Data Organization and Processing

(NDBI007)

David Hoksza, Petr Škoda
http://siret.ms.mff.cuni.cz/hoksza

Lecture Outline

- **Data storage systems**
  - **primary storage/memory**
  - **secondary storage/memory**
    - magnetic disc
    - disc interface
    - RAID
    - SSD
  - **tertiary storage/memory**
    - optical disc
    - magnetic tape

Memory Classification (1)

- **Mutability**
  - **read only**
    - allows reading but not modification
  - **read/write**
  - **WORM** (Write Once Read Multiple)
  - **slow write/fast read**
    - write operation much slower than the read operation

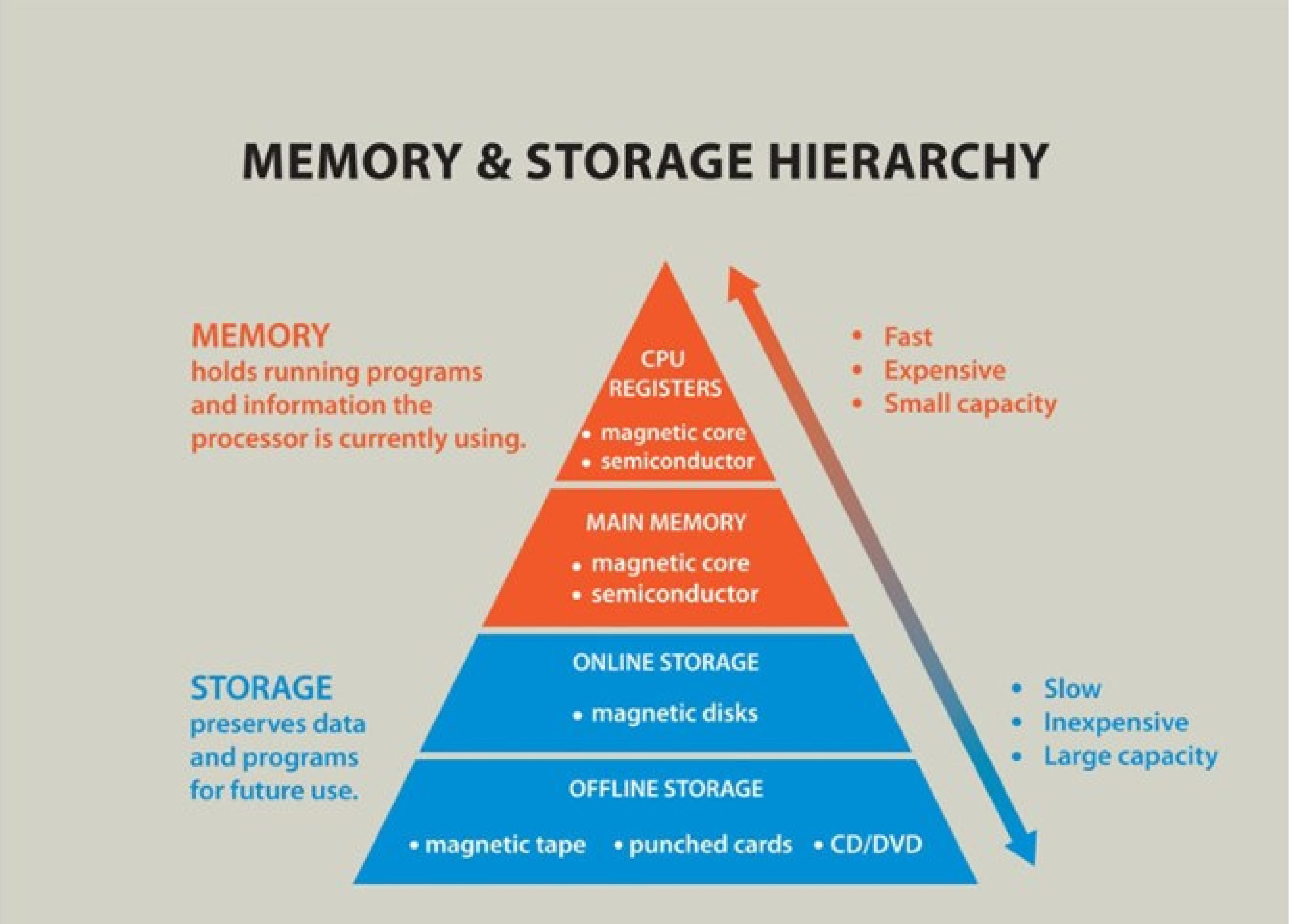
- **Accessibility**
  - **random access**
    - retrieval from **any location** takes approximately the **same** amount of **time**
    - **latency** is **independent** on the location
  - **sequential access**
    - data can only be **accessed** in **sequential manner**, i.e. one has to **scroll** to the desired location

Memory Classification (2)

- **Performance**
  - **latency**
    - access time
  - **throughput**
    - the speed at which data can be transmitted
- **Cost**
  - cost per unit of data
- **Capacity**
  - **raw capacity**
  - **storage density**

- **Reliability / Volatility**
  - **volatile** – data are lost when power supply is withdrawn
    - faster regarding both reading and writing data
  - **non-volatile** – data persist even in the absence of active power source
    - slower
    - long term storage

Memory Hierarchy



Memory Hierarchy

- **Primary** memory
  - fastest, volatile (needs power supply on to keep the information)
  - CPU registers, caches, main memory

- **Secondary** memory
  - moderate access time, non-volatile
  - not accessible by the CPU
  - online storage
  - magnetic disks, SSD disks

- **Tertiary** memory
  - slow access time, non-volatile
  - offline storage (removable)
  - floppy disks, optical disks, magnetic tapes

Primary Memory

- **Register**
  - inside processor
  - **volatile**
  - used by arithmetic and logic unit
  - 32/64 bit (word of data)
  - fastest and most costly
- **Cache**
  - inside processor or disk
  - **volatile**
  - most often used data from main memory are stored in a CPU cache
  - managed by HW or operating system
  - can be hierarchized
- **Main memory**
  - general-purpose machine instructions operate on data resident in the main memory
  - fast access, but generally too small to store the entire data set
  - **volatile**
  - connected to the processor

Secondary Memory

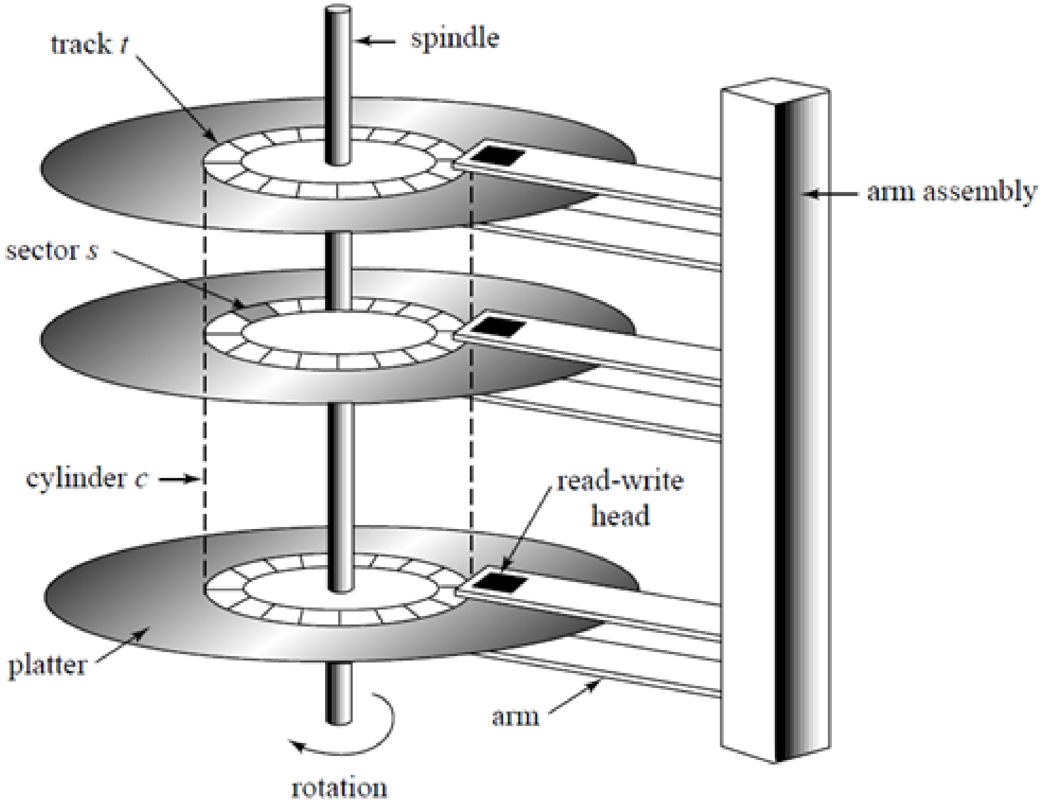
- **Magnetic disk**
  - **non-volatile**
  - data must be moved from disk to main memory for access and written back to storage
  - random access
- **Flash memory**
  - **non-volatile**
  - memory cards, USB disks, solid-state drives (SSD)

