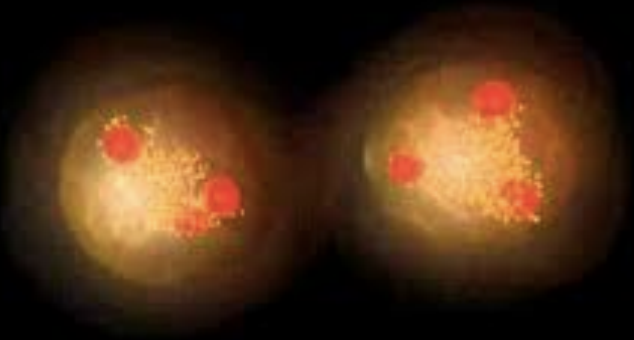
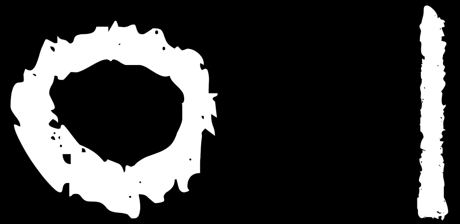




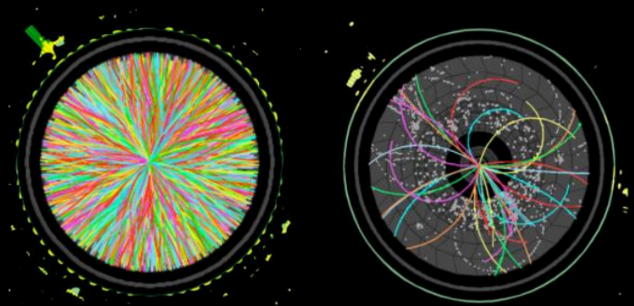
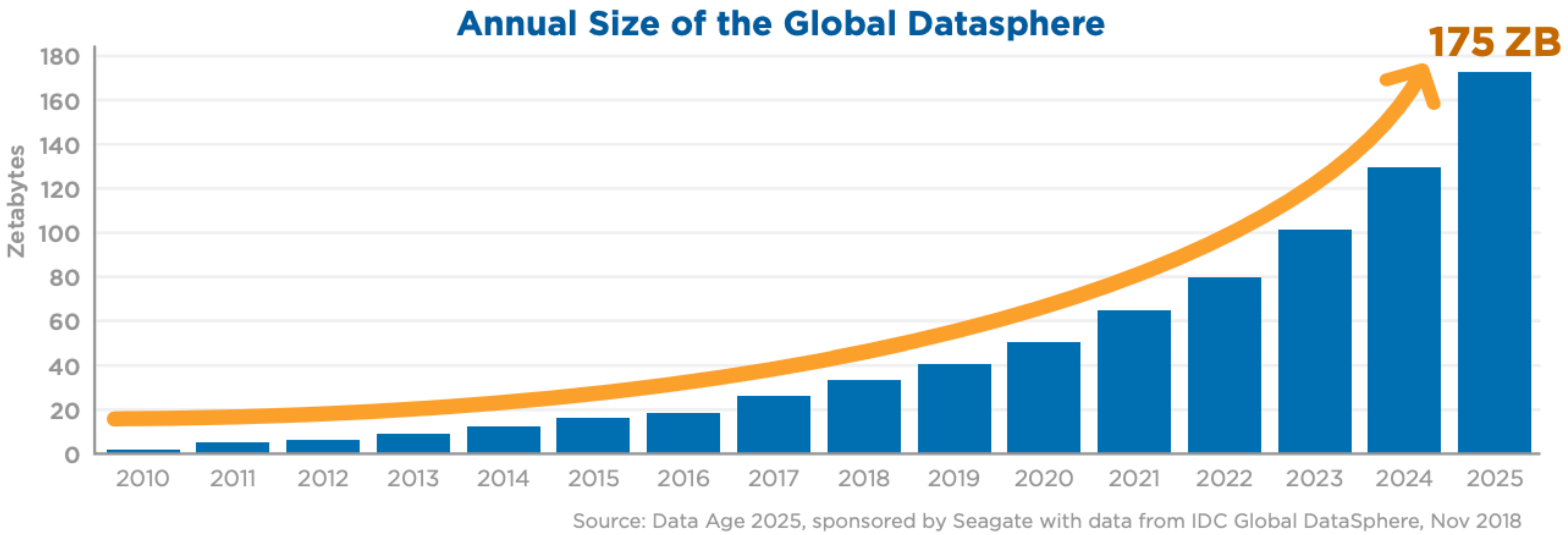


photo @ https://seanmunger.com/



https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

LHC: pushing computing to the limits

The Large Hadron Collider produced unprecedented volumes of data during its two multi-year runs, and, with its current upgrades, more computing challenges are in store

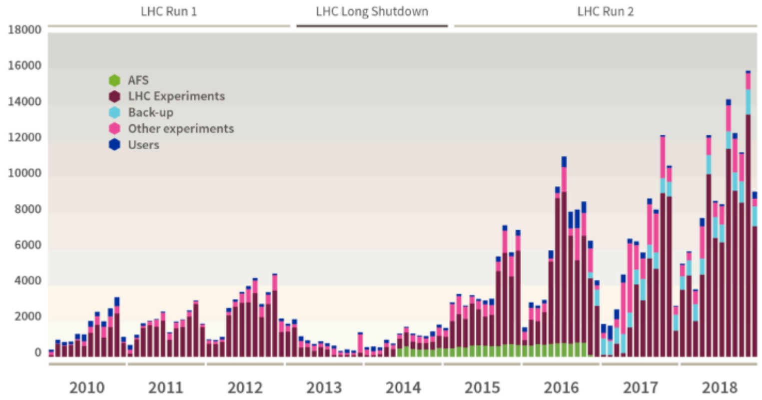
1 MARCH, 2019 | By Esra Ozcesmeci

Racks of computers in CERN's computing centre are just a fraction of the hardware needed to store and process the data from the LHC (Image: CERN)

At the end of 2018, the Large Hadron Collider (LHC) completed its second multi-year run ("Run 2") that saw the machine reach a proton–proton collision energy of 13 TeV, the highest ever reached by a particle accelerator. During this run, from 2015 to 2018, LHC experiments produced unprecedented volumes of data with the machine's performance exceeding all expectations.

This meant exceptional use of computing, with many records broken in terms of data acquisition, data rates and data volumes. The CERN Advanced Storage system (CASTOR), which relies on a tape-based backend for permanent data archiving, reached 330 PB of data (equivalent to 330 million gigabytes) stored on tape, an equivalent of over 2000 years of 24/7 HD video recording. In November 2018 alone, a record-breaking 15.8 PB of data were recorded on tape, a remarkable achievement given that it corresponds to more than what was recorded during the first year of the LHC's Run 1.

The distributed storage system for the LHC experiments exceeded 200 PB of raw storage with about 600 million files. This system (EOS) is disk-based and open-source, and was developed at CERN for the extreme LHC computing requirements. As well as this, 830 PB of data and 1.1 billion files were transferred all over the world by File Transfer Service. To face these computing challenges and to better support the CERN experiments during Run 2, the entire computing infrastructure, and notably the storage systems, went through major upgrades and consolidation over the past few years.

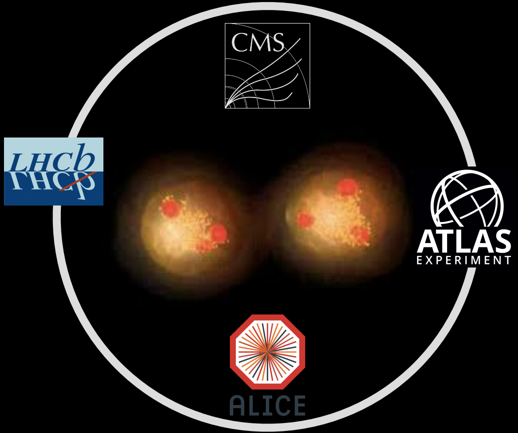


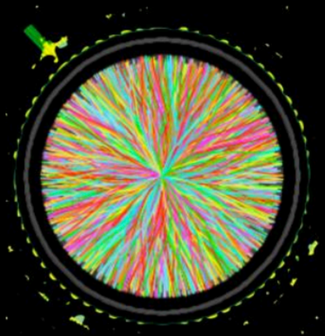
Data (in terabytes) recorded on tape at CERN month-by-month. This plot shows the amount of data recorded on tape generated by the LHC experiments, other experiments, various back-ups and users. In 2018, over 115 PB of data in total (including about 88 PB of LHC data) were recorded on tape, with a record peak of 15.8 PB in November (Image: Esma Mobs/CERN)

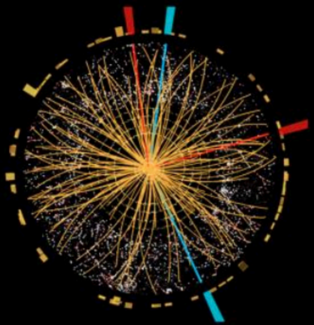
New IT research-and-development activities have already begun in preparation for the LHC's Run 3 (foreseen for 2021 to 2023). "Our new software, named CERN Tape Archive (CTA), is the new tape storage system for the custodial copy of the physics data and a replacement for its predecessor, CASTOR. The main goal of CTA is to make more efficient use of the tape drives, to handle the higher data rate anticipated during Run 3 and Run 4 of the LHC," explains German Cancio, who leads the Tape, Archive & Backups storage section in CERN's IT department. CTA will be deployed during the ongoing second long shutdown of the LHC (LS2), replacing CASTOR. Compared to the last year of Run 2, data archival is expected to be two-times higher during Run 3 and five-times higher or more during Run 4 (foreseen for 2026 to 2029).

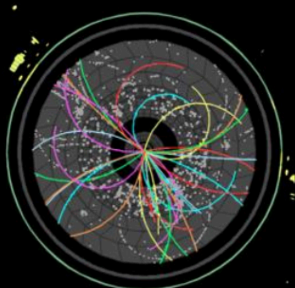
The LHC's computing will continue to evolve. Most of the data collected in CERN's data centre is highly valuable and needs to be preserved and stored for future generations of physicists. CERN's IT department will therefore be taking advantage of LS2, the current maintenance and upgrade of the accelerator complex, to perform the required consolidation of the computing infrastructure. They will be upgrading the storage infrastructure and software to face the likely scalability and performance challenges when the LHC restarts in 2021 for Run 3.

