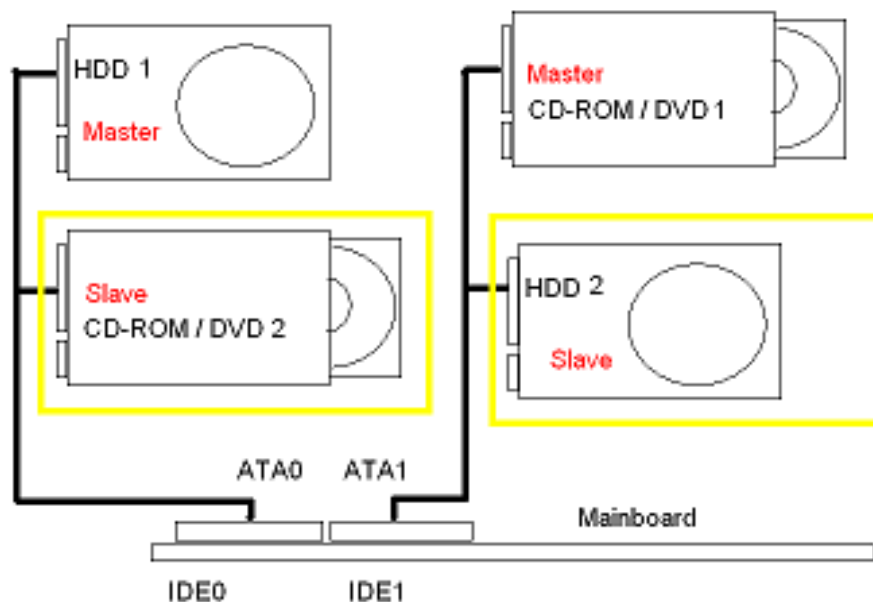
# Интерфейсы дисков

---

- PATA
- SCSI
- SAS
- NVMe

# PATA

- **ATA** (*Advanced Technology Attachment*) или **IDE** (*Integrated Drive Electronics*)
- До 133 Мб/с
- НЕ актуален



# SCSI

- **SCSI** (*Small Computer System Interface*) — представляет собой набор стандартов для физического подключения и передачи данных между компьютерами и периферийными устройствами. SCSI-стандарты определяют команды, протоколы и электрические и оптические интерфейсы.
- Устарел

# SCSI



PC-Information-Guide.ru



# SCSI

## Current SCSI Standards

Technology name	Maximum cable length (meters)	Maximum speed (MBps)	Maximum number of devices
SCSI-I	6 Single Ended	5	8
Fast SCSI	3 Single Ended	10	8
Fast Wide SCSI	3 Single Ended	20	16
Ultra SCSI	1.5 Single Ended	20	8
Ultra SCSI	3 Single Ended	20	4
Wide Ultra SCSI	-	40	16
Wide Ultra SCSI	1.5 Single Ended	40	8
Wide Ultra SCSI	3 Single Ended	40	4
Ultra2 SCSI	12 LVD/25 HVD	40	8
Wide Ultra2 SCSI	12 LVD/25 HVD	80	16
Ultra3 SCSI (Ultra160 SCSI)	12 LVD	160	16
Ultra320 SCSI	12 LVD	320	16
Ultra640 SCSI	10 LVD	640	16



# SATA

- SATA (*Serial ATA*) — последовательный интерфейс обмена данными с накопителями информации
- **SATA Revision 1.0 - до 1,5 Гбит/с - 2003**
- **SATA Revision 2.0 - до 3 Гбит/с - 2005**
- **SATA Revision 3.0 - до 6 Гбит/с - 2008**

# SATA

- последовательный интерфейс обмена данными с накопителями информации.
- SATA является развитием параллельного интерфейса ATA (IDE)
- SATA работает в полудуплексном режиме
- SATA поддерживает Hot Plug
- Работает по протоколу AHCI (Advanced Host Controller Interface )





# SAS (Serial Attached SCSI)

- последовательный компьютерный интерфейс, разработанный для подключения различных устройств хранения данных, например, жёстких дисков и ленточных накопителей.
- Протокол SAS обеспечивает полнодуплексную передачу данных.
- SAS разработан для замены параллельного интерфейса SCSI и основывается во многом на терминологии и наборах команд SCSI.
- Совместим с SATA
- SAS поддерживает большое количество устройств ( $> 16384$ ), в то время как интерфейс SCSI поддерживает 8, 16, или 32 устройства на шине.

