Introduzione ai Big Data

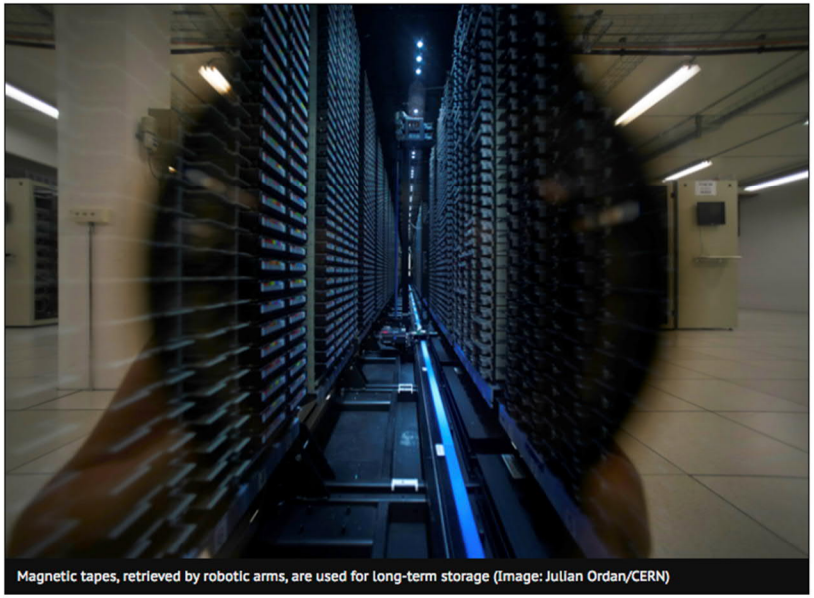
La gestione dell'informazione dal Data Taking al Cloud

Giuseppe Lo Presti
CERN IT Department

Italian Teachers Programme 2024 - Discovery

Breaking data records bit by bit

by Harriet Jarlett



Magnetic tapes, retrieved by robotic arms, are used for long-term storage (Image: Julian Ordan/CERN)

This year CERN's data centre broke its own record, when it collected more data than ever before.

During October 2017, the data centre stored the colossal amount of 12.3 petabytes of data. To put this in context, one petabyte is equivalent to the storage capacity of around 15,000 64GB smartphones. Most of this data come from the Large Hadron Collider's experiments, so this record is a direct result of the outstanding LHC performance, the rest is made up of data from other experiments and backups.

"For the last ten years, the data volume stored on tape at CERN has been growing at an almost exponential rate. By the end of June we had already passed a data storage milestone, with a total of 200 petabytes of data permanently archived on tape," explains German Cancio, who leads the tape, archive & backups storage section in CERN's IT department.

STORAGE

CERN swells storage space beyond 1EB for LHC's latest ion-whacking experiments

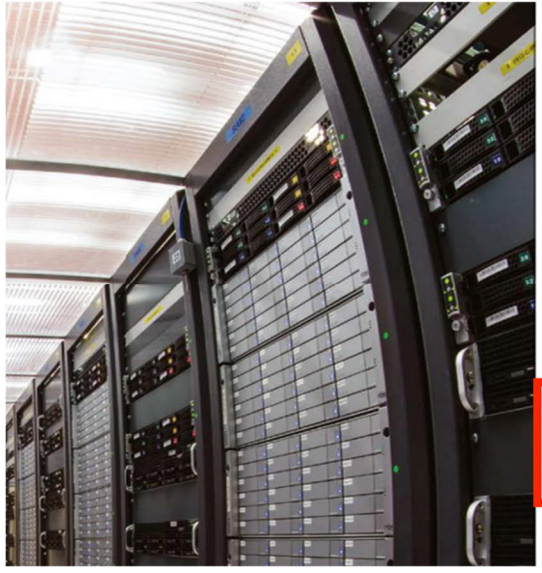
A petabyte or more a day of readings? No problem, pal

Tobias Mann

Mon 2 Oct 2023 | 19:48 UTC

In preparation for its latest round of ion-smashing tests, CERN boosted its storage array for the experiments to more than one million terabytes in total size.

The facility's data store now exceeds an exabyte of raw capacity — with much of it on hard disk drives and an "increasing fraction of flash drives," the European super-lab's team explained in a report.



