
CERN

Computación en el CERN

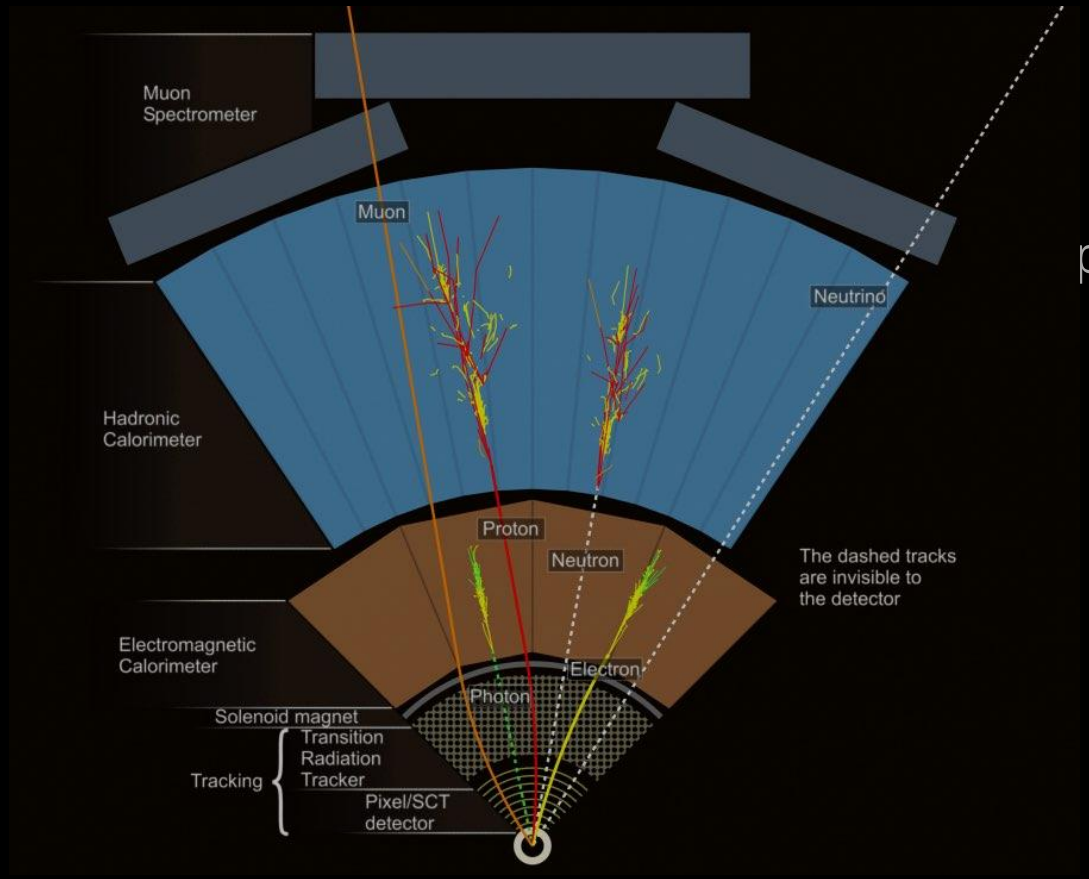
Maite Barroso Lopez
CERN-IT
(slides from Xavier Espinal y Alejandro Alvarez)


