POLITECNICO DI MILANO

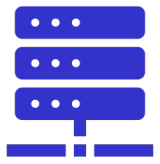
Computing Infrastructure





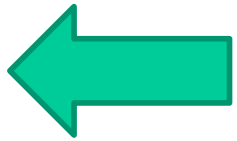
Storage

The topics of the course



A. HW Infrastructures:

- **System-level**: Computing Infrastructures and Data Center Architectures, Rack/Structure;
- Node-level: Server (computation, HW accelerators), Storage (Type, technology), Networking (architecture and technology)
- **Building-level**: Cooling systems, power supply, failure recovery



B. SW Infrastructures:

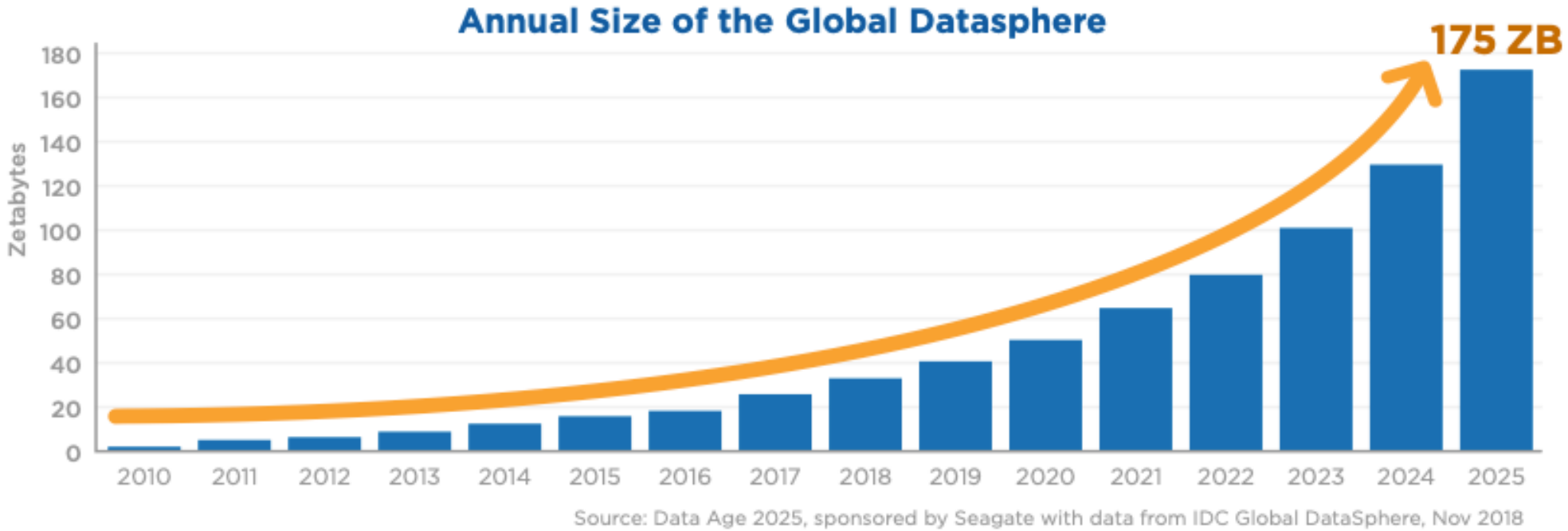
- **Virtualization**: Process/System VM, Virtualization Mechanisms (Hypervisor, Para/Full virtualization)
- **Computing Architectures**: Cloud Computing (types, characteristics), X-as-a service, Edge/Fog Computing
- **Machine and deep learning-as-a-service**

C. Methods:

- **Reliability and availability of datacenters** (definition, fundamental laws, RBDs)
- **Disk performance** (Type, Performance, RAID)
- **Scalability and performance of datacenters** (definitions, fundamental laws, queuing network theory)