Tertiary Memory

- **Optical disk**
  - non-volatile
  - CD ROM, DVD ROM, Blu-ray, …

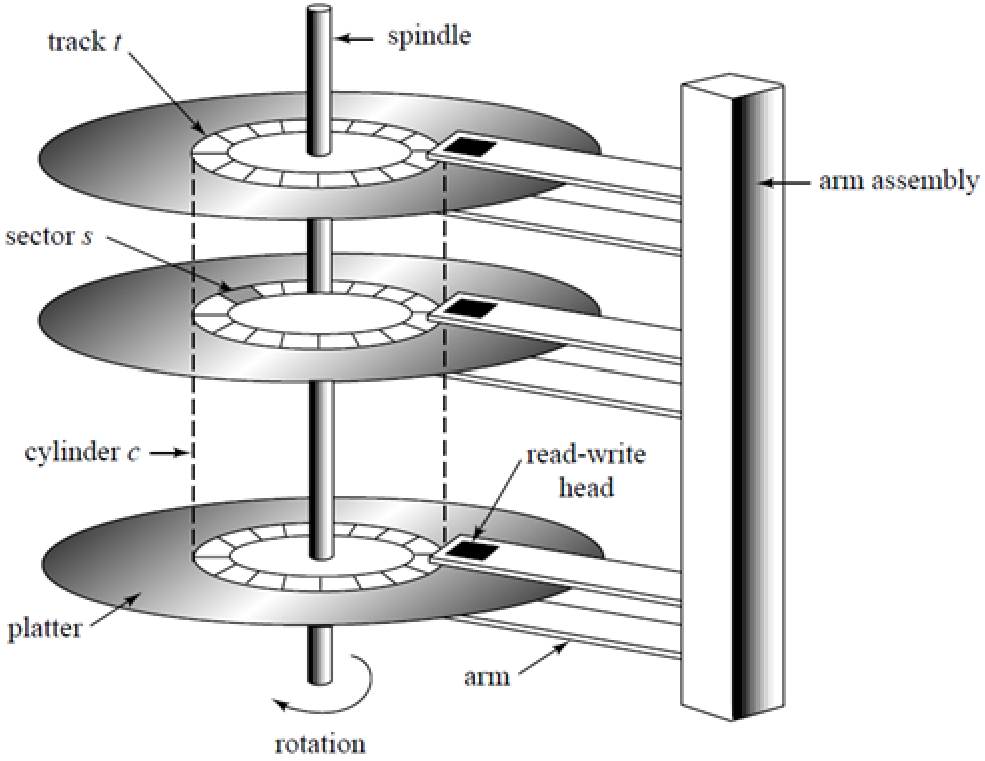
- **Magnetic tape**
  - non-volatile
  - sequential access
  - very high capacity and persistence
  - cheap
  - used for backup

Magnetic Disk (1)



- **Disk pack** (*disková sestava*) consists of multiple **platters** (*disky/plotny*) on a **spindle** (*osa*)
  - platters are usually double-sided
- Data read by read-write **head** (*hlava)*
  - kept on an **arm** (*raménko*)
  - arms kept on the **arm assembly** (*vystavovací mechanismus*)
    - 1 disk - 2 read-write heads (1 head per surface)

Magnetic Disk (2)



- Surface of platters divided into **tracks** (*stopy*)

- **Tracks** are divided into **sectors** (*sektory*)

- Set of all tracks with the same diameter form a **cylinder** (*válec*)

Magnetic Disk (3)

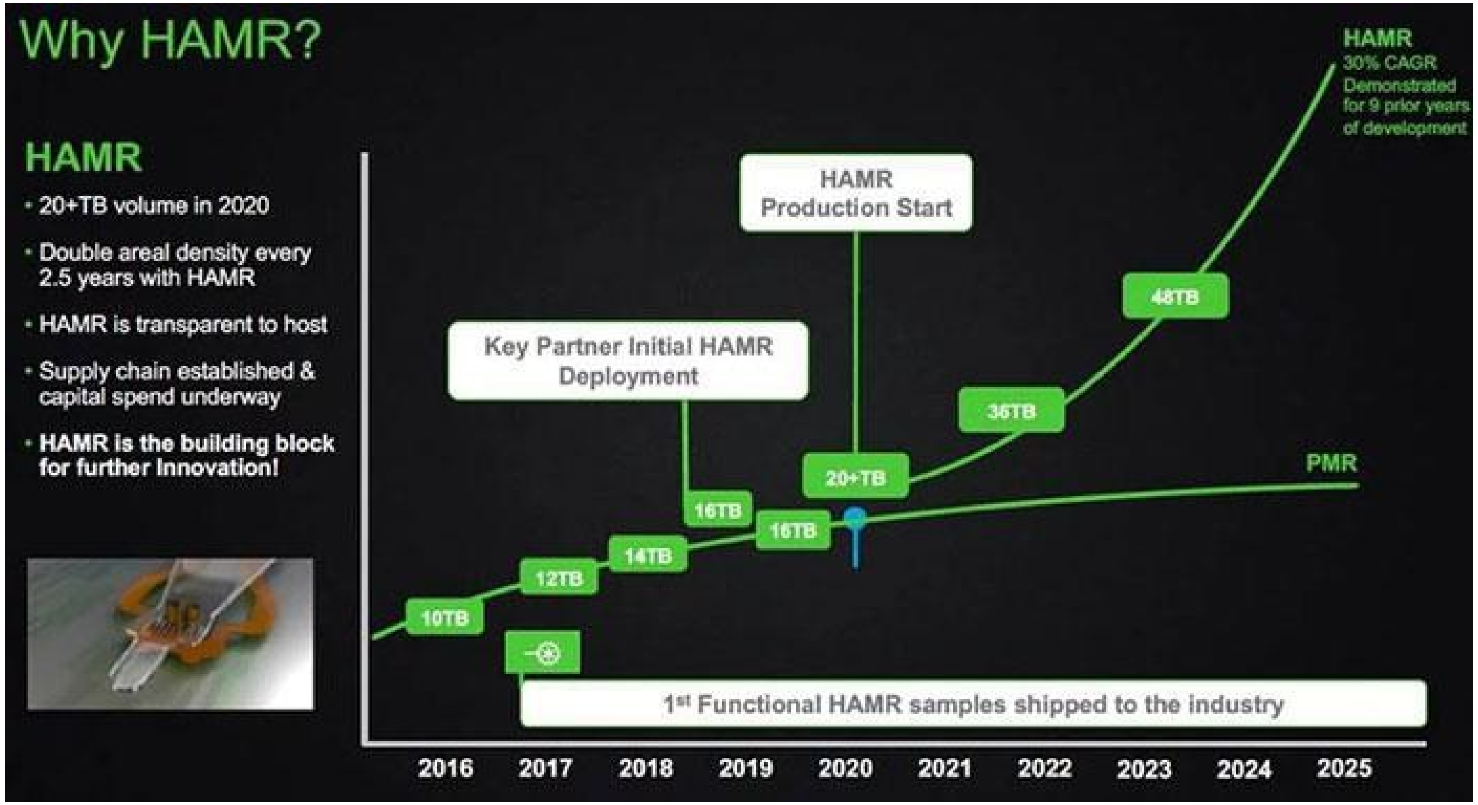
- **Sector**
  - define **minimal** amount of **information** to read or write
  - smallest addressable unit
  - 512B, 4KB (common nowadays)
  - different number of sectors on inner and outer tracks

- **Head**
  - Heads flow close to the magnetic surface on air cushion created by the spinning disks
  - information is **magnetically coded by magnetizing a region** by generating strong local magnetic field

Magnetic Disk

2019.01.08 https://toshiba.semicon-storage.com/ap-en/product/storage-products/enterprise-hdd/mg08.html

Formatted Capacity	16 TB
Buffer Size	512 MiB
Data Transfer Speed (Sustained)	262 MiB/s
Rotation Speed	7,200 rpm
Sector	4K native (4Kn) 512 emulation (512e)



Why HAMR?

HAMR

- 20+TB volume in 2020
- Double areal density every 2.5 years with HAMR
- HAMR is transparent to host
- Supply chain established & capital spend underway
- **HAMR is the building block for further Innovation!**

HAMR
30% CAGR
Demonstrated for 9 prior years of development
HAMR Production Start
Key Partner Initial HAMR Deployment
48TB
36TB
20+TB
16TB
16TB
14TB
12TB
10TB
PMR
1st Functional HAMR samples shipped to the industry
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

2018.06.11 https://www.guru3d.com/news-story/new-storage-roadmap-shows-100tb-hdds-in-2025.html

Magnetic Disk (3)

Each of the drives attached to this computer has a block size of 4096; each time we read one byte, the drive gives us a block containing 4096 bytes. If we organize our data structure carefully, this means that each disk access could yield 4096 bytes that are helpful in completing whatever operation we are doing.

http://opendatastructures.org/versions/edition-0.1e/ods-cpp/14_External_Memory_Searchin.html1.7

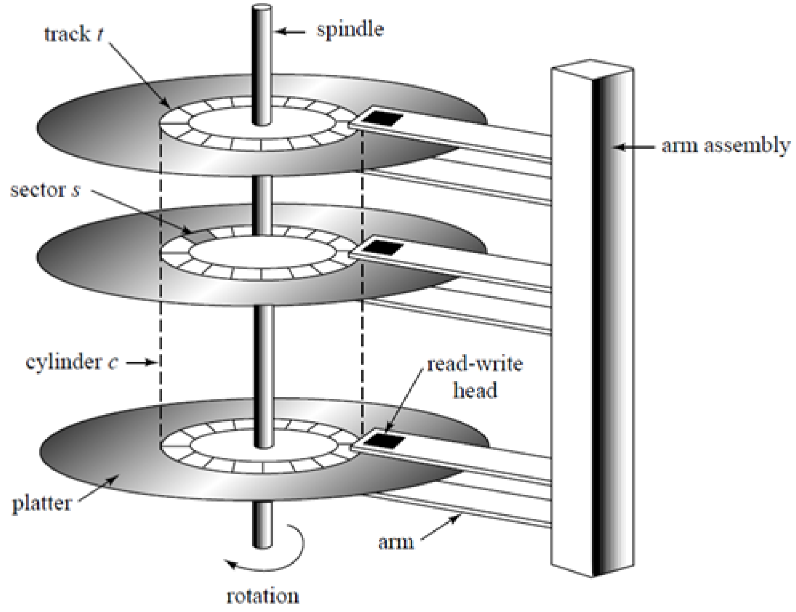
Even when a single bit is required, the whole sector needs to be transfered → data structures need to be adapted to such environment.