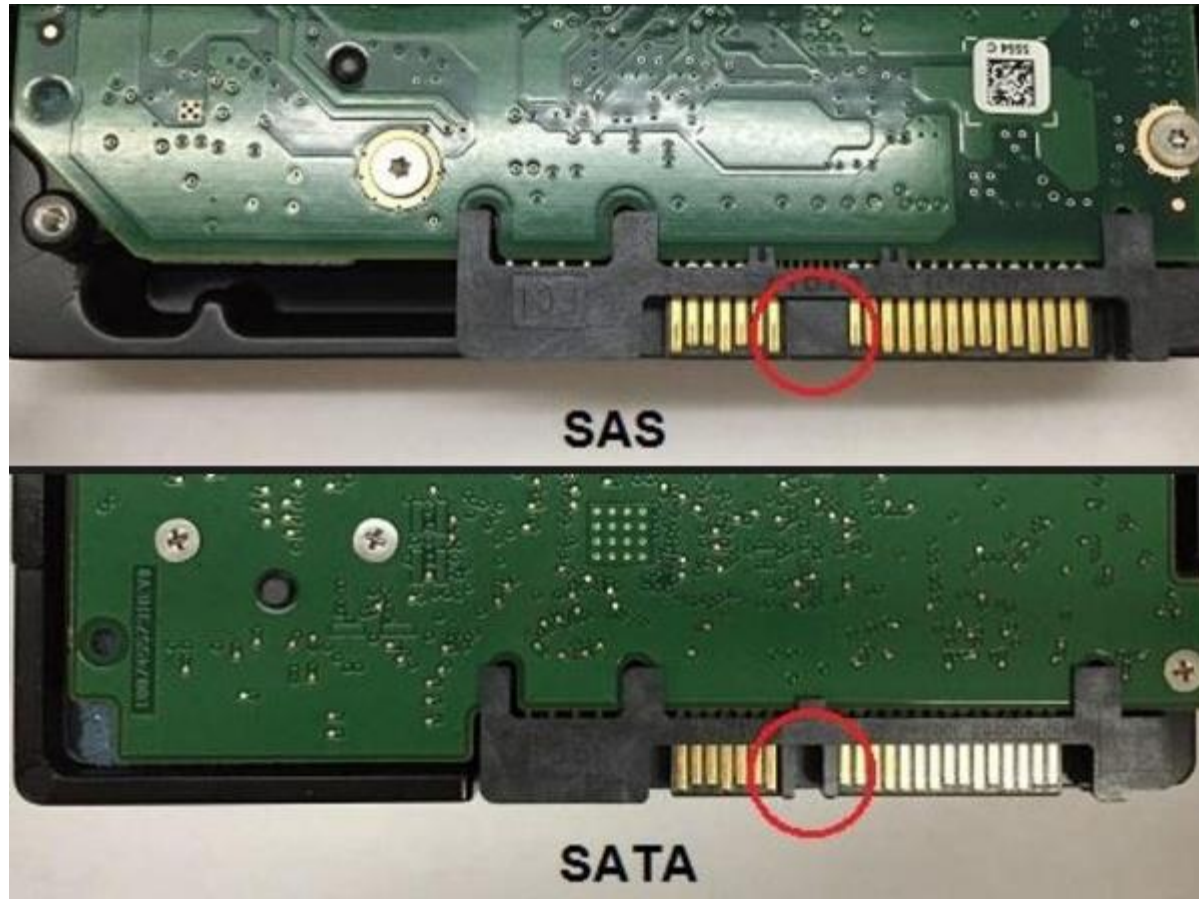
# SAS

- SAS-1 - 3.0 Gbit/s - 2004
- SAS-2 - 6.0 Gbit/s - 2009
- SAS-3 - 12.0 Gbit/s - 2013
- SAS-4 - 22.5 Gbit/s (24G) - 2017

## SAS Standards

	3Gb/s SAS (SAS-1)	6Gb/s SAS (SAS-2)	12Gb/s SAS (SAS-3)
			
Raw bandwidth (Gbit/s)	3 Gb/s	6 Gb/s	12 Gb/s
Raw bandwidth (Gbit/s) Wide Port (4 Ports)	12 Gb/s	24 Gb/s	48 Gb/s
Encoding	8b/10b 		
Transfer speed (MB/s)	300 MB/s	600 MB/s	1.200 MB/s
Transfer speed (MB/s) Wide Port (4 Ports)	1.200 MB/s	2.400 MB/s	4.800 MB/s
Discovery	by initiators	by SAS Expanders	
SATA compatible (operation of SATA HDDs on SAS bus)	yes		
Zoning	vendor-specific	T10 standardised	
Spread-Spectrum-Clocking		yes	
Connection Multiplexing (3Gb/s in 6Gb/s)		yes <sup>[1]</sup>	

