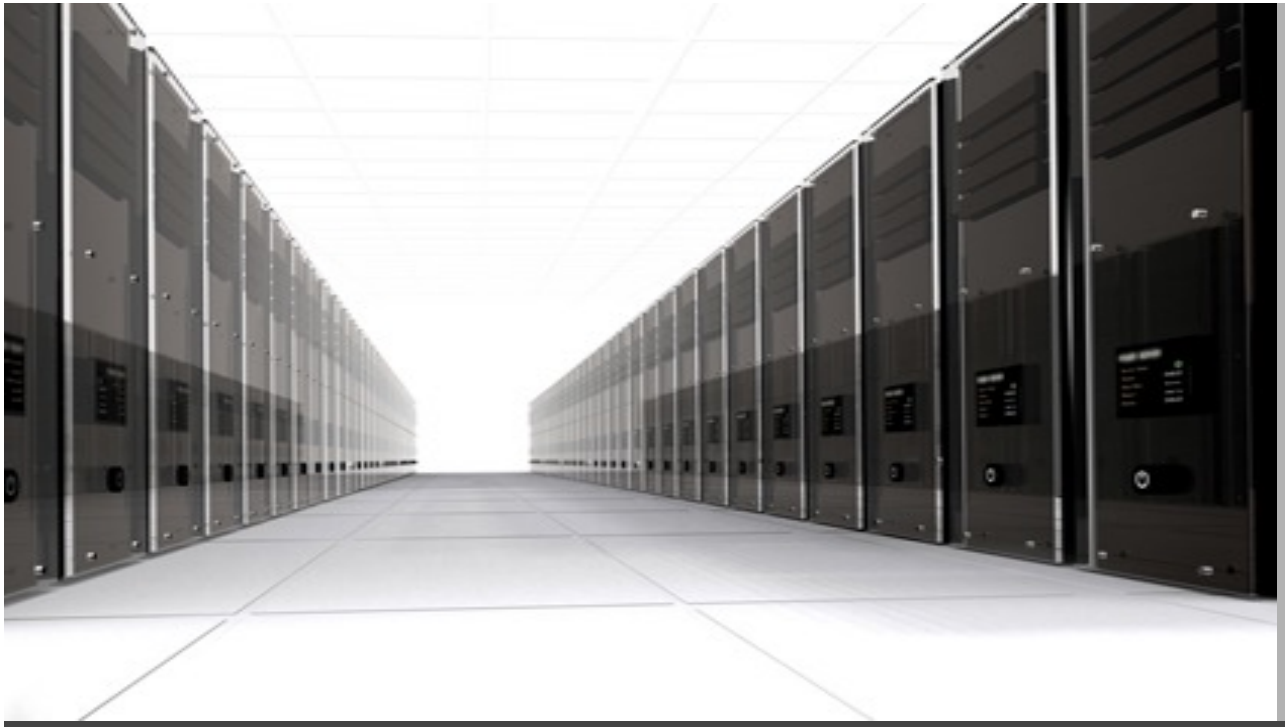


STORAGE (III): NETWORK CONCEPTS FOR DATA CENTERS

David López
V1.0.1
Updated spring 2025

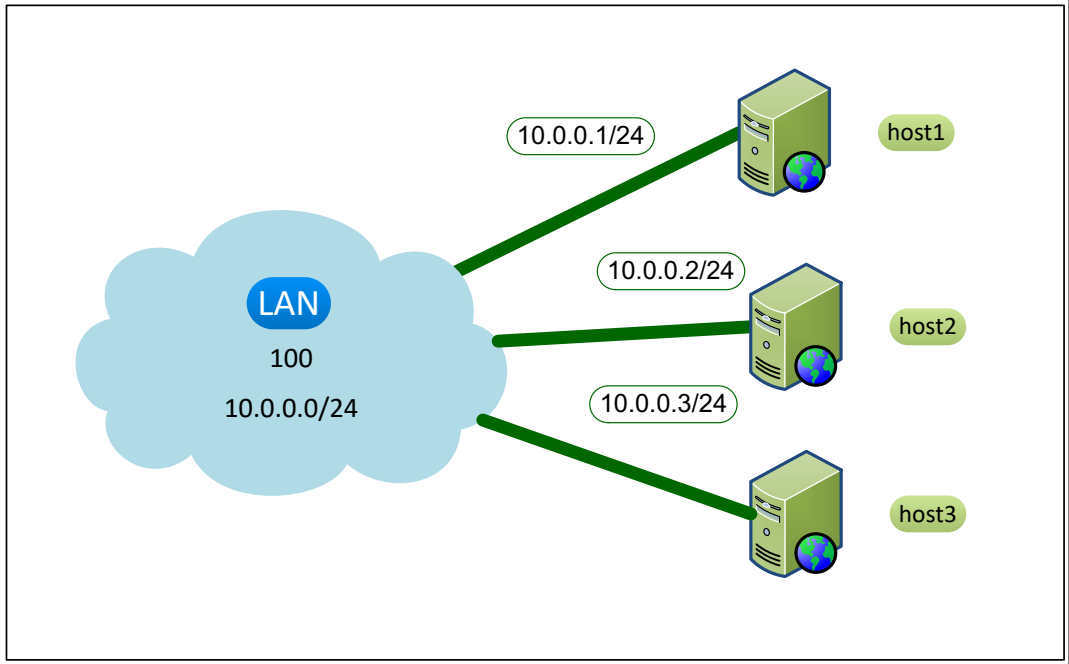
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Some basic ideas (that you probably know)

LAN (Local Area Network)

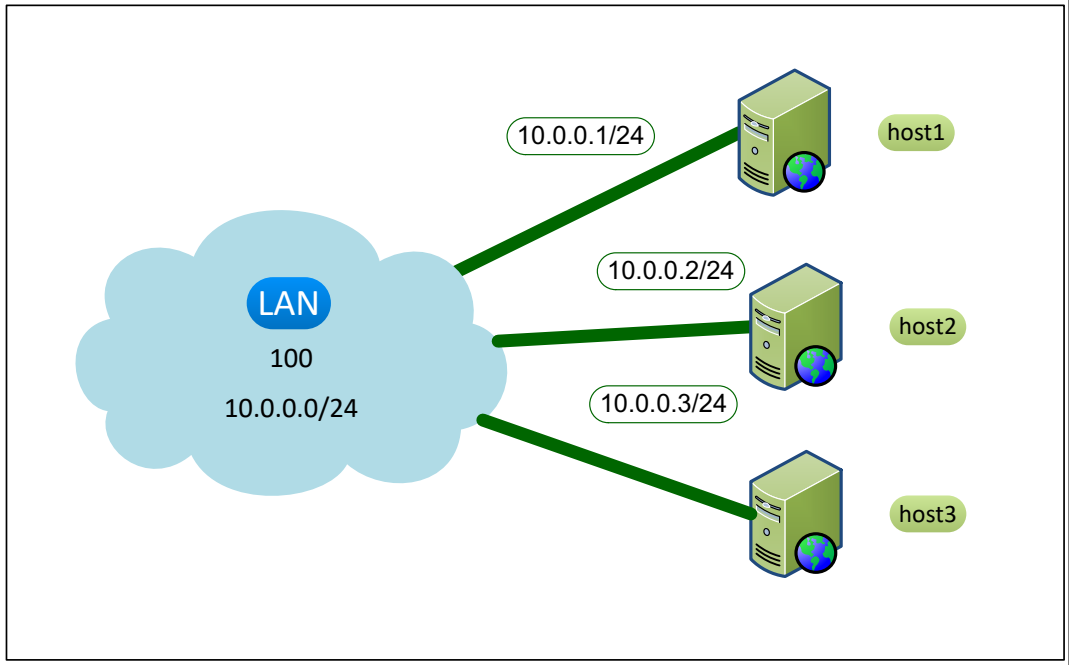
Technology defined by the IEEE 802.x family of standards (Ethernet, Wi-Fi, …)

- All hosts inside the LAN communicate directly (via switch) without using IP router
- Each LAN has an IP range associated, defined by the base IP and a mask. e.g. 10.53.0.0/16
- Switches connect hosts in the same IP range at high velocity