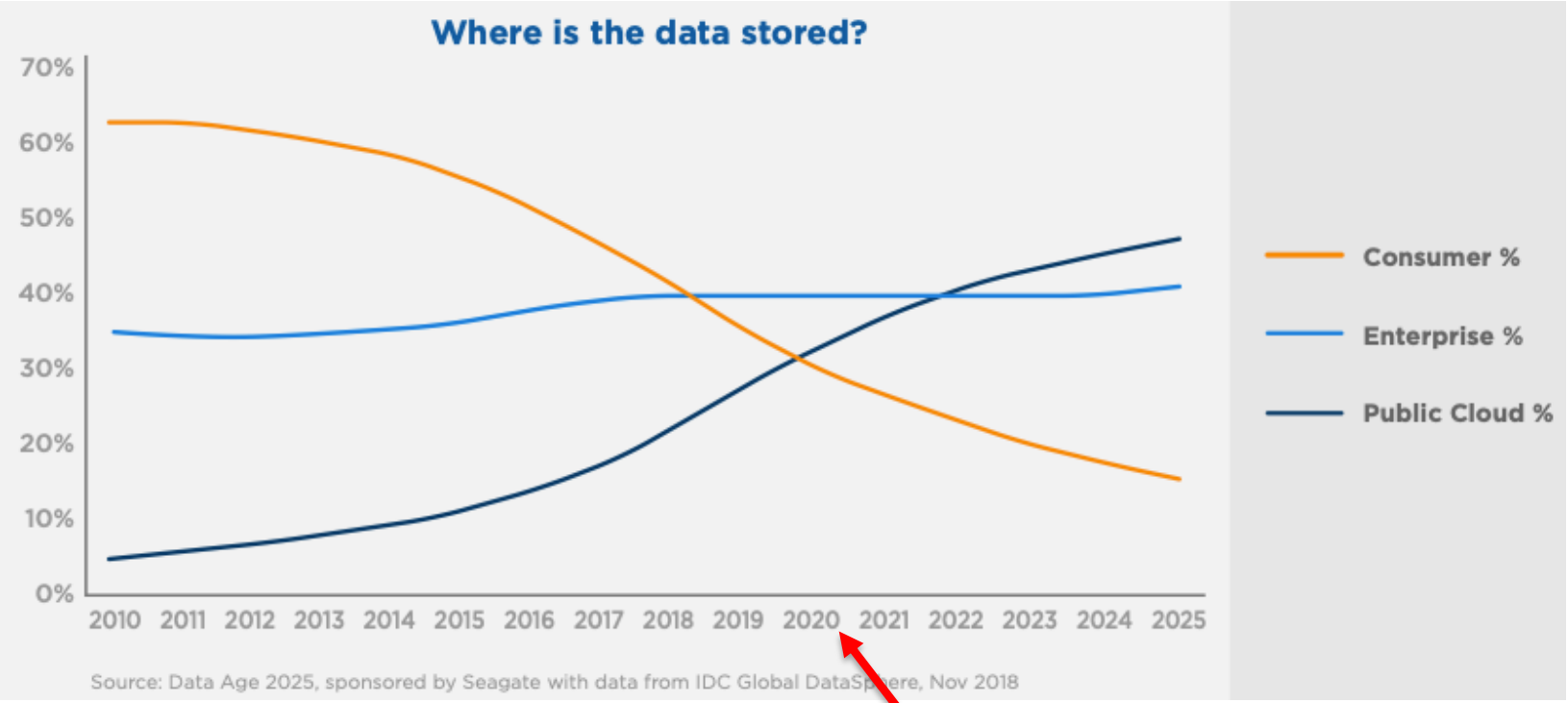
Some Trends…

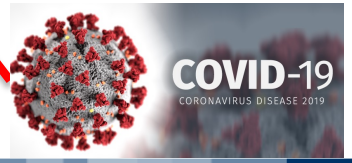
- Data-driven world
  - 80s-90s data was primarily generated by humans
  - Nowadays machines generate data at an unprecedented rate
    - Industry4.0 and AI
    - Sensors, surveillance cameras, digital medical imaging devices…
    - Multiple media (image/video/audio/socialmedia) as big-data source