# SAS vs SATA



# SAS. Основные сущности

- Инициатор (*Initiator*) — устройство, которое порождает запросы на обслуживание для *целевых устройств* и получает подтверждения по мере исполнения запросов. Чаще всего инициатор выполняется в виде.
- Целевое устройство (*Targets*) содержит логические блоки и целевые порты, которые осуществляют приём запросов на обслуживание, исполняет их; после того, как закончена обработка запроса, инициатору запроса отсылается подтверждение выполнения запроса. Целевое устройство может быть как отдельным жёстким диском, так и целым дисковым массивом.
- Подсистема доставки данных (*Service Delivery Subsystem*) Является частью системы ввода-вывода, которая осуществляет передачу данных между инициаторами и целевыми устройствами. Обычно подсистема доставки данных состоит из кабелей, которые соединяют инициатор и целевое устройство. Дополнительно, кроме кабелей в состав подсистемы доставки данных могут входить *расширители SAS*.
- Расширители (экспандеры, *Expanders*) SAS — устройства, входящие в состав подсистемы доставки данных и позволяют облегчить передачи данных между устройствами SAS; например, расширитель позволяет подключить несколько целевых устройств SAS к одному порту инициатора. Подключение через расширитель является абсолютно прозрачным для целевых устройств.



Кеннет Хотон

# HDD vs SSD

- Накопитель на жёстких магнитных дисках, (HDD, HMDD), жёсткий диск, винчестер — запоминающее устройство (устройство хранения информации) произвольного доступа, основанное на принципе магнитной записи.
- Твердотельный накопитель (solid-state drive, SSD) — компьютерное энергонезависимое немеханическое запоминающее устройство на основе микросхем памяти. Кроме микросхем памяти, SSD содержит управляющий контроллер. Наиболее распространённый вид твердотельных накопителей использует для хранения информации флеш-память типа NAND.



# HDD vs SSD

Устройство	Тип	IOPS	Интерфейс
7,200 об/мин SATA-диски	HDD	~75-100 IOPS	SATA 3 Гбит/с
10,000 об/мин SATA-диски	HDD	~125-150 IOPS	SATA 3 Гбит/с
10,000 об/мин SAS-диски	HDD	~140 IOPS	SAS
15,000 об/мин SAS-диски	HDD	~175-210 IOPS	SAS

Устройство	Тип	IOPS	Интерфейс
Intel X25-M G2 MLC	SSD	~8 600 IOPS	SATA 3 Гбит/с
OCZ Vertex 3 MAX IOPS	SSD	~75 000 IOPS	SATA 6 Гбит/с
OCZ Vertex 4	SSD	~120 000 IOPS	SATA 6 Гбит/с