LAN LIMITATIONS

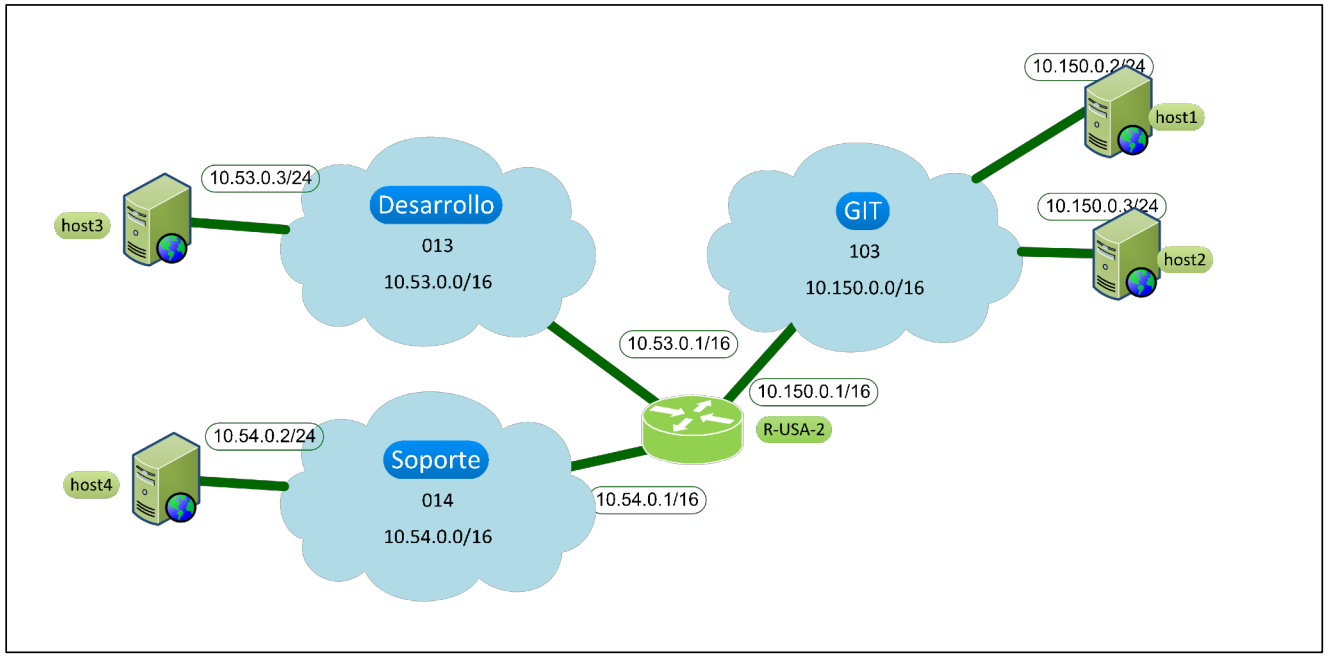
- Can connect thousands of hosts ☺ but not tens of thousands ☹
- Unsecure: velocity is key (and it is really fast) but if one host is compromised, all are compromised



Connecting LANs/VLANs

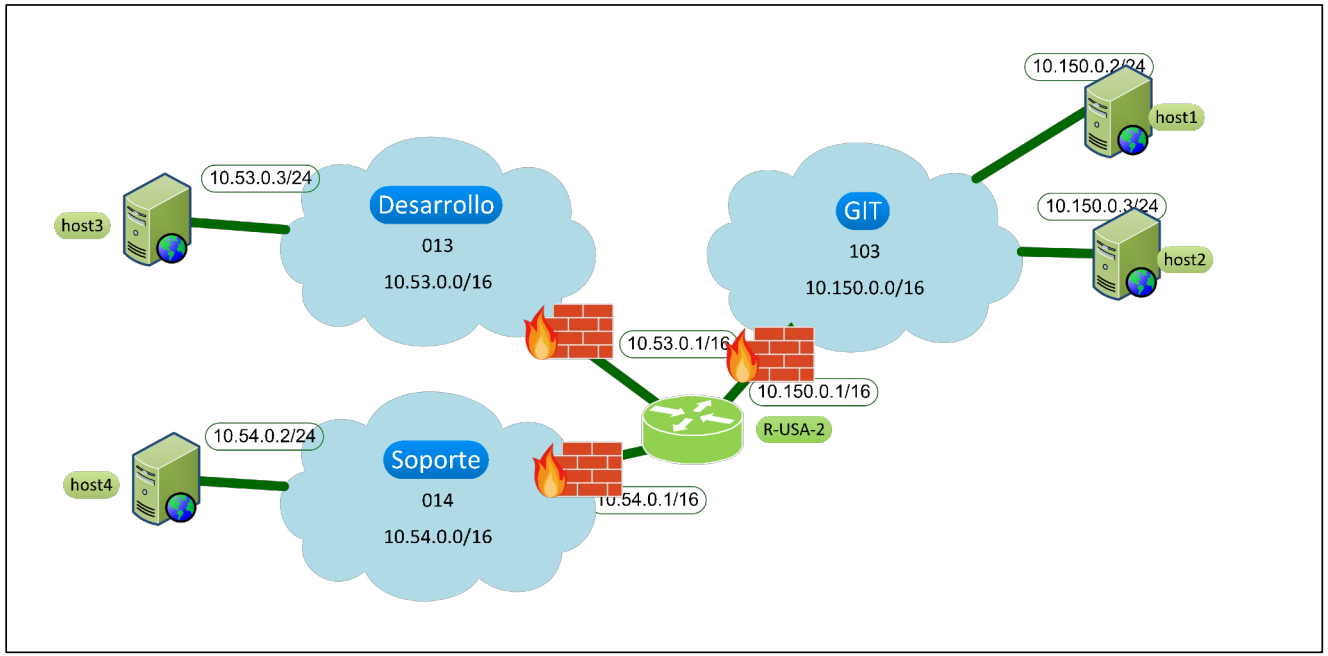
Two LAN can be connected using an **IP router**

- You can have VLAN (virtual LAN): logically independent LAN in the same physical LAN (configured using software) or different physical LANs for security and administration
- Each network interface of the router owns one IP belonging to the IP range of the LAN/ VLAN to which it connects



Security in **IP router**

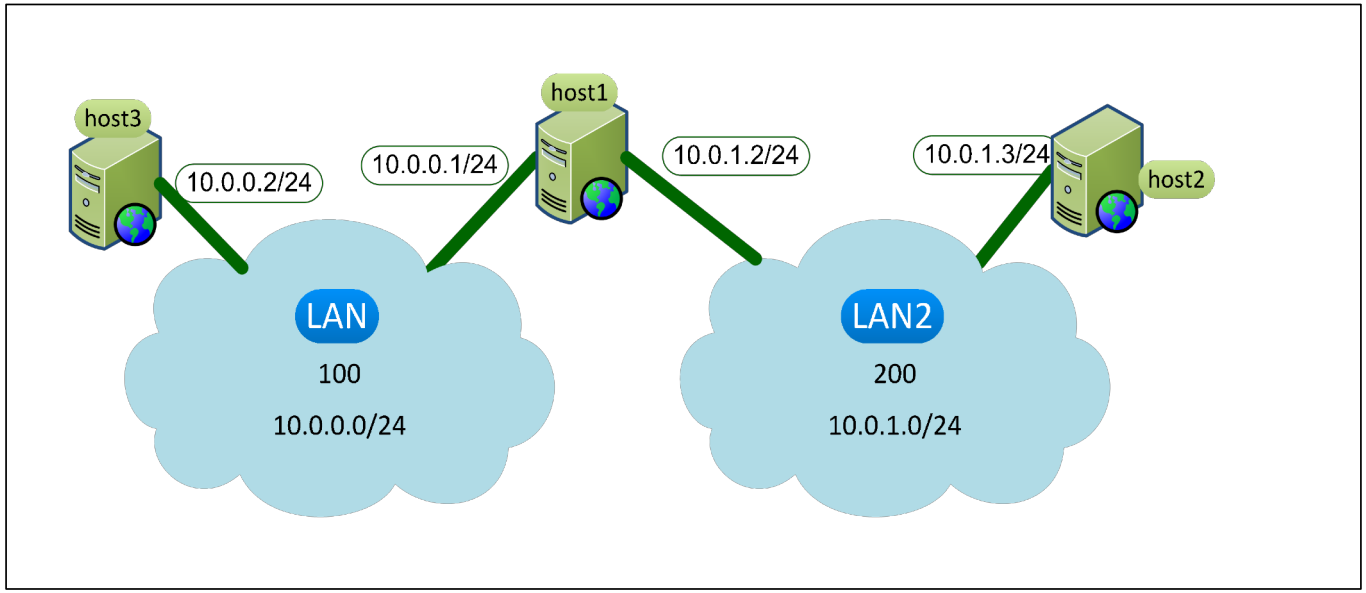
- An IP router **forwards** IP packets coming from one host to another in a different VLAN
- But you can install **Firewalls**