LHCb
ATLAS
CERN Meyrin
CERN Prévessin
SPS 7 km
SUISSE
FRANCE
CMS
ALICE
LHC 27 km

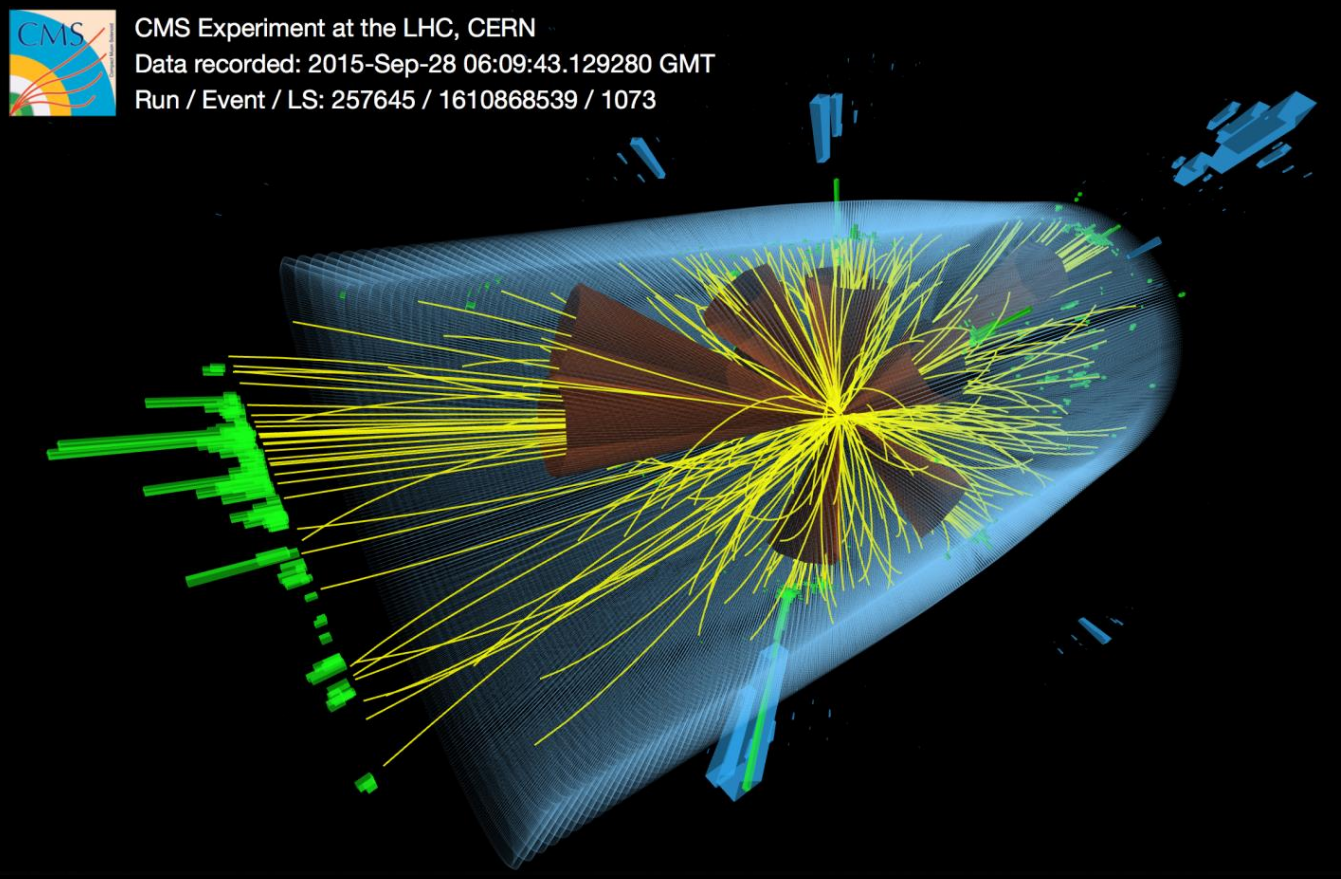
Adquisición de datos

- 40 millones de events (“imágenes”) por segundo
- ~1 Petabyte (1024 Terabytes) por segundo
  - Imposible de almacenar
  - Hace necesario filtrar, y conservar sólo datos “interesantes”

From the Hit to the Bit: DAQ

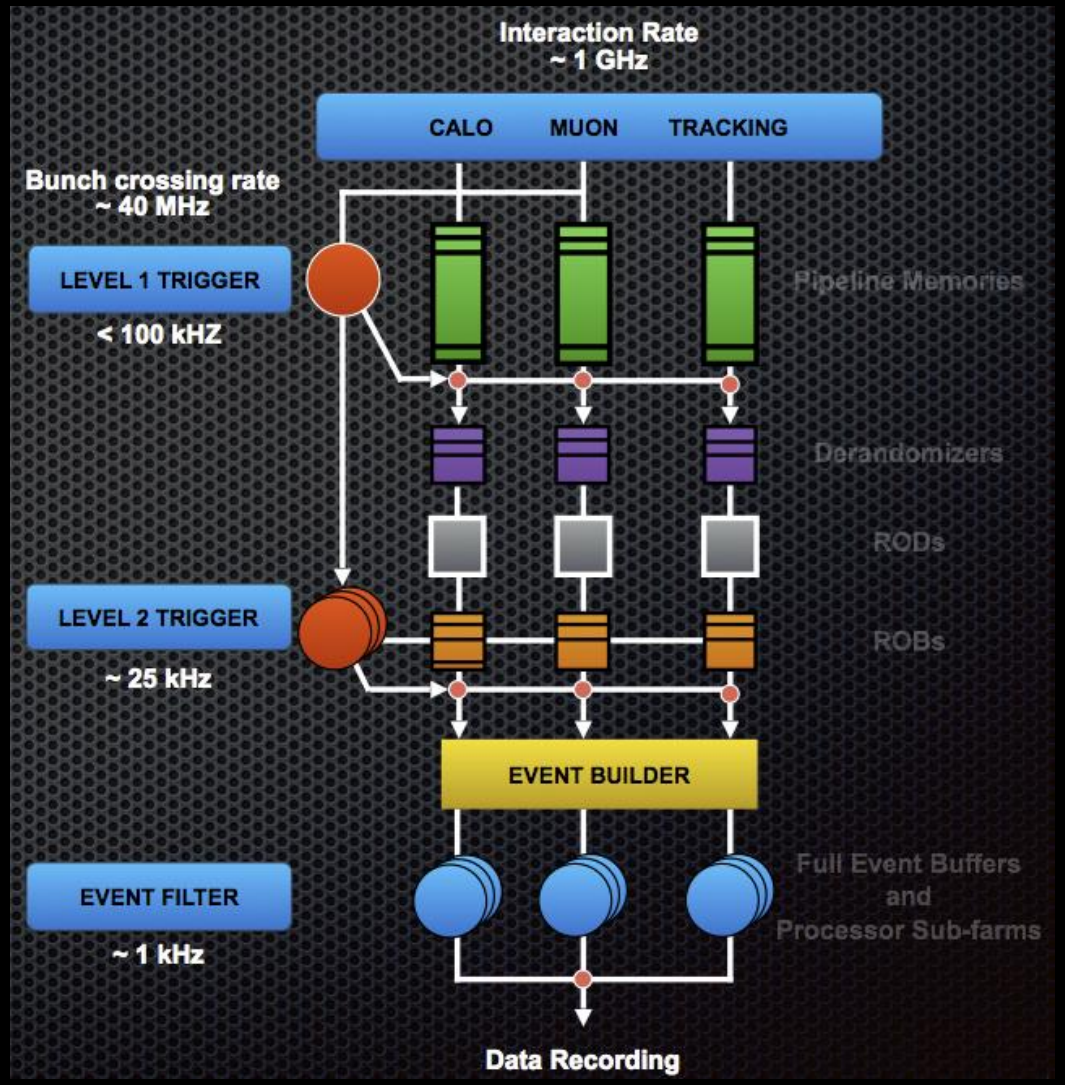


parts



From the Hit to the Bit: event filtering (1/2)