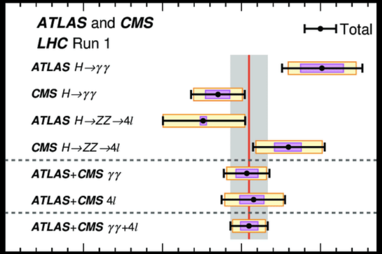
Computing at CERN: The Big Picture





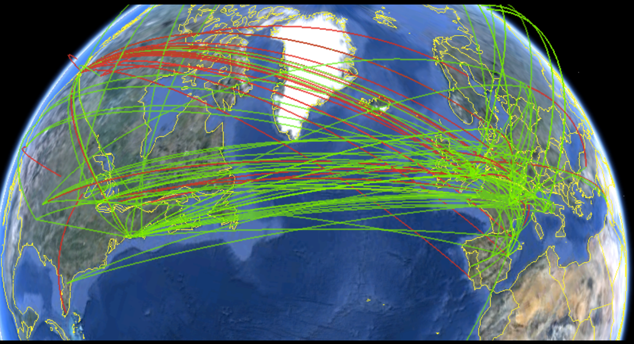






Data Storage - Data Processing - Event generation - Detector simulation - Event reconstruction - Resource accounting

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



GAUDI-LHCb

ATHENA-ATLAS

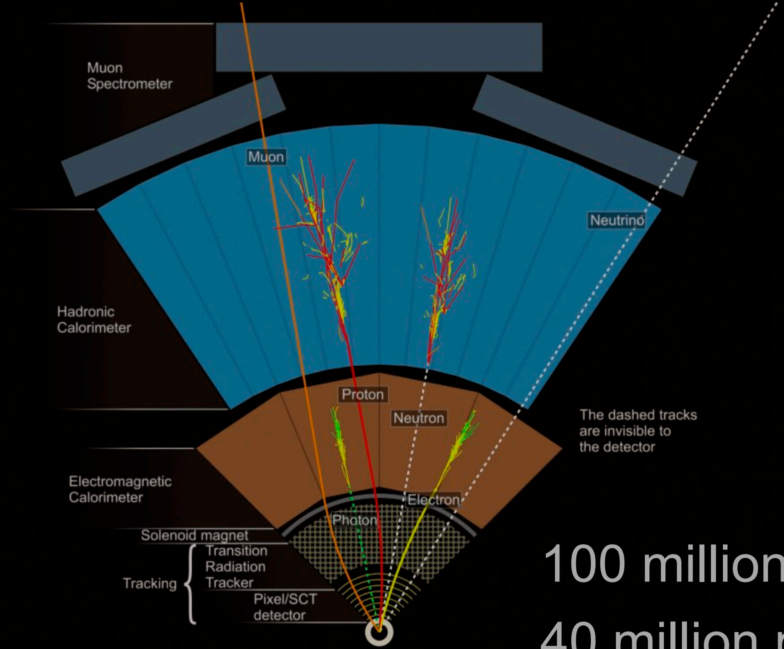
CMSSW-CMS

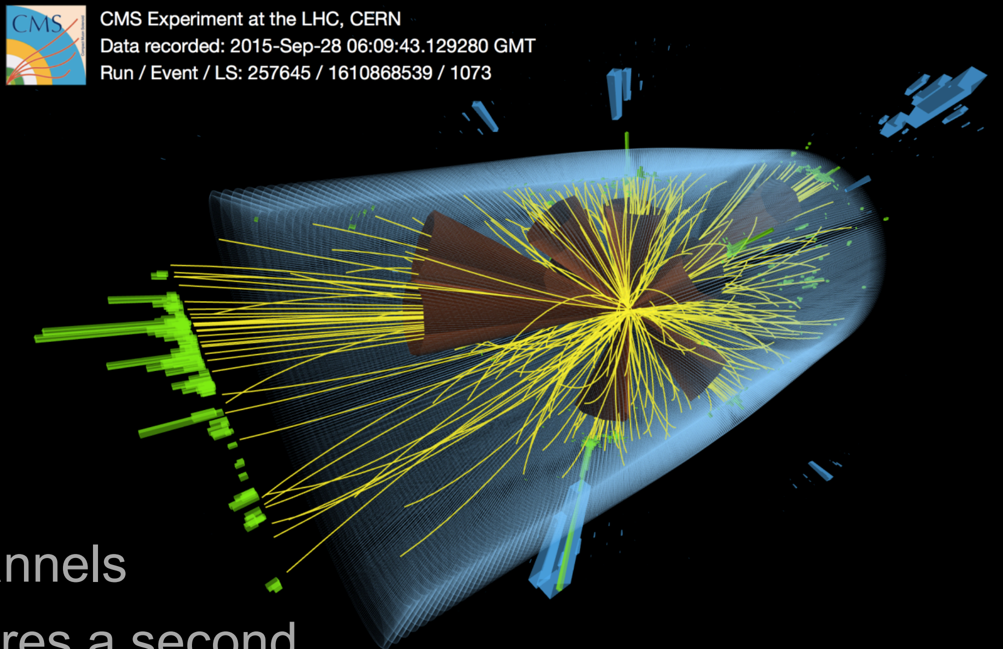
GEANT4 A SIMULATION TOOLKIT

SHERPA

PYTHIA

From the Hit to the Bit: DAQ





100 million channels

40 million pictures a second

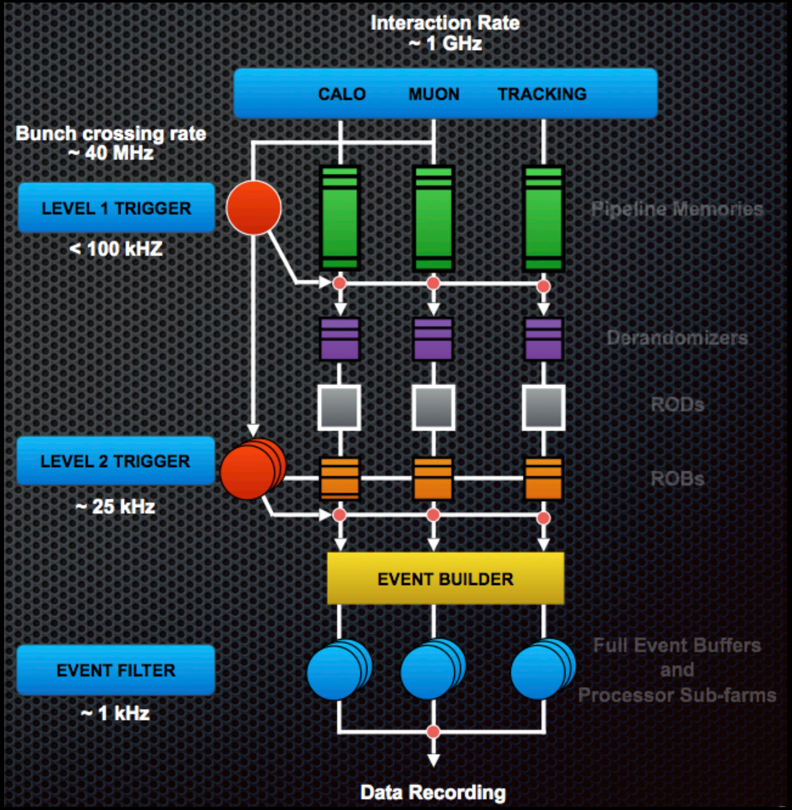
Synchronised signals from all detector parts

From the Hit to the Bit: event filtering

- L1: 40 million events per second
  - Fast, simple information
  - Hardware trigger in a few micro seconds

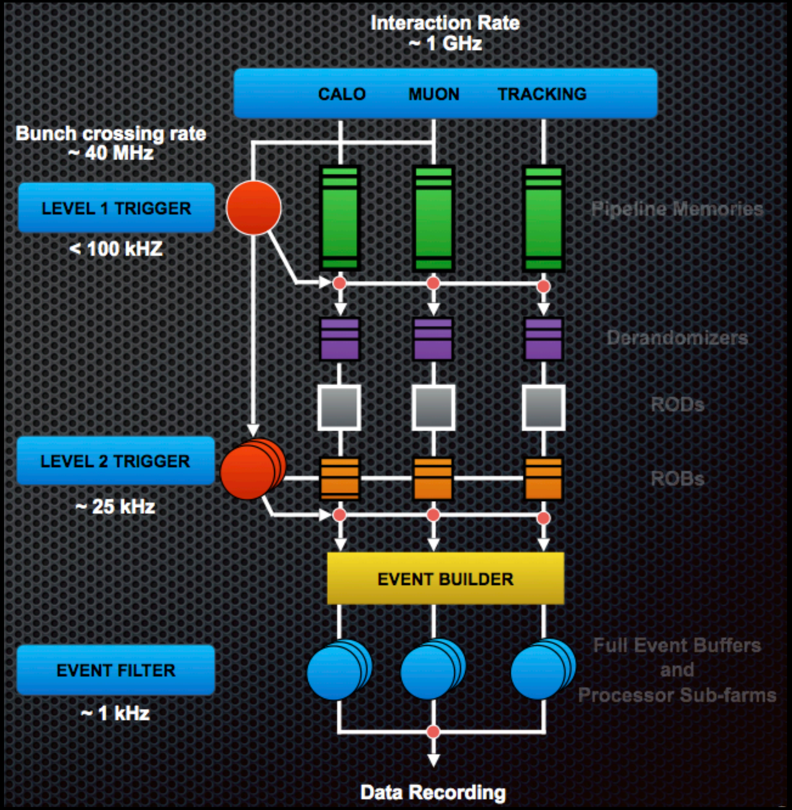
- L2: 100 thousand events per second
  - Fast algorithms in local computer farm
  - Software trigger in <1 second

- EF: Few 100 per second recorded for study



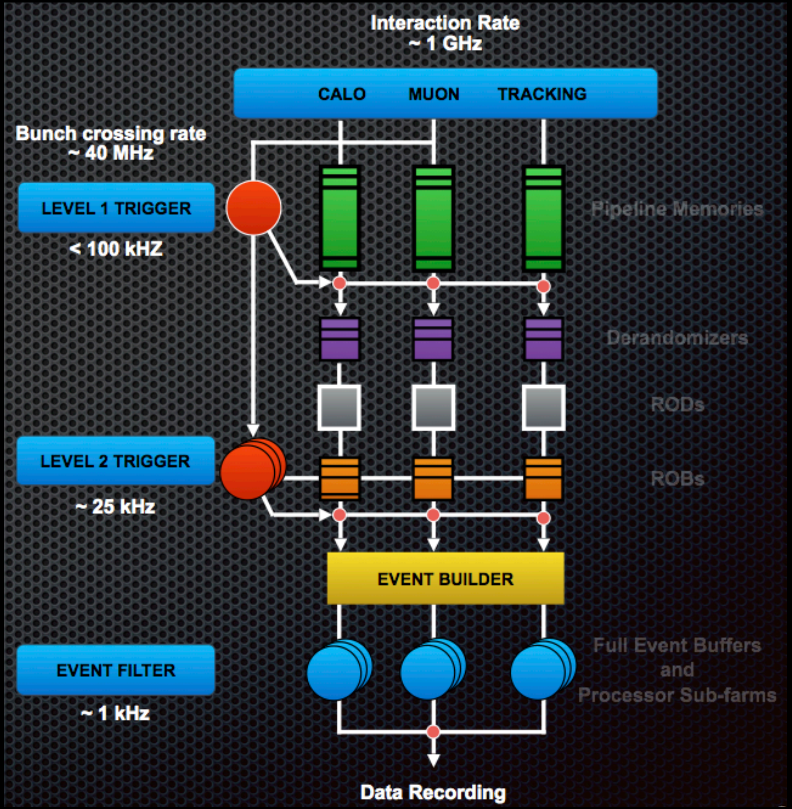
From the Hit to the Bit: event filtering

- L1: this is ~1 Petabyte per second!
  - Cannot afford to store it
  - 1 year’s worth of LHC data at 1 PB/s would cost few hundred trillion euros
- L2: 100 thousand events per second
  - Fast algorithms in local computer farm
  - Software trigger in <1 second
- EF: Few 100 per second recorded for study



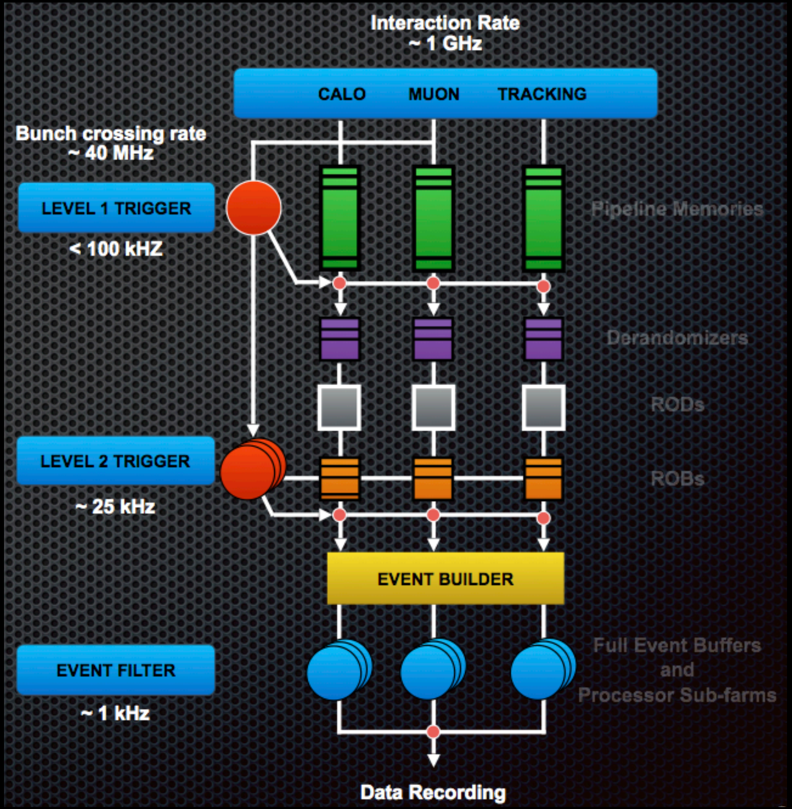
From the Hit to the Bit: event filtering