Addressing (CHS)

- Addressing matching the physical make up of early drives → **geometry-based** access

- **CHS** address (**cylinder-head-sector** address)

cylinder number	10 bit → 0 .. 1023
head number	8 bit → 0 .. 254
sector number	6 bit → 1 .. 63

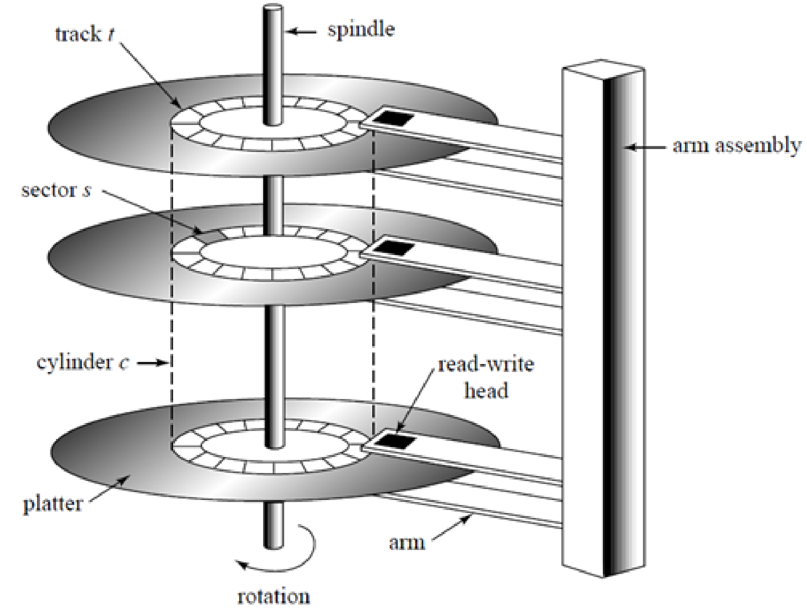


- maximum active primary partition size = 8GiB $(2^{24} \times 512)$

Addressing (LBA)

- **CHS** addressing has **low maximum** size limit and **does not map well** to **non-magnetic** disk types (tape, SSD, …)

- **LBA (logical block addressing)**
  - **linear** addressing scheme
  - each **sector** is assigned a **unique** sector **number**
    - **sequential** numbering starting from 0
  - in order to work, must be supported by the disk, BIOS and O



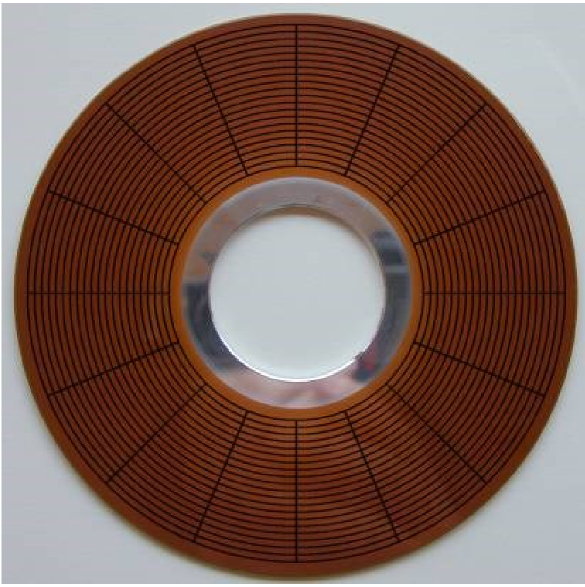
- **CHS to LBA** conversion

$$\boldsymbol{LBA = (C \times N_{heads} + H) \times SPT + (S}$$

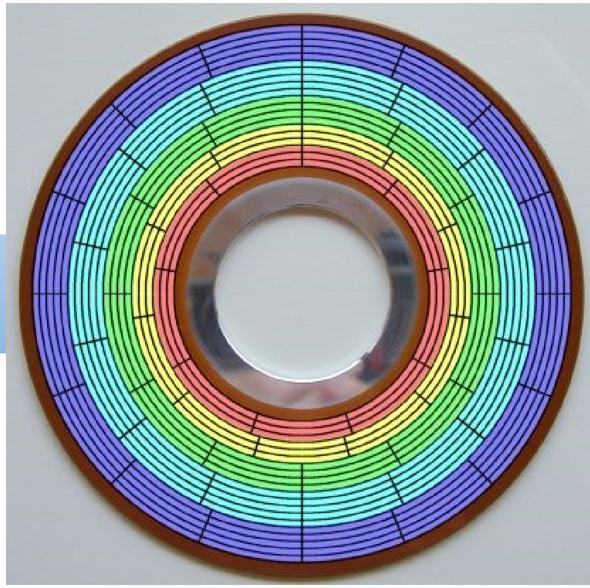
  - N_{heads} number of heads
  - SPT sectors per track

Zoned Bit Recording

- Earlier disks had the **same number of sectors per track**
  - controllers couldn't handle complicated arrangements and all tracks had the same number of sectors
  - **inner tracks** as **dense** as possible
  - **outer tracks underutilized** by reducing bit density



- **Zoned bit recording (ZOR)**
  - **tracks grouped** into **zones**
  - **each zone** assigned a number of **sectors per track**
    - tracks **close to the outer edge** contain **more sectors per track**
  - **data transfer rate** for **outside** cylinders is **higher**, since the **angular velocity** is **constant** regardless of which track is being read
    - hard disks are filled from the outside in, the fastest data transfer occurs when the drive is first used → possible problems when benchmarking



Magnetic Disk Parameters (1)

- **Access time components**
  - **s – seek**
    - average seek time from one random track (cylinder) to any other
  - **r - rotational delay** (latency)
    - one revolution equals 2r (r is **average** latency)
  - **btt (block transfer time)**

- **Random access**
  - **set heads** → wait for the disk to roll the correct position → data transfer
  - **s + r + btt**

Magnetic Disk Parameters (2)

- Seek time
  - 3ms – 15ms
  - usually between 8 and 12ms
- RPM (Revolutions Per Minute)
  - 4,200 – 15,000
  - more revolutions → more energetically demanding
- Rotational latency

Speed (RPM)	Average latency (ms)
15,000	2
10,000	3
7,200	4.16
5,400	5.55
4,800	6.25

Magnetic Disk Parameters (3) – Transfer Rates

- Depends on cylinder position
- **(average) media transfer rate**
  - refers *only* to the speed of **reading** or **writing bits to a *single* track** of one surface of the disk (no positioning, no track or head switching, only inside of the disk)
- **interface/external transfer rate**
  - speed with which bits can be moved to **(from) the hard disk platters from (to) the hard disk's integrated controller**
  - purely electronic operation, which is typically much faster than the mechanical operations involved in accessing the physical disk platters
- **(average) sustained/sequential transfer rate**
  - **real-world transfer rate** when a file spans multiple platters and cylinders
  - **media transfer rate + head switch time** (electronic operation) **+ cylinder switch time**
  - 100-200MB/s

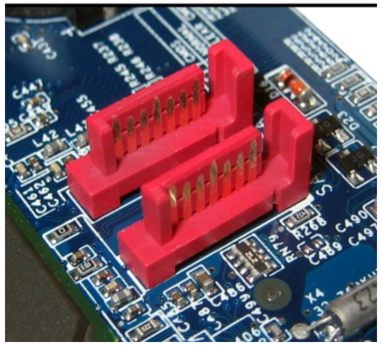
Disk Subsystem

- **Bus** (*sběrnice*)
  - bus is physical and logical **infrastructure** for **transferring data** between components (such as drives) and PC
  - PATA, SATA, Fiber Channel, SCSI, …

- **Controller** (*řadič*)
  - **interface** between **disk** and the **system**
  - accepts instructions from the system to read and write data
  - controller on the side of the motherboard's bus is called **host bus adapter** (HBA) and controller on the side of a disk is **disk controller**
  - todays disks controllers include logic for checksum, validation, remapping bad sectors

Disk Interface (1)

- **PATA** (Parallel Advance Technology Attachment)
  - originally called **ATA** (until the introduction of SATA)
  - Integrated Drive Electronics (IDE) → AT Attachment (ATA) + AT Attachment Packet Interface (ATAPI) → Parallel ATA (PATA)
  - **parallel**
    - multiple bits can be transferred simultaneously
  - **up to 167 MB/s**
  - **ATAPI (ATA Packed Interface)**
    - introduced to allow ATA to be used with various devices by including additional commands
- **SATA** (Serial ATA)
  - **replacement of PATA**
  - faster data transfer
  - enables **hotplug**
  - modifications for different device types
    - eSATA (external SATA) – external devices
    - mSATA (mini SATA) – netbooks, SSD, …
  - **up to 600 MB/s**



Disk Interface (2)

- **SCSI** (Small Computer System Interface)
  - set of standards for transferring data between computer and devices
    - commands, protocols, HW interfaces, …
    - magnetic disks, optical drives, printers, …
  - allows to connect up to 16 devices to single bus
  - **high-end**, more expensive solution used in server environment
  - **up to 640 MB/s**