- 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

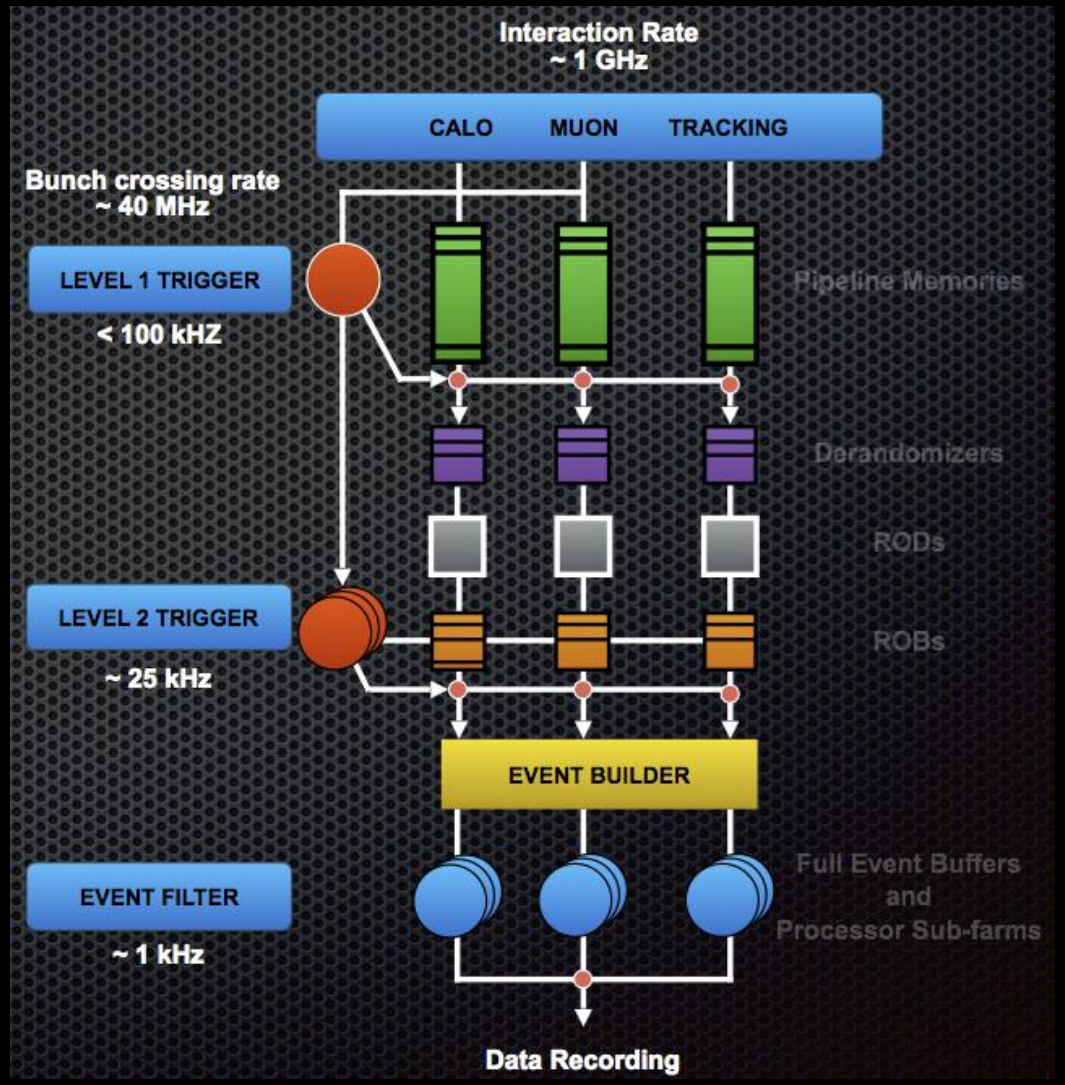


From the Hit to the Bit: event filtering (2/2)

- 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

- Have to filter in real time to keep only "interesting" data
  - We keep ~1 event in a million
  - Yes, 99.9999% is thrown away