Some More Trends…

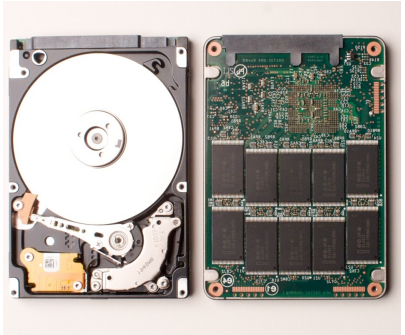
- The growt favors the centralized storage strategy
  - Limiting redundant data
  - Automatizing replication & backup
  - Reducing management costs

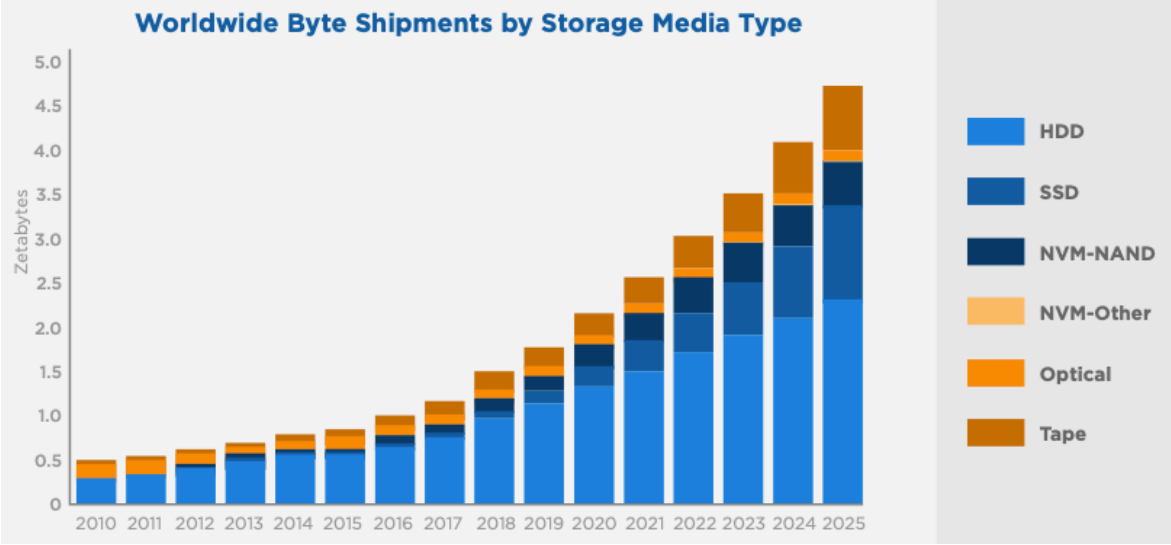




Storage Technologies

- Storage technology is dominated by HDDs
  - Magnetic disks with mechanical interactions
- «Recent» technology advancement brought SSDs
  - No mechanical or moving parts
  - Built out of transistors (NAND flash-based devices)
- NVMe - Non-Volatile Memory Express
  - Industry-standard to run PCIe SSDs
- Tapes … *will never die*





Tapes … will never die

Amazon S3 vs. Glacier

- Same durability (11 nines), different performance in data retrieval

S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data

First 50 TB / Month	$0.023 per GB
Next 450 TB / Month	$0.022 per GB
Over 500 TB / Month	$0.021 per GB

S3 Glacier Flexible Retrieval (Formerly S3 Glacier)***- For long-term backups and archives with retrieval option from 1 minute to 12 hours

All Storage / Month	$0.0036 per GB

S3 Glacier Deep Archive* - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours (tapes)

All Storage / Month	$0.00099 per GB

Hybrid solutions (HDD + SSD)

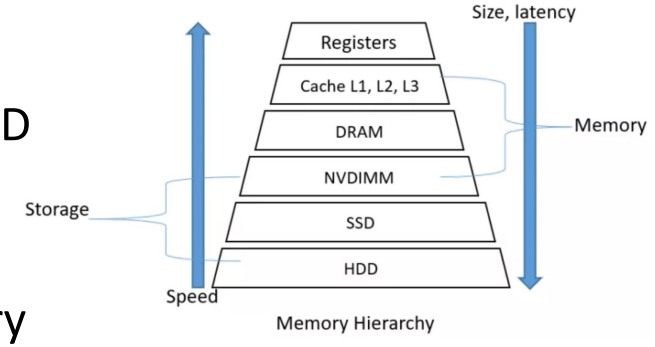
Some large storage servers use SSD as a cache for several HDD. Some mainboards have the same feature: they combine a small SSD with a large HDD to have a faster disk