Gateways

One host can be connected to more than one LAN/VLAN

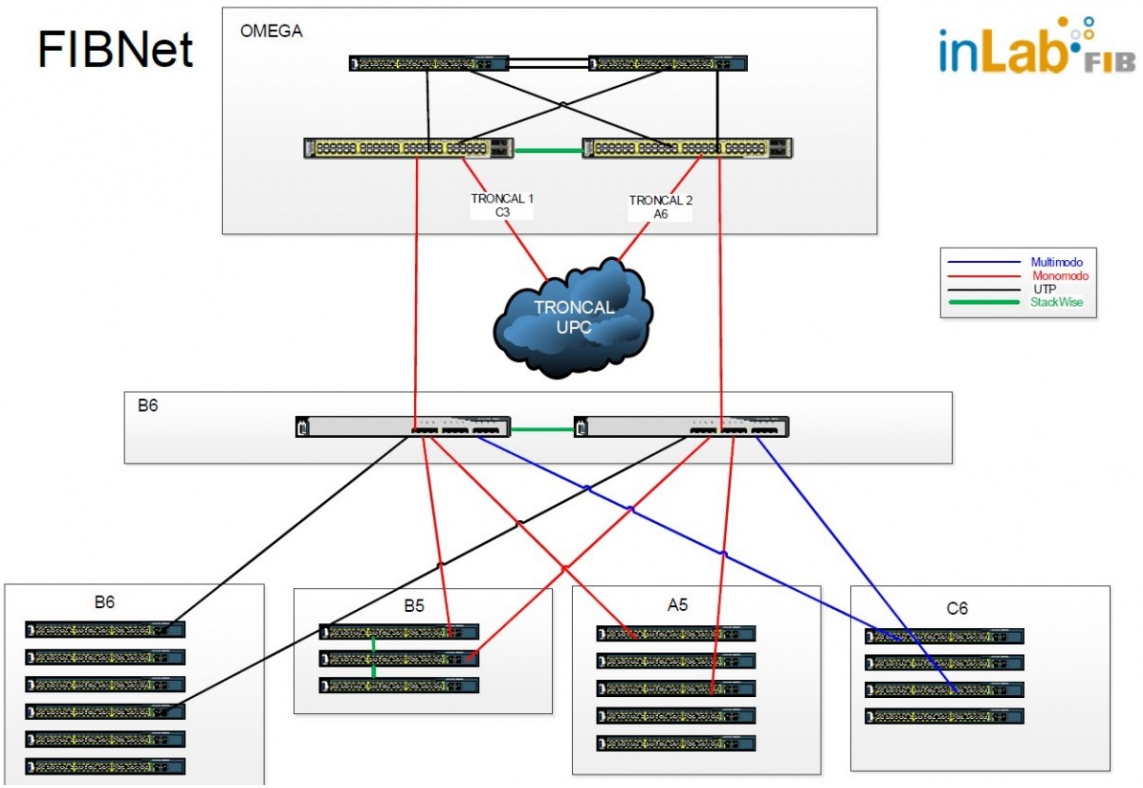
- Each Network Interface Controller (NIC) has an IP belonging to the IP range of each LAN/VLAN
- A host **never forwards** IP packets from one VLAN to the other
- This is a **Gateway**, quite common in Data Centers
- It connects layers (User Interface, Business Logic, Data layer)



And the external connection?

Wide Area Network (WAN)

- You can connect sites in your property (e.g. D6 & BSC, or FIBNet)



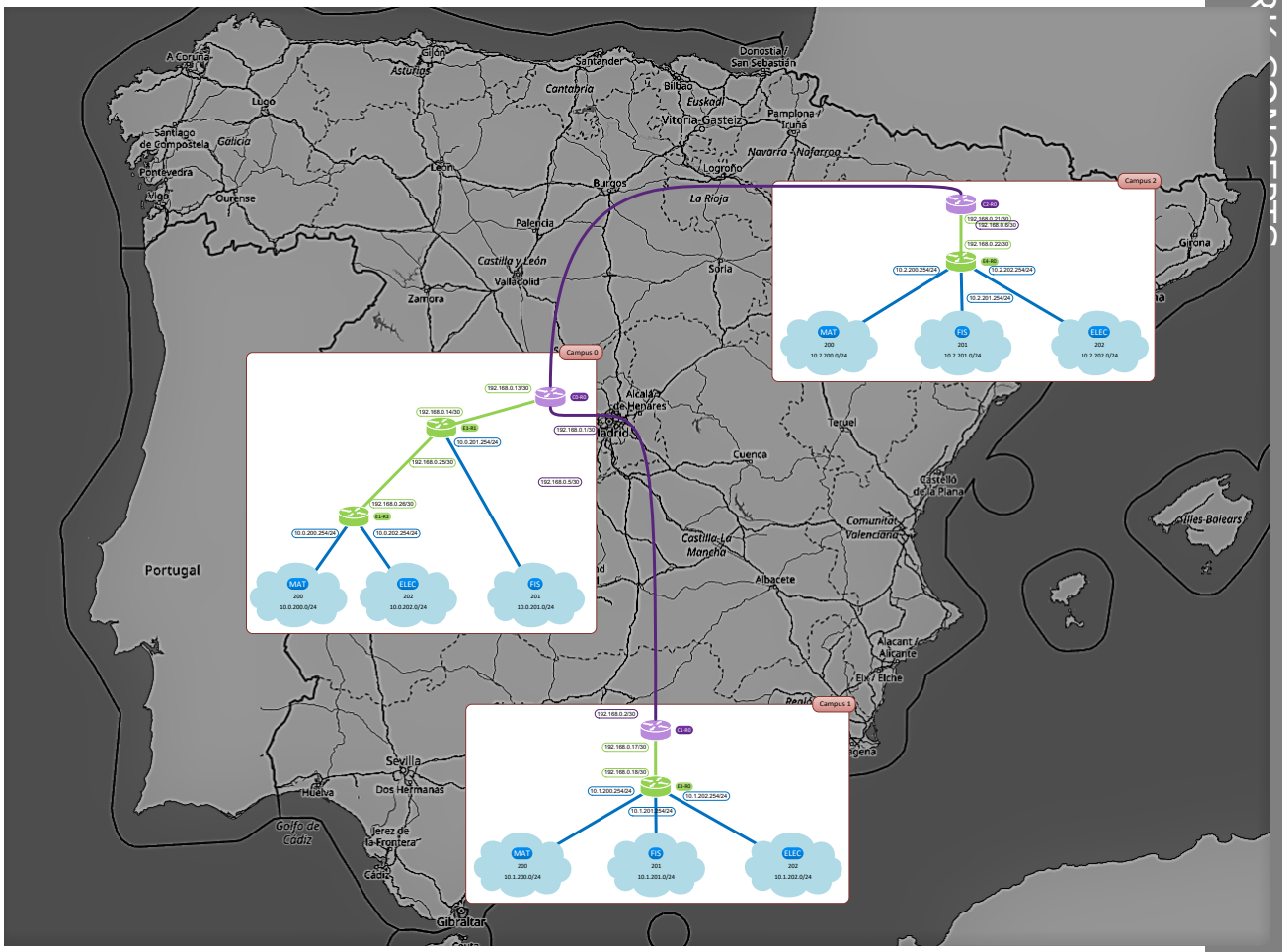
And the external connection?

Wide Area Network (WAN)

- You can connect sites in your property (e.g. D6 & BSC, or FIBNet)
- But, what when you have distant sites?
  - You have to pay crossing rights of cables (expensive), Telecom operators do that
  - They offer two services: private point-to-point links and Internet access

Wide Area Network (WAN)

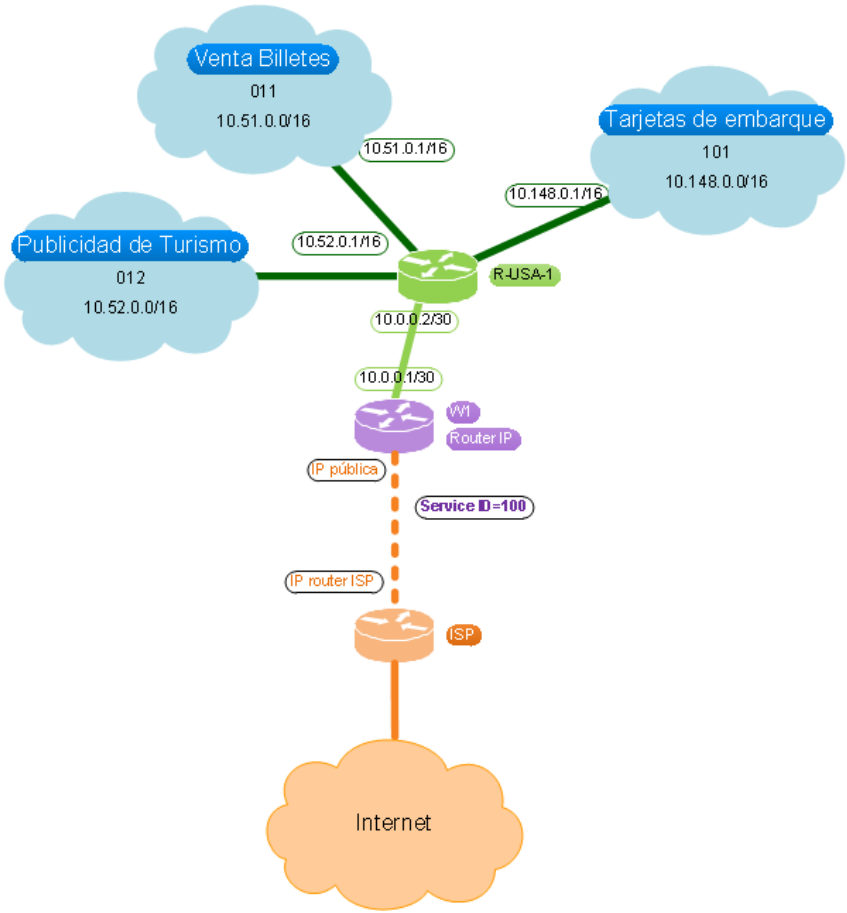
- Private connection between sites
- WAN links and routers are property of the Telecom operator (rented)
- Works as a virtual cable between site routers (green)
- There are WAN routers at each end (violet)
- Transparent to the user (it seems a physical point-to-point cable)
- **Not connected to the Internet!**