- Final rate is O(Gigabyte per second)*

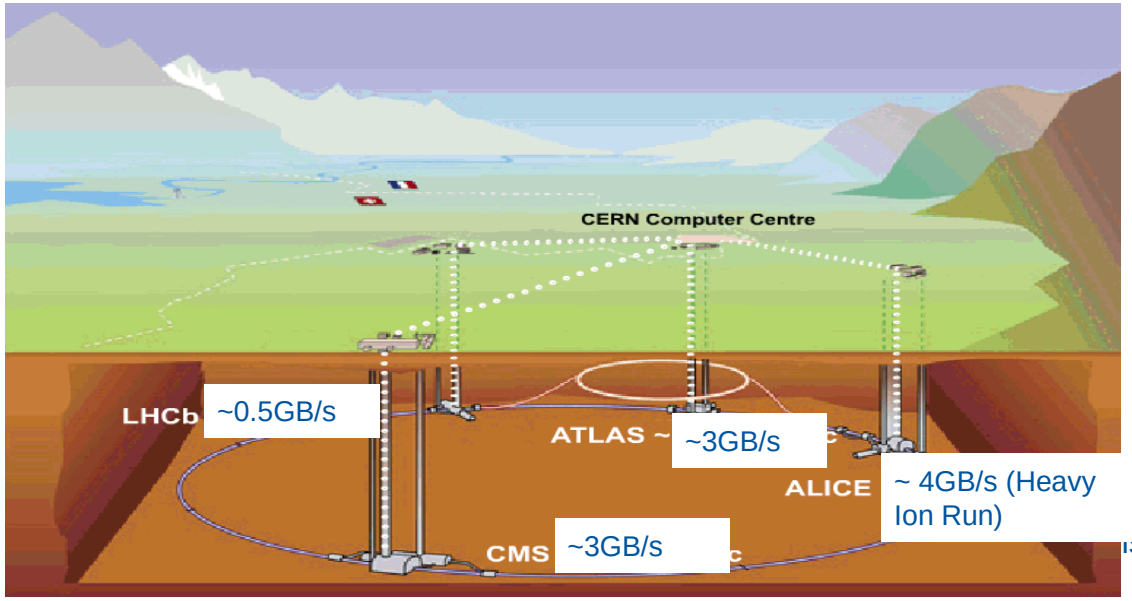


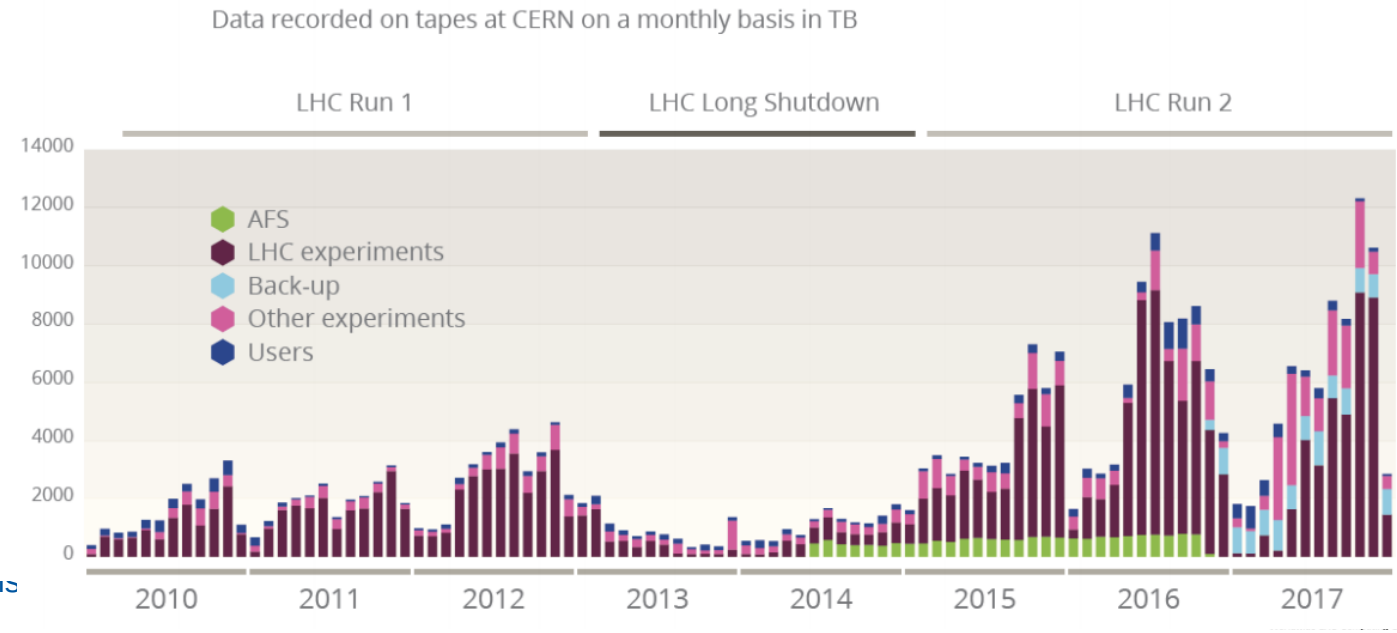
Adquisición de datos

- Al Data Center llega ~1 evento de cada millón
- ~1 Petabyte/segundo => ~1 Gigabyte/segundo
- Se descartan eventos “no interesantes”
  - Es necesario pre-establecer qué es “interesante”: Simulación

Data Processing

- Experiments sent 70 Petabytes of data in 2017 year
  - 40 Petabytes from the four LHC experiments
- The LHC data is aggregated at the CERN data centre to be stored and processed





CERN Data Center

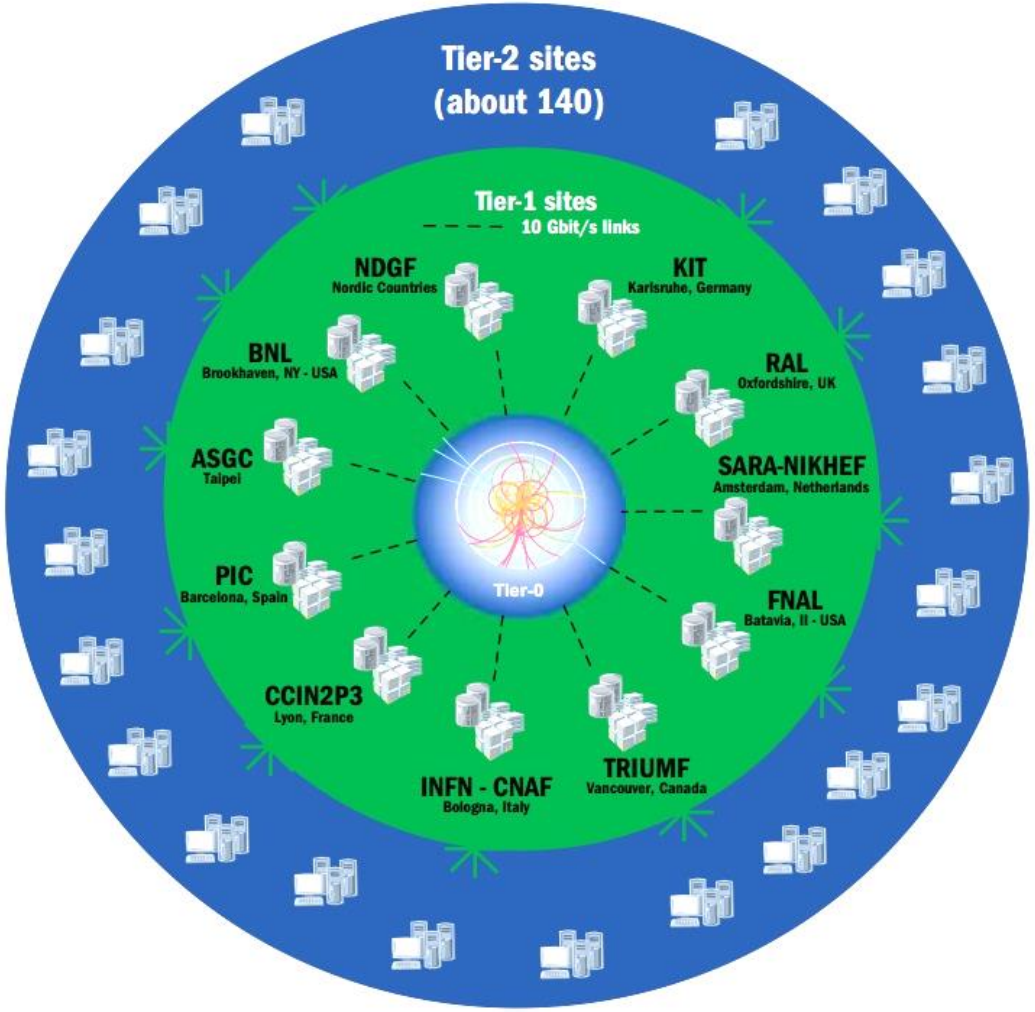
- Built in the 70s on the CERN site (Meyrin-Geneva)
  - 3.5 MW for equipment
- Extension located at Wigner (Budapest)
  - 2.7 MW for equipment
  - Connected to the Geneva CC with 3x100Gb links (21 and 24 ms RTT)
- Hardware generally based on commodity
- **15,000** servers, providing **230,000** processor cores
- **90,000** disk drives providing **280 PB** disk space
- **30,000** tapes drives, providing **400 PB** capacity