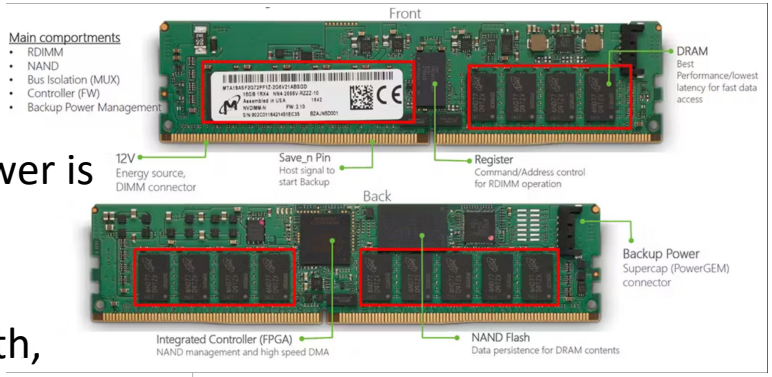
Some HDD manufacturers produce Solid State *Hybrid Disks* (SSHD) that combine a small SDD with a large HDD in a single unit

Hybrid solution: NVDIMM

NVDIMM (Non-Volatile Dual In-line Memory Module)

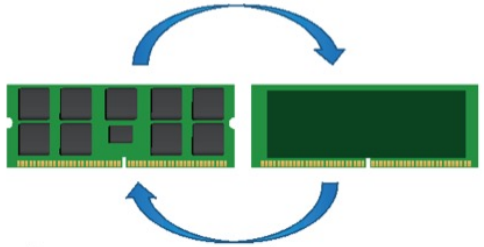
- Integrates DRAM with non-volatile memory, typically NAND flash, on a single module
- High performance while ensuring data persistence during power outages (onboard backup power source, e.g. battery or capacitor)
- Key features:
  - Data Persistence: NVDIMMs retain data even when power is lost, making them ideal for applications requiring high reliability and fast recovery times
  - Performance: They offer low latency and high bandwidth, similar to traditional DRAM, but with the added benefit of non-volatility
  - Byte-Addressable: NVDIMMs support direct CPU access, allowing for efficient data manipulation without needing traditional storage interfaces
  - Cost: more expensive than traditional DRAM due to the inclusion of non-volatile memory and backup power, but significant performance and reliability advantage