- L1: this is ~1 Petabyte per second!
  - Cannot afford to store it
  - 1 year's worth of LHC data at 1 PB/s would cost few hundred trillion euros
- L2: Real time to keep only "interesting" data
  - We keep ~1 event in a million
  - Yes, 99.9999% is thrown away
- EF: Few 100 per second recorded



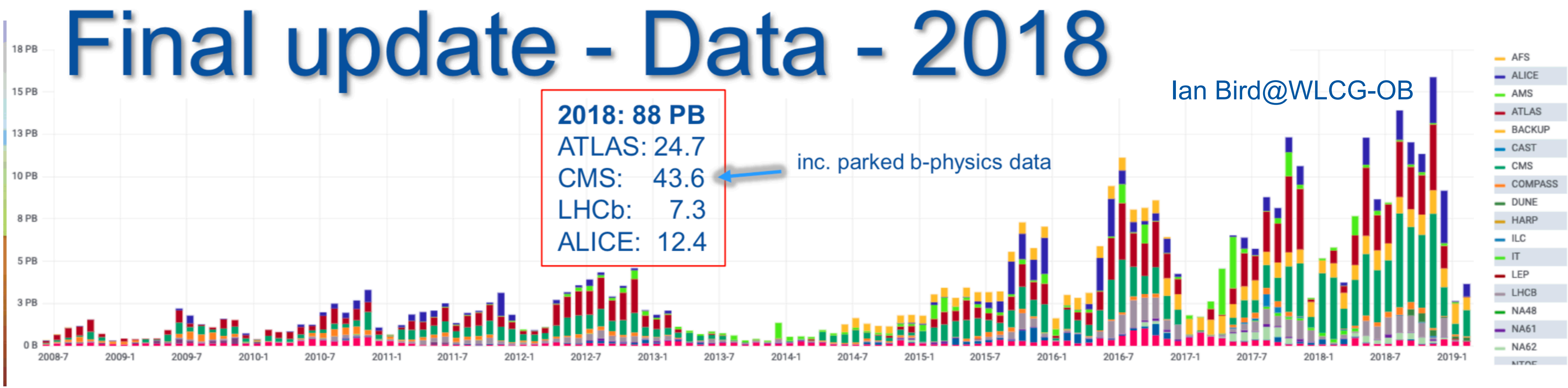
From the Hit to the Bit: event filtering

- L1: this is ~1 Petabyte per second!
  - Cannot afford to store it
  - 1 year's worth of LHC data at 1 PB/s would cost few hundred trillion euros
- L2: Real time to keep only "interesting" data
  - We keep ~1 event in a million
  - Yes, 99.9999% is thrown away
- EF: Final rate is O(Gigabyte per second)*



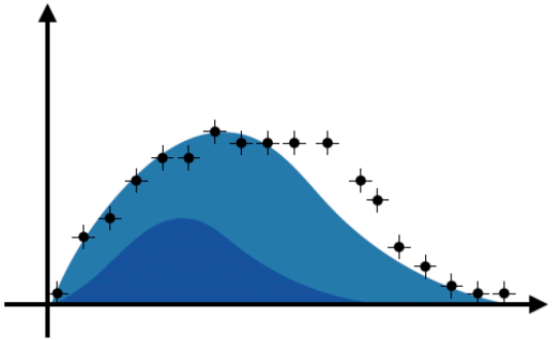
Data Processing

- Experiments recorded 88 Petabytes of data in 2018
  - 15.8PB in a single month (November)
- The LHC data is aggregated at the CERN data centre to be stored, processed and distributed

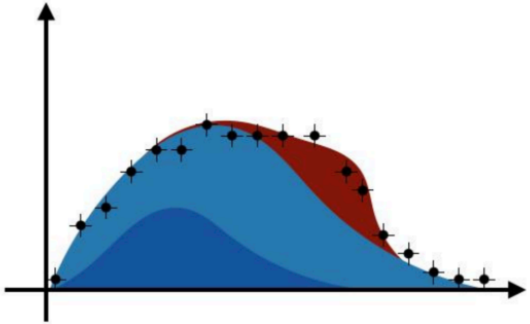


Discovery!

Baseline only

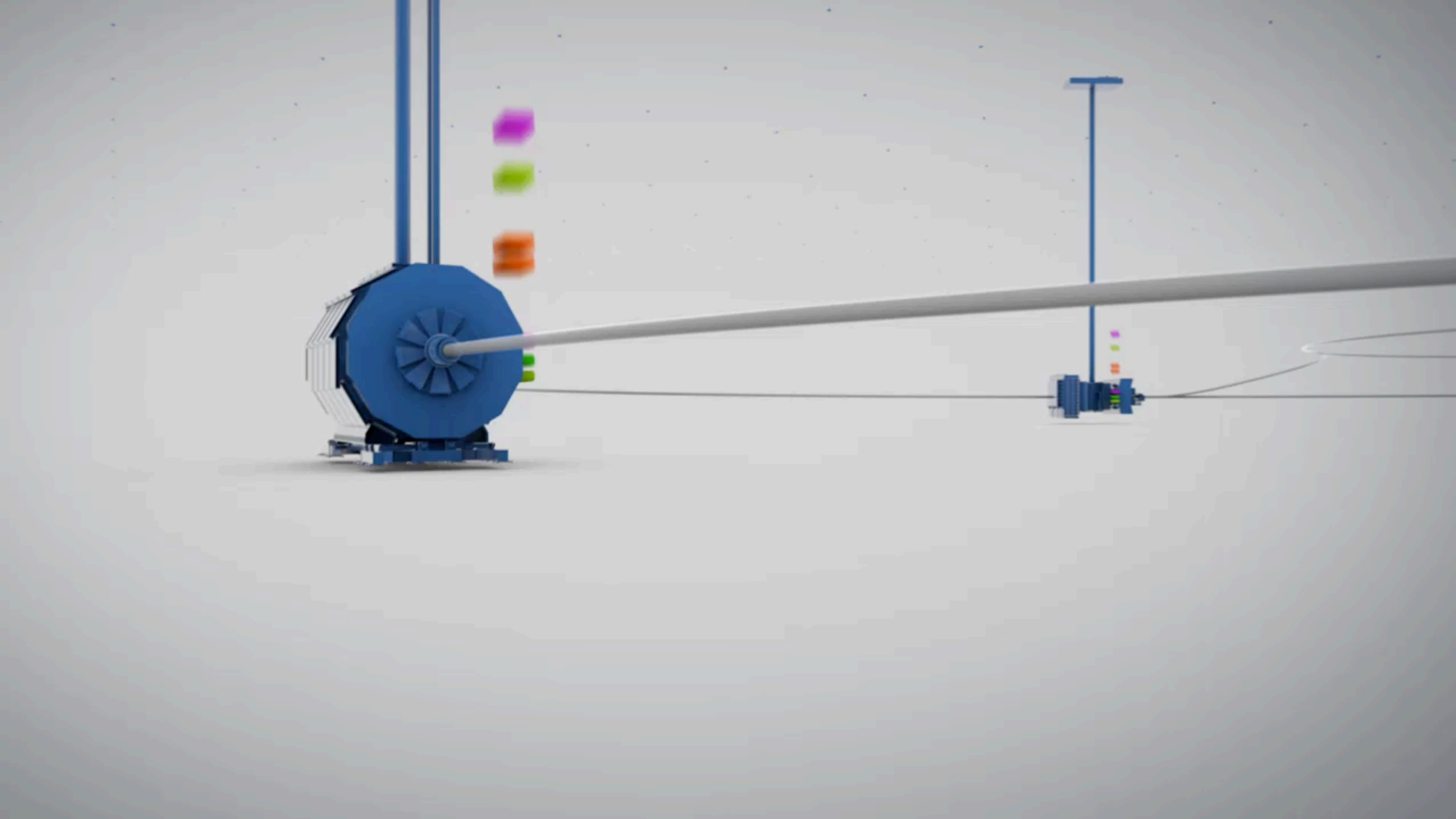


Baseline + New Physics



Baseline cannot describe data
… but baseline + new physics theory does -> Discovery!

Heinrich, Rocha @Kubecon and @CERN-ITTF
Keynote: Re-performing a Nobel Prize Discovery on Kubernetes
https://www.youtube.com/watch?v=CTfp2woVEkA (14:30)

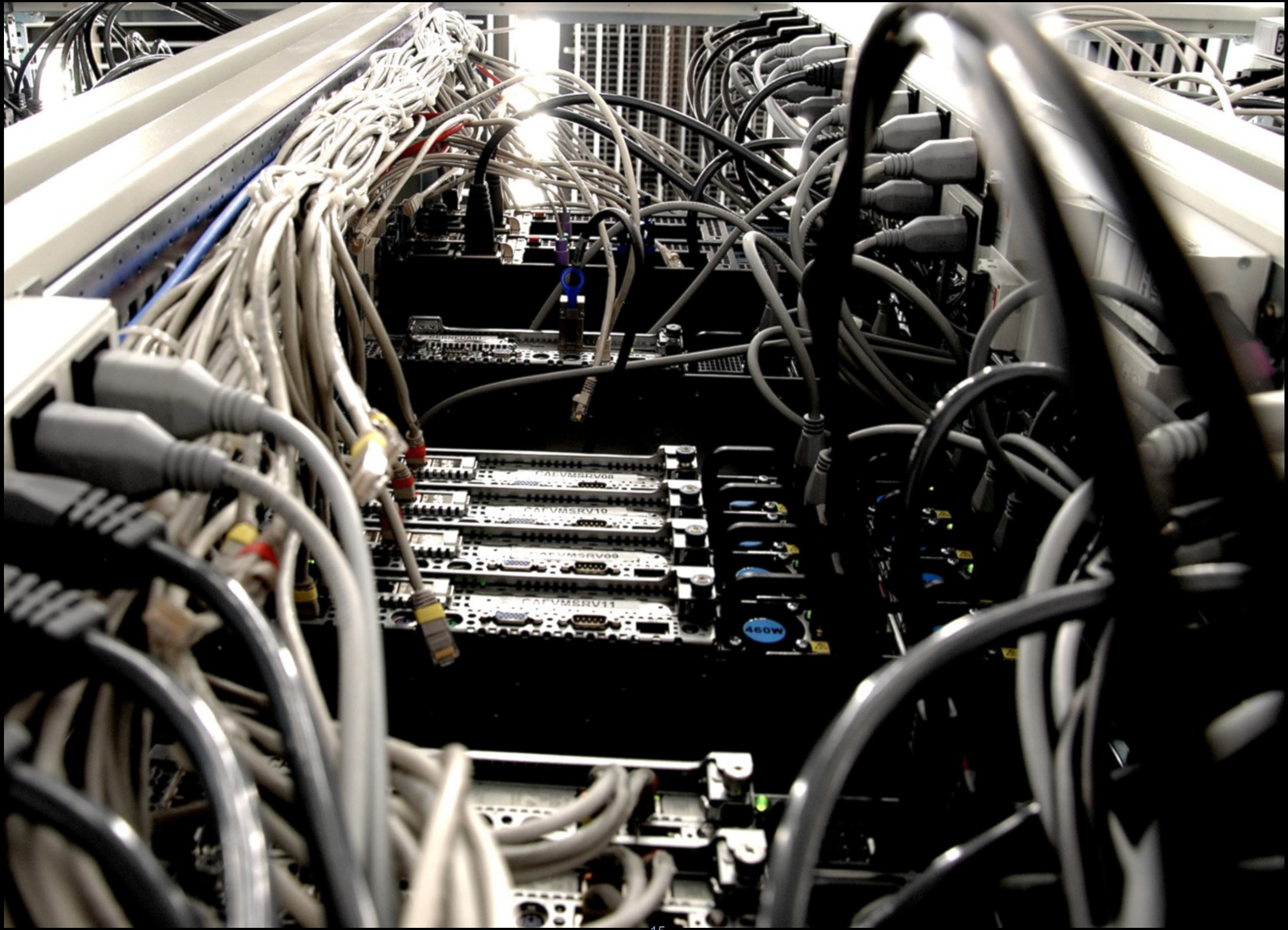


Take-away #1

- LHC data rates range from the PB/sec to the GB/sec after filtering
- 90PB of LHC data in 2018 (15.8PB in Nov only)
- 1EB data transferred world-wide
- Scientific data entering Exabyte scale:
  - 1EB = 1.000PB = 1.000.000TB = 1.000.000.000 GB