CERN Data Center

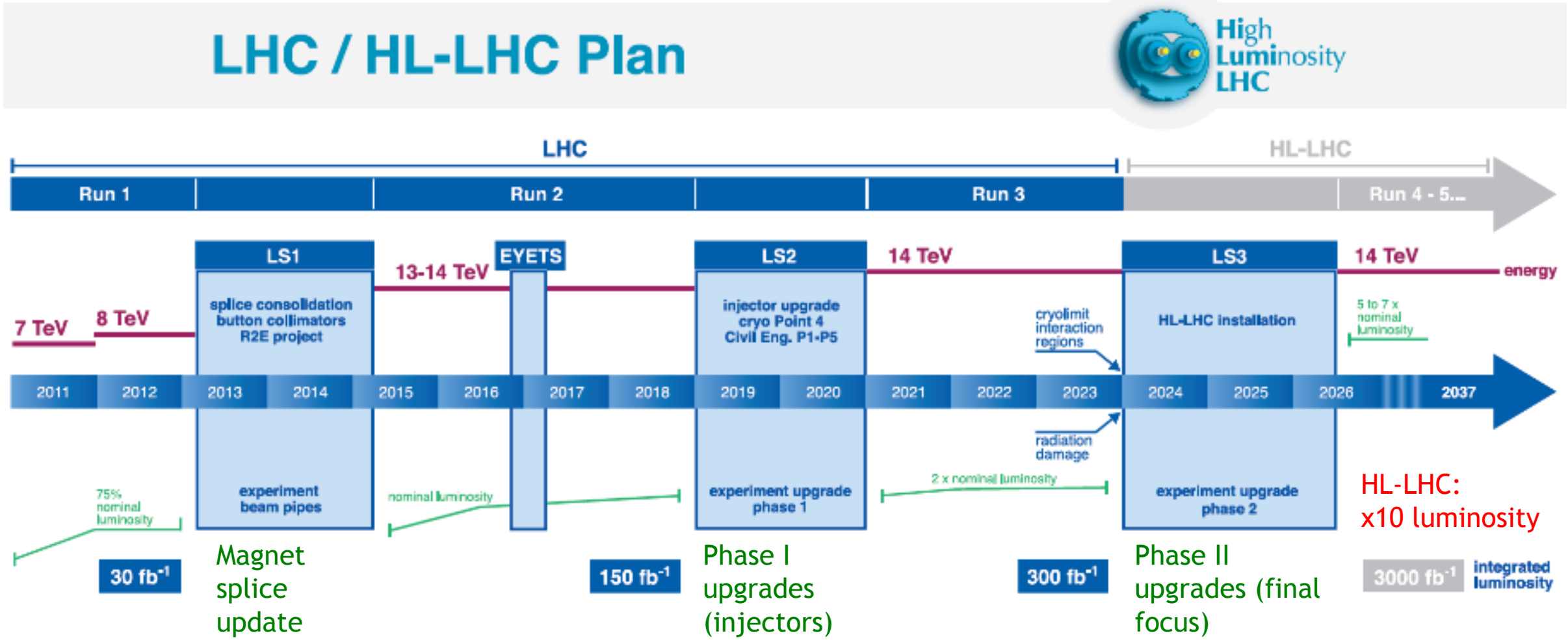
- Data agregation and Initial reconstruction
- Copy to long term storage (tape)
- Data distribution to the Worldwide LHC Computing Grid (WLCG)

Worldwide LHC Computing Grid (WLCG)