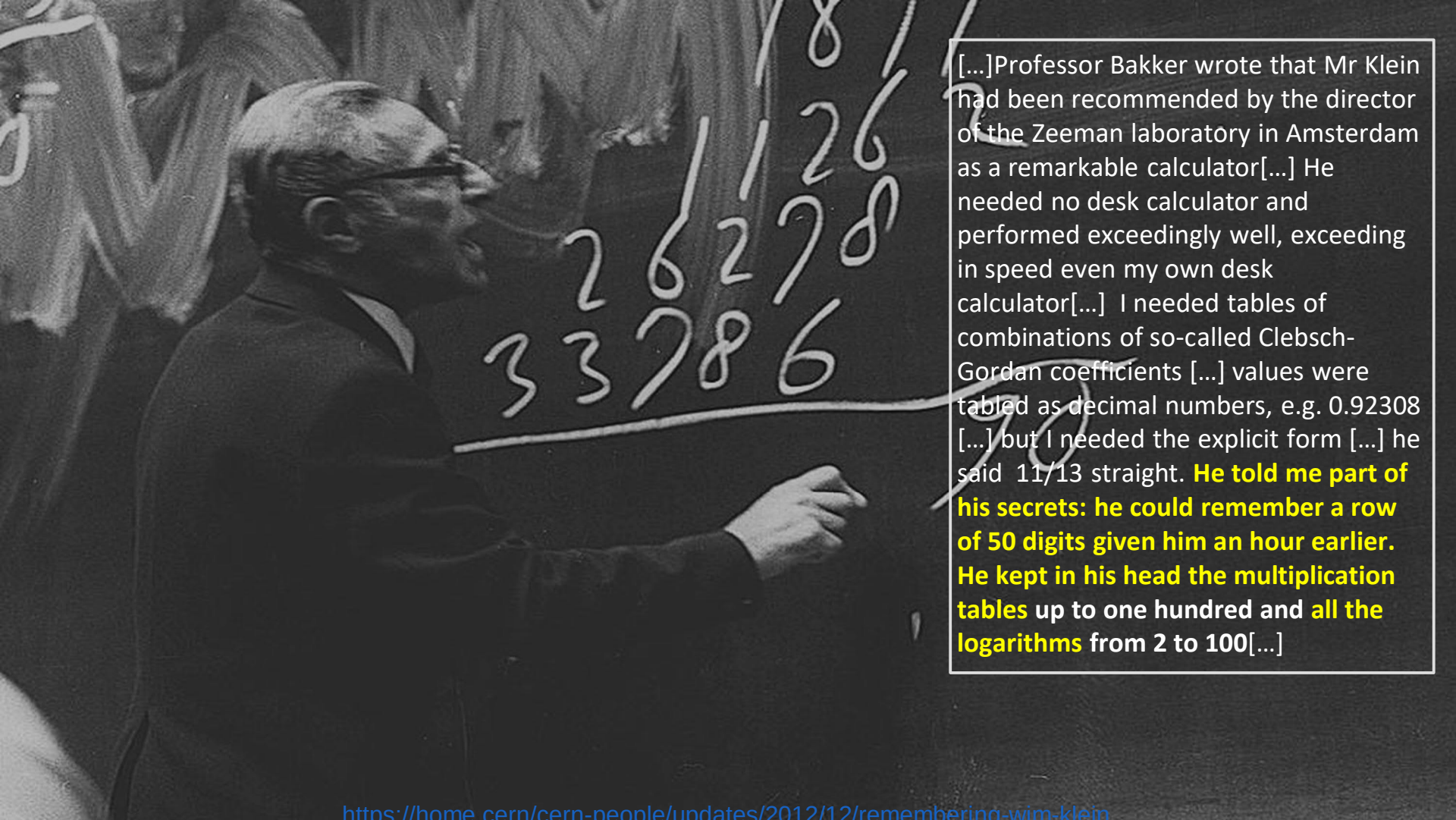
CERN Courier April 2018

Software and computing

Time to adapt for big data

Radical changes in computing and software are required to ensure the success of the LHC and other high-energy physics experiments into the 2020s, argues a new report.

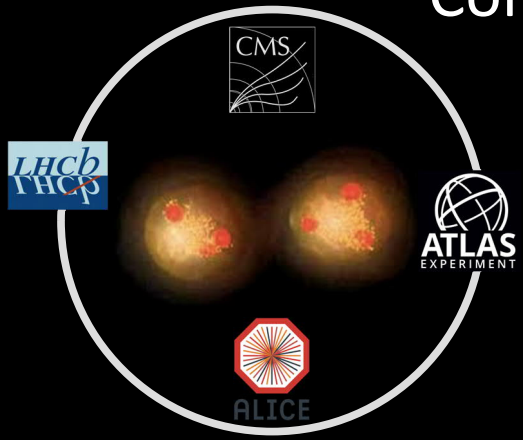
It would be impossible for anyone to conceive of carrying out a particle-physics experiment today without the use of computers and

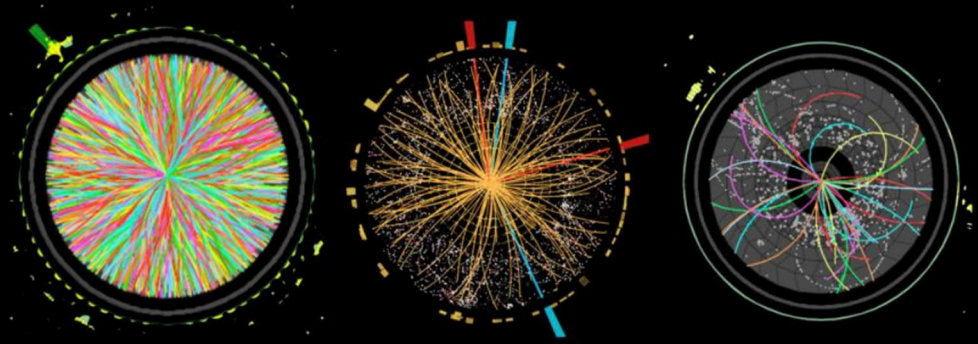


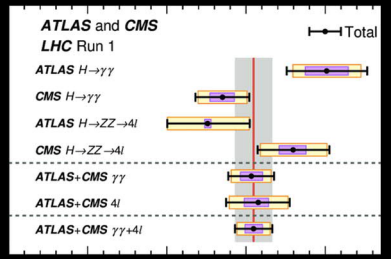
[…]Professor Bakker wrote that Mr Klein had been recommended by the director of the Zeeman laboratory in Amsterdam as a remarkable calculator[…] He needed no desk calculator and performed exceedingly well, exceeding in speed even my own desk calculator[…] I needed tables of combinations of so-called Clebsch-Gordan coefficients […] values were tabled as decimal numbers, e.g. 0.92308 […] but I needed the explicit form […] he said 11/13 straight. **He told me part of his secrets: he could remember a row of 50 digits given him an hour earlier. He kept in his head the multiplication tables up to one hundred and all the logarithms from 2 to 100**[…]

https://home.cern/cern-people/updates/2012/12/remembering-wim-klein

Computing at CERN: The Big Picture

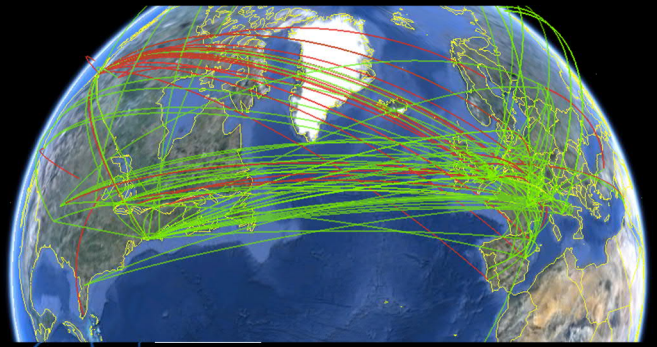






Data Storage - Data Processing - Event generation - Detector simulation - Event reconstruction - Resource accounting

Distributed computing - Middleware - Workload management - Data management - Monitoring



