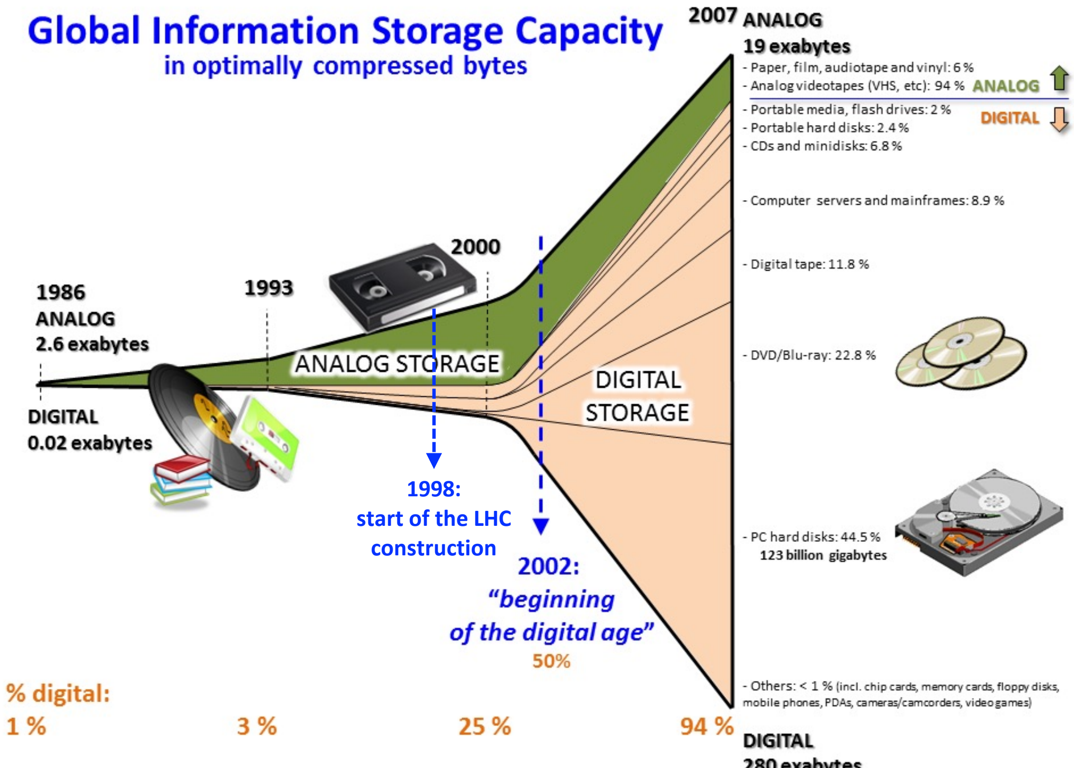
WLCG Lessons in Big Data Processing

Andrey Kiryanov, NRC KI – PNPI

Workshop on Information Technology in Natural Sciences,
27 October 2023

What is Big Data?

"Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software" — *Wikipedia*



Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60 –65. http://www.martinhilbert.net/WorldInfoCapacity.html