- **Fiber Channel**
  - mainly for storage networking (**SAN** – storage area network)
  - communication using SCSI commands (FCP – Fiber Channel Protocol)
  - **high-end** solution
  - **up to 1600 MB/s**

RAID

- **Redundant Arrays of Inexpensive (Independent) Disks**
  - consists of multiple disk forming a logical unit
  - **Inexpensive**
    - original motivation
    - utilization of higher number of inexpensive disks
    - alternative to high-capacity expensive disks
  - **Independent**
    - present-day motivation
    - higher reliability – redundancy
    - higher bandwidth – parallelism
- RAID **levels** express different **cost, performance and reliability characteristics**
- Must be **supported by the controller**
  - responsible for the data distribution within the array

RAID – Reliability (1)

- Mean time to failure (**MTTF**) **of a system** is much **lower than** MTTF of an **individual device**
  - system with 100 disks each with MTTF 100,000 hours (11 years) will have system MTTF 1000 hours (41 days)

- **Redundancy** of information can help by storing multiple copies of data which are then used in case of failure

- **Mirroring** *(zrcadlení)*/**shadowing**
  - keeps copies of a disk → each **write is carried out on multiple disks**
  - data are read from one disk – if one goes awry, the backup disk can be utilized
  - the probability that the second disk will break down before repairing the first one is very low (if mean time to repair (MTTR) is 10 hours then for a two disk system Mean Time To Data Loss (MTTDL) is 57,000 years)

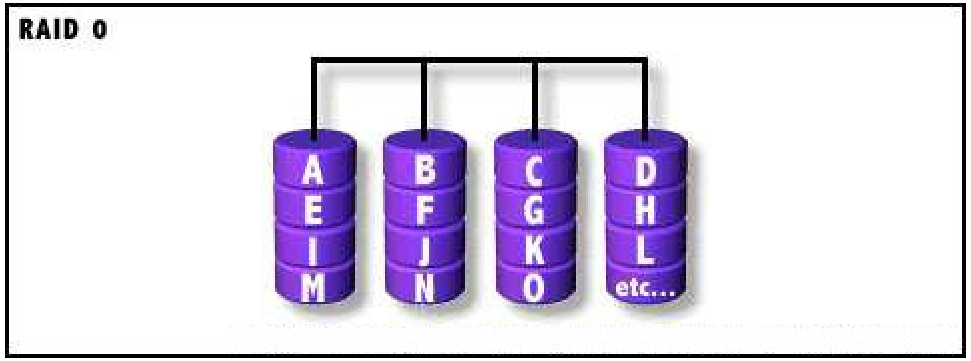
RAID – Reliability (2)

- **Parity**
  - uses **redundancy** to be able to **recover** missing data
  - example with 3 disks (D1, D2, D3), one parity disk (DP) and one hot spare disk (HS) used for recovery purposes
  - to calculate parity XOR operation can be used

```
D1: 00100101
D2: 11101001
D3: 10101101
DP: 01100001 (= D1 XOR D2 XOR D3)
```

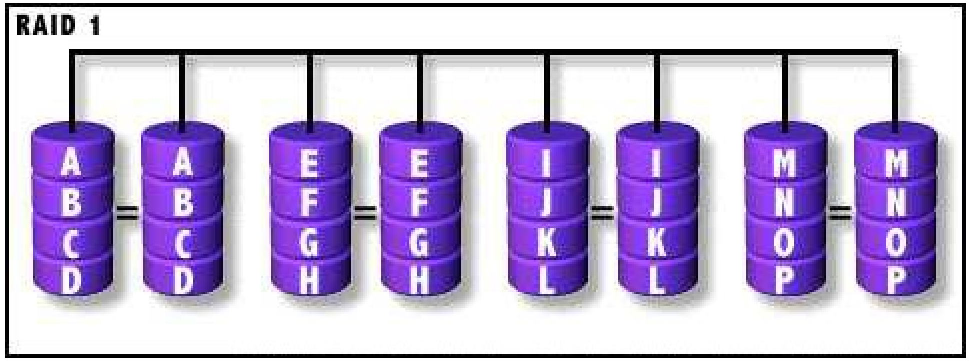
  - D2 breaks down
  - now the original values of D2 can be obtained from the parity information from DP (D1 XOR D3 XOR DP) and can be written to HS which can serve as a new D2

RAID 0



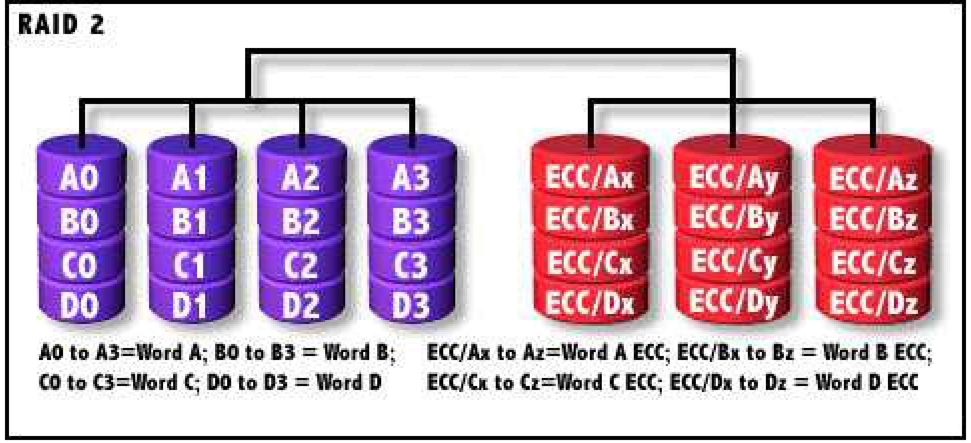
- Non-redundant **striping** *(prokládání)* (**block level**)
- Data striped – **each block on one disk**
  - block ***i*** written to disk (***i mod n***)
- Advantages
  - superior I/O performance
  - no overhead by parity controls
  - all storage capacity used
  - easy to implement
- Disadvantages
  - No mirroring or parity checking
    - → **no redundancy**
    - → **not fault-tolerant** (failure of one disk means loosing of all data)
- Suitability
  - high-performance but non-critical applications

RAID 1



- **Disk mirroring**
- Each write is done in parallel
- Data are read from the disk which can access the data faster
  - if supported by the controller, parallel reads can be supported
- Advantages
  - I/O speed comparable to single disk environment
  - in case of disk failure, data are only copied
  - easy to implement
- Disadvantages
  - high redundancy (50%)
  - usually does not allow hot swap of the failed disk
- Suitability
  - data-critical applications (storing log files, accounting systems, …)

RAID 2

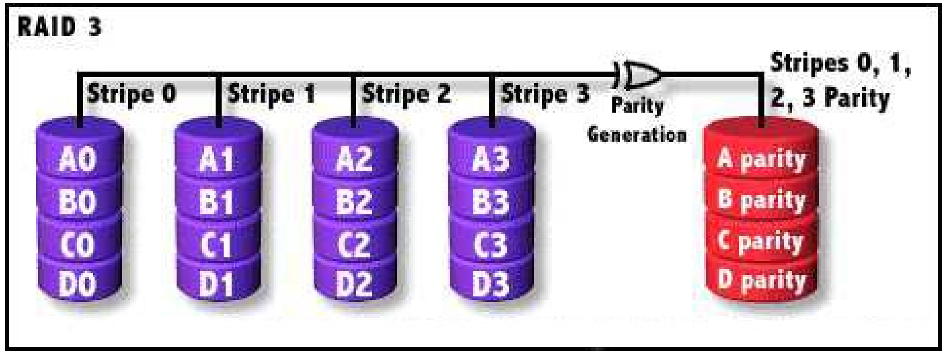


A0 to A3=Word A; B0 to B3 = Word B; C0 to C3=Word C; D0 to D3 = Word D

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

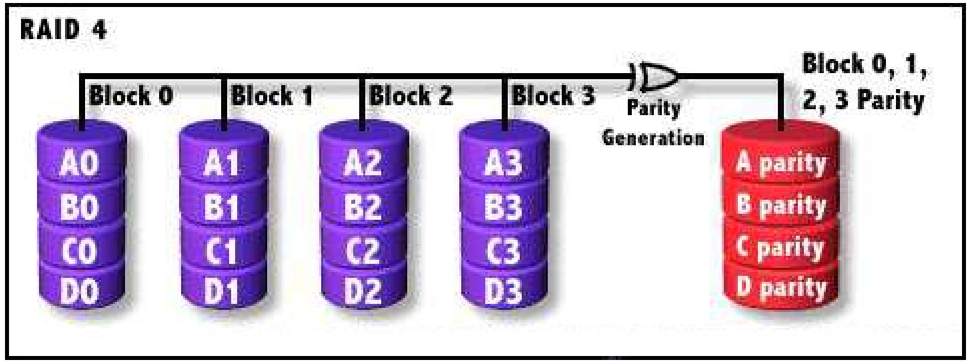
- **Bit-level striping**
- Hamming code parity
  - can detect up to two and correct up to one bit errors
  - the number of redundant disks is proportional to the log of the total number of the disks on the system
- Rarely used

RAID 3



- **Byte-level striping**
- **One parity disk** (XOR)
  - each write requires one more write (and computation) of the parity bit
  - parity disk checked on read
  - **performance bottleneck**
- Advantages
  - high-throughput for large I/O
    - the size of a request is always higher than the number of stripes (minimum size is the size of the sector)
- Disadvantages
  - resource intensive (at least 3 disks)
  - slow for small I/O operations
  - I/O requires activity on every disk
- Rare – **substituted by RAID 5**