GAUDI-LHCb

ATHENA-ATLAS

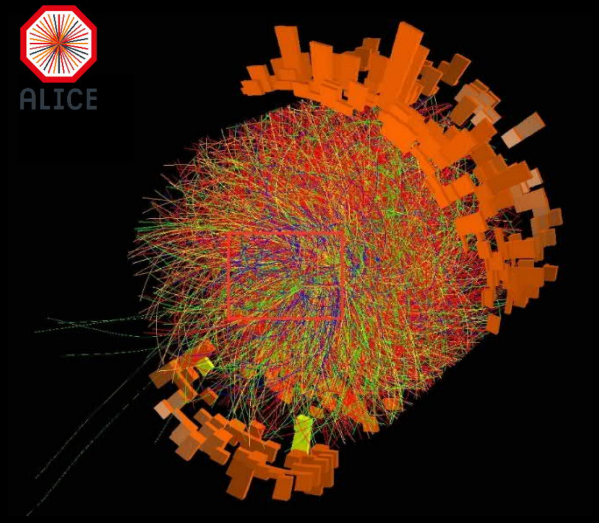
CMSSW-CMS

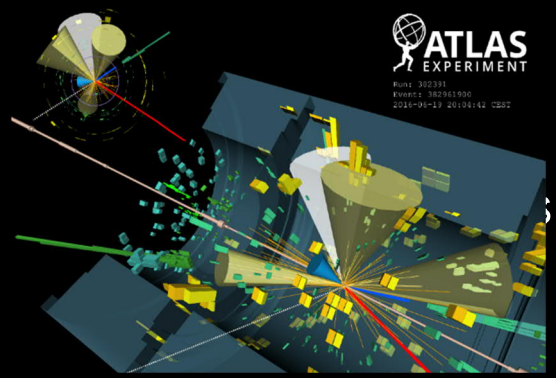
GEANT4
A SIMULATION TOOLKIT

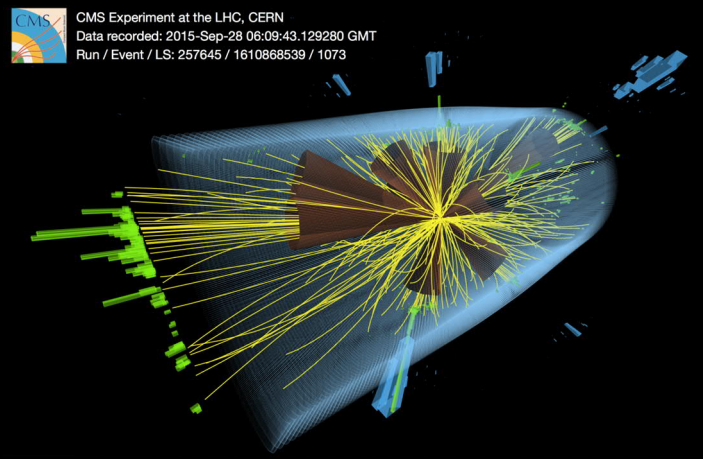
SHERPA

PYTHIA

From the Hit to the Bit: Data Acquisition



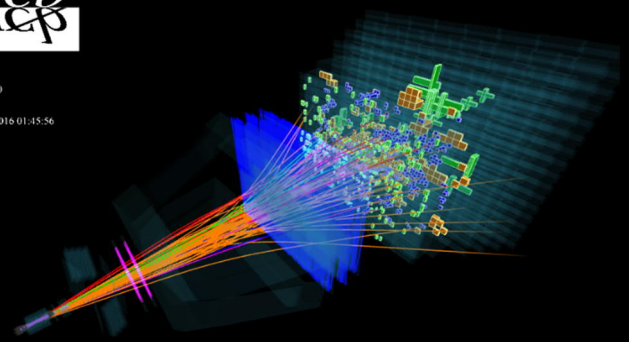




100 million channels

40 million pictures a second

Synchronised signals from all detector parts



From the Hit to the Bit: Event Filtering

L1: 40 million bunch cross per second

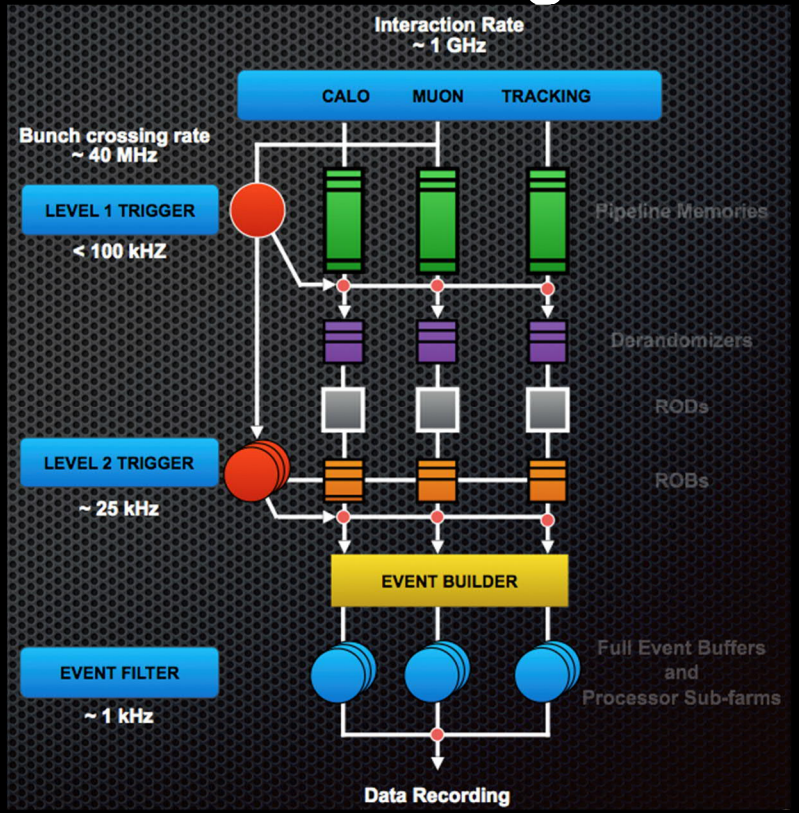
- Fast, simple information
- **Hardware** trigger in a few micro seconds