- 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



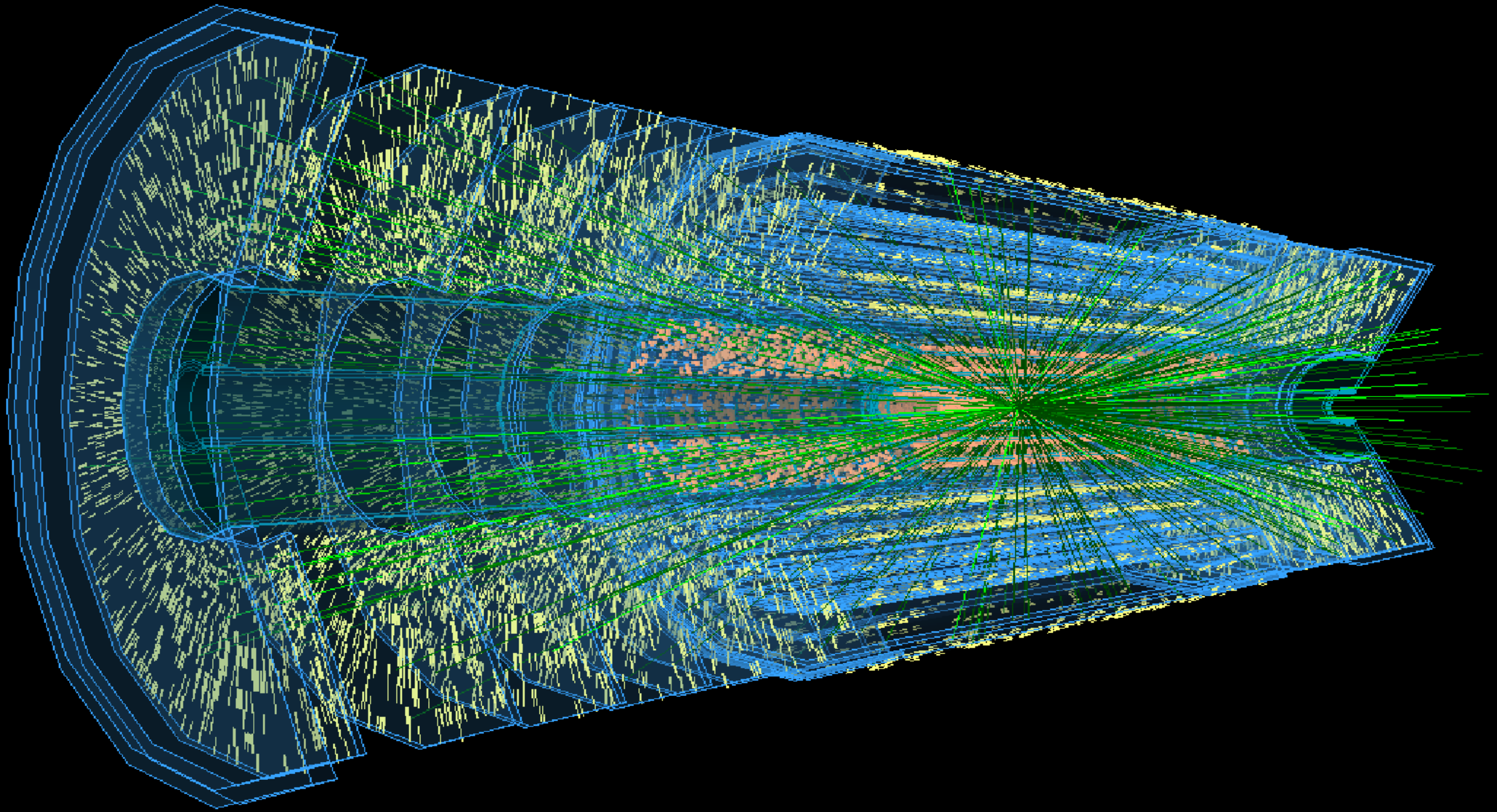
HL-LHC: a computing challenge



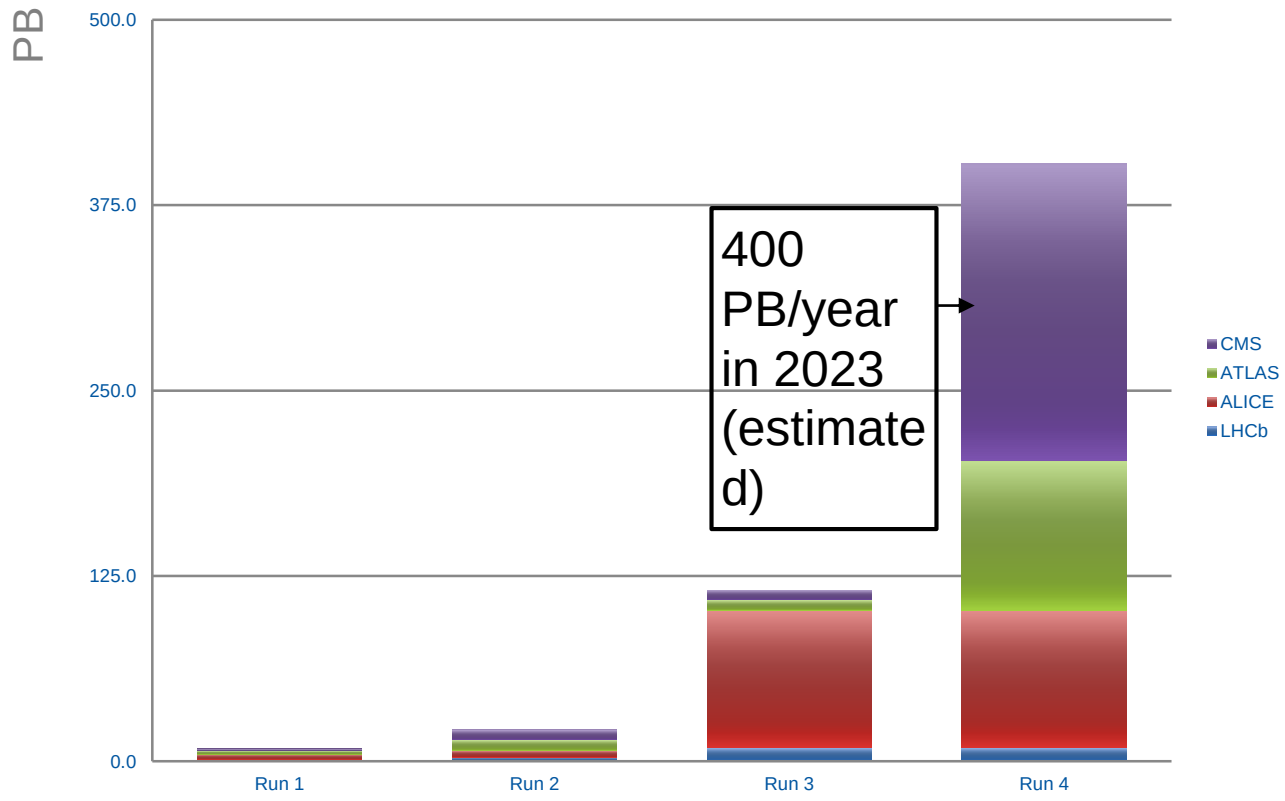
Magnet splice update

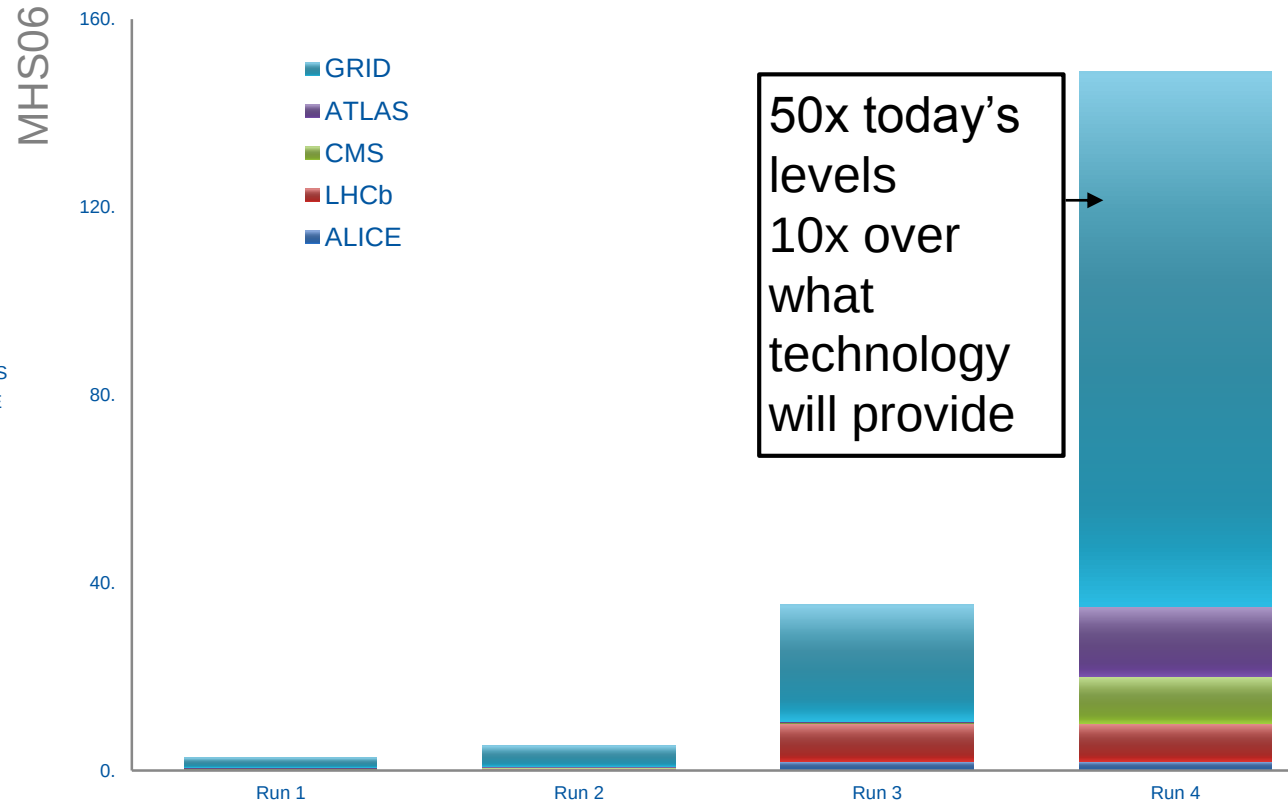
Phase I upgrades (injectors)

Phase II upgrades (final focus)



HL-LHC: a computing challenge





Future in Computing

- Manage storage and computing continuous growth (budget and infrastructures)
  - Storage technologies evolution: HDD, SSD, Tapes
  - CPU speed and multicore/vector exploitation
  - Data Center engineering: optimize energy consumption (PUE and green IT)
- Provide to the experiments and the users the computing requirements while optimizing resources
  - Worldwide LHC Computing Grid (WLCG)
- Improve software performance
  - HEP (High Energy Physics) Software Foundation, HSF
- Data preservation:
  - “We are nonchalantly throwing all of our data into what could become an information black hole without realising it.” Vint Cerf, vice president of Google and an early internet pioneer, February 2015
  - How to ensure that all the data collected and published is still readable by the next generations … and how to make sense out of it
- CERN is leading a global effort for HEP, that others will inevitably face soon or later