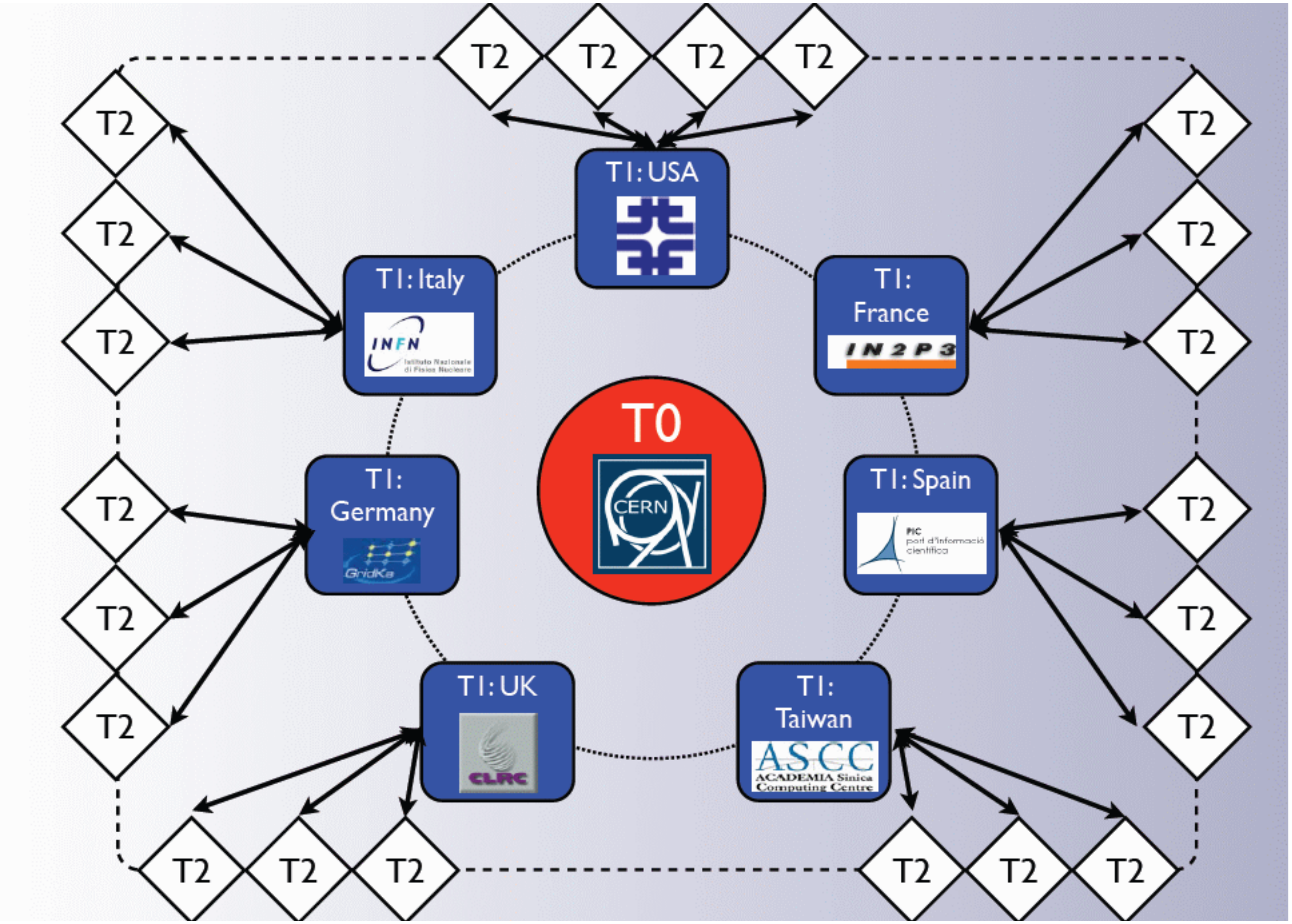
The LHC Precedent

- LHC construction started in 1998
- There was no Google, no Facebook, no YouTube back then
- Collectively, the LHC experiments were expected to produce about **10 petabytes** of raw data each year that must be stored, processed, and analyzed
- LHC Computing Grid (LCG) was a pioneer in a field where no proven technology or experience existed yet

The MONARC Approach

- Strictly hierarchical model
- Three-tier
  - 0 is CERN where all raw data is stored on tape
  - 1 is a large computing center (~10 sites) with partial copies of raw data on tape for redundancy
  - 2 is a medium-size computing center with no raw data
- Dedicated Optical Private Network (OPN) for data transfers between T0 and T1s
- Compute where the data is
- CERN provides about 20% of all computing resources