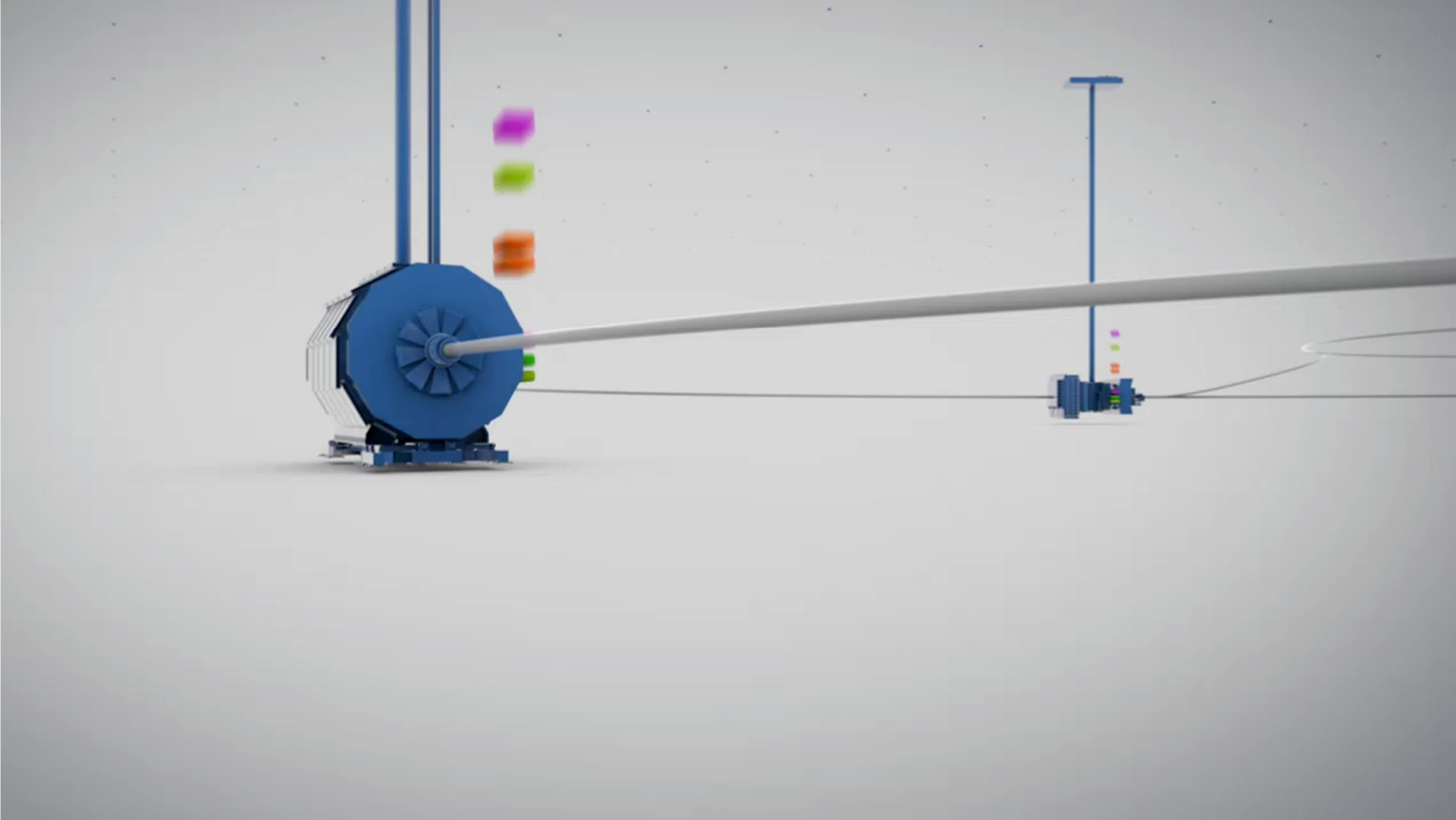
L2: 100,000 events per second

- Fast algorithms in local computer farm
- **Software** trigger in <1 second
- ***Which OS for such task?***

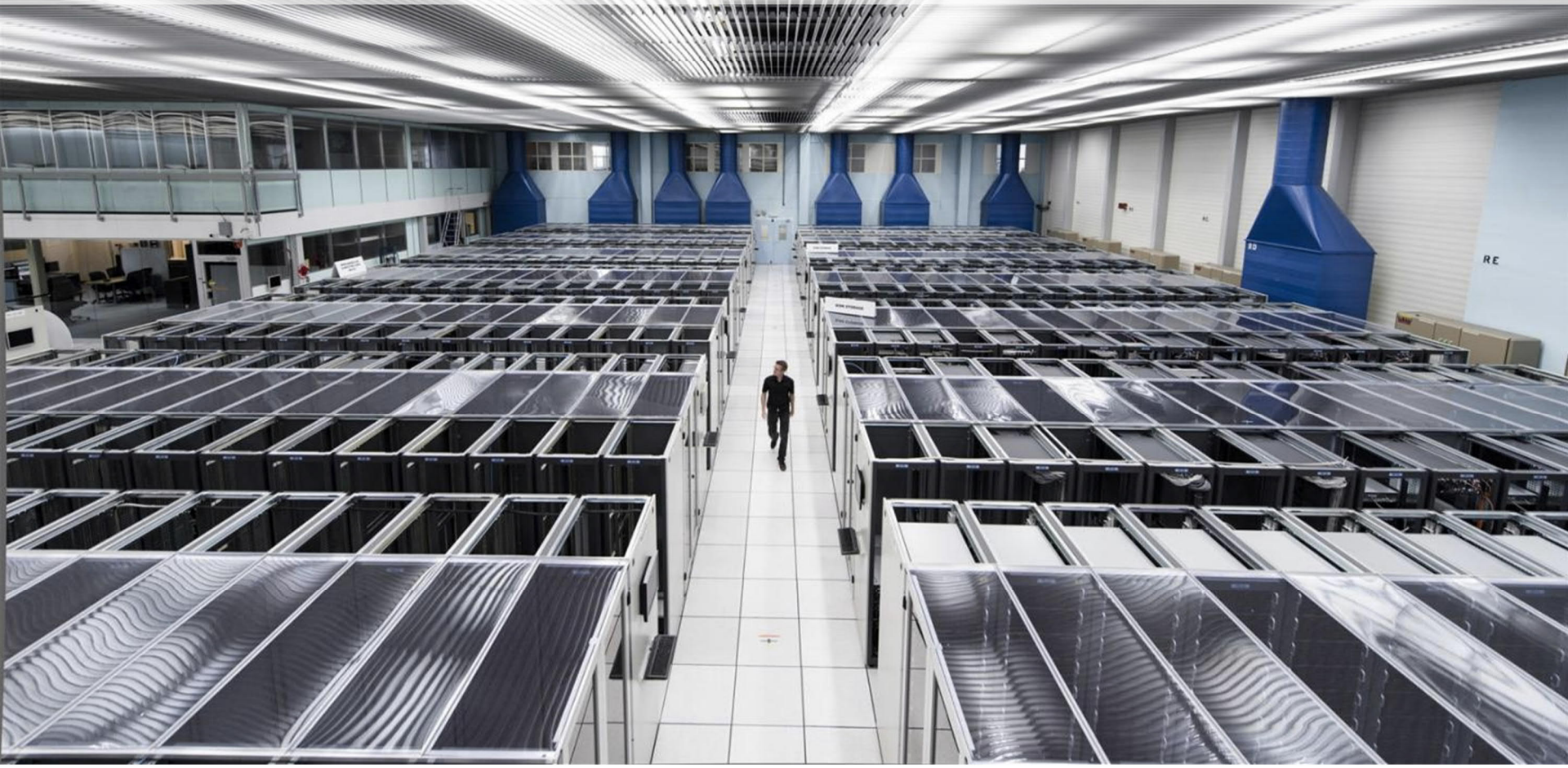
EF: Up to 10,000 per second recorded for offline analysis

- By each experiment!