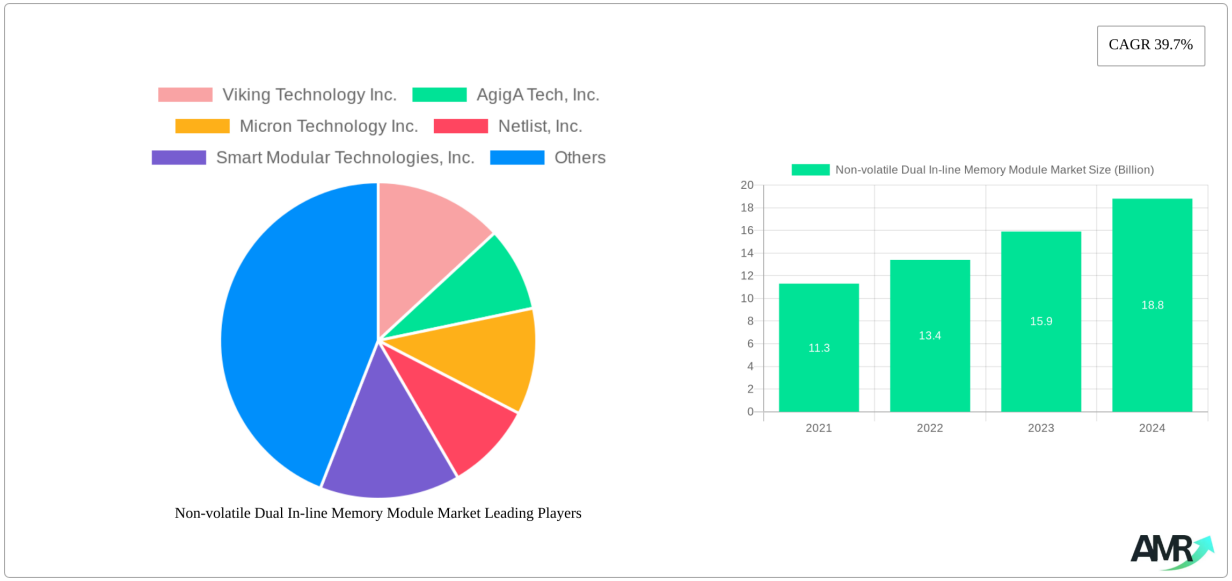
How It Works

If there is a power failure, the supercap module powers NVDIMM while it copies all data from the DDR-3 to on-module flash



When power is restored NVDIMM copies all data from flash to DDR-3 and normal operation resumes

Hybrid solution: NVDIMM



https://www.archivemarketresearch.com/reports/non-volatile-dual-in-line-memory-module-market-6237

NVDIMM Type	Access Method	Latency	Capacity	Technology	Use Case	Application Domain
NVDIMM-N	Byte-addressable	10^{-8} s	Limited (DRAM-like)	DRAM + NAND Flash	Fast data access, write logging	**Databases, Big Data Analytics, In-Memory Databases**
NVDIMM-F	Block-addressable	10^{-5} s	High (TB)	NAND Flash (requires separate DRAM)	SSD-like performance, high capacity	**Cloud Storage, Data Centers, Tiered Storage**
NVDIMM-P	Both byte and block	10^{-7} s	High (TB)	Persistent Memory Technology (e.g., DRAM + Flash)	High performance and large capacity	**Cloud Computing, High-Performance Computing, Financial Services**

POLITECNICO DI MILANO

Computing Infrastructure





A server for a datacenter: an example

An example of Server for a Datacenter

Descrizione
RB120: Server 1U Dual Socket Intel GPU – 2 bays SAS/SATA + 2 internal

1 x 1U – 2 x SAS/SATA – Ridondante 2000W
1U Rackmount Black Chassis. 2000W Redundant Power Supplies. 43mm (H) x 437mm (W) x 894mm (D). N. 2 Hot-swap 2.5" SAS/SATA drive bays, n. 2 Internal 2.5" drive bays.

1 x Dual Xeon Scalable – C621 – Server GPU
Proprietary Motherboard. Intel® C621 chipset. Dual Socket P (FCLGA3647). Support up to 165W TDP. N. 12 DIMM Slots supported Memory Types: 2666/2400/2133MHz RDIMM, LRDIMM and 3DS ECC LRDIMM modules. Optimal memory configuration: Six memory channels per CPU.

E4 COMPUTER ENGINEERING

2 x Xeon 8-Core 4110 2,1Ghz 11MB
Intel® Xeon® Silver 4110 Processor. 8Cores. 16Threads. FCLGA3647 Socket.11MB L3. 2,1Ghz Base Frequency. 85W max. TDP. DDR4-2400 Memory type.

6 x DDR4-2666 Reg. ECC 16 GB module
Full brand memory, tested and certified by manufacturer for thorough compatibility with proposed system. The real operating speed depends on the processor's model and on the number of the installed modules. Better performances are achieved through a proper channel configuration.

1 x Intel C621 SATA III 4 ports #

1 x SEAGATE 2TB 2,5" SATA III 7.200RPM
Seagate Enterprise Capacity hard disk drive. Form factor: 2,5''. Capacity: 2TB. Interface: 512N SATA 6Gb/s. Buffer: 128MB. Rotational Speed: 7200RPM. Max. Sustained Transfer Rate (MB/s): 136MB/. 2Million-hour MTBF.

1 x Intel S4500 240GB 2,5" SSD SATA III
Intel® SSD DC S4500 Series. Sequential Read (up to): 500MB/s. Sequential Write (up to): 190MB/s. Random Read (100% Span): 69000 IOPS. Random Write (100% Span): 16000 IOPS. Endurance Rating (Lifetime Writes): 0.62 PBW.

1 x Backplane 2 bays SAS/SATA