CERN-IT: pushing boundaries

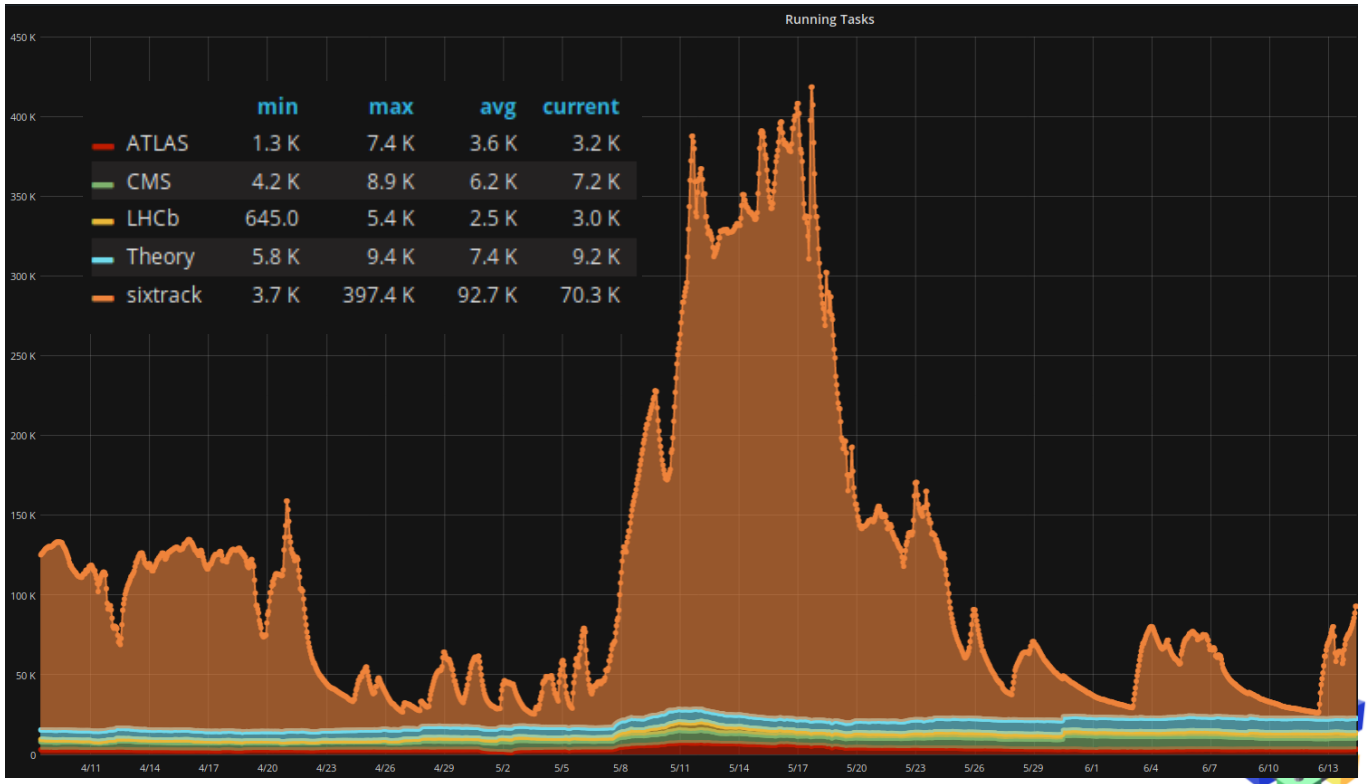
- CERN-IT impact on society through computing:
  - Need for collaboration tools for Global Science led to invent the **World Wide Web**
  - Need for collaboration of computing resources for the Global LHC led to adopt **Grid Computing** and first concept of **Computing Clouds**
  - Need for sharing the results had led CERN to pave to way to open access to documents and now data: **LHC@home** and **CERN Opendata Portal**

- Could these technologues have been originated somewhere else?
  - Probably. But often we are faced to challenges 5 to10 years before others, pushed by the needs of the detectors and accelerators. Pushed by fundamental science.

LHC@Home

- 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

@ Volunteer computing for the LHC



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

opendata CERN

About

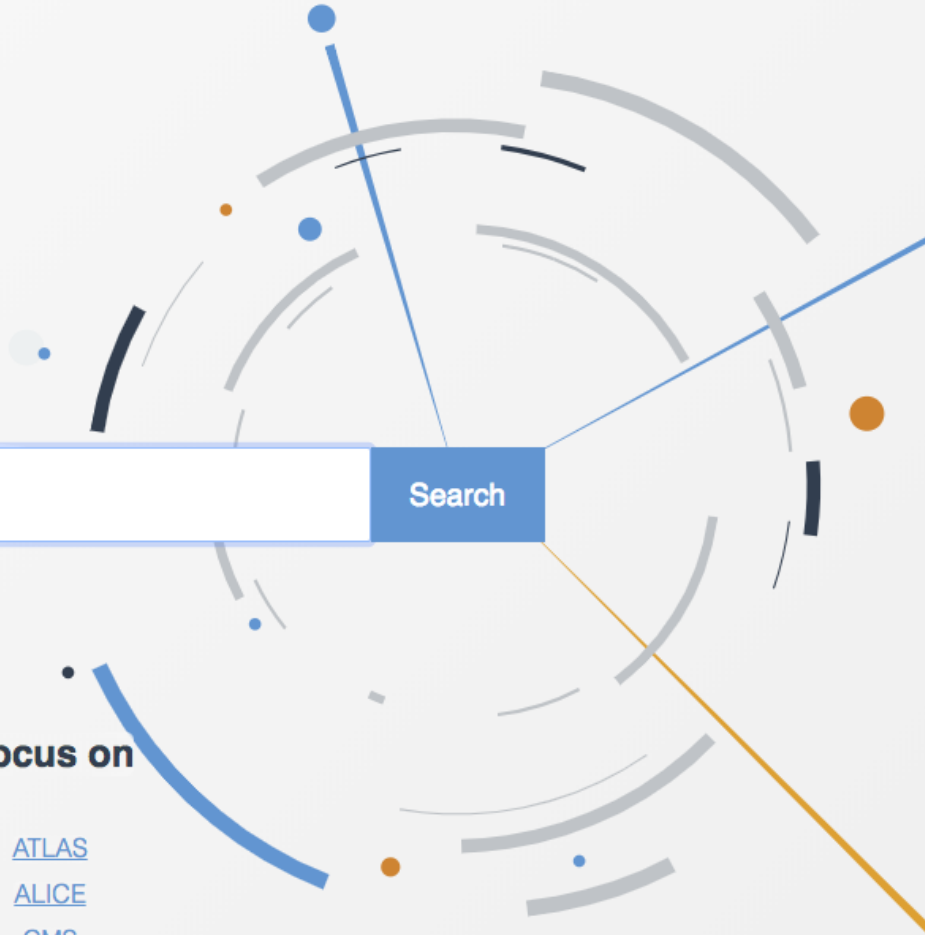
CERN OPENDATA

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

Muchas gracias!