LHC Computing Grid



The reasoning behind T1/T2 split

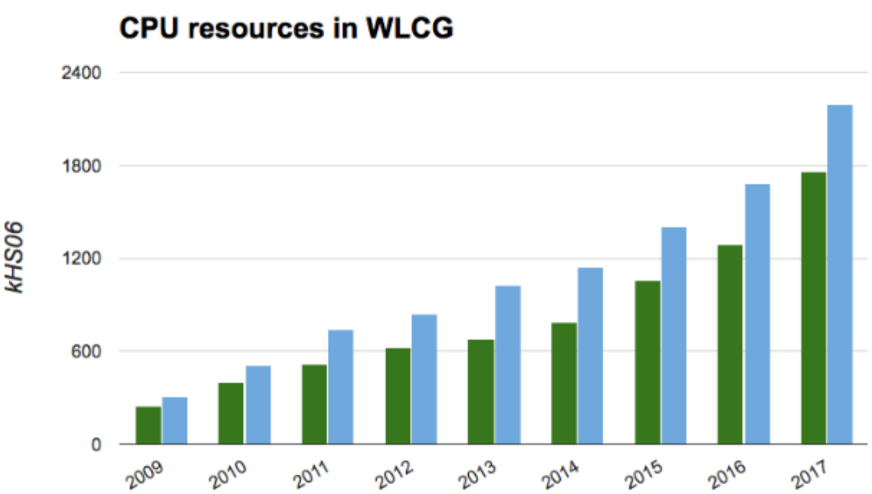
- Wide area networks were rather slow
- Spinning disks were rather small in size (<1 TB)
- The only way to store a large volume of data was tapes
  - Tape drive is not a random-access device
  - Data stage-in/stage-out can take days
- Reliability of T1 was expected to be higher than of T2
  - Diesel-backed UPS mandatory

2007

2017

Resource growth

- At the start of the LHC operation cumulative disk resources of T1s were almost twice as large as of T2s



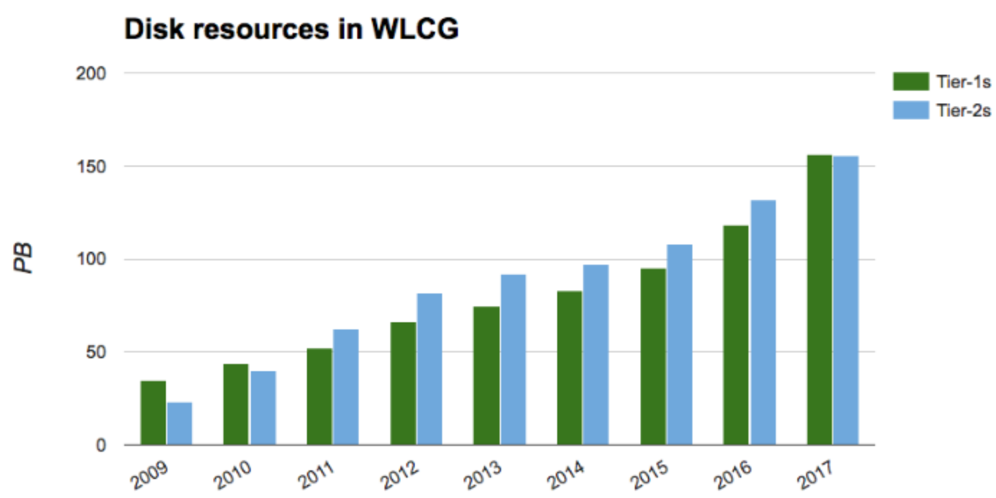


Chart by J. Flix – PIC/CIEMAT – jfix@pic.es
March 2017 GDB – ISGC2017 - Taipei

Why do we need a copy elsewhere? (lesson 1)

INFN underwater Tier-1 in 2018

Conclusions made after the Run-1

- Network technology growth was significantly underestimated (lesson2)
  - Available bandwidth utilization is sub-optimal
- Lack of flexibility wrt CPU⇄Storage mapping
- Dark data and excessive replication waste disk resources (lesson 3)
- Data popularity matters
- Issues with Resource Brokers cause resource under-utilization and high job failure rates (lesson 4)

Most significant changes

- What was initially a hierarchical model transformed into a mesh
- Data processing was no longer limited by what’s available at a local storage
- Local file catalogs were replaced by several central registries (experiment-specific)
- Data movement was automated
- Pilot-based jobs made Resource Brokers redundant