And the Internet connection?

- Is Internet a WAN? The Internet is, strictly speaking, a WAN. However, when we talk about WANs, we usually mean private or semi-private networks
- It is a service provided by an ISP (Internet Service Provider) with a WAN router
- Two services: private WAN link between the WAN router and the ISP router (dotted orange line) and a forwarding service to Internet
- An organization can:

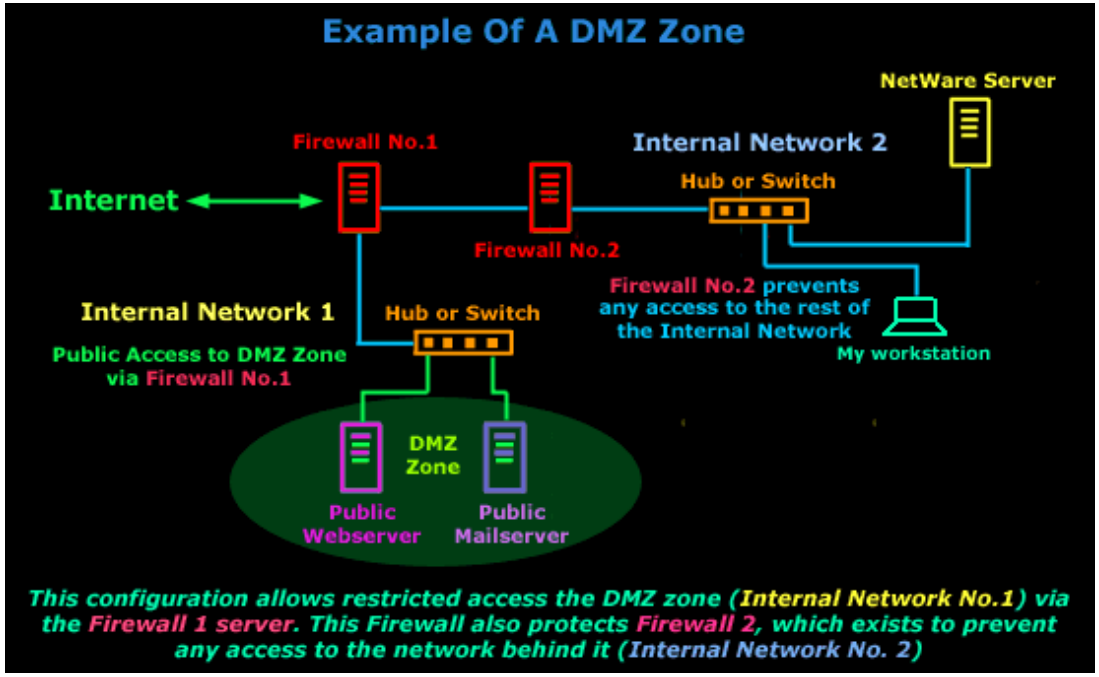
- Have a single internet access or several to guarantee performance
- Have several internet accesses to different ISP to guarantee the service
- Have WAN with other sites and also Internet access
- Use VPN (Virtual Private Network)



Intranet & Demilitarized Zone (DMZ)

Intranet
- It is a security concept
- Part of the company with **no connection to the Internet**
- If sites are connected to Private WAN links is a network spanning across sites

Demilitarized Zone (DMZ)
- The part connected to the Internet
- It is considered unsecure by design
- LANs and servers offering services to the Internet are part of the DMZ



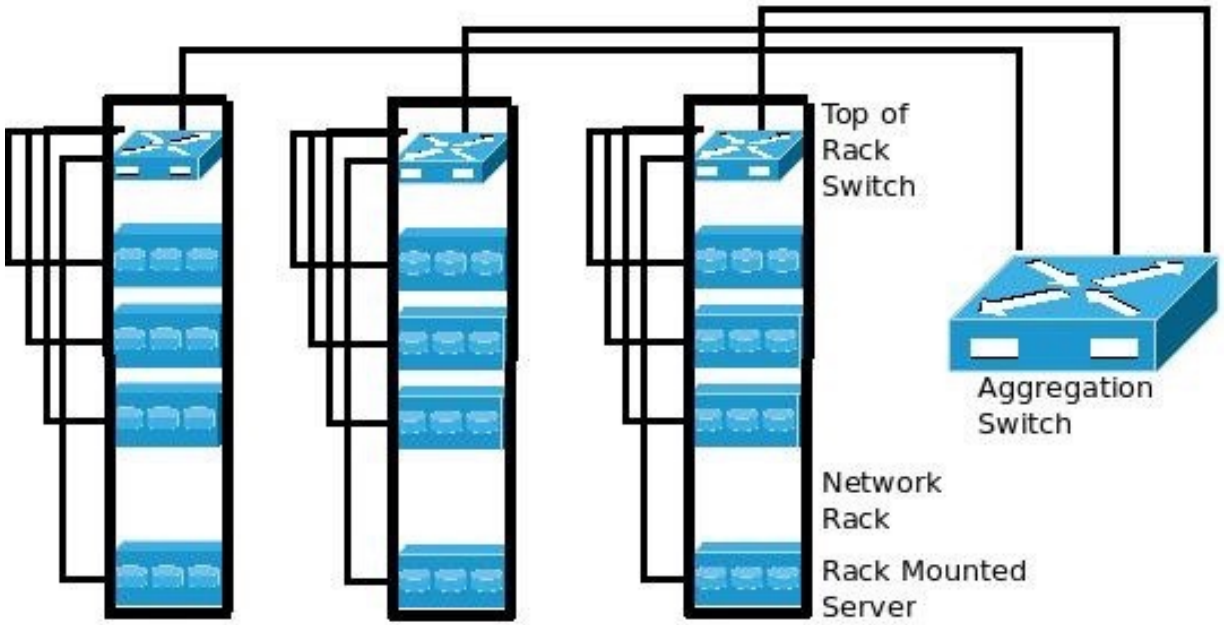
OK, and in Data Centers?

First of all: we work with hosts (physical servers), placed in racks
- Usually some racks of application servers for each rack of communication

Top of Rack (ToR)