(1TB = your computer) (10GB = your smartphone)

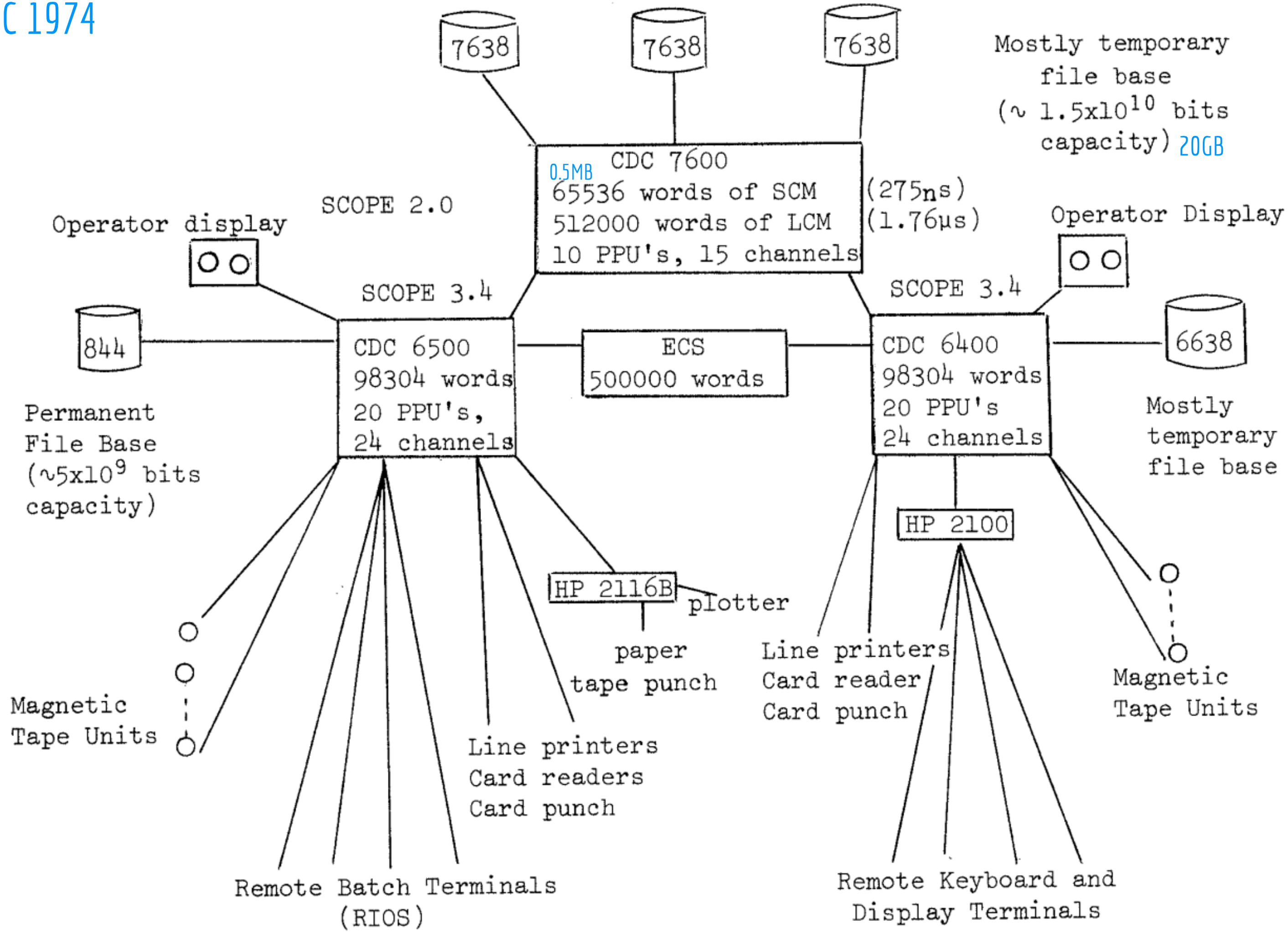
CERN Data Center



CERN Data Center

- Built in the 70s on the CERN site (Meyrin,Geneva)
  - 3.5 MW for equipment
- Extension located at Wigner (Budapest)
  - 2.7 MW for equipment
  - Connected to the Geneva CC with 3x100Gbps links (21 and 24 ms RTT)
- Hardware generally based on commodity
  - ~**15,000** servers, providing **230,000** processor cores
  - ~**90,000** disk drives providing **280PB** disk space
  - ~**30,000** tapes drives, providing **0.4EB** capacity (1EB=1000PB)

CERN-DC 1974

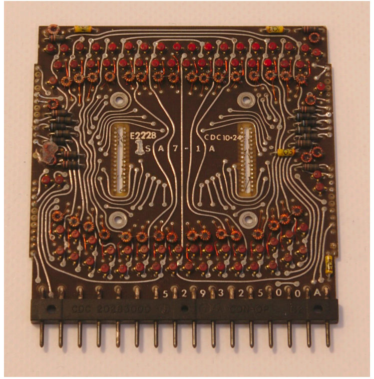


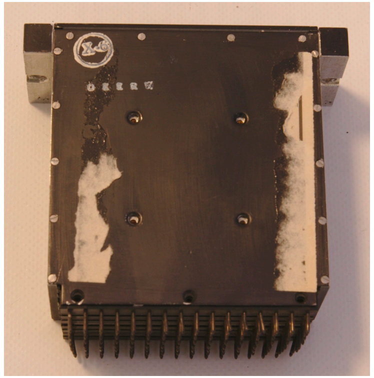
7638
7638
7638
Mostly temporary file base (∿ 1.5x10^10 bits capacity) 20GB
0.5MB
CDC 7600
65536 words of SCM (275ns)
512000 words of LCM (1.76μs)
10 PPU's, 15 channels
SCOPE 2.0
Operator display
SCOPE 3.4
844
Permanent File Base (∿5x10^9 bits capacity)
CDC 6500
98304 words
20 PPU's,
24 channels
ECS
500000 words
Operator Display
SCOPE 3.4
CDC 6400
98304 words
20 PPU's
24 channels
6638
Mostly temporary file base
HP 2116B
plotter
paper tape punch
HP 2100
Line printers
Card reader
Card punch
Magnetic Tape Units
Magnetic Tape Units
Line printers
Card readers
Card punch
Remote Batch Terminals (RIOS)
Remote Keyboard and Display Terminals

CDC7600 SUPERCOMPUTER (1972)

60-bit word size and 36MHz processor

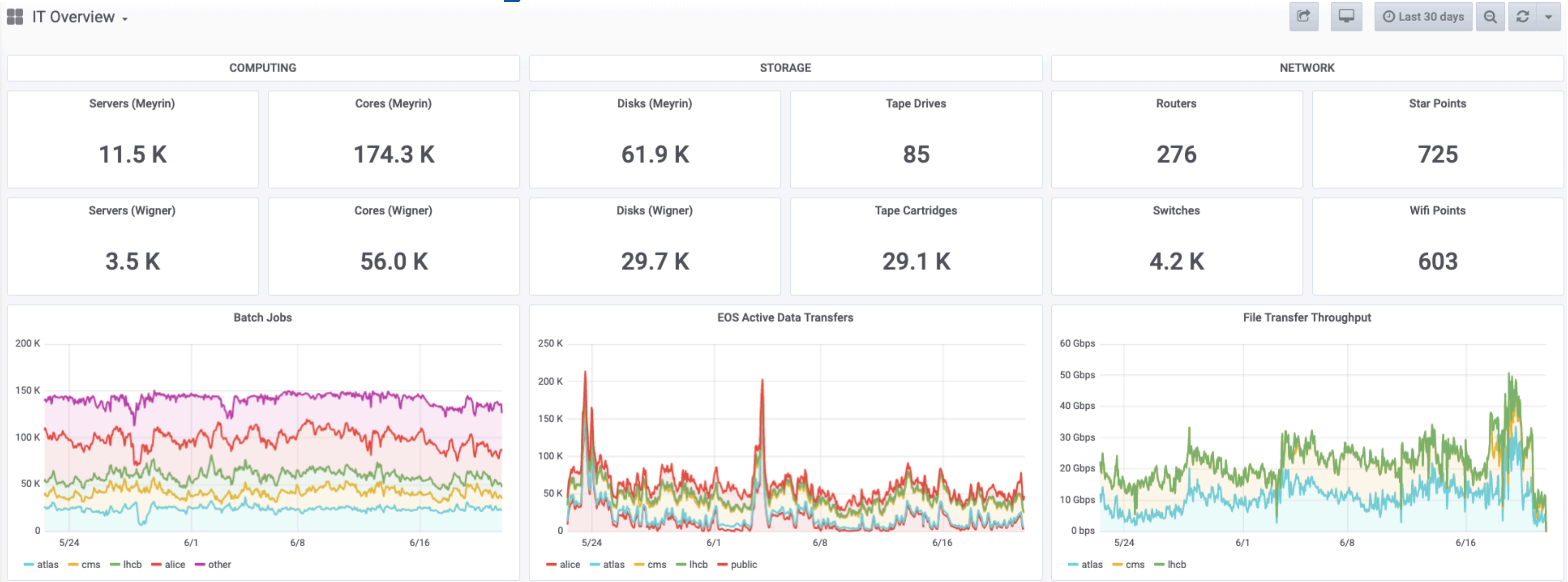
The computing centre was enlarged with the installation of a CDC 7600, which is shown here during its assembly in February 1972.





http://cerncourier.com/cws/article/cnl/24597
https://en.wikipedia.org/wiki/CDC_7600
https://videos.cern.ch/record/43172
https://videos.cern.ch/record/43113