# SSD SATA vs SSD SAS

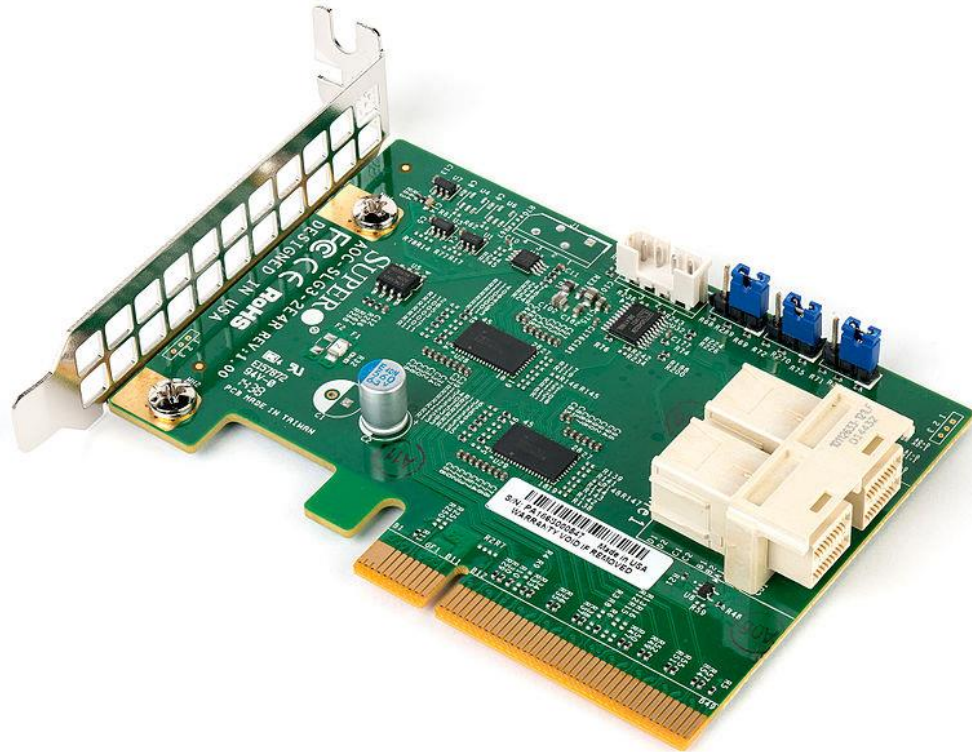
- Потолок производительности SAS 6Gb при интенсивной нагрузке в половину выше потолка SATA 6Gb - только за счет полнодуплексной передачи и особенностей протокола. Накладные расходы у SAS всегда ниже;
- Переход к SAS 12Gb увеличивает отрыв;
- SAS работает с множеством устройств как единой сетью;
- Длина очереди SATA всего 32 команды, а SAS – от 256;

# NVMe

**NVM Express (NVMe, NVMeHCl** — *OT Non-Volatile Memory Host  
Controller Interface Specification)*

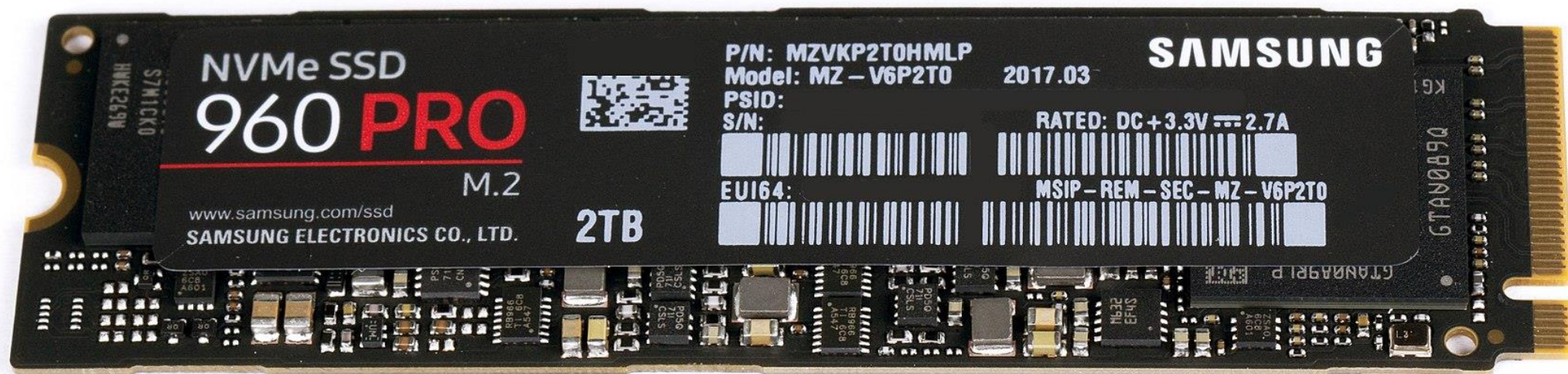


# NVMe





# NVMe (в форм-факторе M.2)



# NVMe



PCI   
**EXPRESS**<sup>®</sup>

[AHCI]



***nvm***  
**EXPRESS**

# NVMe

- NVMe 1.1b — 2014
- NVMe 1.2 — 2014; для устройств потребительского класса: расширенное управление питанием и поддержка накопителей без динамической памяти; для устройств корпоративного класса: возможность обновления прошивки без остановки работы накопителя, снижение задержек в топологиях с большим количеством NVMe-накопителей и коммутаторами PCIe
- NVMe 1.2a — 2015
- NVMe 1.3c — 2018
- Типичные скорости около 2000..2500 Мб/с

# NVMe



## **INTEL Optane 900P SSDPED1D480GASX 480Гб**

Этот накопитель на 480 Гб  
подключается уже в слот PCIe.  
Скорости чтения и записи у него  
2500 / 2000 Мб/с.