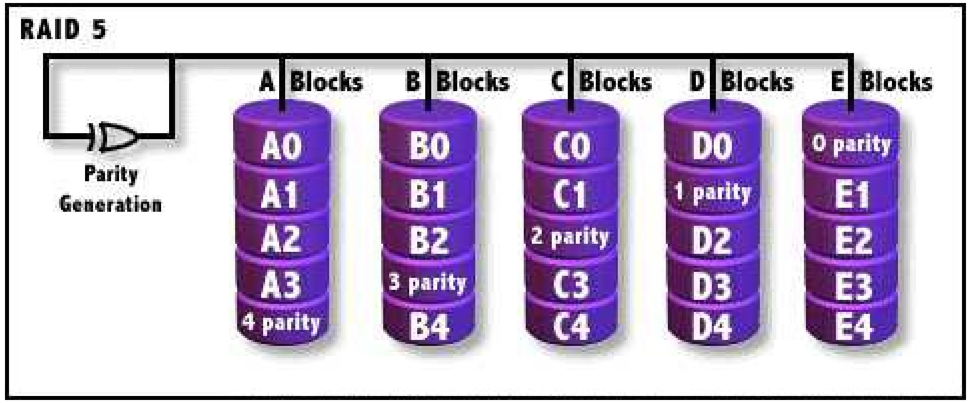
RAID 4



- **Block-level striping** with **parity on a separate disk**
- Dedicated parity disk
  - **performance bottleneck**
  - can handle single disk failure
- Resembles
  - RAID 0 + parity disk
  - RAID 3 but block-level
  - RAID 5 but without distributed parity
- Rare – **substituted by RAID 5**

- Advantages
  - **I/O requests can be carried out in parallel** if the blocks are on different disks
    - if supported by the controller
- Disadvantages
  - lot of small write operations can be problematic (parity disk)
    - four disk IOs: 1 to write the new data, 2 to read the old data and old parity for computing the new parity, 1 to write the new parity
  - complex controller design

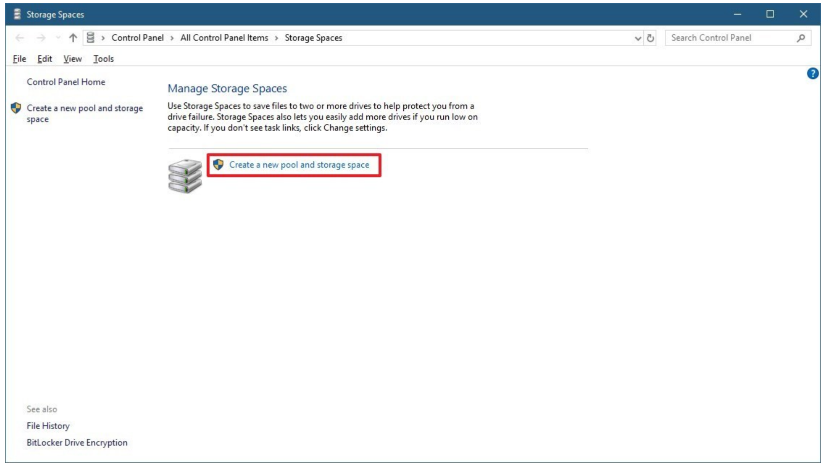
RAID 5



- **Block-level striping with distributed parity**
  - low redundancy
- Most common
- Reads do not check the parity block (too expensive)
- Can handle single disk failure

- Advantages
  - high-throughput read operation
    - Even the “parity disk” is utilized (unlike RAID 4)
  - good aggregate transfer rate
- Disadvantages
  - write operation is slower (parity computation)
  - at least 3 disks

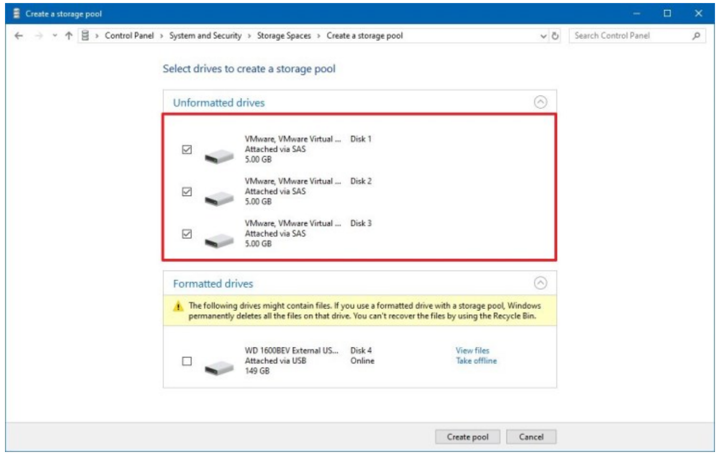
RAID on Win10


Storage Spaces
Control Panel › All Control Panel Items › Storage Spaces
Search Control Panel
File
Edit
View
Tools
Control Panel Home
Create a new pool and storage space
See also
File History
BitLocker Drive Encryption

Manage Storage Spaces

Use Storage Spaces to save files to two or more drives to help protect you from a drive failure. Storage Spaces also lets you easily add more drives if you run low on capacity. If you don't see task links, click Change settings.

Create a new pool and storage space


Create a storage pool
Control Panel › System and Security › Storage Spaces › Create a storage pool
Search Control Panel

Select drives to create a storage pool

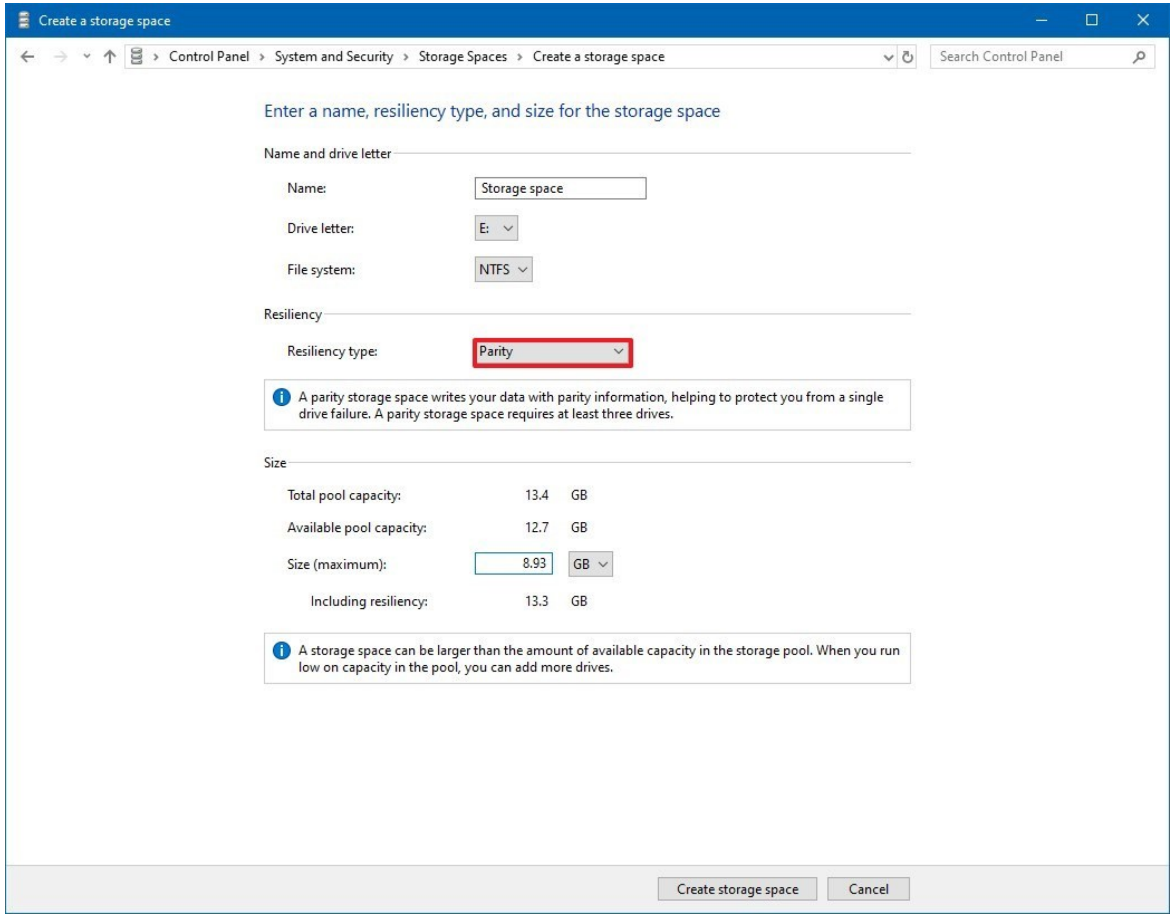
Unformatted drives

- ☑ VMware, VMware Virtual ... Disk 1, Attached via SAS, 5.00 GB
- ☑ VMware, VMware Virtual ... Disk 2, Attached via SAS, 5.00 GB
- ☑ VMware, VMware Virtual ... Disk 3, Attached via SAS, 5.00 GB

Formatted drives

The following drives might contain files. If you use a formatted drive with a storage pool, Windows permanently deletes all the files on that drive. You can't recover the files by using the Recycle Bin.