An ordinary week at the CERN DC



Meyrin

Metric	Avg
Total Memory (TB)	1038
Disk Space (TB)	150135

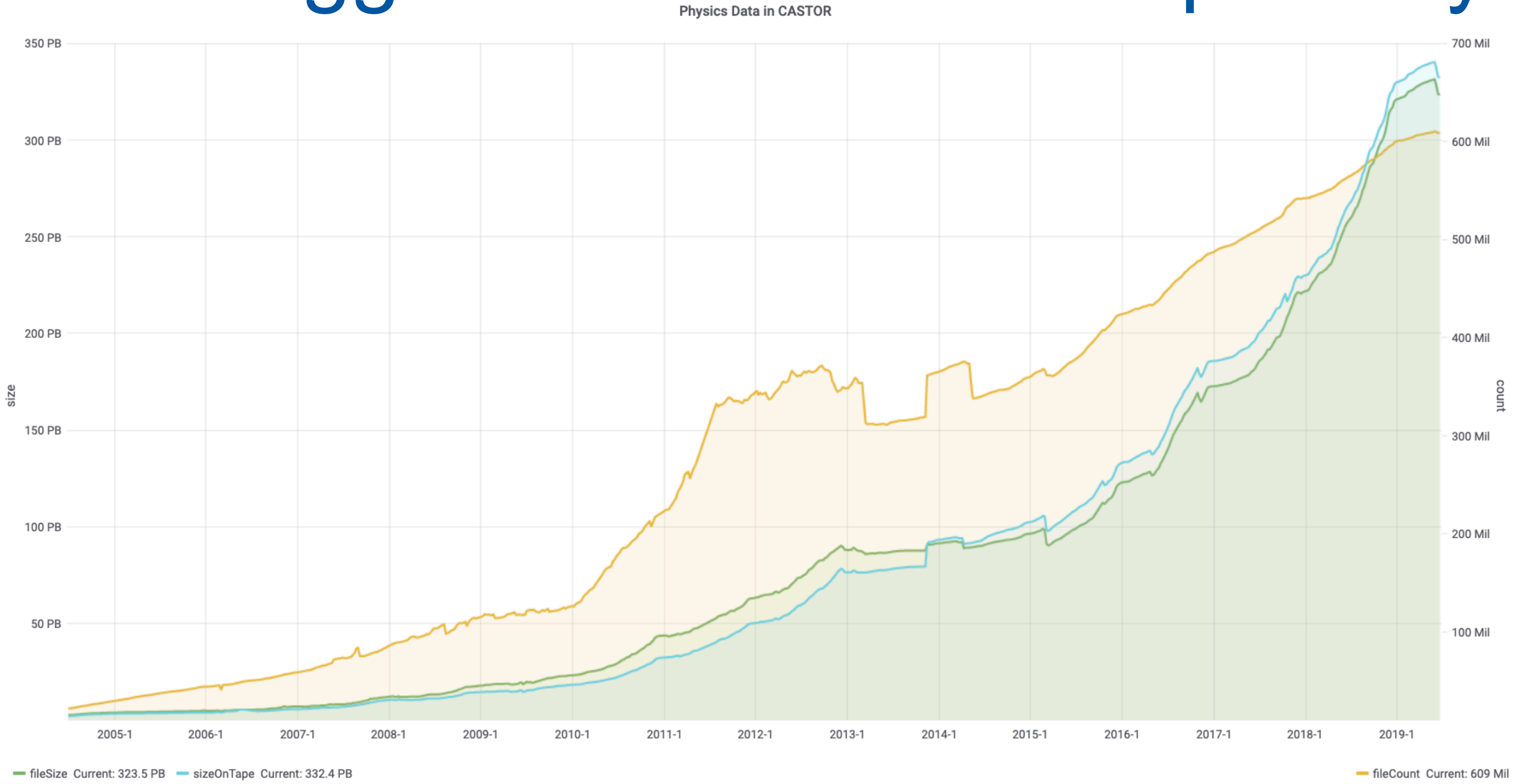
Wigner

Metric	Avg
Disk Space (TB)	97252
Total Memory (TB)	221

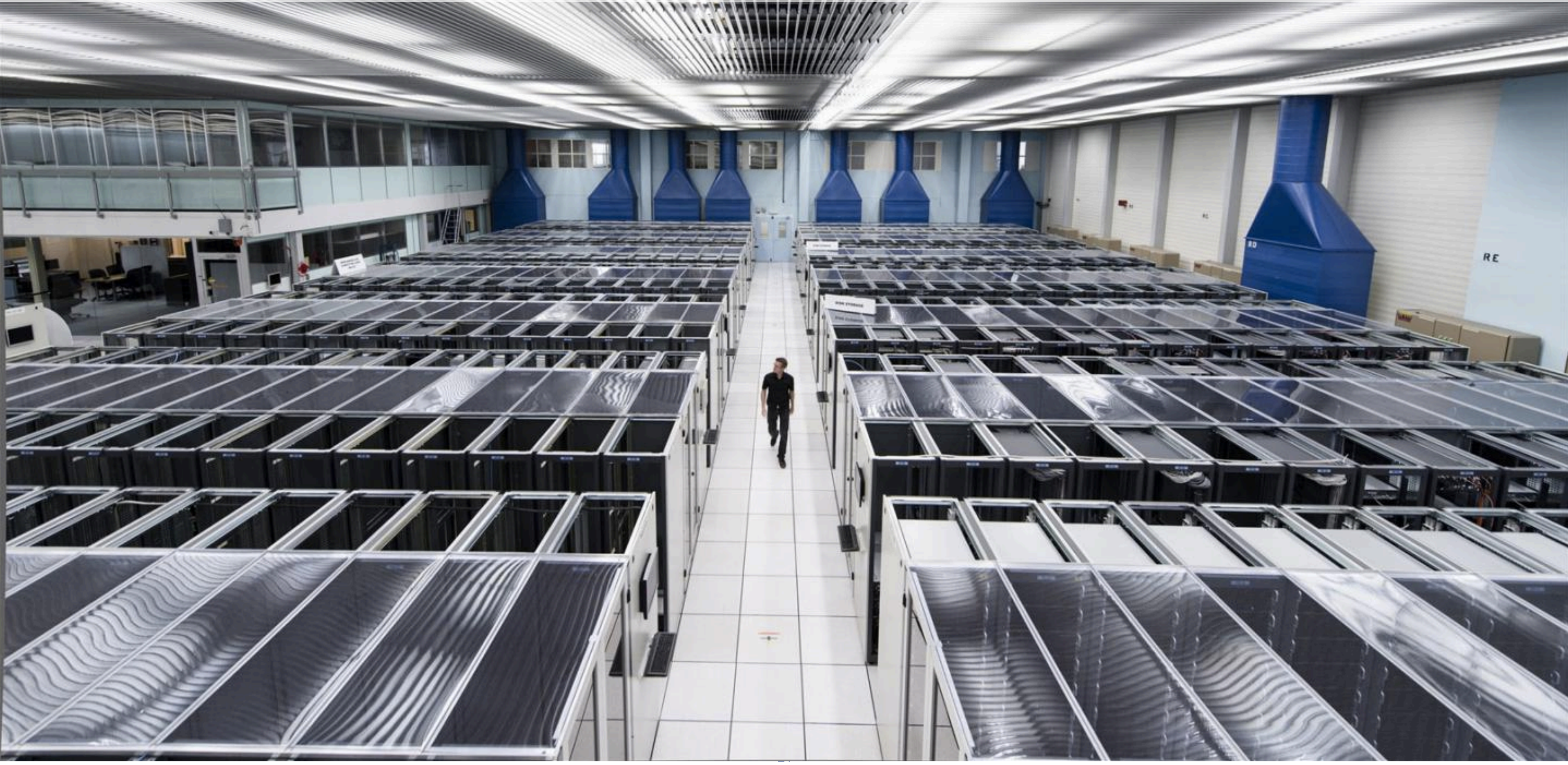
Tape Storage

Metric	
Drives	85
Cartridges	29134

World biggest scientific data repository



CERN Data Center

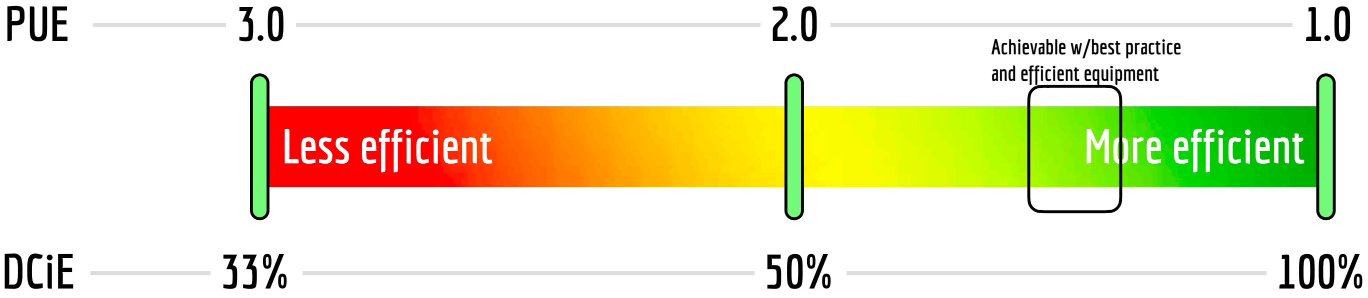


CERN Data Center

Shepherds Flat Wind Farm, second largest wind farm at 845 MW

Green IT

$$\text{PUE} = \text{Power Usage Effectiveness} = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$$



$$\text{DCiE} = \text{Data Center Infr. Efficiency} = \frac{1}{\text{PUE}}$$

The Arctic World Archive: 300 meters below the ground

https://news.microsoft.com/features/under-the-sea-microsoft-tests-a-datacenter-thats-quick-to-deploy-could-provide-internet-connectivity-for-years/

June Council - DG presentation to personnel / Conseil de juin - présentation de la DG au personnel

View mode: Slides Camera **Camera & Slides** PIP

4th July 2019

Example of new expenses: new Computing Centre in Prévessin (PCC)

This MTP provides resources (~ 20 M) for a new Computing Centre in Prévessin → needed to fulfil Tier-0 obligations for end of Run 3 (needed resources: ~ 1.5 x 2018) and HL-LHC (~3 x 2018)

Currently:
- ❑ Computing Centre in Meyrin: 2.9 MW for computing equipment
  ~ No room for expansions (lack of space, inefficient cooling)
- ❑ Wigner (Budapest): 1 MW
  Contract terminates end 2019; 4 M/year operating costs (facility+network) → very fruitful partnership!

→ PCC designed for high Power-Usage Effectiveness (ratio of total energy used by centre to energy used for computing equipment) → lights-out facility minimizing energy losses, no office space, efficient cooling → PUE: ~ 1.1 (compared to 1.5 for Meyrin centre) → cost-effective
-- 4 MW upgradable to 12 MW → will centralize all future computing needs at CERN
-- turnkey building from specialized company (à la Green Cube in GSI)

Construction 2021-2022
→ until then: use 1 MW spare capacity in new LHCb containers at Point 8

Information

Speaker: Fabiola Gianotti

Room: 500/1-001

Event link: indico

stract Share

Take-away #2

- Power and Heat management: PUE and Green-IT
- Data centers run on commodity hardware
- CERN remains largest scientific repository in the world
- Commercial cloud and personal computing providers getting larger and large: Amazon, Microsoft, Google, Dropbox