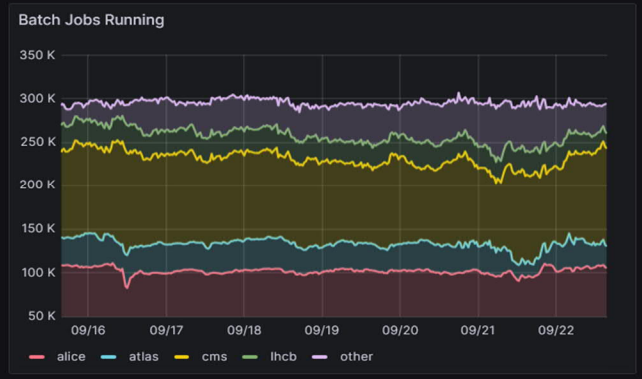


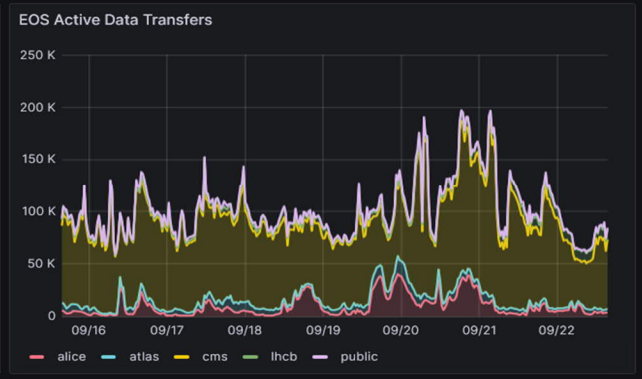
The CERN Data Centre

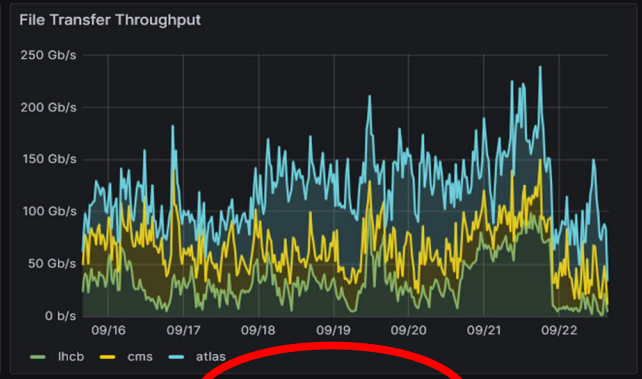


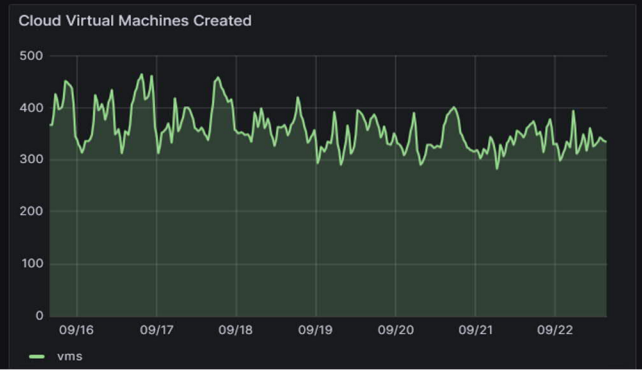
CERN DC: an ordinary week in numbers

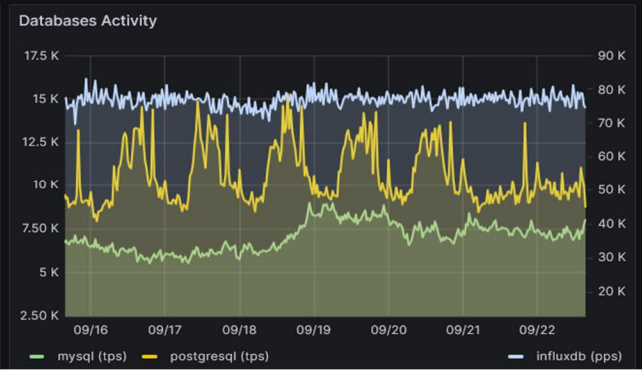
Servers	Cores	Disks	Tape Drives	Routers	Wifi Points
12.1 K	358.2 K	221.2 K	180	302	5.1 K