- □ WD 1600BEV External US... Disk 4, Attached via USB, Online, 149 GB — View files, Take offline


Create a storage space
Control Panel › System and Security › Storage Spaces › Create a storage space
Search Control Panel

Enter a name, resiliency type, and size for the storage space

Name and drive letter

Name: Storage space

Drive letter: E:

File system: NTFS

Resiliency

Resiliency type: Parity

A parity storage space writes your data with parity information, helping to protect you from a single drive failure. A parity storage space requires at least three drives.

Size

Total pool capacity: 13.4 GB

Available pool capacity: 12.7 GB

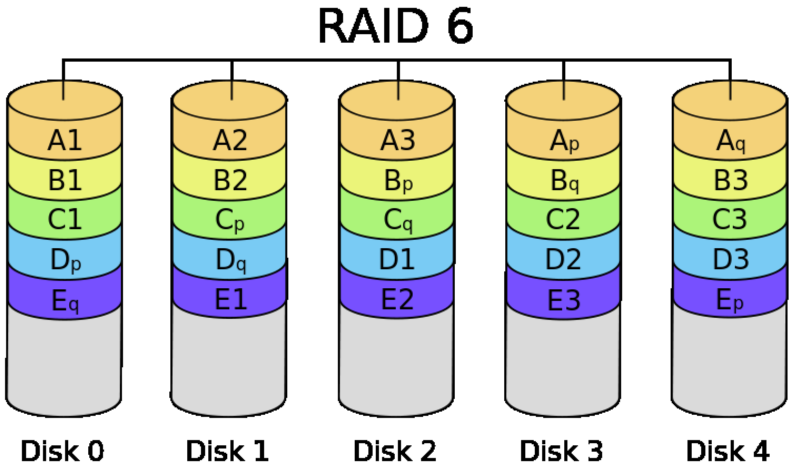
Size (maximum): 8.93 GB

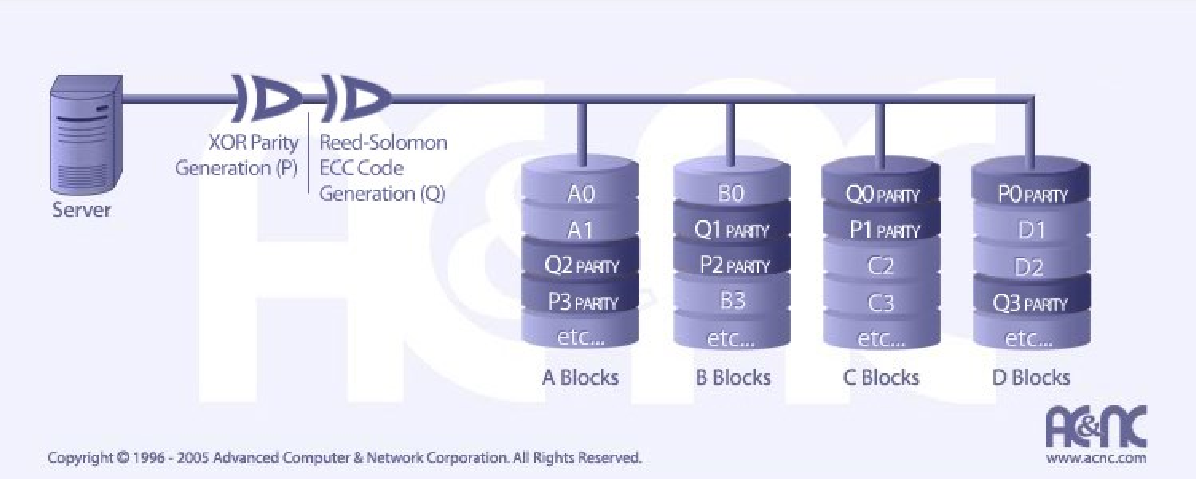
Including resiliency: 13.3 GB

A storage space can be larger than the amount of available capacity in the storage pool. When you run low on capacity in the pool, you can add more drives.

RAID 6

- **Block-level striping with 2 distributed parity blocks**
  - extension of RAID 5
- Can handle failure of 2 drives
  - RAID 5 for enterprise environments
- Advantages
  - additional fault tolerance
- Disadvantages
  - expensive write
- Suitability
  - mission critical applications

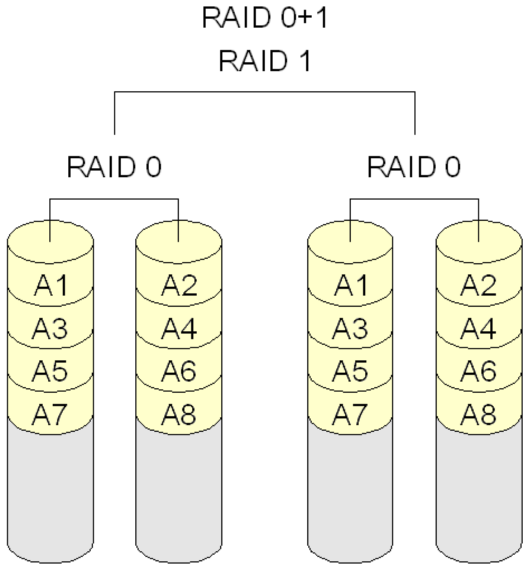




RAID 0+1 – Mirrored Stripes

- Combination of **RAID 0** (high performance) with **RAID 1** (high reliability)
- **Mirrored array whose segments are RAID 0 arrays**
- **6 disks**
  1. **2 sets of 3 disks (RAID 0)**
  2. **turn each set into RAID 0 array and mirror the two sets (RAID 1)**

- Advantages
  - additional fault tolerance
  - high data transfer rate
  - reliability as in RAID 5 without parity computation (more than one disk can fail but only in one RAID 0 array)
- Disadvantages
  - high overhead
  - limited scalability
  - requires 4 disks

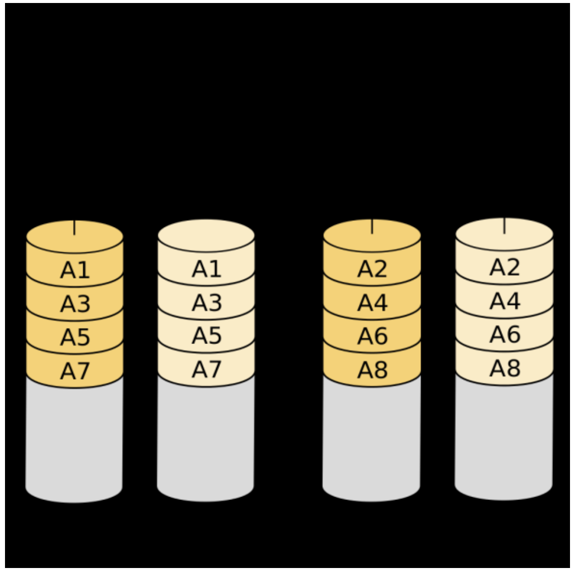


RAID 1+0 – Striped Mirrors

- Combination of **RAID 0** (high performance) with **RAID 1** (high reliability)

- **Striped array whose segments are RAID 1 arrays**
- **6 disks**
  1. **3 sets of 2 disks (RAID 1)**
  2. **stripe data across the 3 sets (RAID 0)**

- Advantages
  - very high reliability (in each RAID 1 array, 1 disk can fail)



- Disadvantages
  - high overhead
  - limited scalability
  - requires 4 disks

- Suitability
  - Database server requiring high performance and fault tolerance

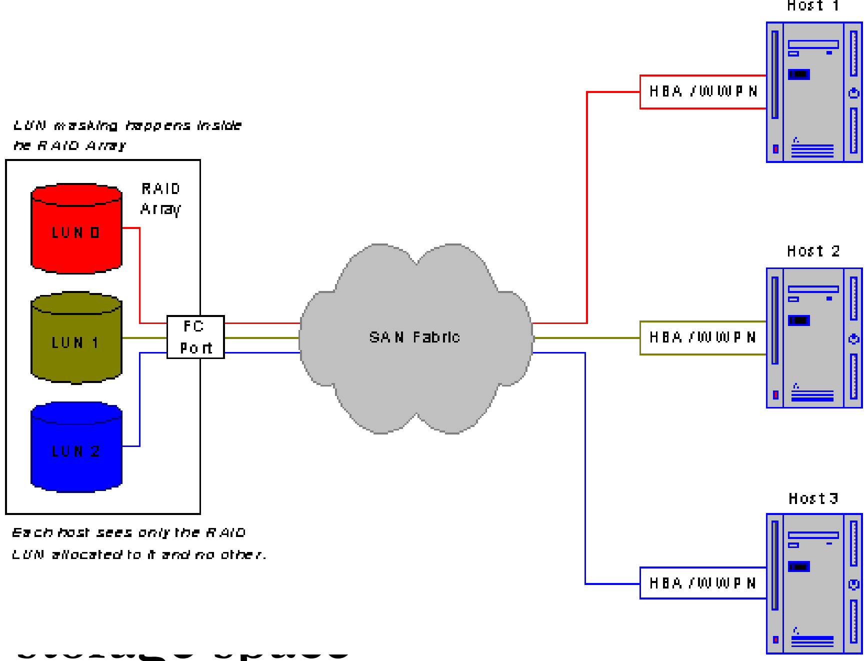
Disk Attachment Strategies

- **DAS** (Direct Attached Storage)
  - **block-level storage**
  - data storage connected directly to a server or a computer (no network device between)
  - protocols – ATA, SATA, Fibre Channel, …
- **NAS** (Network Attached Storage)
  - **file-level storage**
  - single data storage device (**file server**) connected to network providing storage capacity for a set of heterogeneous clients
  - accessed by mapping (\\NAS\share) – NAS addresses data by filename and offset
  - file system managed by NAS OS
- **SAN** (Storage Area Network)
  - **block-level storage**
  - accessed as drive (E:\) – SAN addresses data by logical block number
  - multi-server multi-storage network
  - local network of multiple data storages
  - file system managed by server