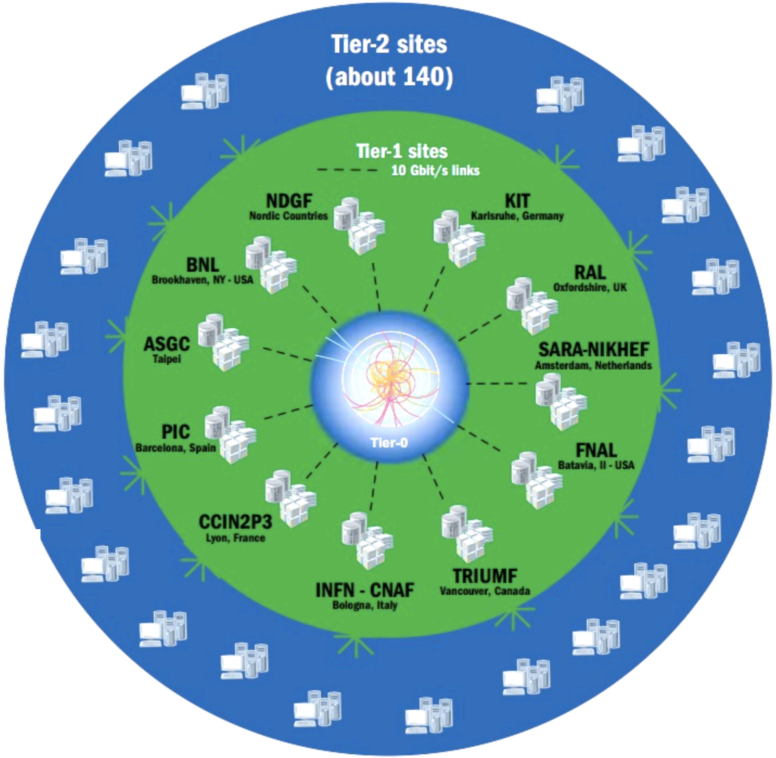
CERN IT Department

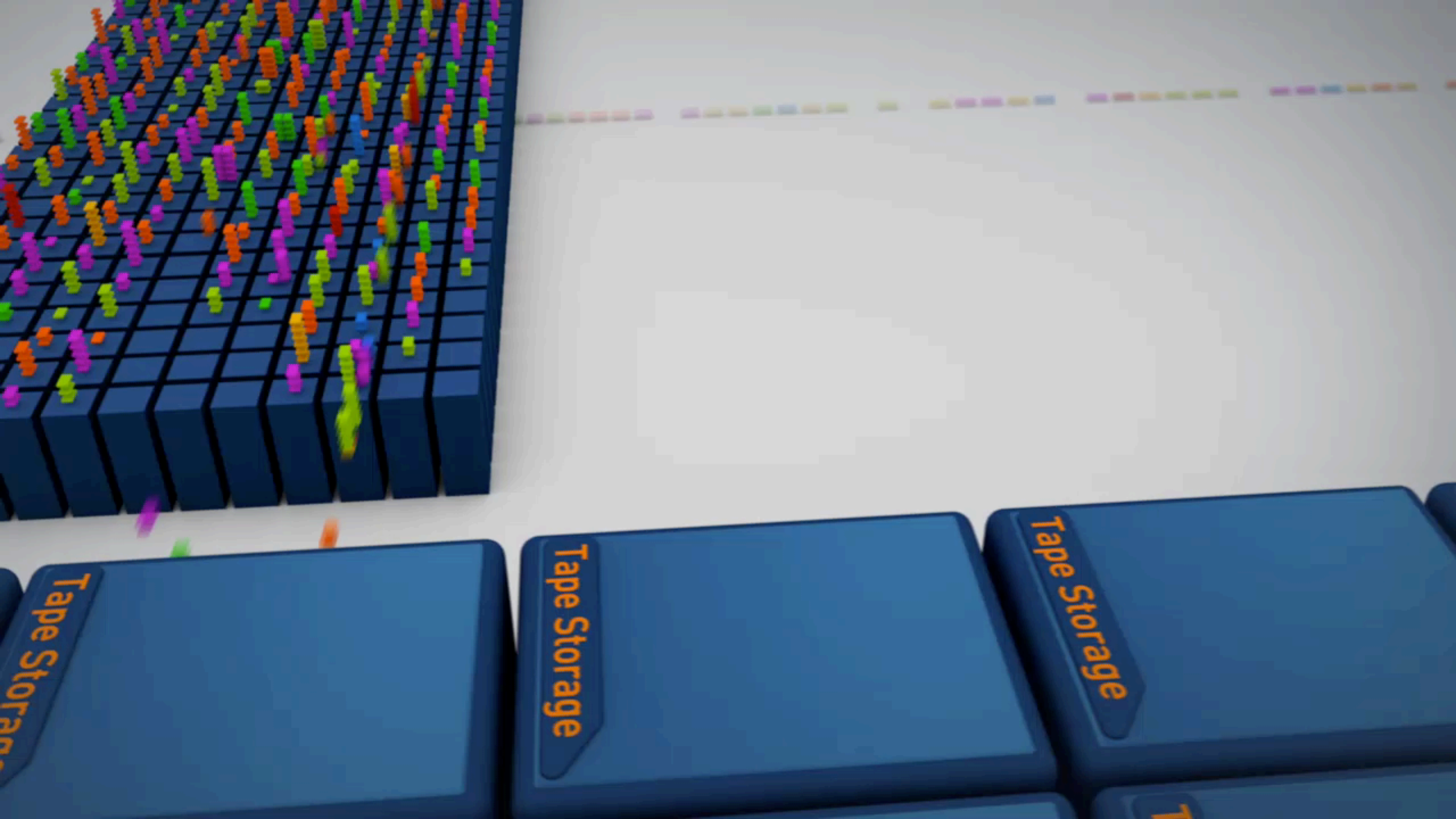
CERN IT Department

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
- The resources are distributed – for funding and sociological reasons

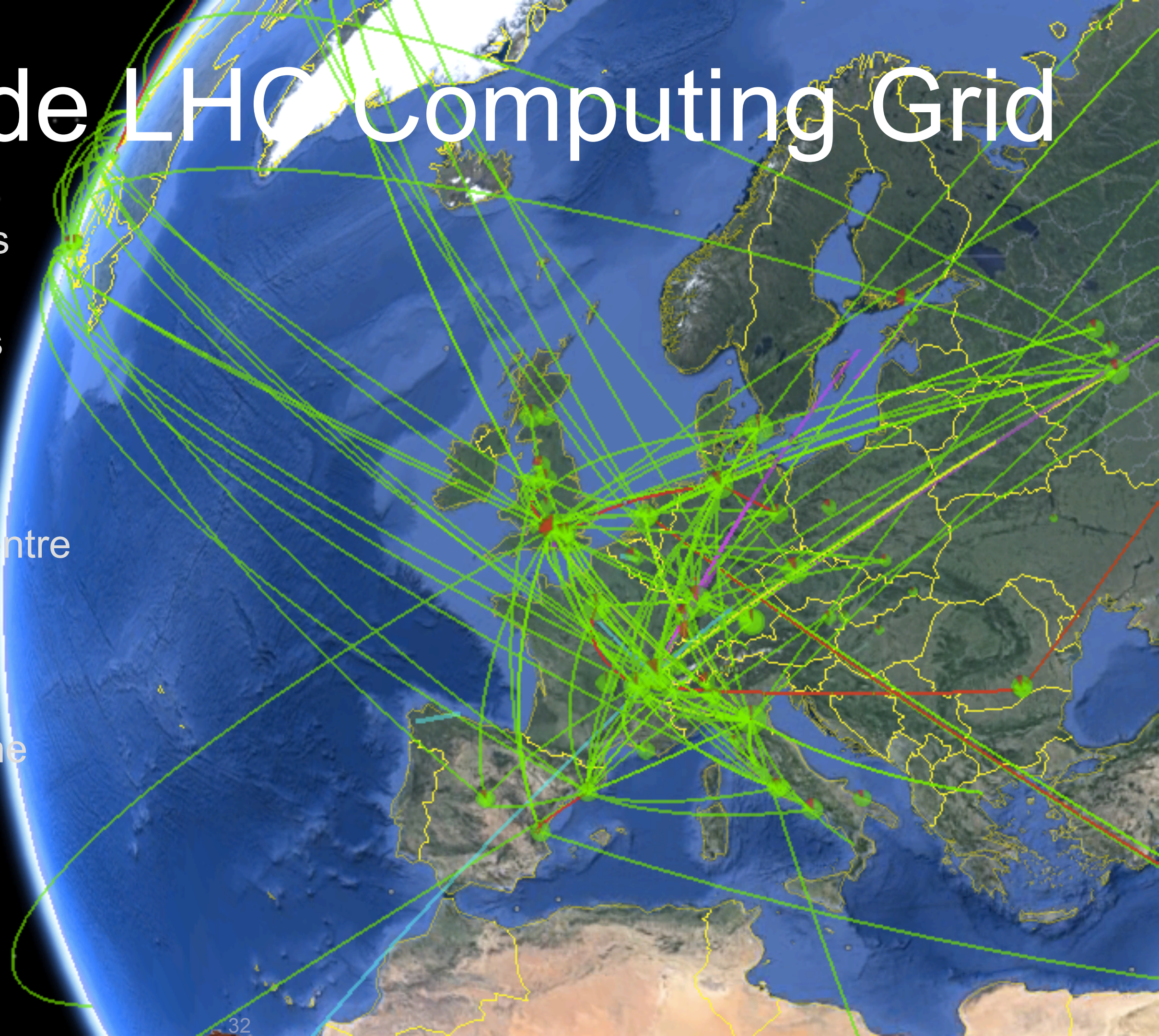
WLCG
Worldwide LHC Computing Grid

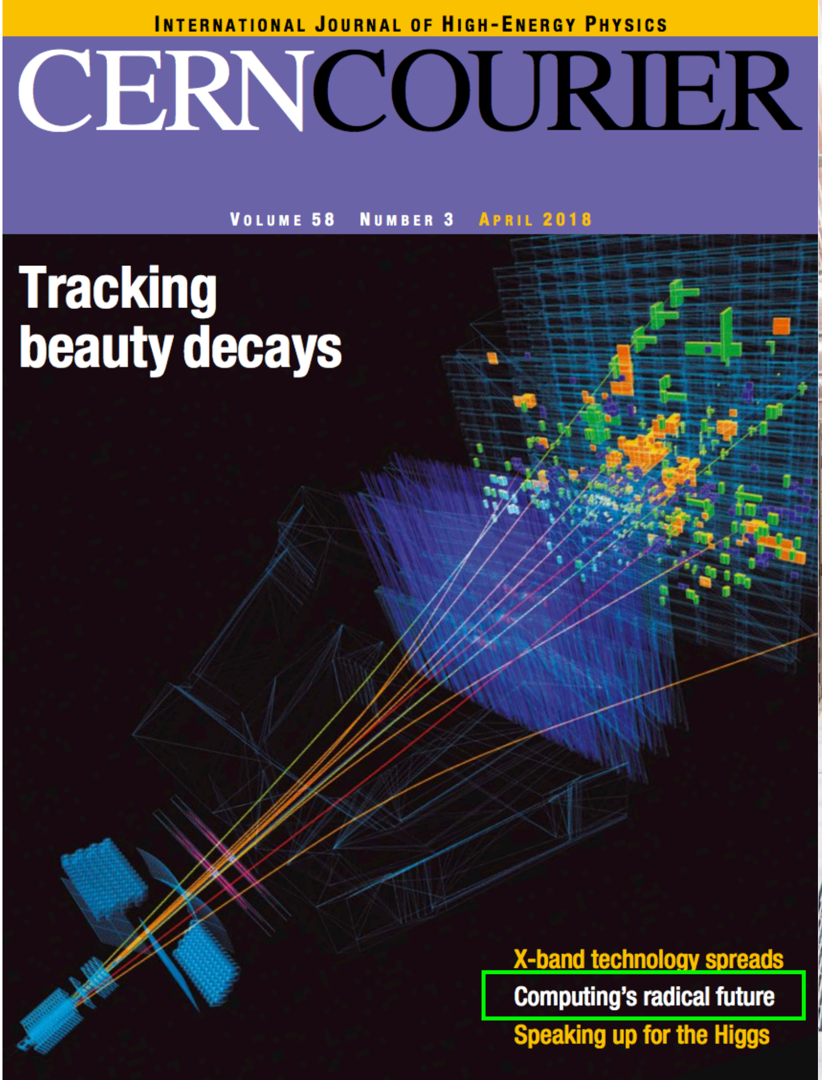


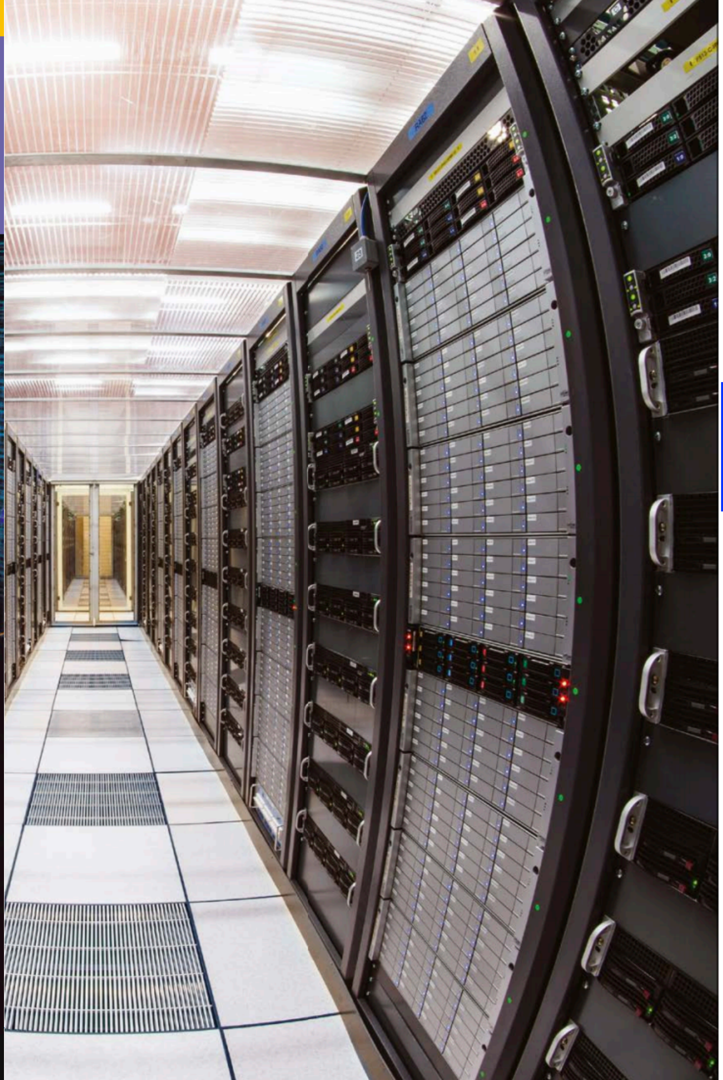

Tape Storage
Tape Storage

The Worldwide LHC Computing Grid

- A community of 10,000 physicists are the WLCG users
- On average around 250,000 jobs running concurrently
- 600,000 processing cores
- 15% of the WLCG computing resources are at CERN's data centre
- 500 petabytes storage available worldwide
- 20-40-80-100 Gbit/s optical fibre links connect CERN to each of the 13 Tier-1 institutes







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.