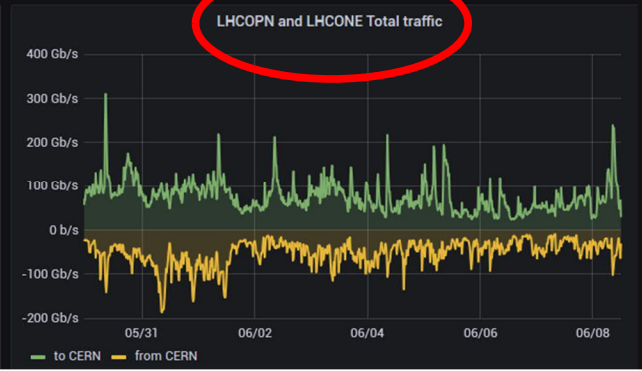






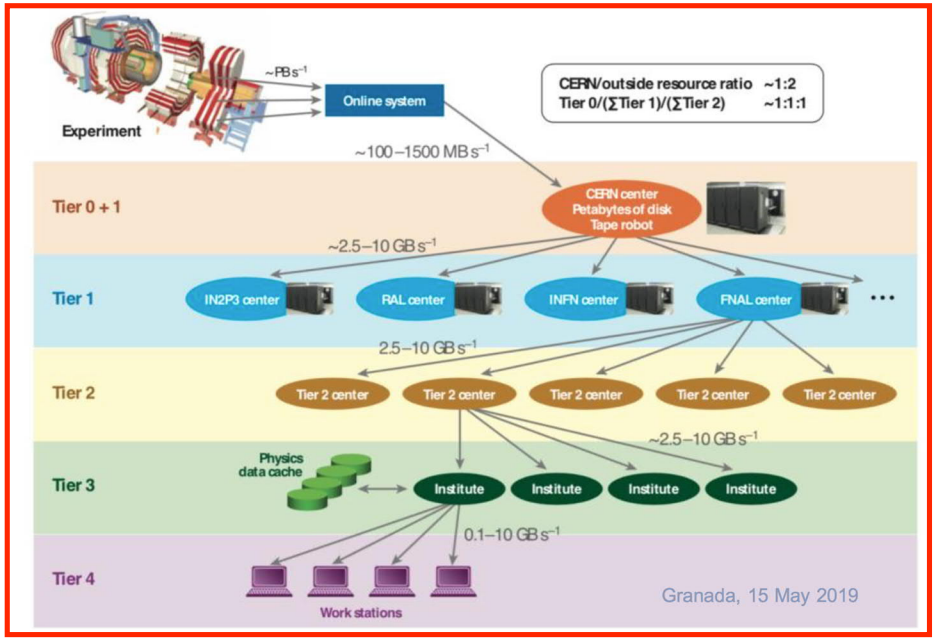






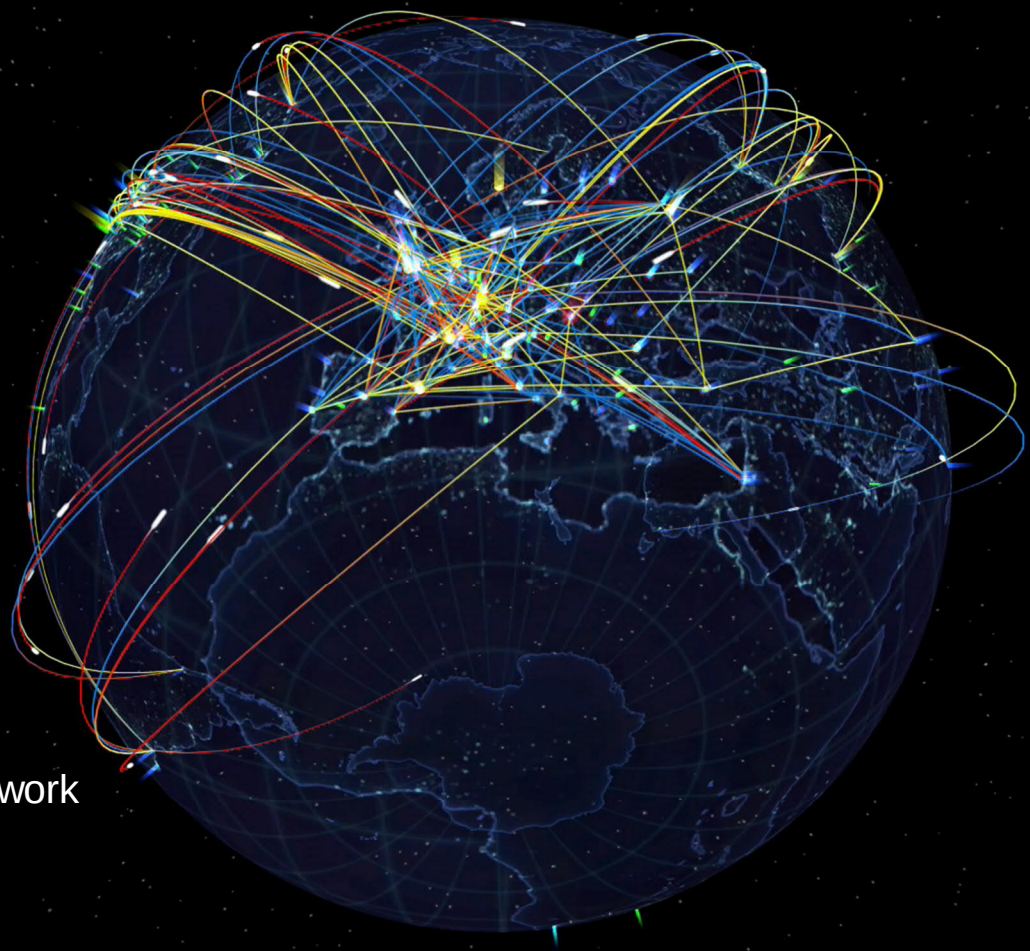
The Worldwide LHC Computing **Grid**

- The Worldwide LHC Computing Grid (WLCG) is a global collaboration of more than 170 data centres around the world, in 42 countries
- The CERN data centre (Tier-0) distributes the LHC data worldwide to the other WLCG sites (Tier-1 and Tier-2)
- WLCG provides global computing resources to store, distribute and analyse the LHC data
  - CERN = only 15% of CPU resources
  - Distributed funding
  - "Sociological" reasons



Data Distribution in WLCG

- Global transfer rates regularly exceeding **80 GB/s**
- **1+ EB** and 1.1B files transferred yearly in Run 3
- Main **challenge** is to have the **useful data close** to available computing resources
  => match storage/compute/network



Software Platforms for HEP

- Home made solutions vs. integrating software platforms from the (open source) market
  - Infrastructure moving towards the latter as industry grew in front of us!