NAS

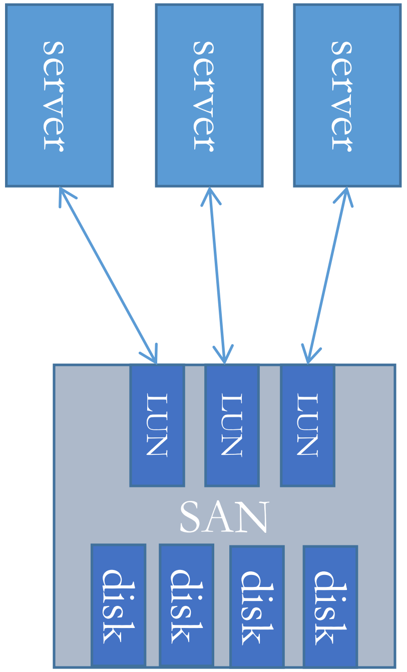
- Computer connected to a network **designed to store and provide data**
  - **File-level storage**
  - file system slimmed down
  - **self-contained solution**
    - usually comes as a specialized device (although also a standard computer can be used)
  - NAS storage **can contain multiple standard disks** (can be in RAID)
- Often used for simply sharing files between multiple devices (Linux, Windows, iOS, Android)
- Can handle many network connections
- Access through **network file sharing protocols**
  - NFS (Unix), SMB/CIFS (Windows), …
  - over TCP/IP
  - performance highly influenced by the network

SAN


LUN masking happens inside he RAID Array

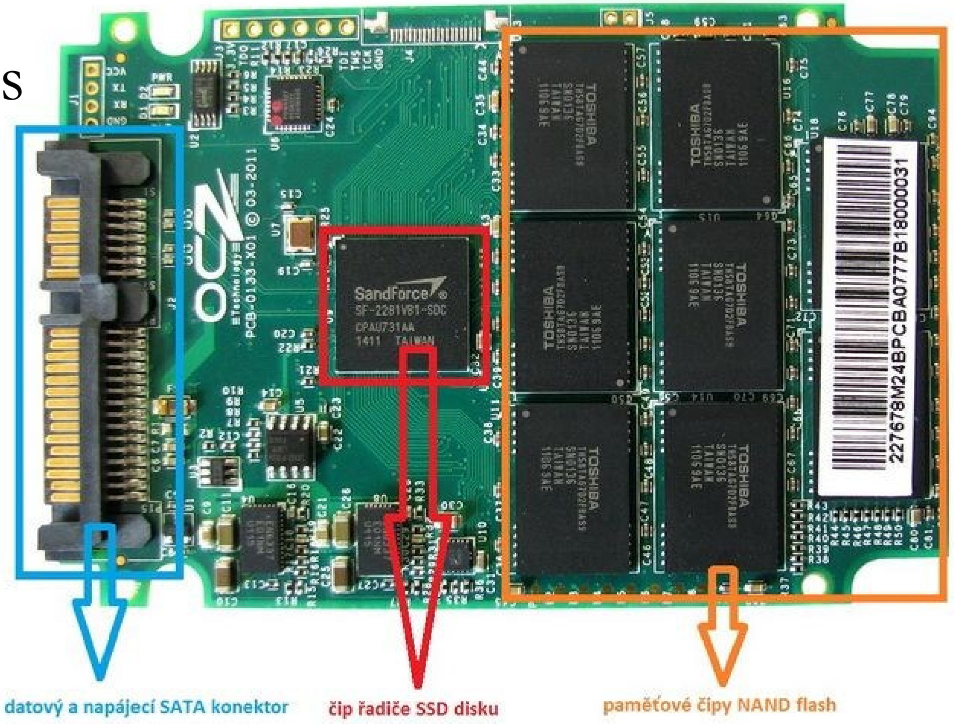
Each host sees only the RAID LUN allocated to it and no other.

- **Block-level storage**
- Dedicated network for data storage
- Usually only server accesses SAN (not clients)
- Advantage over dedicated storage is in the ability to **reallocate storage space**
    - data in SAN are divided into LUNs (Logical Unit Number), i.e. virtual partitions
- Protocols
    - iSCSI (Internet Small Computer System Interface) [:aiskazi:]
        - TCP/IP-based protocol for establishing and managing connections between IP-based storage devices, hosts and clients
        - uses existing ethernet network to connect to SAN
        - lower performance and cost
    - **Fibre Channel**
        - requires dedicated switch → also card in the server
        - higher performance and cost
    - **FCoE (Fibre Channel over Ethernet)**



Solid State Drive (SSD)

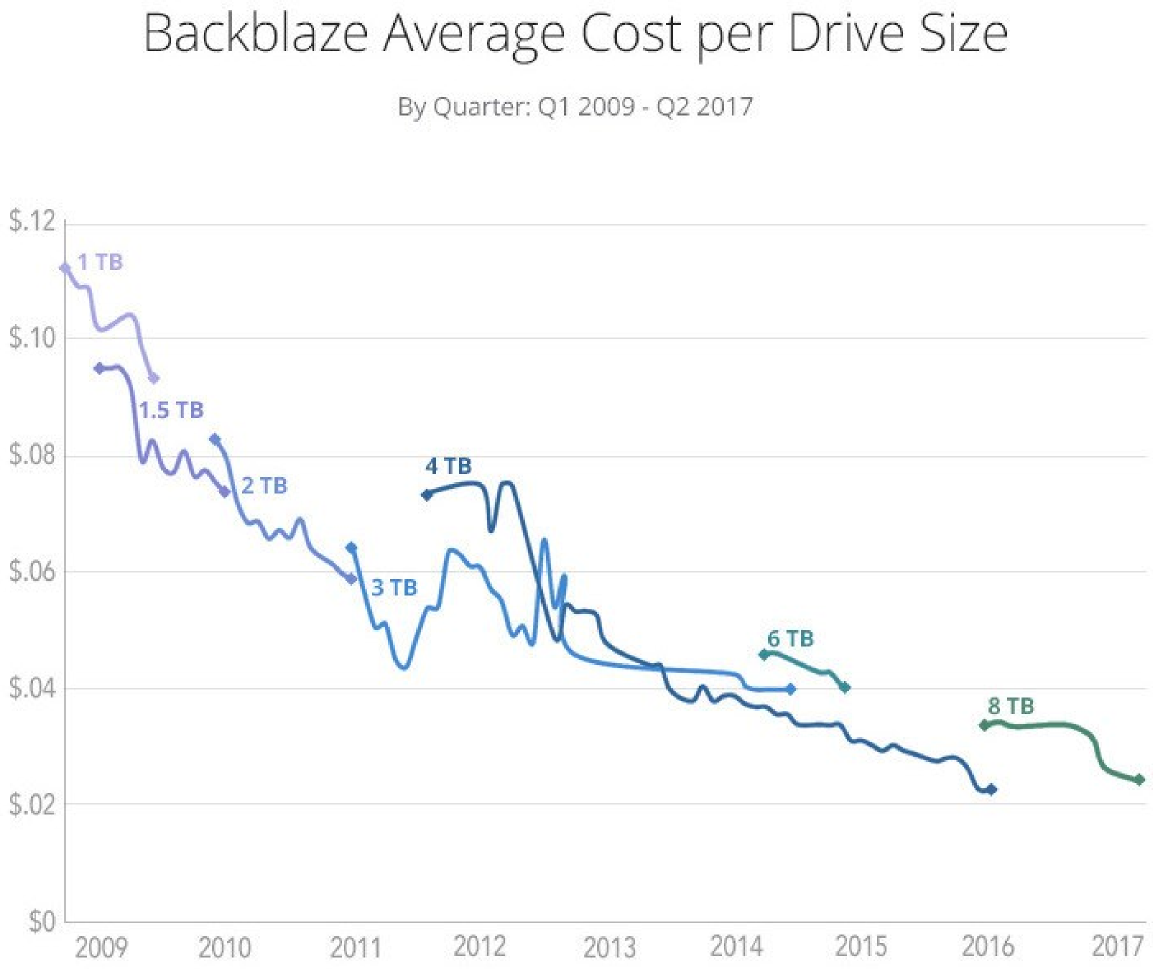
- Does not contain moving mechanical components
- Stores data to
  - flash memory
    - most common
    - non-volatile
    - lower cost than DRAM-based
  - DRAM
    - volatile
    - faster than flash-based
- Controller
  - embedded processor
  - can highly increase performance by, e.g., data striping, data compression or caching
- Interface emulates HDD interface



SSD - Types of NAND flash memories

- SLC (Single-Level Cell)
    - much higher endurance regarding number of writes
    - highest performance and price
    - lower capacity
    - enterprise-level
- MLC (Multi-Level Cell)
- TLC (Triple-Level Cell)
- QLC (Quad-Level Cell)
- PLC (Penta-Level Cell)

2019.08.25
https://www.tomshardware.com/news/toshiba-5-bit-per-cell-flash-plx-xl-nand-pcie-4.0-ssd,40237.html
Started a research …



2018.11.26 https://www.techpowerup.com/249972/ssds-are-cheaper-than-ever-hit-the-magic-10-cents-per-gigabyte-threshold

SSD vs Mechanic Drives

- Advantages of SSDs
  - silent
  - lower consumption
  - more resistant to shock and vibration
  - lower access time (no need to move heads)
  - higher transfer rates (up to 500MB/s or even higher in enterprise-level solutions)
  - does not require cooling