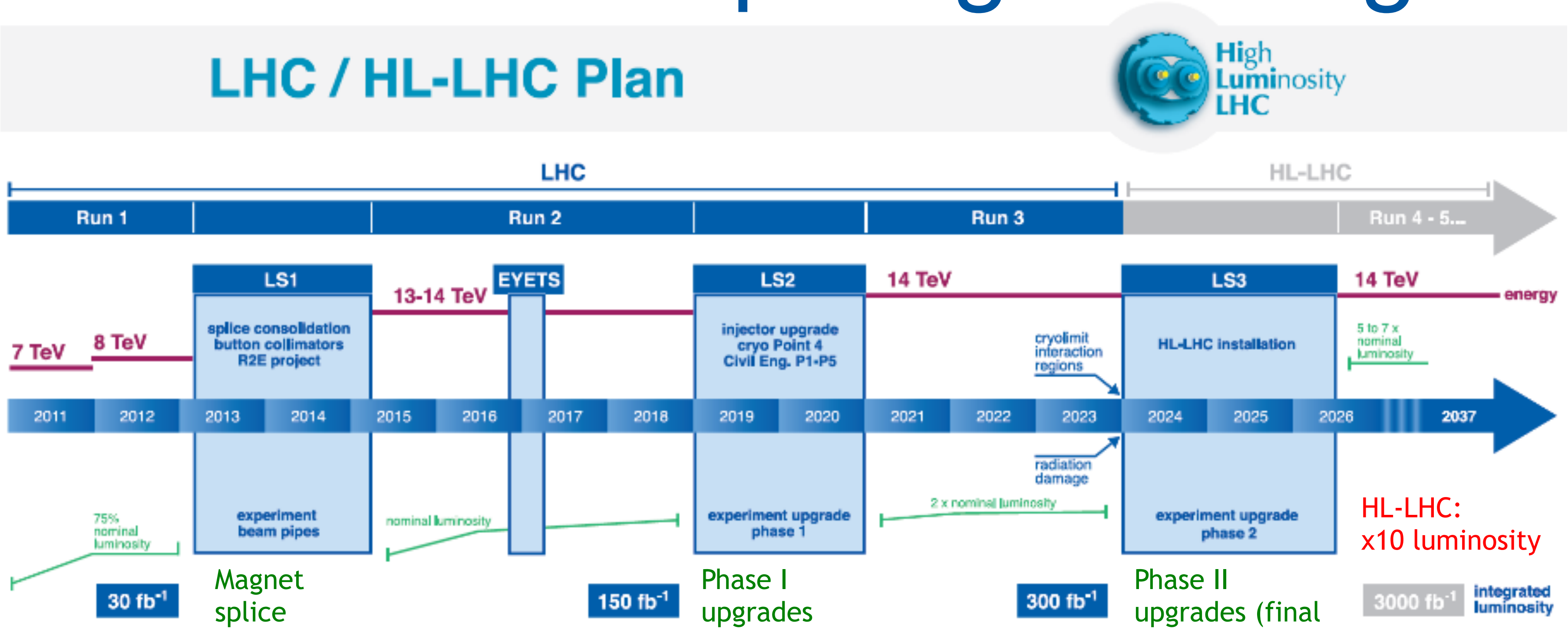
It would be impossible for anyone to conceive of carrying out a particle-physics experiment today without the use of computers and software. Since the 1960s, high-energy physicists have pioneered the use of computers for data acquisition, simulation and analysis. This hasn't just accelerated progress in the field, but driven computing technology generally – from the development of the World Wide Web at CERN to the massive distributed resources of the Worldwide LHC Computing Grid (WLCG) that supports the LHC experiments. For many years these developments and the increasing complexity of data analysis rode a wave of hardware improvements that saw computers get faster every year. However, those blissful days of relying on Moore's law are now well behind us (see panel overleaf), and this has major ramifications for our field.

The high-luminosity upgrade of the LHC (HL-LHC), due to enter operation in the mid-2020s, will push the frontiers of accelerator and detector technology, bringing enormous challenges to software and computing (*CERN Courier* October 2017 p5). The scale of the HL-LHC data challenge is staggering: the machine will collect almost 25 times more data than the LHC has produced up to now, and the total LHC dataset (which already stands at almost 1 exabyte) will grow many times larger. If the LHC's ATLAS and CMS experiments project their current computing models to Run 4 of the LHC in 2026, the CPU and disk space required will jump by between a factor of 20 to 40 (figures 1 and 2).

Even with optimistic projections of technological improvements there would be a huge shortfall in computing resources. The WLCG hardware budget is already around 100 million Swiss francs per year and, given the changing nature of computing hardware and slowing technological gains, it is out of the question to simply throw

Inside the CERN computer centre in 2017. (Image credit: J Ordan/CERN.)

HL-LHC: a computing challenge

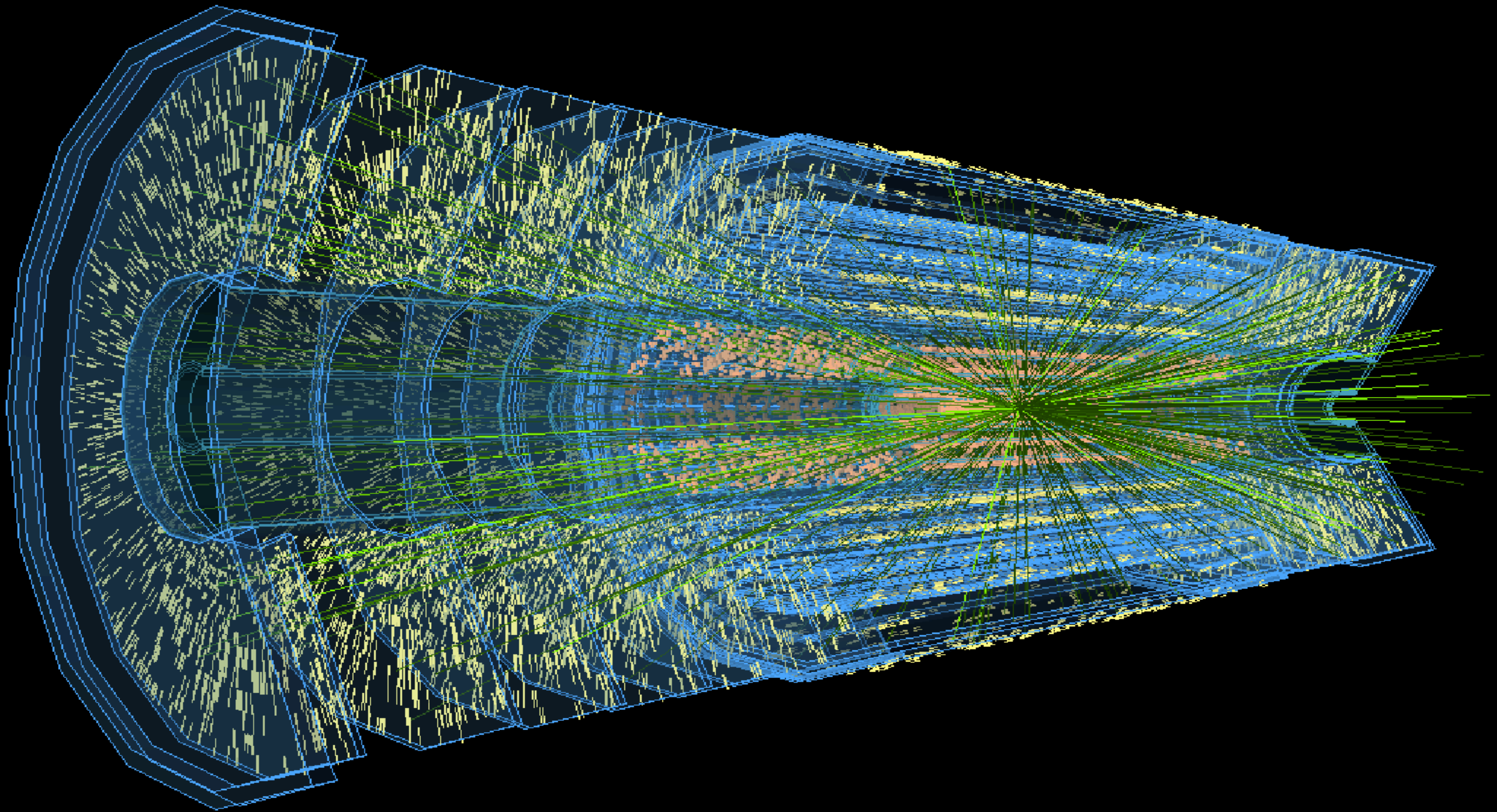


Magnet splice update

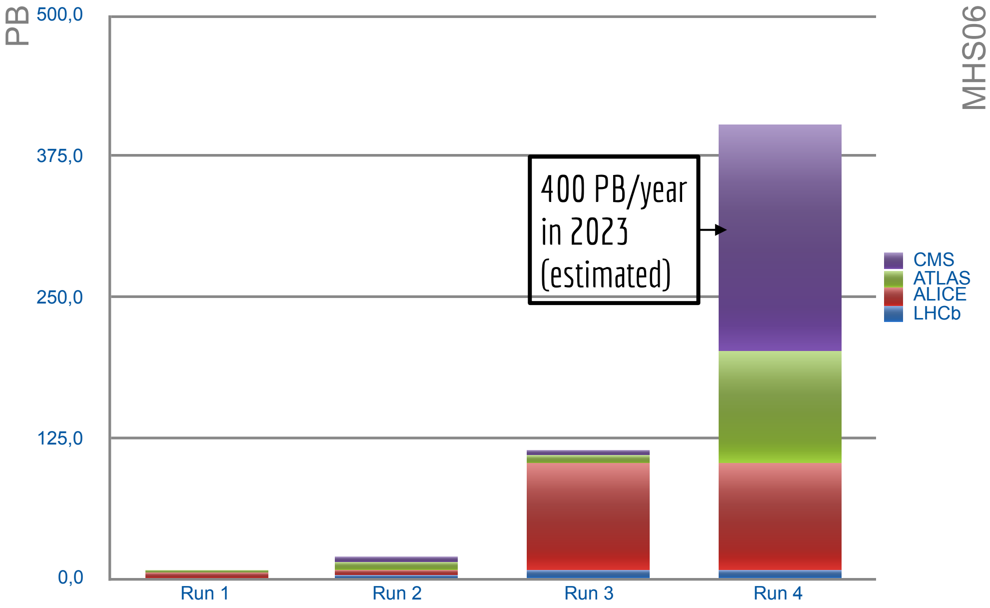
Phase I upgrades (injectors)