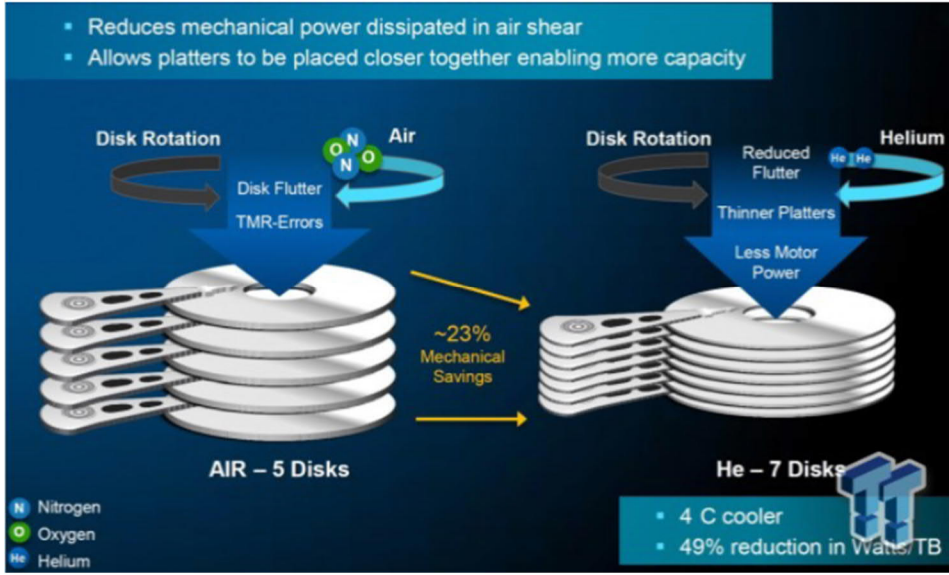
Software Platforms for HEP

- Home made solutions vs. integrating software platforms from the (open source) market
  - Infrastructure moving towards the latter as industry grew in front of us!
  - Yet, **high-level storage software customized** for our **specific** access patterns

Largest scientific data repository

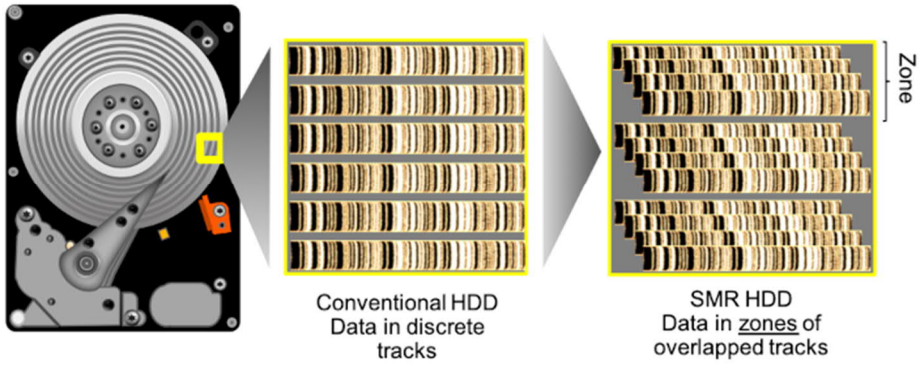
Which devices for so much data?

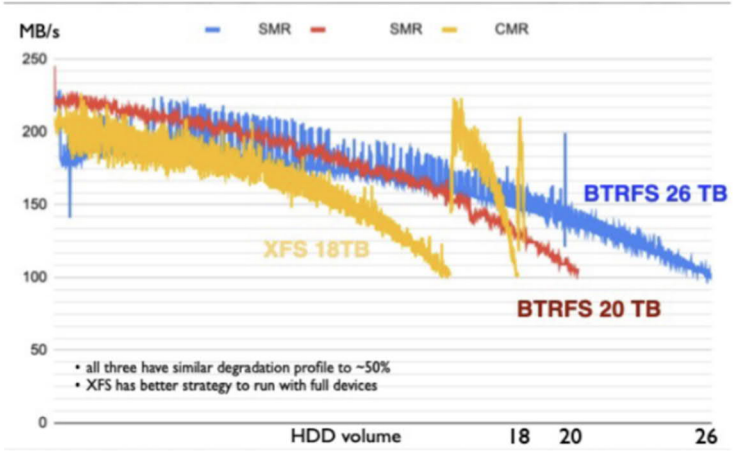
- The demand for storage keeps increasing
  - The “AI era” only makes things worse
- Manufacturers forced to try new tricks to increase bit density
  - Starting from the easy ones…



Which devices for so much data?

- And moving to more sophistications:
  From Conventional to **S**hingled Magnetic Recording
  - But this does not come for free…





Heat-Assisted Magnetic Recording Enables Continued Capacity Growth for Hard Drives

Media Technology

High coercivity material enables smaller, thermally stable grains

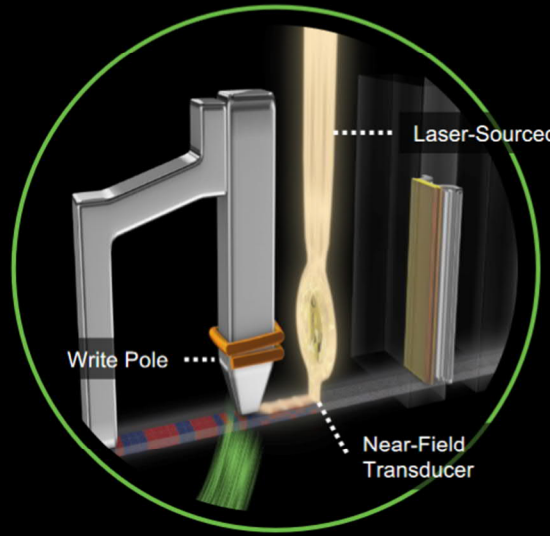
Glass substrate enables required sputter temperatures

Head Technology

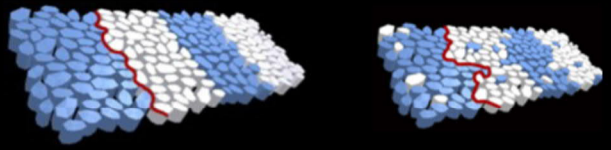
Focused heat provides local reduction in coercivity

Write pole sets magnetic bit

Heats and cools in less than 2 nanoseconds



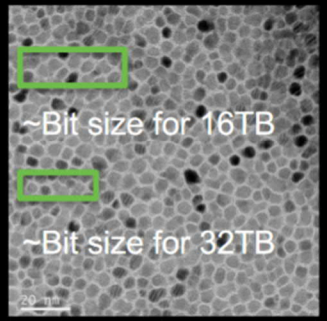
Areal Density Media Challenge:



Hold #Grains / Bit constant

Higher areal density requires smaller grains

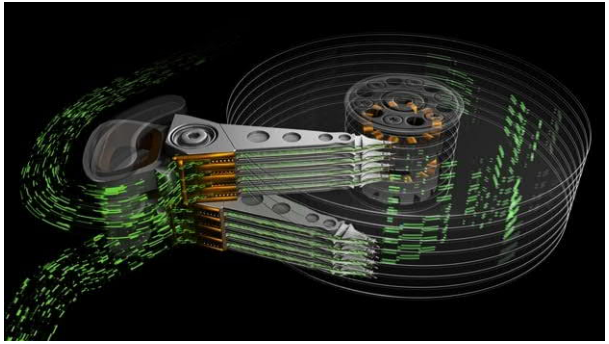
5.6nm HAMR Media



[HEPiX Spring 2024, Paris]

What about throughput?