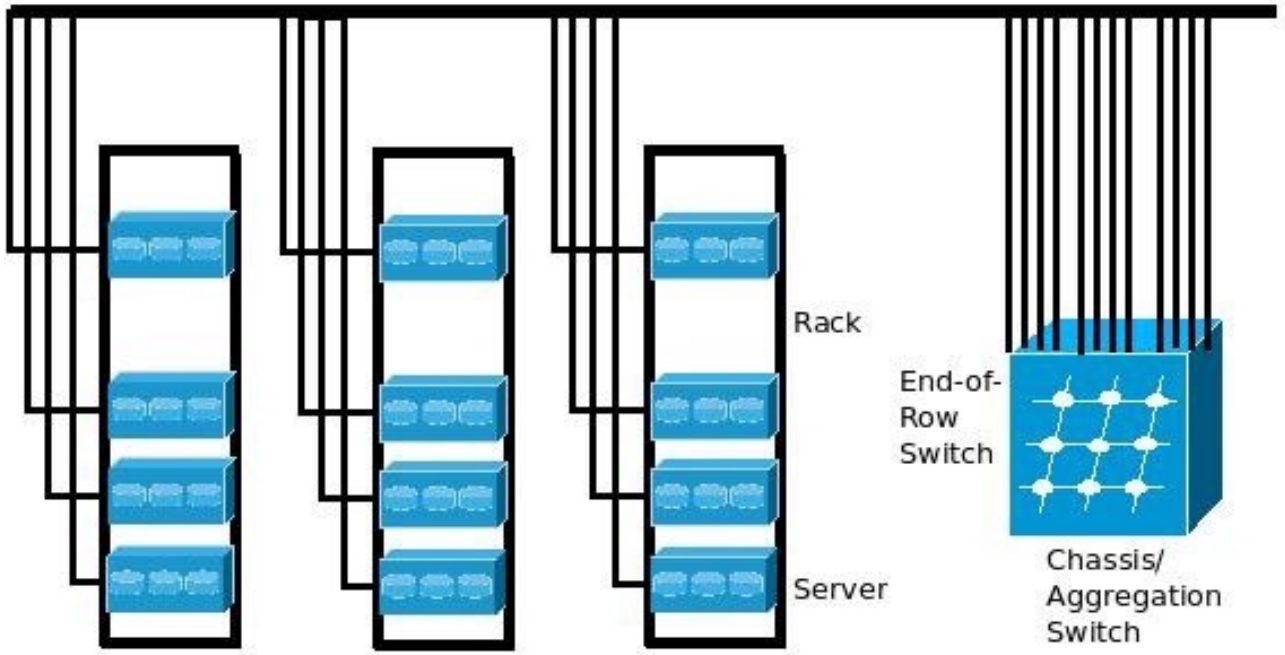
Switch placed in each rack

- Allows fast communication between hosts in the rack
- Can have aggregation switches in the communications rack
- Also routers if we want separate VLANs
- **Redundancy is key**



End of Row (EoR)

- Alternative to ToR
- Similar idea, but just the number of switches necessary to interconnect all hosts in the row, not at least one switch per rack

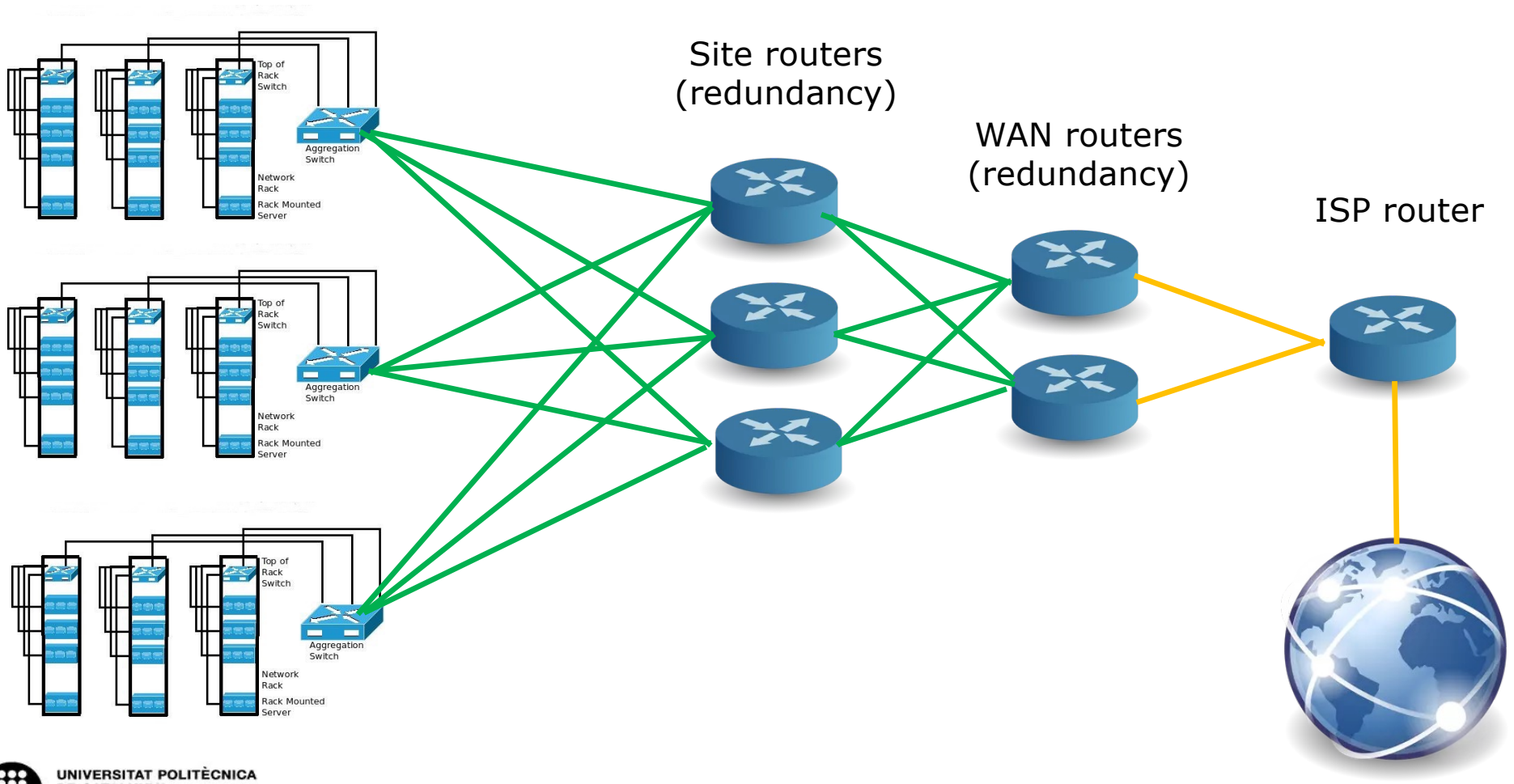


ToR vs EoR

	ToR design	EoR Design
Network deployment	- Minimum 1 switch per rack	- Switches centralized in communication rack.
Required services	- More switches - Less cables	- Less switches - More cables
Power & Cooling	- Underutilization of switches - Higher power consumption - Greater need for cooling	- Effective use of switches - Less power consumption - Lesser need of cooling
Network expansion	- Greater layer 2 data traffic - Network expansion easy	- Lesser layer 2 data traffic - Network expansion is difficult

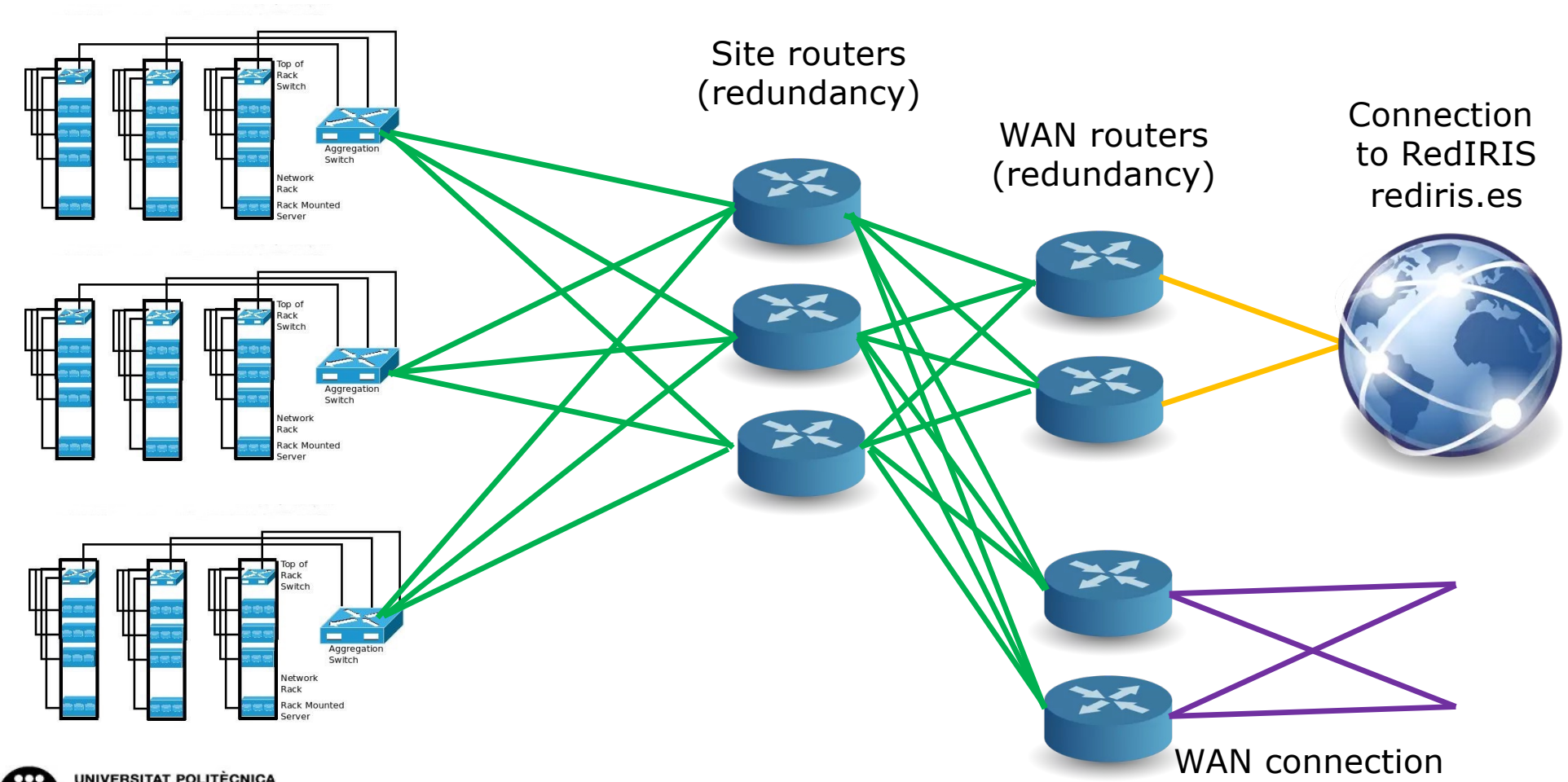
So, the final picture

ToR racks with switches



UPC Omega Data Center case (connected to RedIRIS)

ToR racks with switches



And the storage connection?