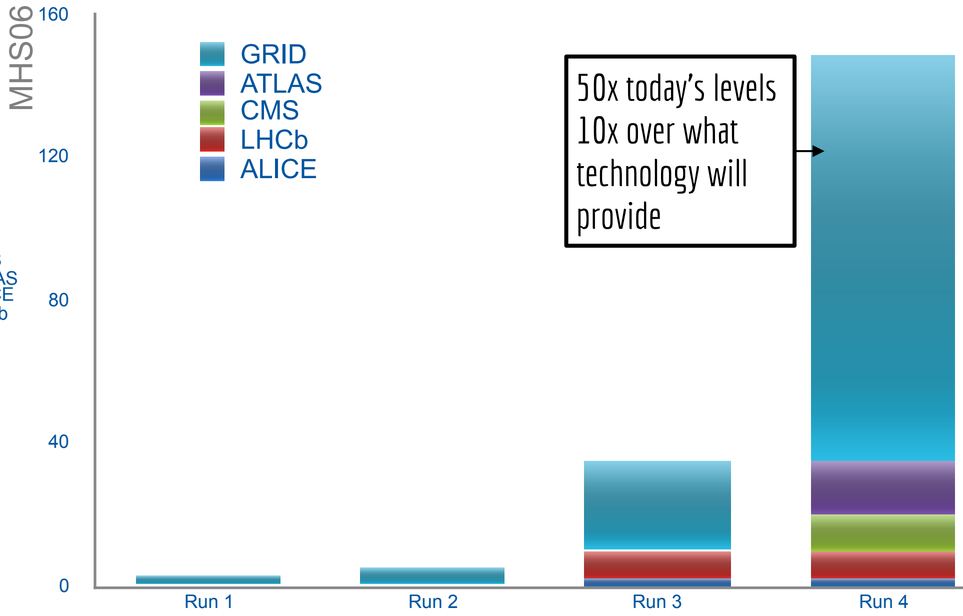
Phase II upgrades (final focus)

HL-LHC: x10 luminosity



HL-LHC: a computing challenge





Take-away #3

- WLCG defined new ways to federate scientific computing, instrumental for the past LHC Run-I&II and for the upcoming Run-III&IV
  - Formed by 150+ sites working together around the world
  - Used by the 4 main LHC experiments, its concepts and tools are **being adopted** by non-LHC and **non-HEP** experiments
- New challenges for High Luminosity LHC need to be addressed
  - New paradigms, new scenarios. It is time for R&D!

Future in Computing (1/2)

- Manage storage and computing continuous growth (budget and infrastructures)
  - Storage technologies evolution: HDD, SSD, Tapes
  - CPU speed and multicore/vector exploitation
  - Data Center engineering: optimise energy consumption (PUE and green IT)

- Provide to the different experiments and its scientists the required computing infrastructure while optimising resources
  - Worldwide LHC Computing Grid (WLCG)
  - ESCAPE EU project

Future in Computing (2/2)

- Improve software performance
  - HEP (High Energy Physics) Software Foundation, **HSF**
- Make use of a different type of resources: HPC (GPUs) and Cloud providers
- Data preservation:
  - How to ensure that all the data collected and published is still readable by the next generations
- CERN is leading a global effort for HEP and science, that others will inevitably face soon or later

CERN-IT: pushing boundaries

- CERN-IT impact on society through computing:
  - Need for collaboration tools for Global Science led to invent the **World Wide Web**
  - Need for collaboration of computing resources for the Global LHC led to adopt **Grid Computing** and first concept of **Computing Clouds**
- Open access to science
  - Need for sharing the results had led CERN to pave to way to open access to documents and now data: **LHC@home** and **CERN Opendata Portal**
- Openlab
  - “CERN openlab is a unique public- private partnership that accelerates the development of cutting-edge solutions for the worldwide LHC community and wider scientific research”
    - Testing software and hardware
    - Large student internship programme
- EC projects:
  - ESCAPE, Archiver, HNSciCloud (cloud computing resources)

From CERN to the world

- Fundamental Science push boundaries and the revenues are immense.
- In computing CERN R&D lead for instance to:
  - Invention of the Web (1990)
    - Key contribution on the INTERNET infrastructure (80% of the total european capacity in 1991)
  - Touch screens (1972)
    - Super Proton Synchrotron control system required complex controls and developed capacitive touch screen
    - It was based on open standards and moved into industry

World Wide Web

The WorldWideWeb (W3) is a wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an executive summary of the project, Mailing lists , Policy , November's W3 news , Frequently Asked Questions .

What's out there?
: Pointers to the world's online information, subjects , W3 servers, etc.

Help
: on the browser you are using

Software Products
: A list of W3 project components and their current state. (e.g. Line Mode ,X11 Viola , NeXTStep , Servers , Tools , Mail robot , Library)

Technical
: Details of protocols, formats, program internals etc

Bibliography
: Paper documentation on W3 and references.

People
: A list of some people involved in the project.

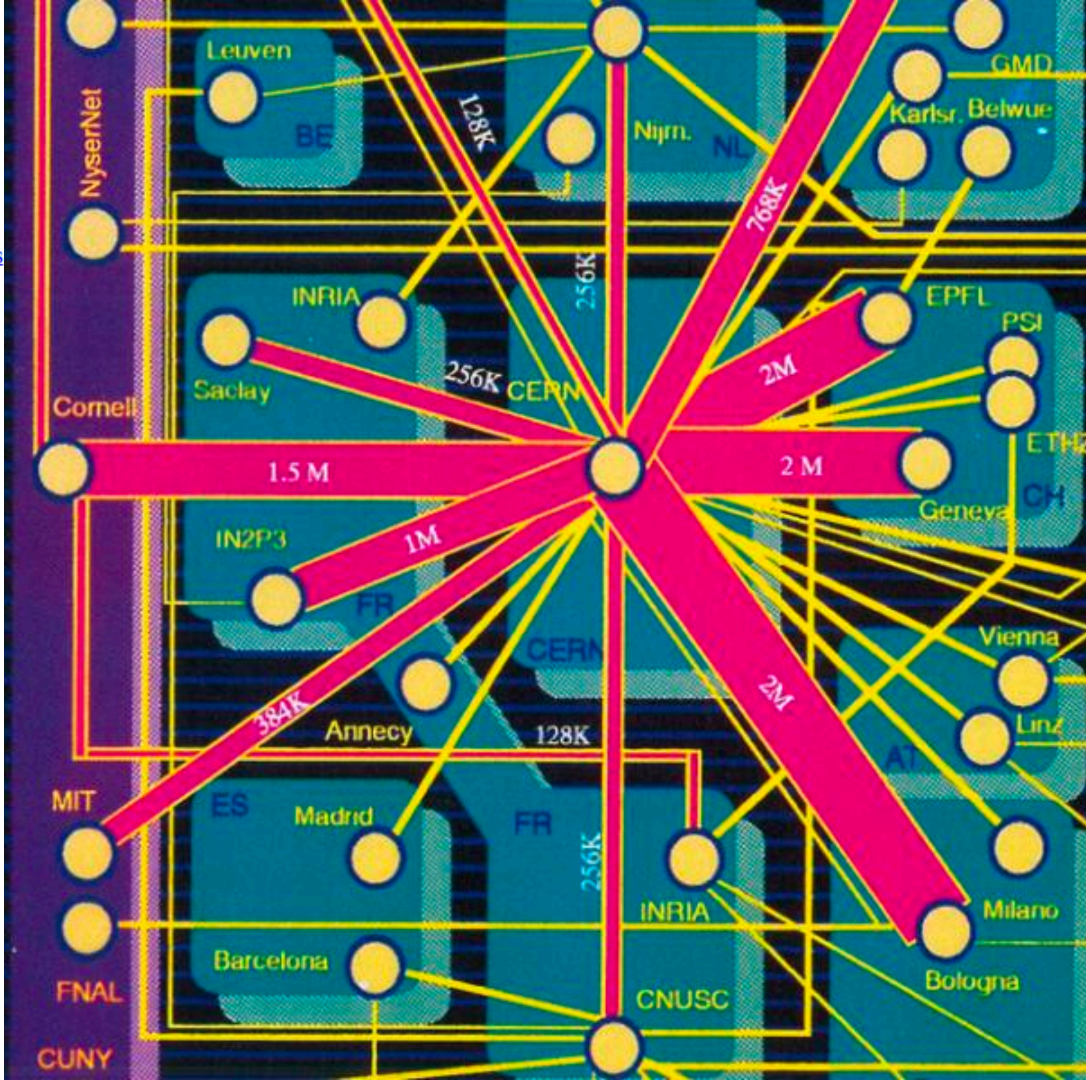
History
: A summary of the history of the project.

How can I help ?
: If you would like to support the web..

Getting code
: Getting the code by anonymous FTP , etc.

https://home.cern/topics/birth-web

https://home.cern/cern-people/opinion/2013/06/how-internet-came-cern



The SPS control room in 1977. Ted Wilson and Rae Stiening are at the desk, with its touch screens; Bent Stumpe and George Shering stand behind.

http://cerncourier.com/cws/article/cern/42092

Take-away #4

- **Fundamental** science continue to be main inspiration for **revolutionary** ideas, due to revolutionary needs

- Industry has well defined offer and demand. We do not. This is the key for **innovation**

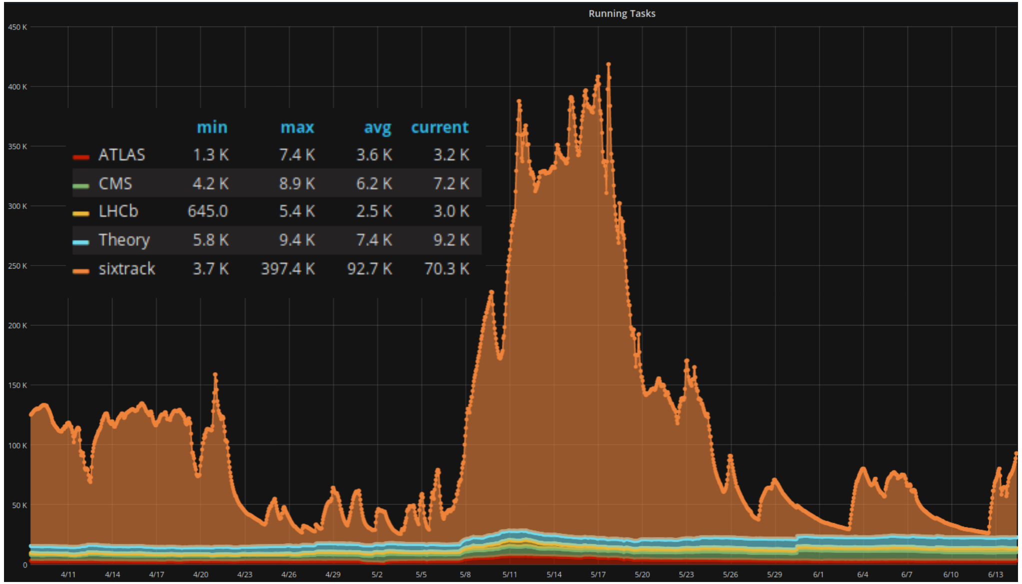
- …and **innovation** foster technological advancements that percolates to the society

Thanks!

- Simulations from ATLAS, CMS, LHCb and Theory running under CernVM and VirtualBox

- You can contribute by running BOINC on your computer outside working hours

- The BOINC client can be configured to run 17:30-8:30 or when your computer is idle



http://lhcathome.web.cern.ch/

opendata CERN

About

CERN OPENDATA

Explore more than **1 petabyte** of open data from particle physics!

Start typing... Search

search examples: collision datasets, keywords:education, energy:7TeV

Explore

datasets
software
environments
documentation

Focus on

ATLAS
ALICE
CMS
LHCb
OPERA