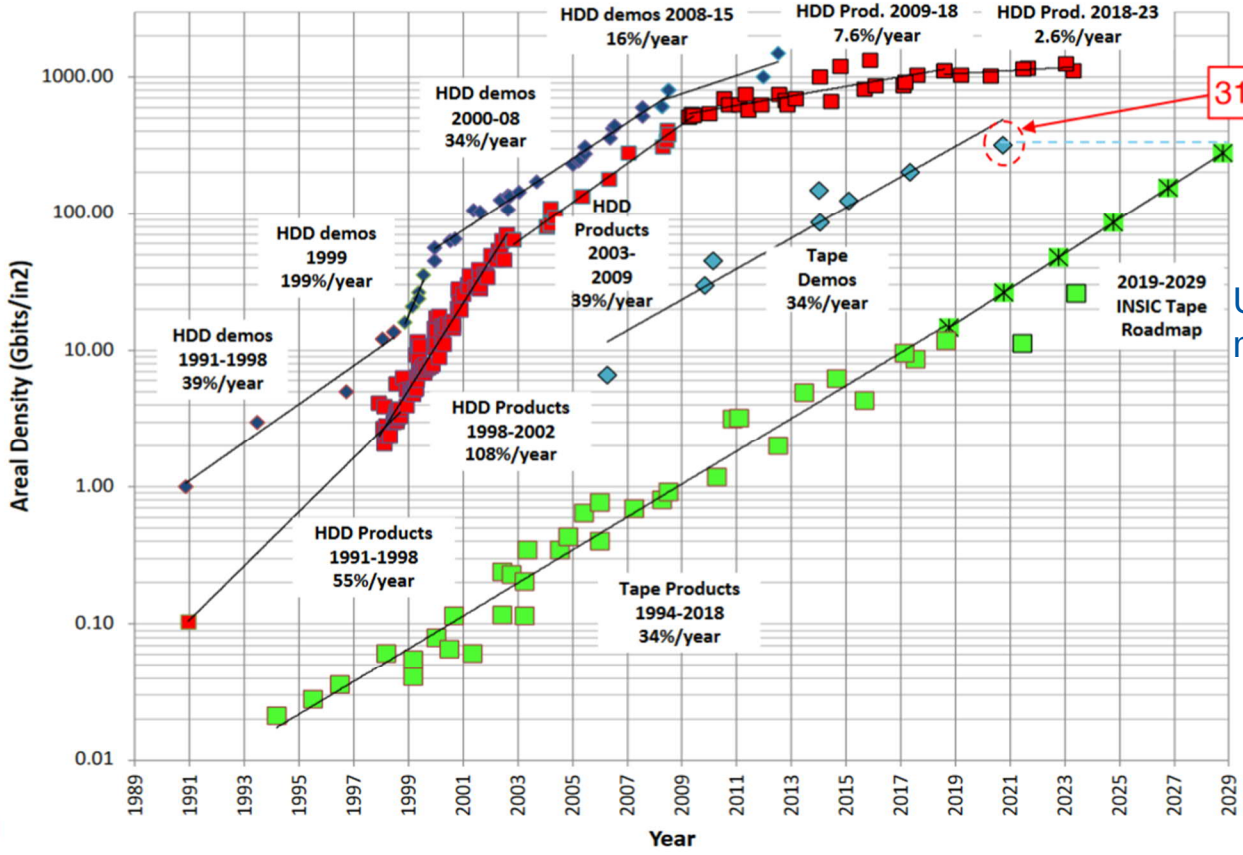
- Most Hard Drives on the market: up to 250 MB/s
- **SMR** and **HAMR/MAMR** do not bring any improvement!
  - A 4 TB drive used to take **4.5h** to read/write fully
  - A 30 TB drive takes **1.5 days**!
- Possible solution:
  double actuator for the head
  - But that's not enough for us…



Tapes vs Disks/SSDs

- Tape technology has a larger growth margin
  - A tape has 1 km of magnetic media

Disk vs Tape Areal Density Scaling

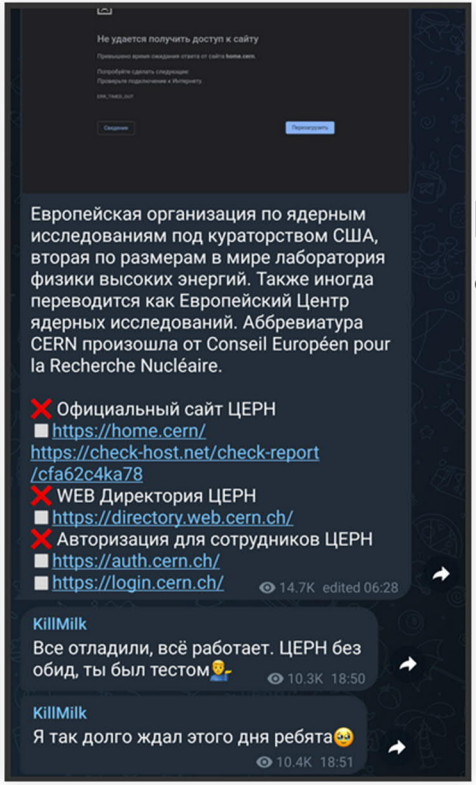


IBM Demo:
580 TB tape

Upcoming on the market: **50 TB** tape

An appealing attack target

- CERN is permanently under cyber attack
- Computer Security is a pillar of the whole IT infrastructure
  - Raising awareness at CERN and at partner institutes
    - It’s not a matter of “if”, but “when”!
    - Phishing campaigns, role games, presentations about real cases and mitigation measures, …
  - Mandatory “Dual-Factor Authentication” (2FA) for IT operators as of 2022, for everyone now
  - Continuous “white hat” penetration testing, in collaboration with the wider scientific community



Take-away #1

- LHC data rates range from the PB/sec at the detector to the GB/sec after filtering
- Scientific data towards Exabyte scale
- Data centres run on commodity hardware and open-source OSes
- *Hyperscalers* are (much) larger
  - They drive the market, including the manufacturers
  - CERN remains the world-largest scientific repository
- …Is this really “Big Data”?

Thanks for your attention! Questions?

Accélérateur de science

Giuseppe.LoPresti@cern.ch
www.linkedin.com/in/giuseppelopresti

Credits to all CERN IT Storage colleagues