First of all, data centralized or distributed (INSIDE a Data Center)

- In some weeks we'll study distribution across sites

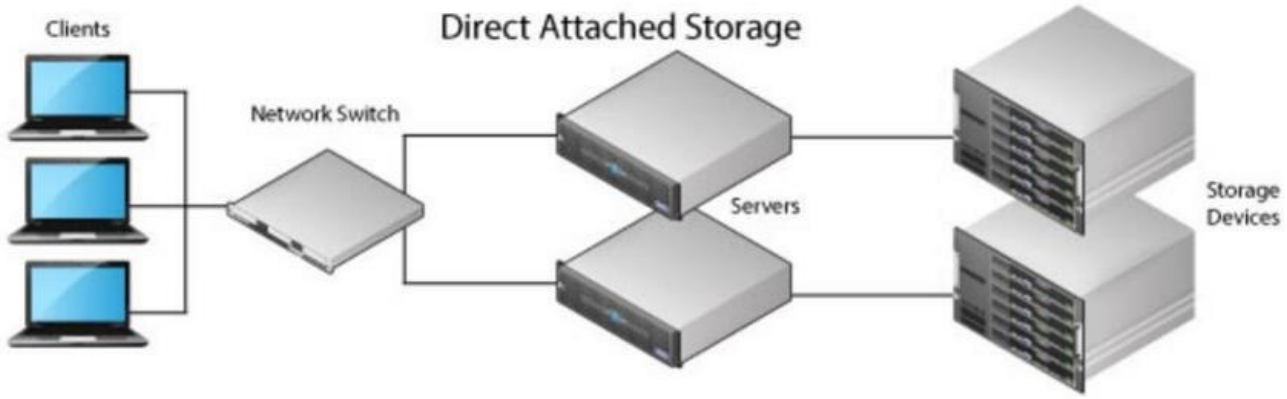
FIRST CONCEPT: Direct-Attached Storage (DAS)

- Attached to a computer directly through a host bus adapter (HBA), with no network
- Typical: a data storage service such as an enclosure with HDD
- Protocols: ATA, SATA, NVMe, SCSI, SAS, …

DAS: Direct-Attached Storage

Can be used in a Data Center:
- Very fast (High Performance)
- Easy to setup and configure
- Low cost

But:
- Limited scalability (<10 servers?)
- Poor performance for shared data
- No central management and backup



Do not confuse DAS with DFS (Distributed File System)

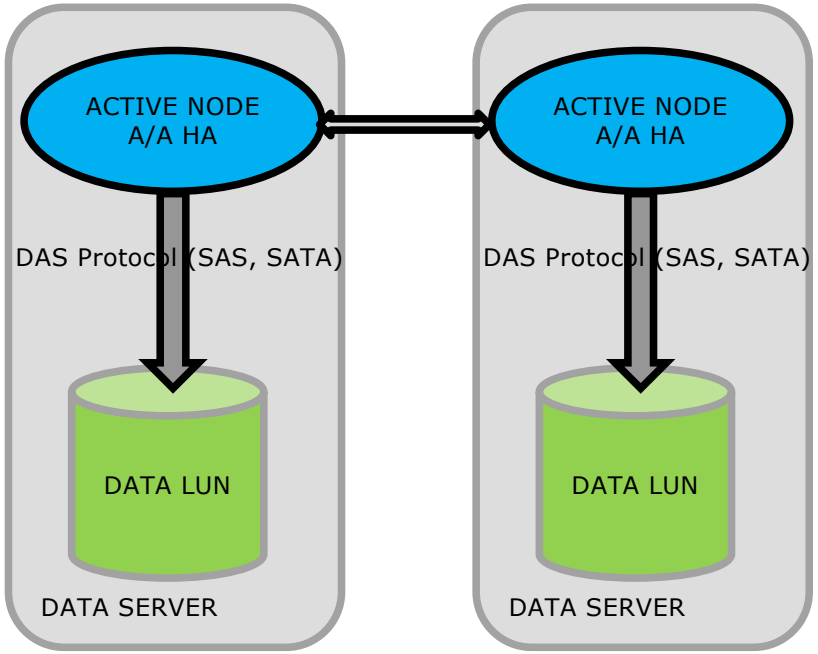
Distributed File System:
- Files are located in different hosts BUT accessible as if they were stored locally

Ideal for
- Transparent local access
- Location independence
- Scale-out
- Fault tolerance

Some models use it, like Big Data and High Performance Parallelism

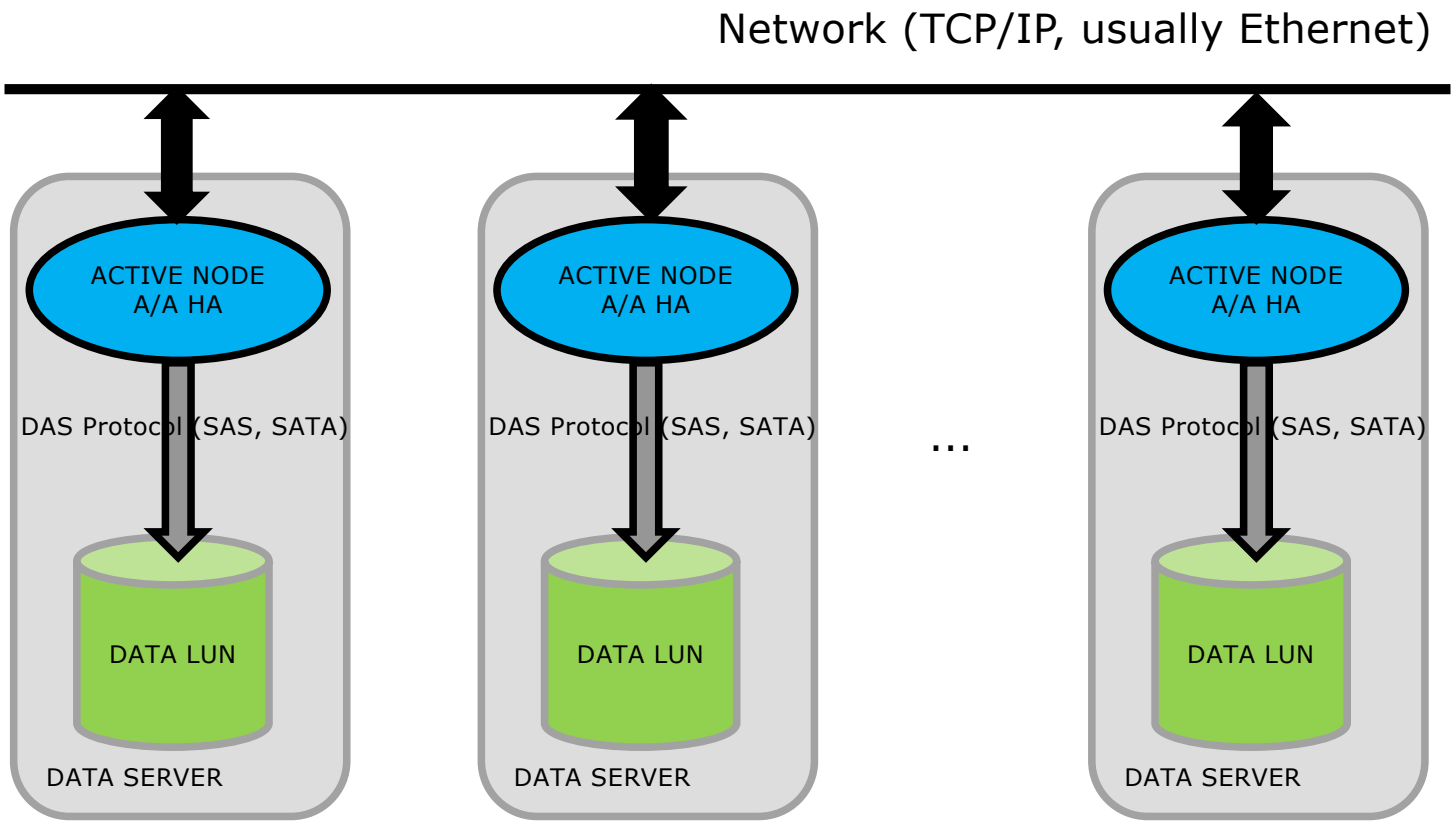
- Each server has an attached local storage system
- Access by asking the server through a network
- High level of replication