CPU side

- There was a large lag between resource state change and the time when it’s reflected in the information system
  - Information system underwent several changes
  - Experiments relied less and less on it
- Broken compute nodes were turning into blackholes
  - Pilot jobs were invented to overcome this
  - They also made Resource Brokers mostly redundant
- Job isolation was pretty bad
  - Pools of user accounts were introduced
- Experiments implemented their own job queues with QoS

Storage side

- In some cases it was faster to just read input data remotely than to stage it to the local storage
  - Xrootd protocol allows fast random reads from large files over the network
  - Storage federations emerged
  - Data lifetime management
- Small site storages created more overhead
  - A per-experiment low limit was introduced
  - Small local storage can be used as a cache
- A new File Transfer Service (FTS) was developed
  - Data movement between endpoints
  - Data stage-in/stage-out from tapes
  - Bulk file deletion

Software side

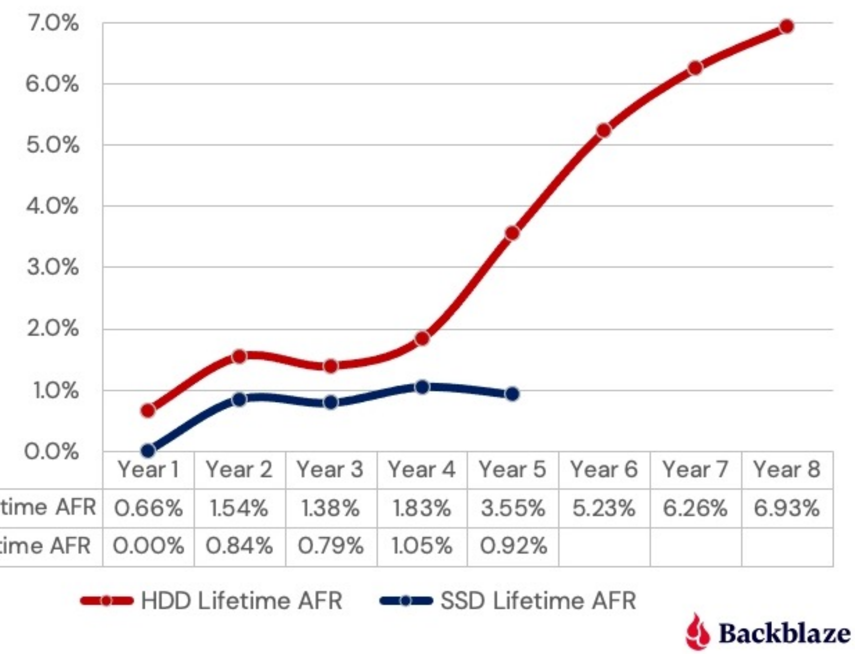
- Software distribution was a major headache
  - Complicated procedures for local deployment and validation
  - Library hell (bundled libraries incompatibility with the system ones)
  - Significant amount of disk storage was required on **every** compute node
  - Shared NFS did not work well with hundreds of nodes
- CVMFS + Docker made life **much** easier
  - Local cache on disk nodes
  - Atomic catalog updates

Infrastructure side

- Complexity shifted from the infrastructure to the experiment's services
- Integration of HPC resources
- HTTP was adopted as a transfer protocol
  - SRM was deprecated
  - HTTPS only for management, plain HTTP for transfers
- New authentication methods
  - Tokens, OAuth2
- Next-gen per-experiment information system (CRIC)
  - No one-size-fits-all approach

Some immediate prospects

- Tape market is not very healthy
- **Ceph** allows for elastic management of disk resources, even the older ones
- SSD expected MTBF surpasses HDD
- Storage federations and Data Lakes absorb individual storages



	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
HDD Lifetime AFR	0.66%	1.54%	1.38%	1.83%	3.55%	5.23%	6.26%	6.93%
SSD Lifetime AFR	0.00%	0.84%	0.79%	1.05%	0.92%			

Thank you!