- Disadvantages of SSDs
  - lower capacity (up to 2TB, but only hundreds of GBs affordable)
  - higher cost
  - limited lifetime (writing to the same spot)
    - as not an issue with a typical IO load

Optical Disk

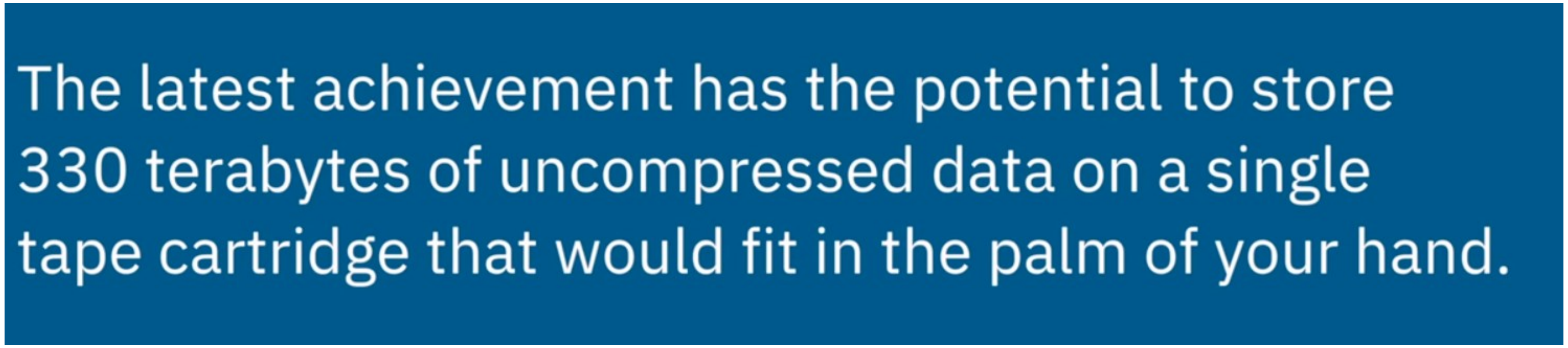
- CD, DVD, Blu-ray
- Based on reflectance
  - pit = 0 (lack of reflection)
  - bump/land = 1 (reflection)
- Data stored by laser and read by laser diode when spinning in the optical disc drive

- Optical drives operate (usually) with constant linear velocity (*konstantní lineární rychlost*)(CLV)
  - magnetic HDDs operate with constant angular velocity (*konstantní úhlová rychlost*) (CAV)
  - audio streaming requires constant bit rate
  - spindle motor has to vary speed to increase revolutions per minute (RPMs) near the center of the disk

Magnetic Tape

- Magnetizable coating on a long, narrow strip of plastic film
- **Sequential access** → well suited for transfer of big continuous blocks
- **Low cost** per bit
  - available surface area on a tape is far greater than for HDD
- Read and written by **tape drive**
- Originally main secondary storage
- Performance
  - **slow access time**
  - transfer rate comparable to magnetic disks
- Automatic change of tapes
  - **single tape drive**
    - autoloaders
  - **multiple tape drives**
    - tape library, tape robot, tape jukebox
    - can store up to hundreds of petabytes of data
- **Usage**
  - **backups of infrequently used information, archiving**
    - longer duration (15-30 yrs)

Magnetic Tape

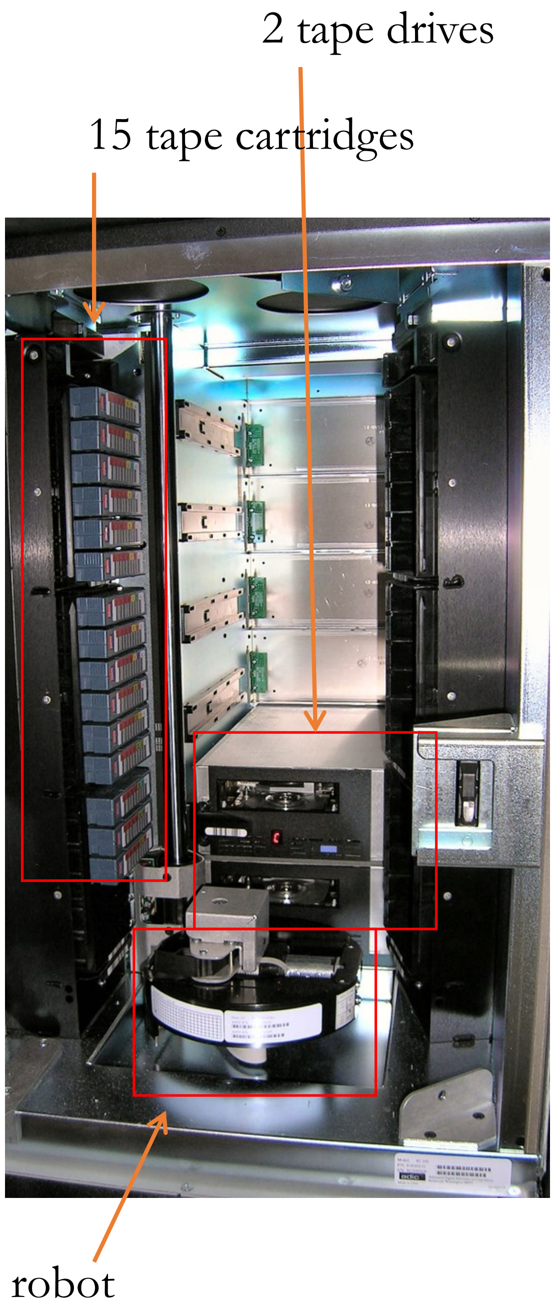


The latest achievement has the potential to store 330 terabytes of uncompressed data on a single tape cartridge that would fit in the palm of your hand.

2018.08.02 IBM Achieves the World's Highest Areal Recording Density for Magnetic Tape Storage

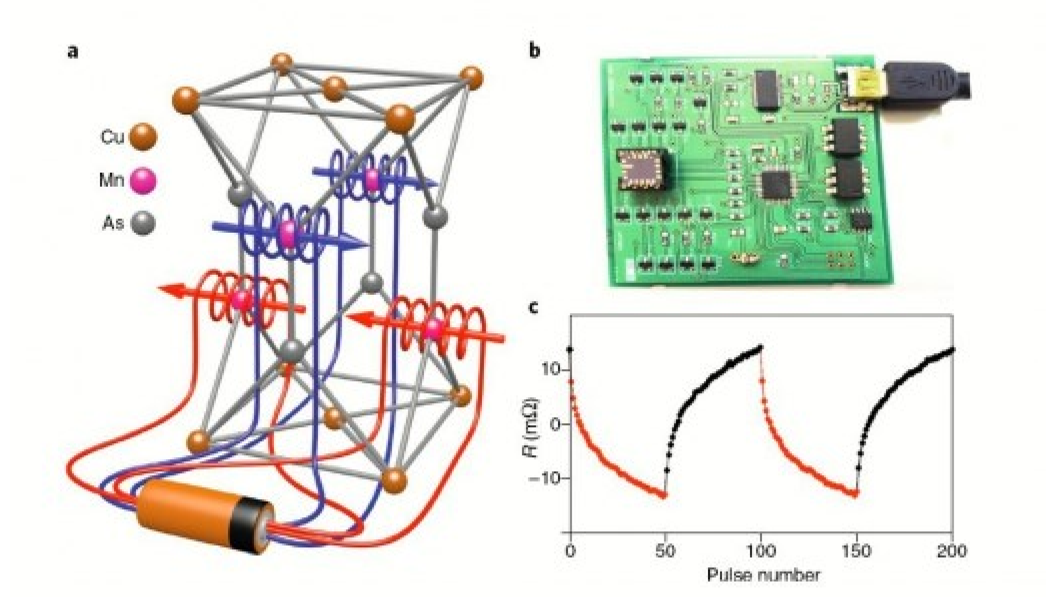
Tape Libraries

- **Tape robot, tape jukebox**
  - **tape drive(s)**
  - **tape cartridges**
  - **robot**
  - **barcode reader**
    - identification of tape cartridges
- **Capacity**
  - up to hundreds of **petabytes** of data
- **Price**
  - up to **$1 million**
- **Autoloaders**
  - small tape libraries with only 1 drive



Future Work: Antiferromagnetism

… They allow for counting and recording thousands of input pulses and responding to pulses of lengths downscaled to hundreds of picoseconds. … To demonstrate the compatibility with common microelectronic circuitry, we implemented the antiferromagnetic bit cell in a standard printed circuit board managed and powered at ambient conditions by a computer via a USB interface. …



https://www.nature.com/articles/ncomms15434/

2018.06.16 https://www.idnes.cz/technet/veda/spintronika-antigeromagneticka-nova-pamet-rozstrel-tomas-jungwirth.A180515_165928_veda_mla

https://www.nfneuron.cz/cs/vedci/tomas-jungwirth

https://www.fzu.cz/novinky/objevy-v-oblasti-antiferomagnetickych-materialu-meni-zpusob-ukladani-dat

Hierarchical storage management

- Using **various types of storages** to increase usable capacity with limited costs
  - Less often used data moved to cheaper storages with higher capacity → **tiers**
    - fast disks
    - MAID - Massive Array of Idle Disks
    - disk libraries
  - Conceptually analogous to the (multi-level) cache
  - Moving of data is managed by a **migration policy**
  - May and may not require special commands

- CESNET (about 21 PB)