Distributed data using DAS



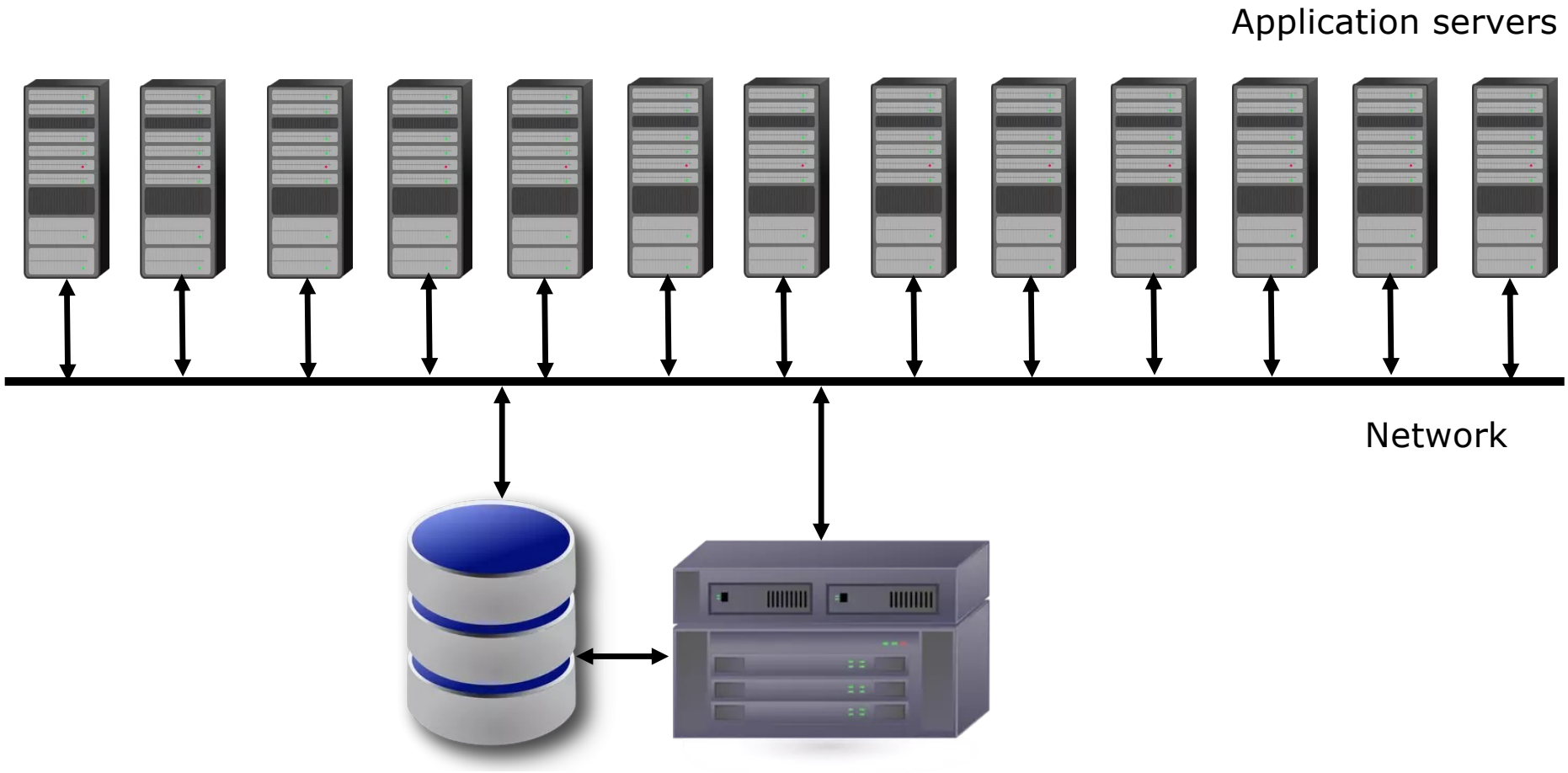
- Both nodes actives (A/A) (not Active/Passive A/P)
- When one node fails, its data are unreachable for the other nodes
- HA (High Availability) requires each piece of data replicated in other nodes (but local disks are cheap)
- Changes are not immediately visible to all nodes.
- Cluster software needed to arbitrate Read & Modify access to replicate data in order to maintain consistency (through Ethernet)

Typical DAS in a Data Center

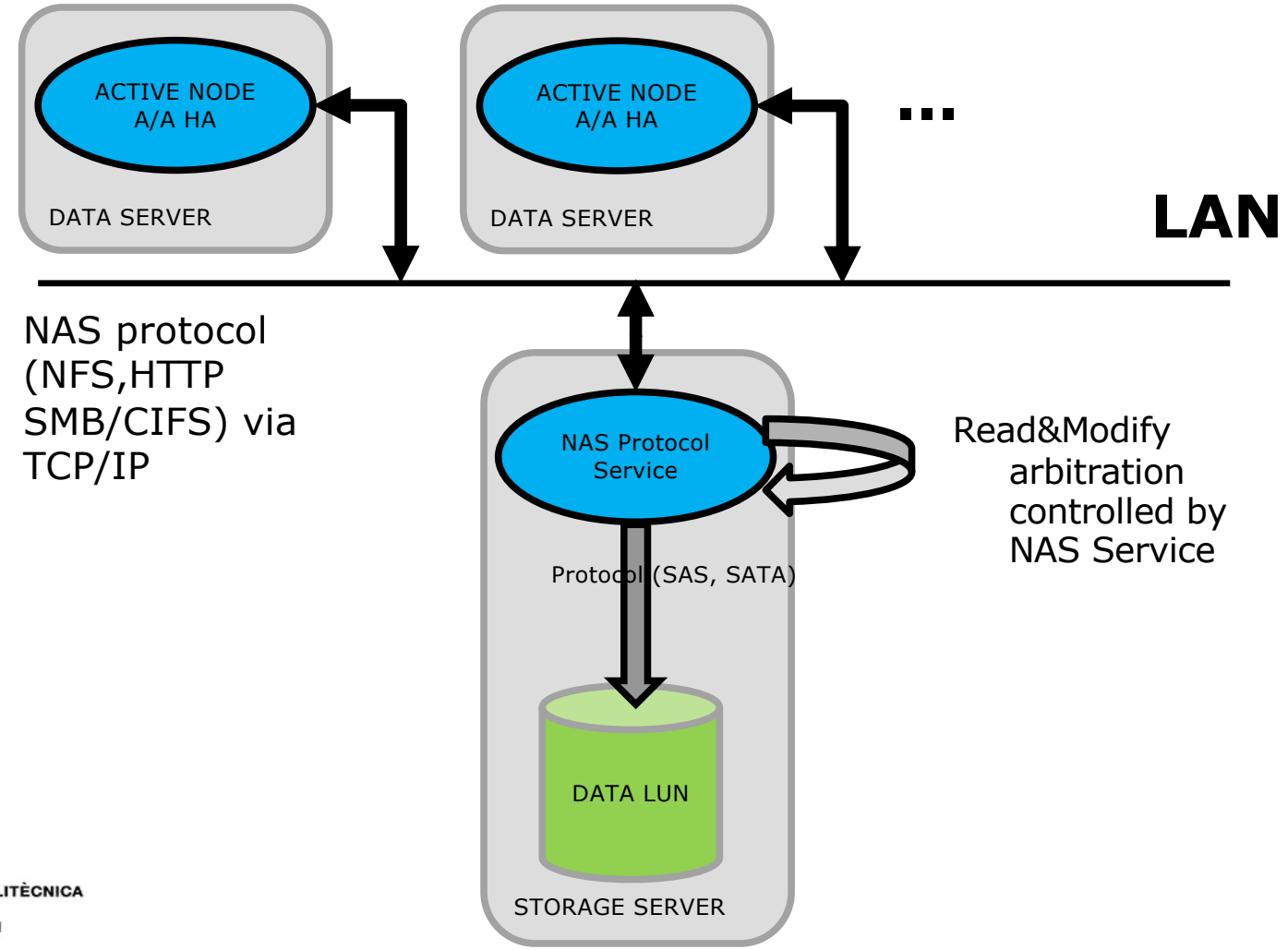


Centralized Storage

INSIDE a site!



NAS (Network Attached Storage)



NAS (Network Attached Storage) details

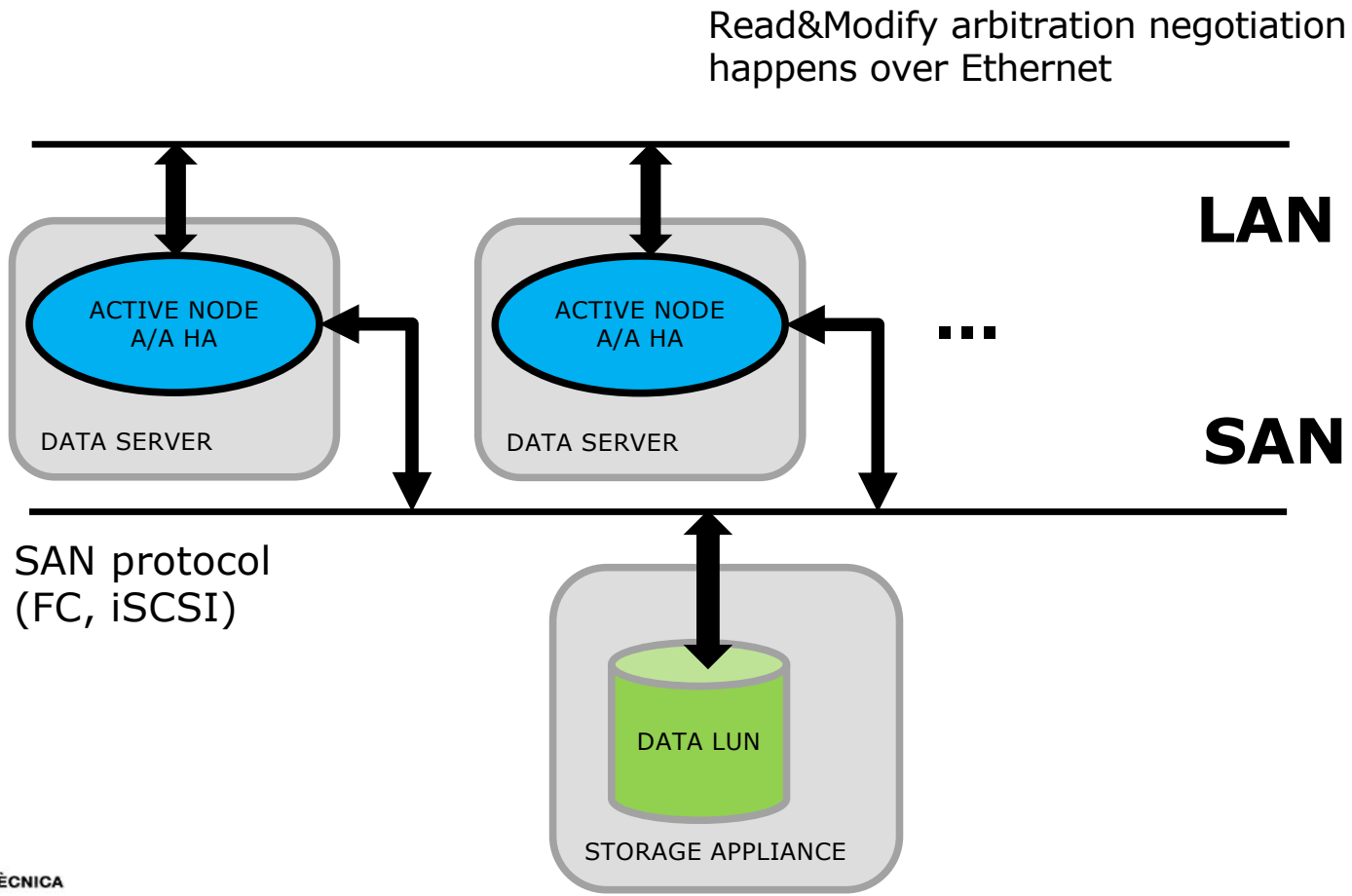
Attached to a TCP/IP network (usually Ethernet), Typically 100 Mbps – 10 Gbps, 2-4µsec latency

- Protocol operates on files – Network File System (NFS), Common Internet File System (CIFS), Server Message Block (SMB), … NAS appears to the OS as a shared folder
- NAS server has its own IP address

Odds:

- NAS is LAN-dependent; if the LAN goes down so does the NAS
- Does not scale very well (in the EBH project we will ignore this)
- **Network congestion and performance degradation**
- Another weakness related with its very nature: Ethernet transfer data via packets, that can be sent out-of-order (or even lost), so the file is not available until all packets has arrived. No problem with small files, problem with large files (video production or consumption)

SAN (Storage Area Network)



SAN (Storage Area Network) details

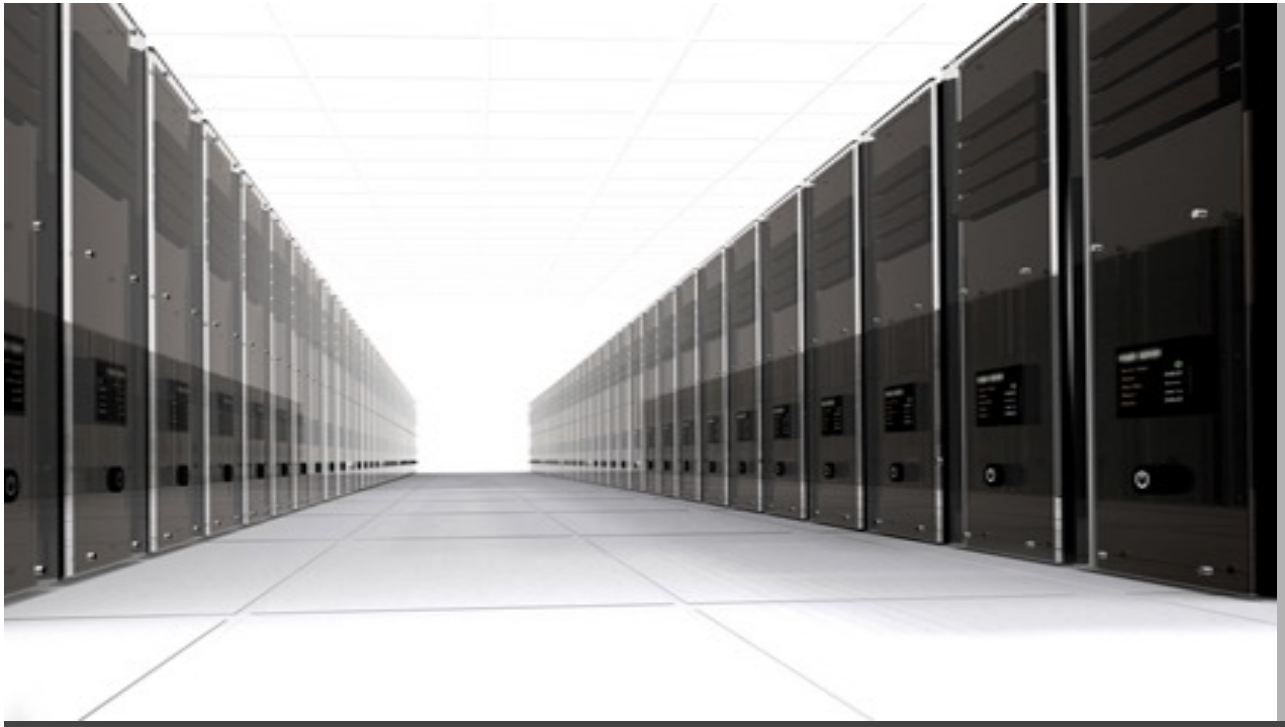
- Dedicated high-performance network for block-level storage
  - Typically 2-200 Gbps, <1µsec latency
- Protocol operates on blocks: multiple clients can access files at the same time with very high performance (as it was a local hard disk). Changes are visible by all nodes
- SAN is LAN-independent; if the LAN goes slow does not affect
- More complex to administrate, more expensive
- Not affected by out-or-order

In the EBH project

We cannot change the LAN

- Technical legacy (more common than you imagine)
- We can add a SAN
  - If there are usual or punctual performance degradation
  - If the very nature of data requires it

REAL LIFE IS ALWAYS MORE COMPLEX



NETWORK CONCEPTS (FOR DATA CENTERS)

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH