



Computing at CERN



Dr Tim SMITH



International Teacher Weeks Programme – 2024/08/12

computing, *n.*

Pronunciation:

Brit. /kəm'pju:tɪŋ/ , U.S. /kəm'pjʊdɪŋ/

1. The action or an instance of calculating or counting; = [computation *n.* 1a.](#)

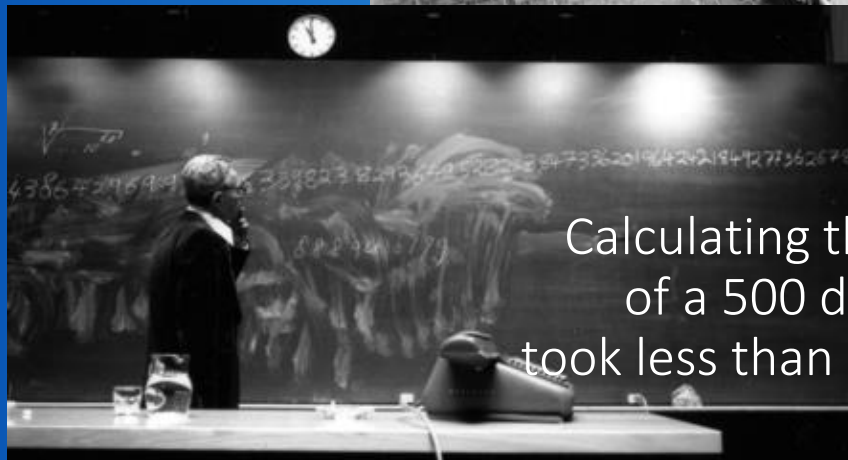
2. The action or practice of using computers, esp. as a professional or expert; the activity or operation of an electronic computer; (also) = [computer science *n.*](#)

Early “Computers”



Katherine Johnson NASA
Hidden Figures!

CERN had 2 British Ladies
from National Physical Lab
Teddington (London)



Wim Klein

Calculating the 73rd root
of a 500 digit number
took less than 3 minutes...

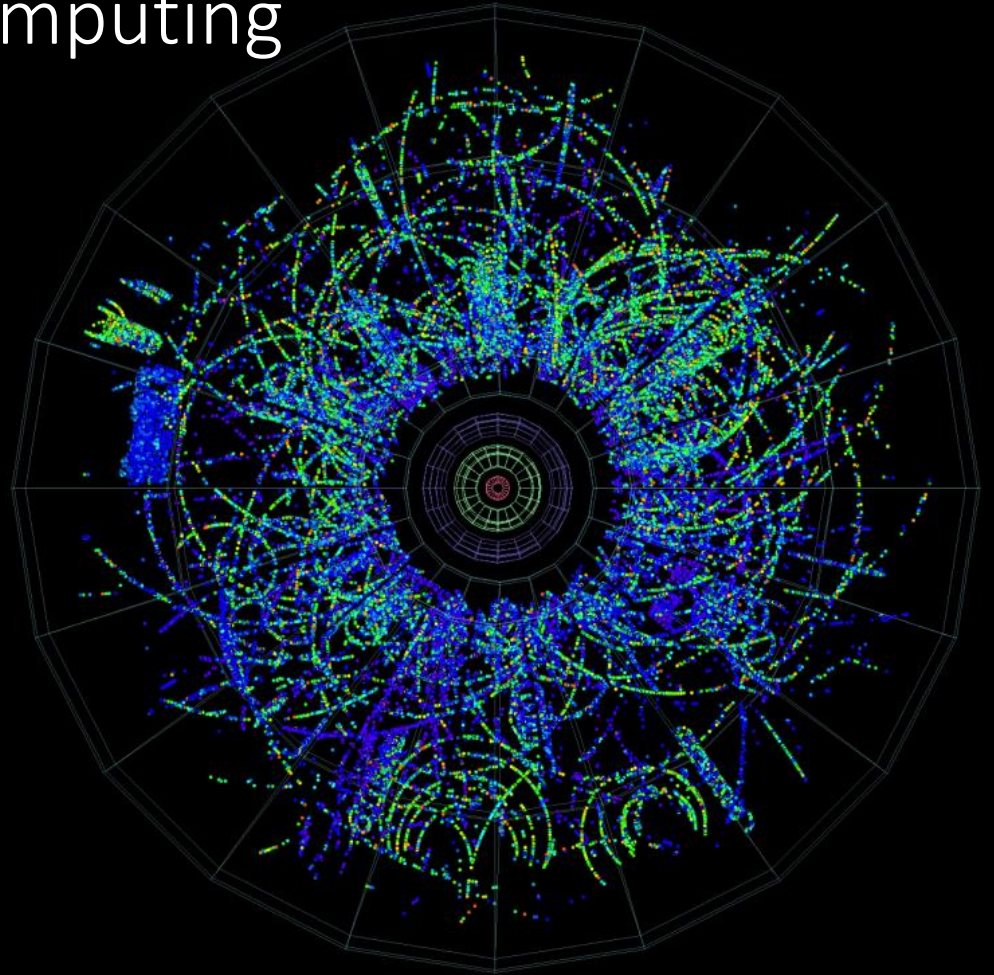
1958: The Ferranti-Mercury arrived!



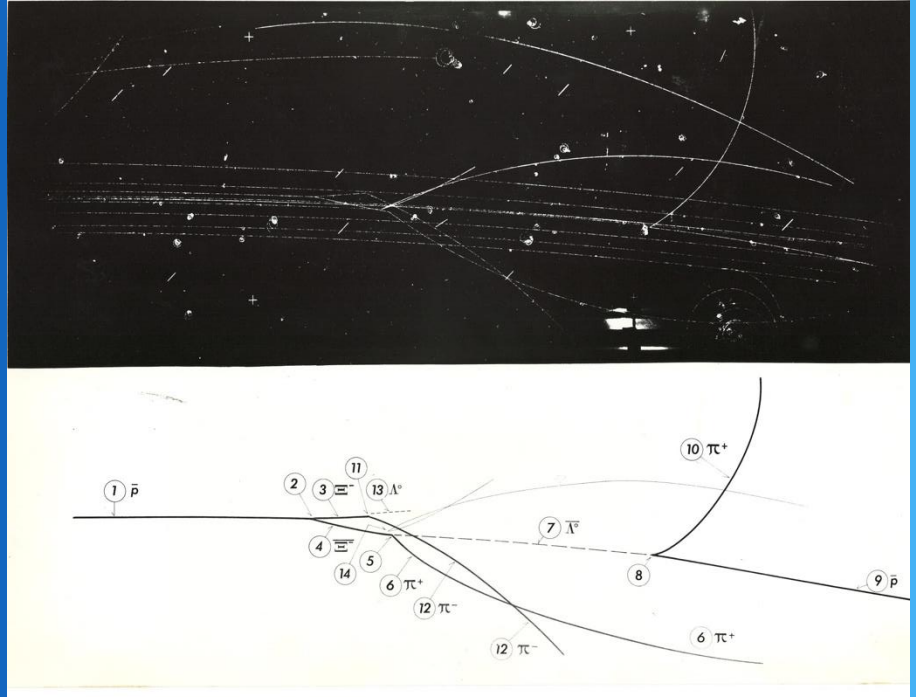
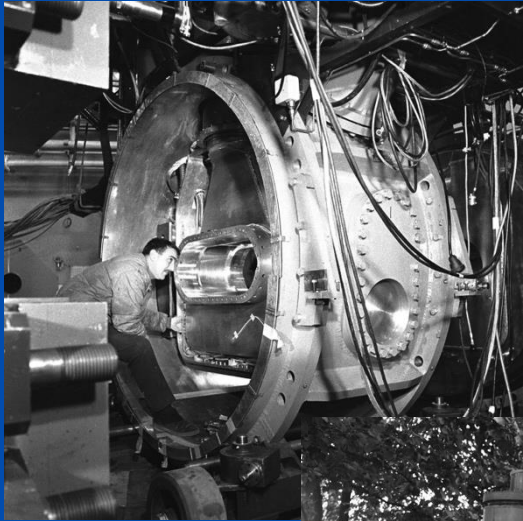
“..courage to go ahead and solve problems that seemed to difficult to do otherwise..”

Computing

- Scientific
 - Compute, Storage, Network
- Technical
 - Design, Operation
- Administrative
 - HR, Finance, Projects
- Desktop
- Collaborative



Bubble Chamber



Bubble Chamber

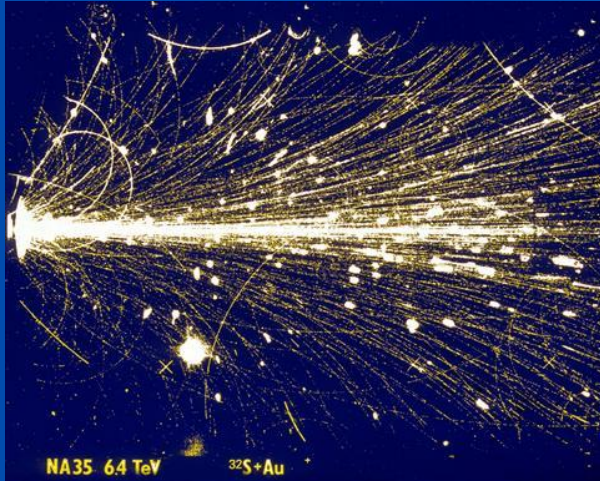


Madeline Znoy
750 photos in a day!

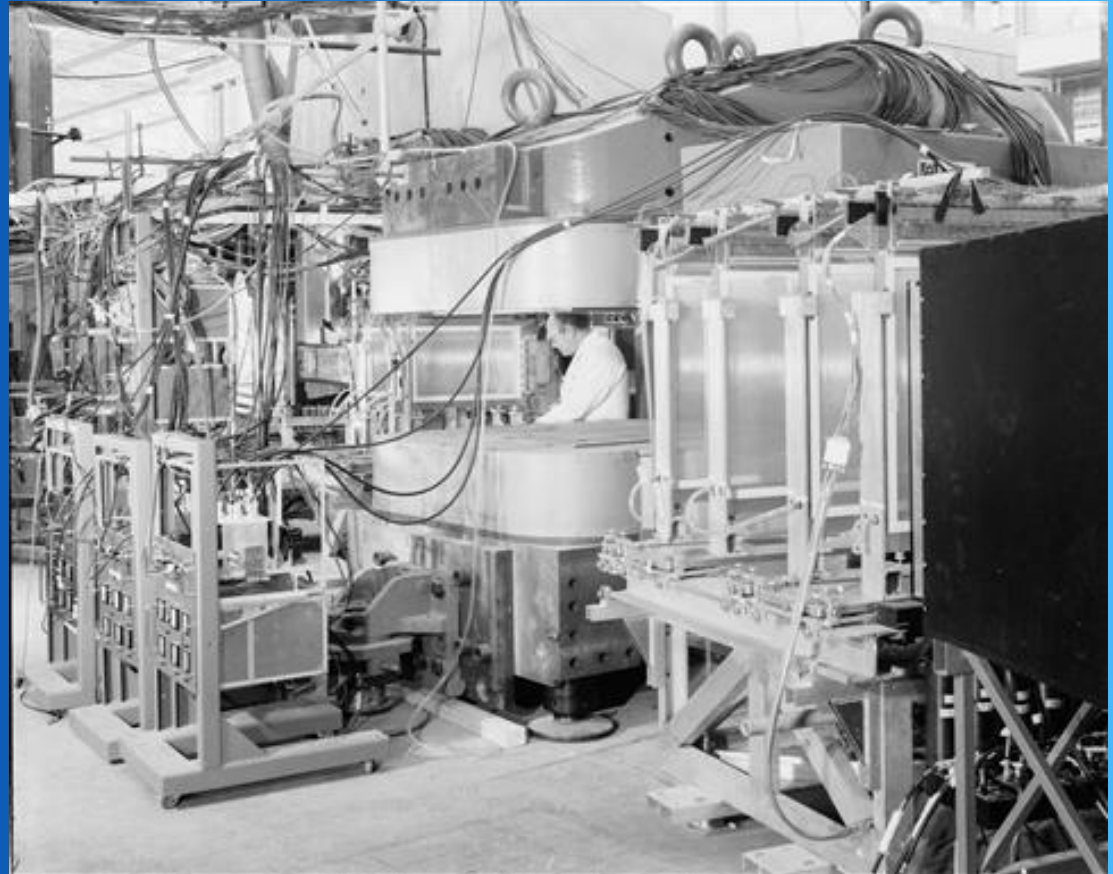


BEBC 1973-1984
6.3 million photos
3000 km of film

Spark Chambers



1965
Magnetostrictive readout



Momentous Events

1960: IBM 709 Vacuum tubes





1965: CDC 6600: #3

1972: New Computer Centre for a New Computer!



Mainframe Era



1983: 2 CDCs and an IBM 3081

Super Computers

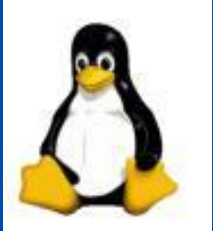


1988: Cray XMP

RISC Workstations



Comodity Computing

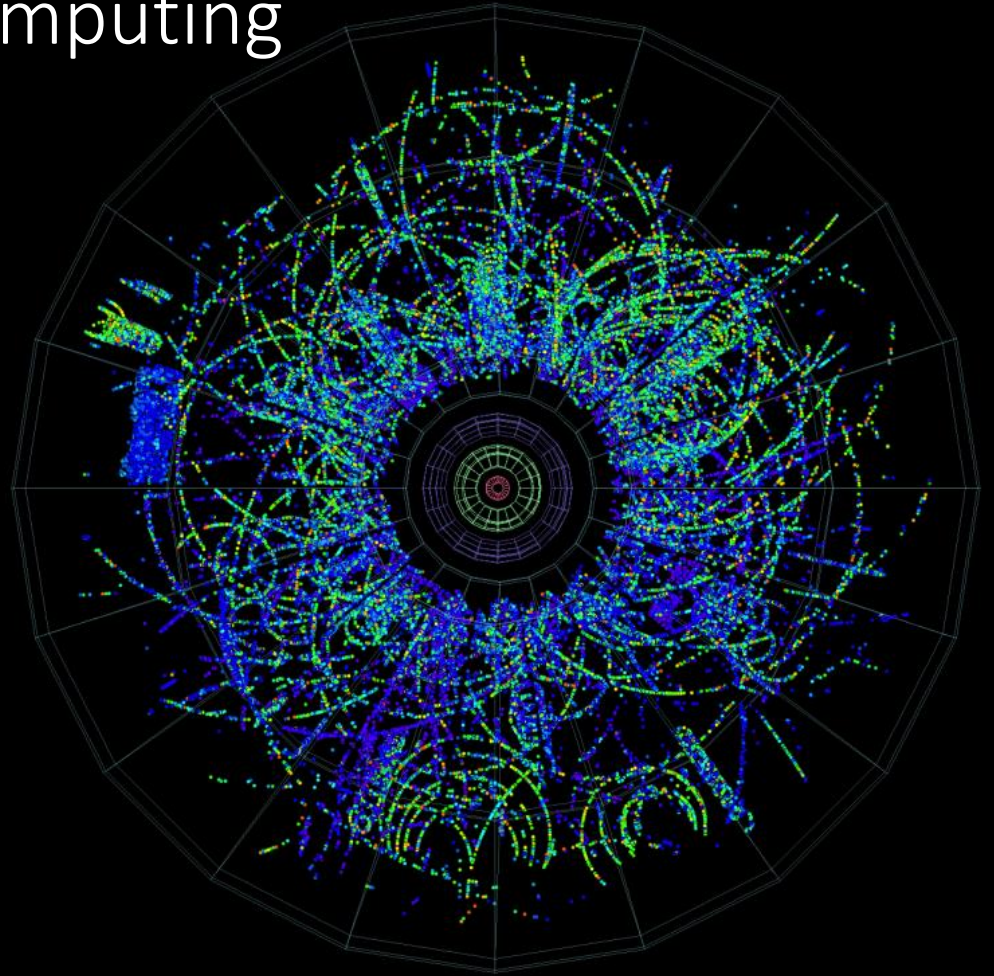


Farming in a Data Centre



Computing

- Networking



Networking

- Packet-switched network
- 1969 ARPANET
 - US DoD sponsored research at US Universities
 - Aim: communications network to survive a nuclear attack
 - Find next best route if one node obliterated
- 70s and 80s proliferation
 - US: NASA Science Net, CSnet, Energy Sciences Net, NSFnet
 - FR: CYCLADES
 - UK: Mark I, SERCnet
 - Commercial: Tymnet, CompuServ, BITnet, DECnet
 - Protocols: NCP, X.25 (1976), TCP/IP (1982)
 - CERnet

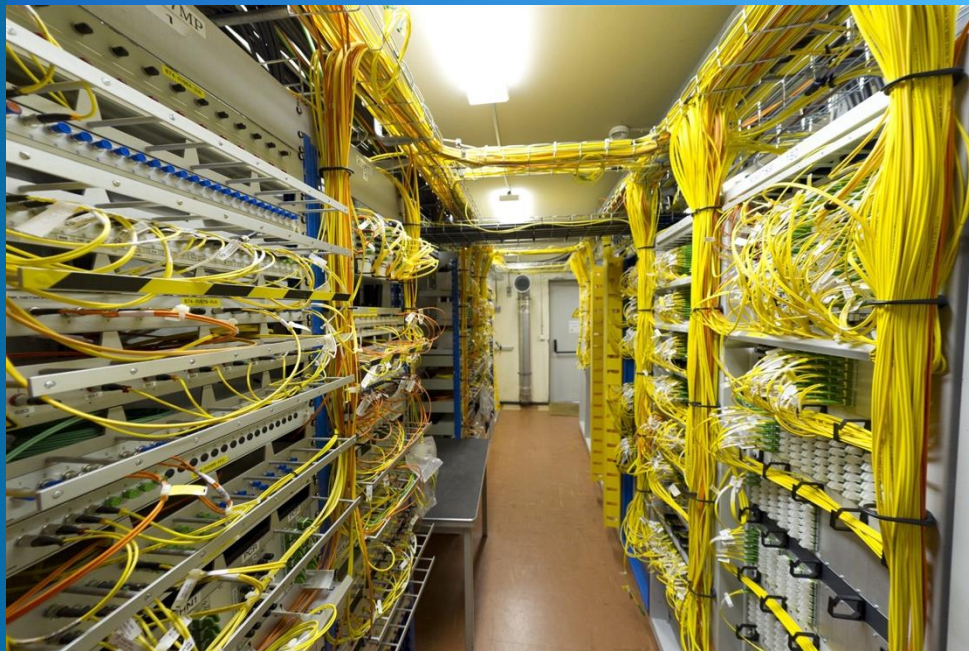
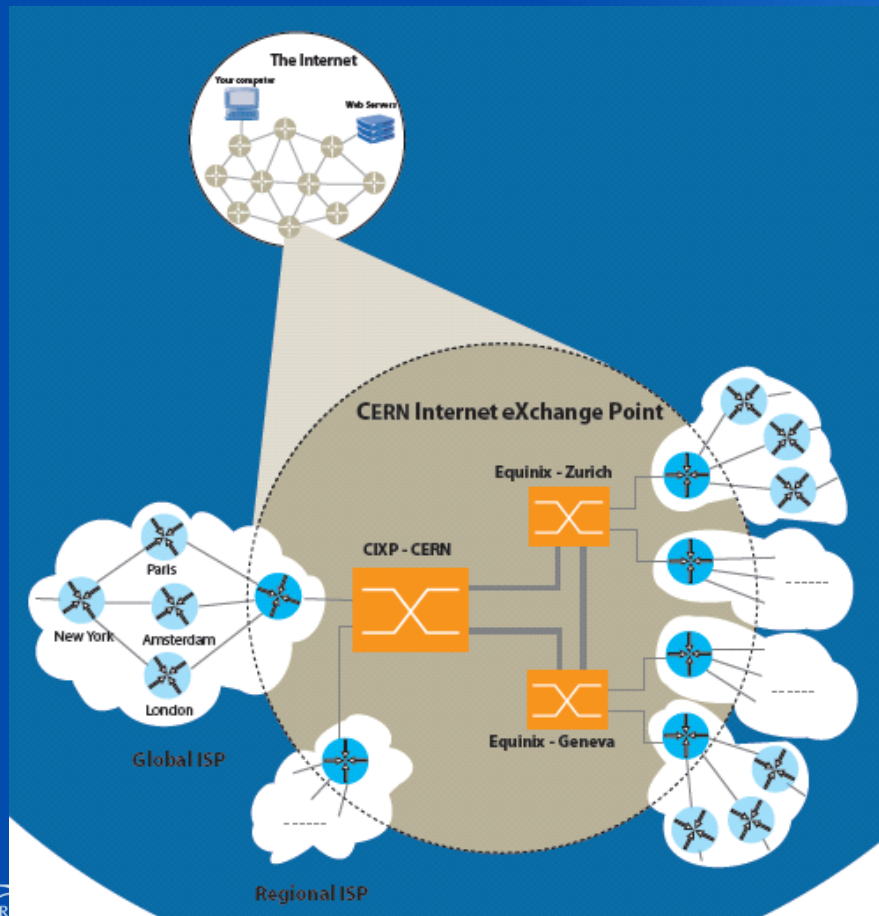


Networking

- Science without borders
 - 1968 Data exchange across the iron curtain – to IHEP in Soviet Union
 - 1988 first data connection between China and scientific world – IHEP to CERN
- Truly international Internet
 - 1989 first external TCP/IP connection
 - 1990 principle link US-EU from CERN
 - (1.5Mb/s)
 - 1991 80% of the internet capacity installed in Europe for international traffic was terminated at CERN



CERN Internet Exchange Point

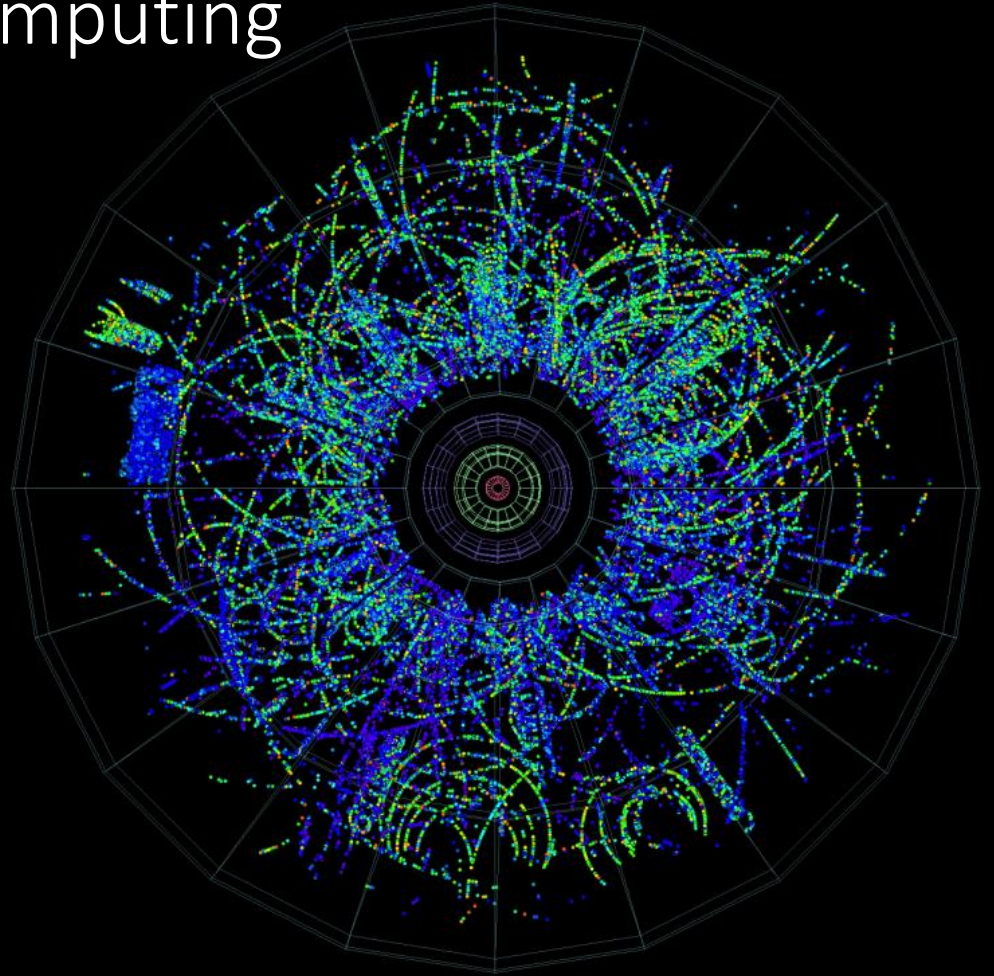


Internet ... Web ... Who?



Computing

- Information Management



Information Management - *circa 1989*

- Keep track of LHC project and CERN?
 - Researchers turnover ~2 years
- Information about CERN and its experiments
 - Not hierarchical, or centrally controlled
 - A multiply connected web
 - Experts store locally, update independently
 - Community is distributed: remote access
- System to link it all together

- CERN is a model in miniature of the rest of world in a few years time

Distributed, Collaborative

CERN DD/OC

Tim Berners-Lee, CERN/DD

Information Management: A Proposal

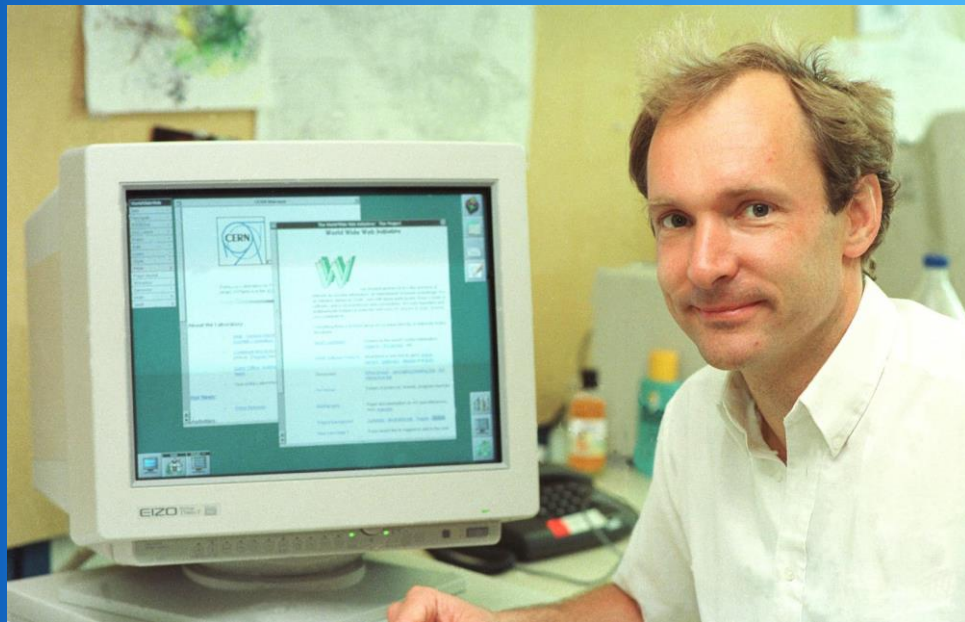
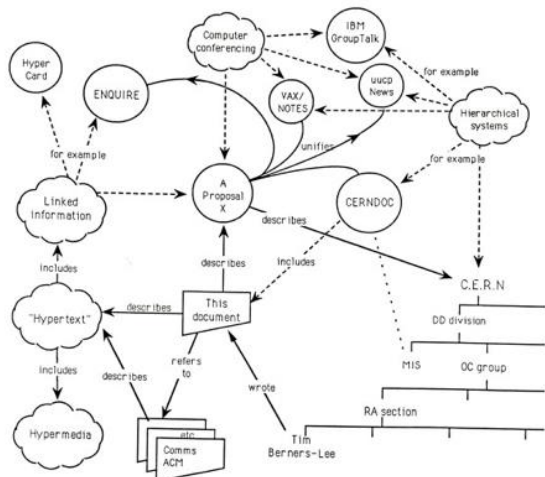
March 1989

Information Management: A Proposal

Abstract

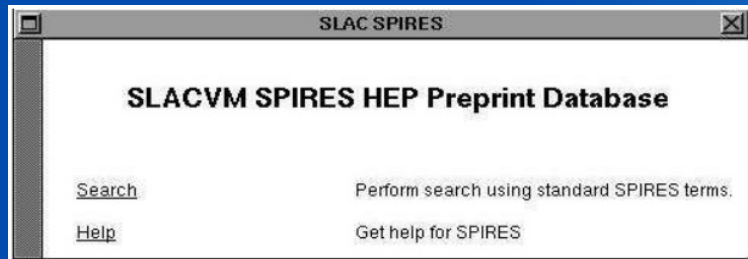
This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control



Growth of the Web

- Aug 1991 went public
 - Tim posted project to alt.hypertext and other internet groups

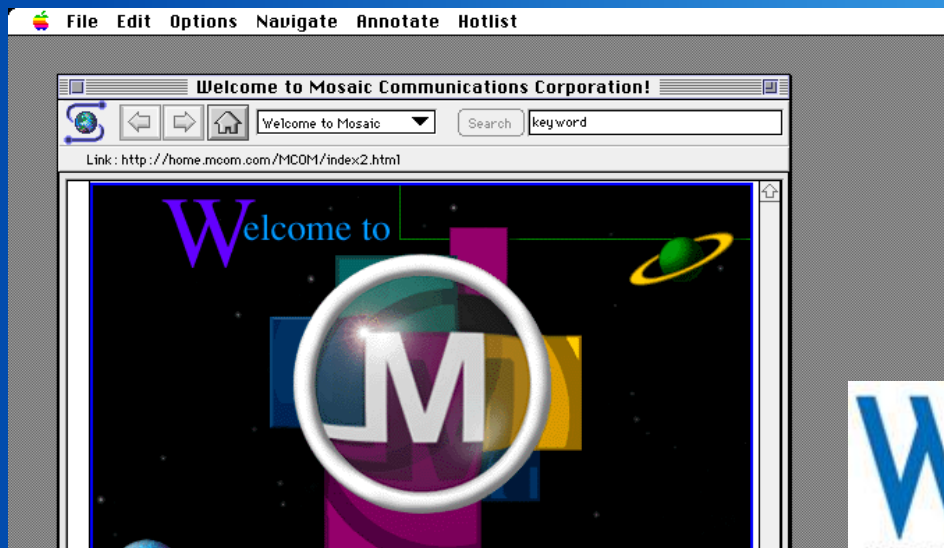


Dec 1991 First web server in US
1992 rapid expansion in HEP
Universities and research institutes

World Wide Web



Growth of the Web



- 1993 rapid expansion across the world
 - National Center for Supercomputing Applications (NCSA) at the University of Illinois released its Mosaic browser

Born in Science



©CERN

Collaborative development of new tools

Universal access to information:
a human right

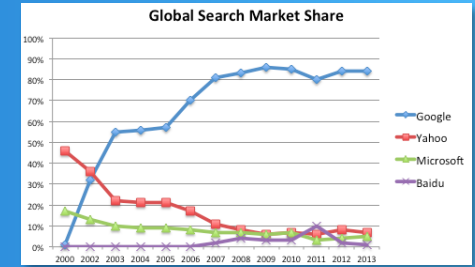


Theme of Concern:
monopolies as gatekeepers
steer for profit, not for humanity

The Concentration of Power

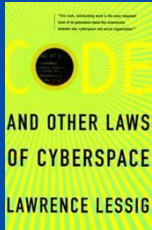


Handful of platforms control which ideas and opinions are seen and shared



Ads

Engagement platforms
Competitive advantage from User giving data



SW creators decide fundamental issues like freedom and privacy
Which content to remove, which users to kick-off
Private Law: EULA

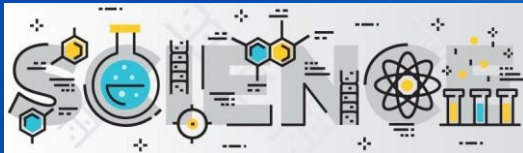
**WEAPONIZED
AT SCALE**

Conspiracy theories trend on social media platforms
Fake Twitter and Facebook accounts stoke social tensions
External actors interfere in elections

The {Mis|Dis}Information Age

- Word-of-the-Year 2018: Misinformation
 - Election tampering
 - Weaponization of falsity
 - Surveillance capitalism
 - Fake news
- Word-of-the-Year 2020: Pandemic
 - Vaccine distrust
 - Climate Change denial
 - Alternative facts

F A C T
K E



- Public mistrust
- Skepticism among government leaders



Evidence Chains



*The research behind
this announcement...*



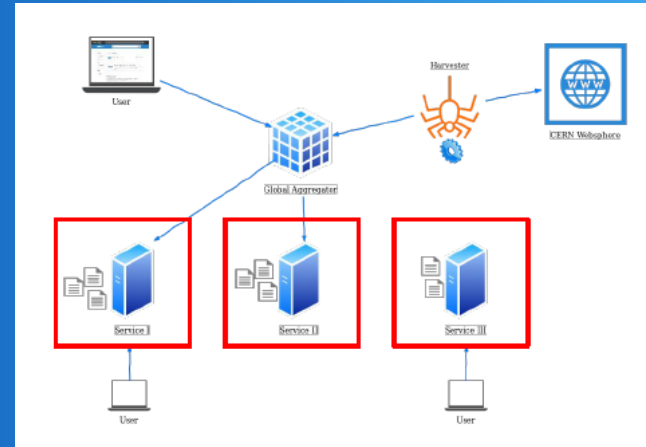
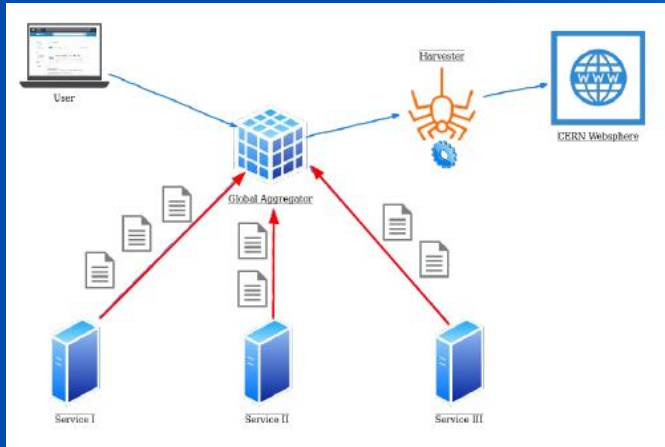
*The data this research
was based on...*



*The dominant theory in academia
The alternatives being discussed*

Search @ CERN

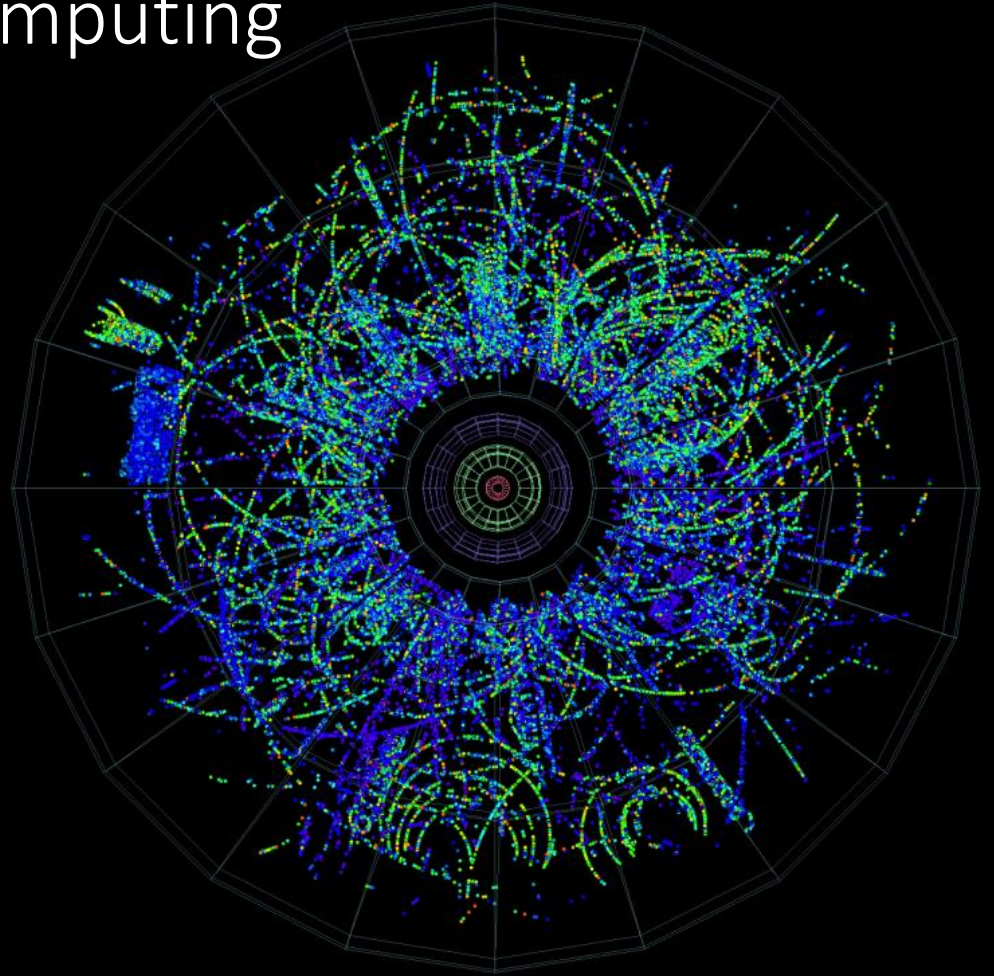
Moving from a centralized commercial Enterprise Search solution ...



... to a distributed Open Source Enterprise Search solution

Computing

- The LHC Era



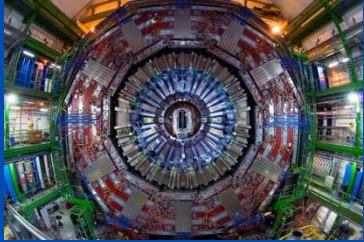
The LHC Data Challenge

Big Data



And bigger
on the way!

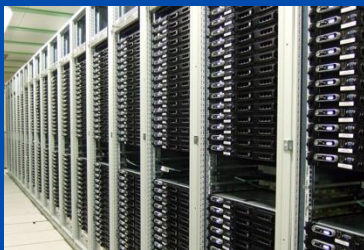
Big Data !



150 million sensors
Generating data 40 million times per second
Hardware trigger in a few microseconds



C.f: Google's computing farm handles
100,000 search queries per second

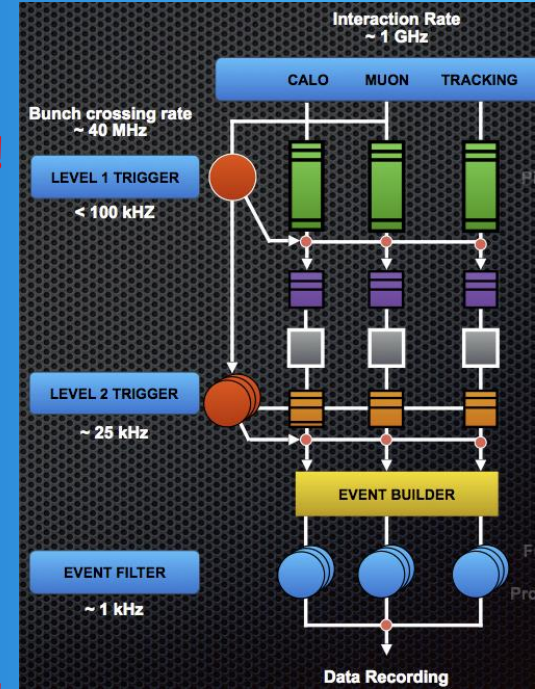


Peta Bytes / sec !

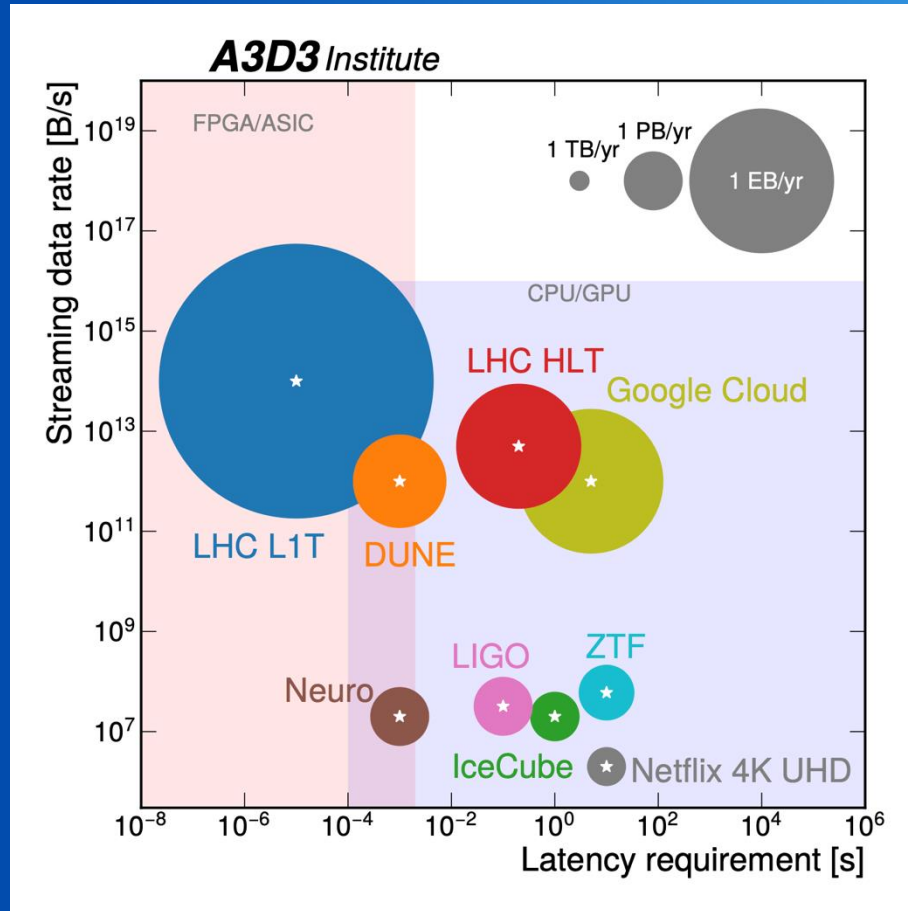
Peta Bytes / sec !

Select 100 per second

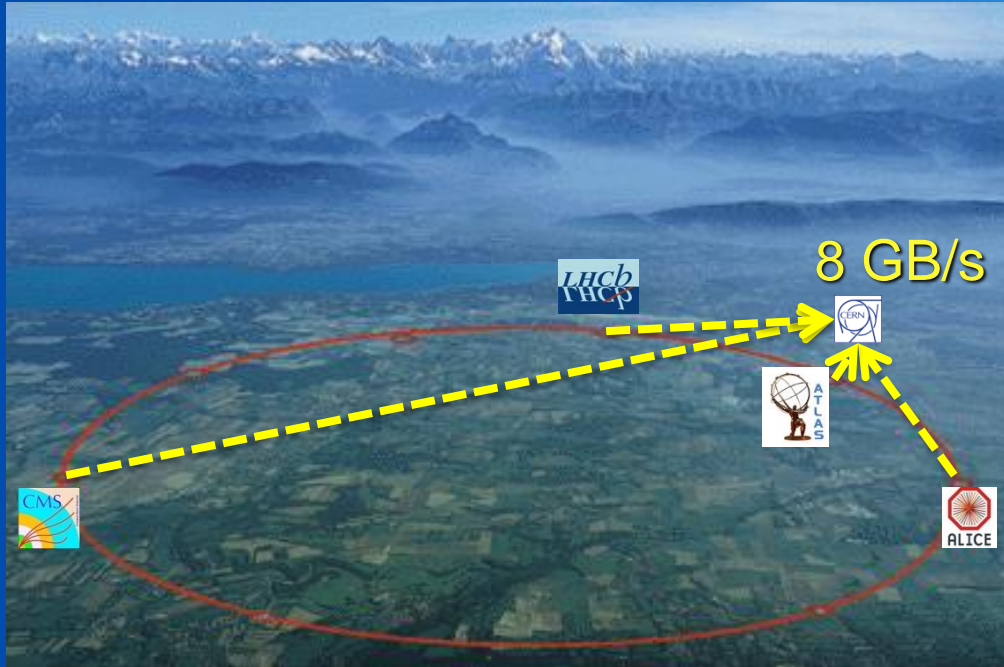
→ Giga Bytes / sec !



Big Data in Context



Primary Storage



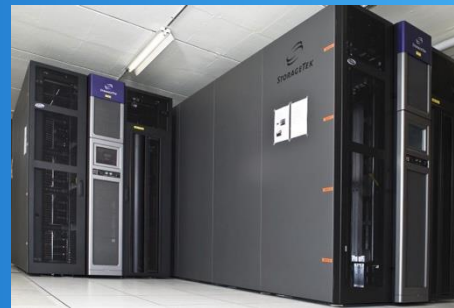
100,000 Disks
450,000 CPU Cores

20,000 1GB NICs
4,400 10GB NICs



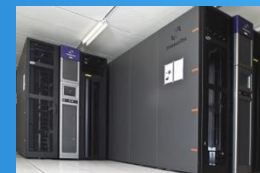
The LHC Data Challenge

- Few places can store it
- Processing needs 3x CERN
- HEP community distributed
 - Local funding for computing
- Distributed solution...



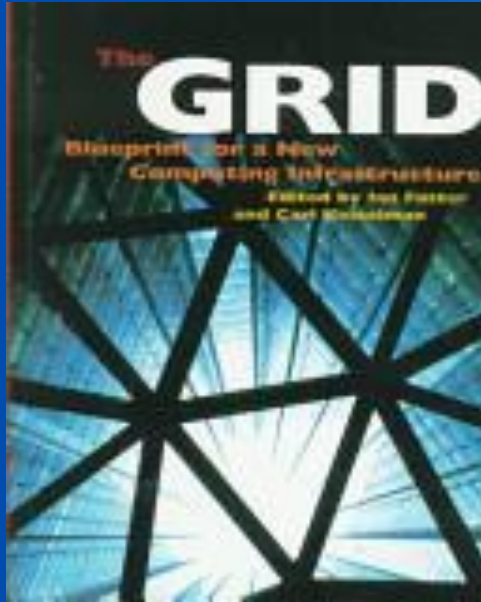
Models of Networked Analysis at
Regional Centres

x 2 locations @ CERN



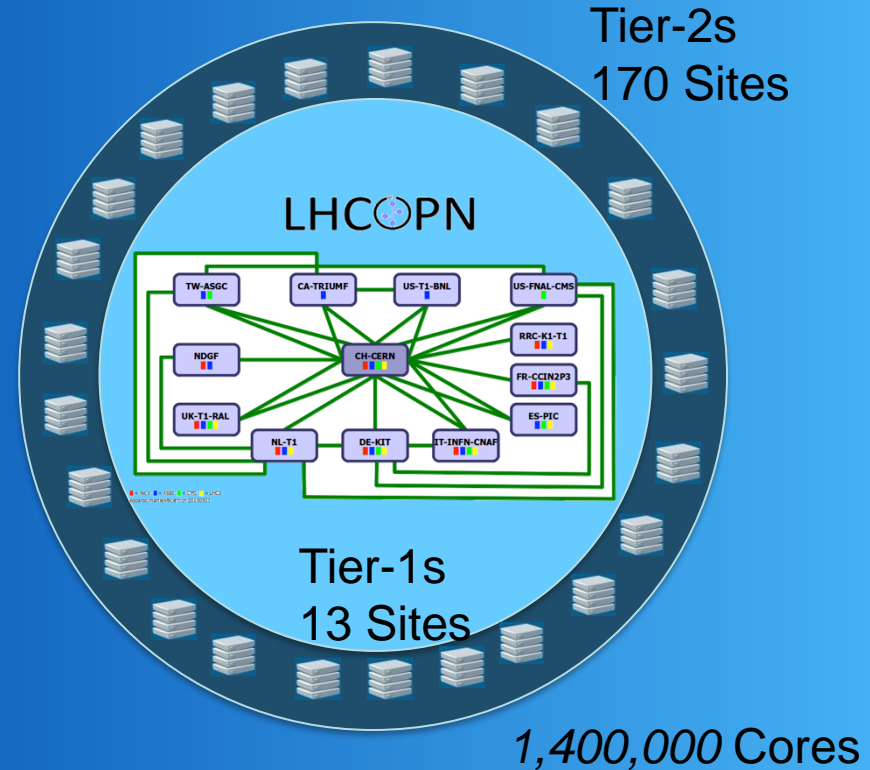
Solution: the Grid

- Use a Grid to unite computing resources of particle physics institutes around the world



Worldwide LHC Computing Grid

- Tier-0 (CERN)
 - Data recording
 - Initial data reconstruction
 - Data distribution
- Tier-1
 - Permanent storage
 - Re-processing
 - Analysis
- Tier-2
 - Simulation
 - End-user analysis



WLCG: The Grid that Never Sleeps

EGL EXPLORER OF GRID LINKS
Concept and production by CERN Media Lab

DATA DISPLAY
08 July 2022 - 14:00

Average number of cores used by LHC VOs: 1177630

Power delivered to LHC VOs: 1332.837 HSD6

LHC VO transfer throughput: 23.26 GB/s

VOLUME TRANSFERS | VOLUME FILES | VOLUME DATA

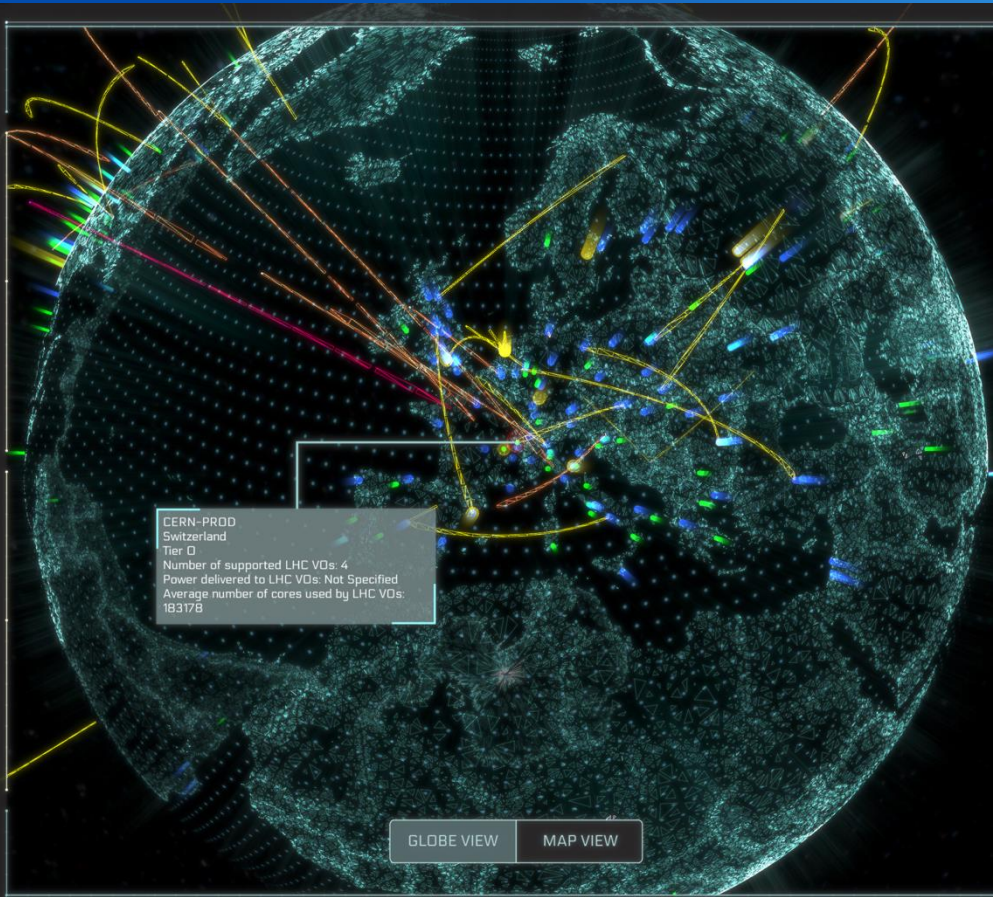


CURRENTLY DISPLAYING
FROM: 08 July 2022 - 14:00
TO: 08 July 2022 - 16:59

LAST DATA UPDATE | LOADING 100 %
Not Specified downloaded on Friday, 08 July 2022
Last transfer was on Friday, 08 July 2022

DATA TRANSFER | CONSOLE

4962281 From 08-12-2022 08:14:10 to 08-12-2022 08:14:10 on Friday, 08 July 2022
44328242 From 08-12-2022 08:14:10 to 08-12-2022 08:14:10 on Friday, 08 July 2022
44328242 From 08-12-2022 08:14:10 to 08-12-2022 08:14:10 on Friday, 08 July 2022
370202448 From CA-670-12 to 08-12-2022 08:14:10 on Friday, 08 July 2022
288201818 From 08-12-2022 08:14:10 to 08-12-2022 08:14:10 on Friday, 08 July 2022
537402877 From 08-12-2022 08:14:10 to 08-12-2022 08:14:10 on Friday, 08 July 2022
228201818 From 08-12-2022 08:14:10 to 08-12-2022 08:14:10 on Friday, 08 July 2022



500PB @ CERN

Stores, distributes,
processes and analyses
LHC experiments' data

1.4 million processing cores
in 170 data centres
and 42 countries

1500 Petabytes
of CERN data stored
world-wide

Cloud?

Cloud

- On Demand
- Dynamically provisioned & metered by e.g. Amazon, Microsoft Azure



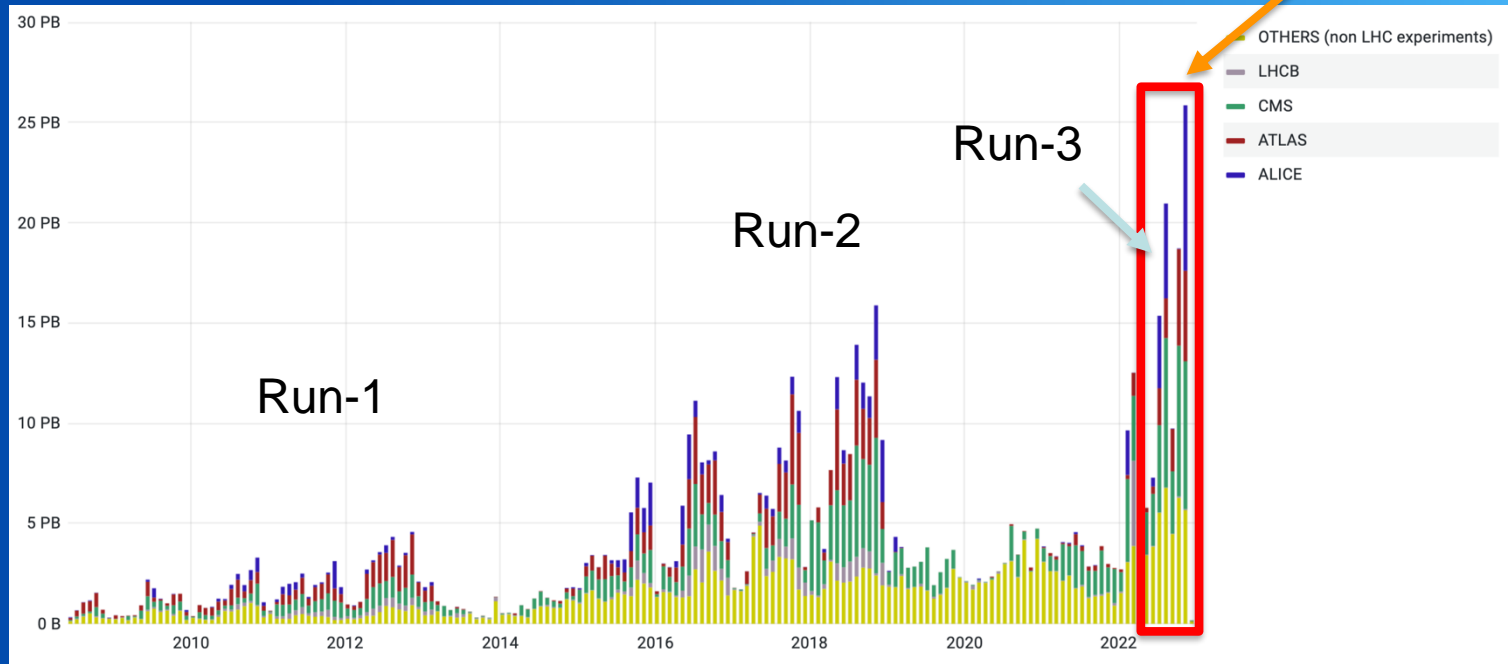
Grid

- Fixed size
- Collaborative, run by community



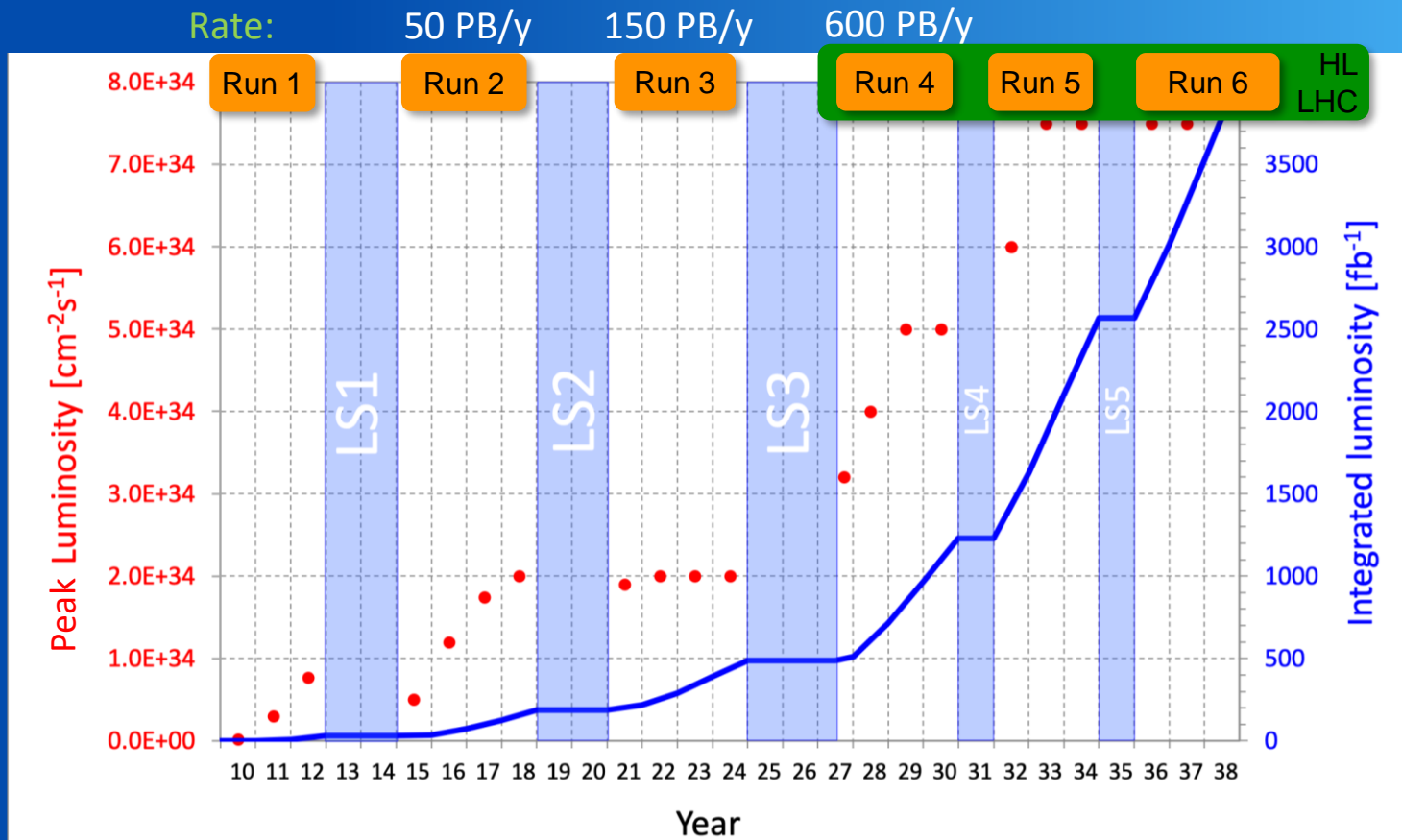
Run-3 data taking

Data written in the CERN tape storage per month

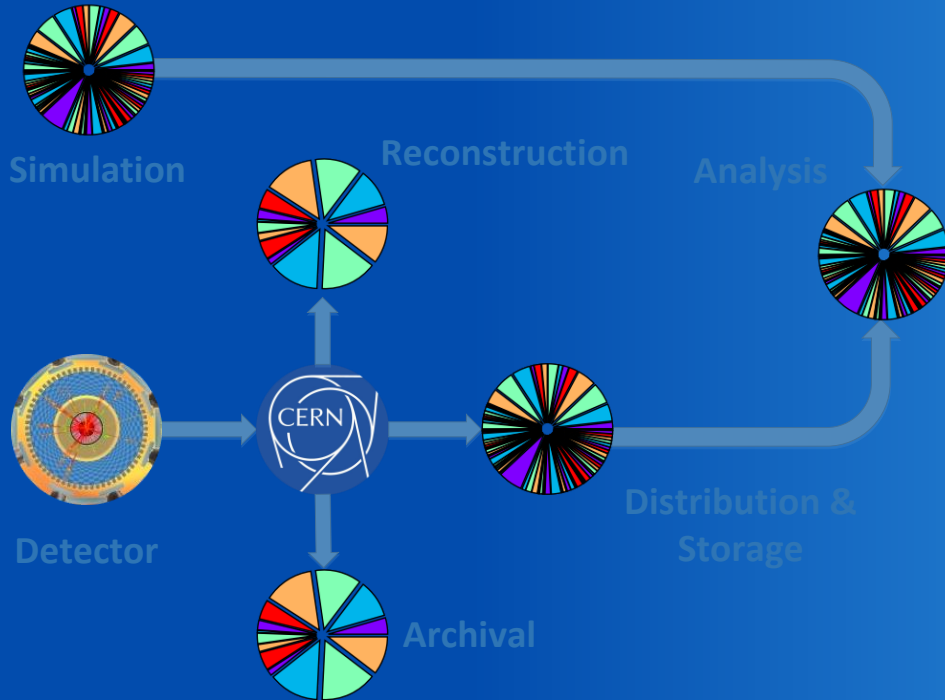


15PB 27PB

The LHC Data Challenge

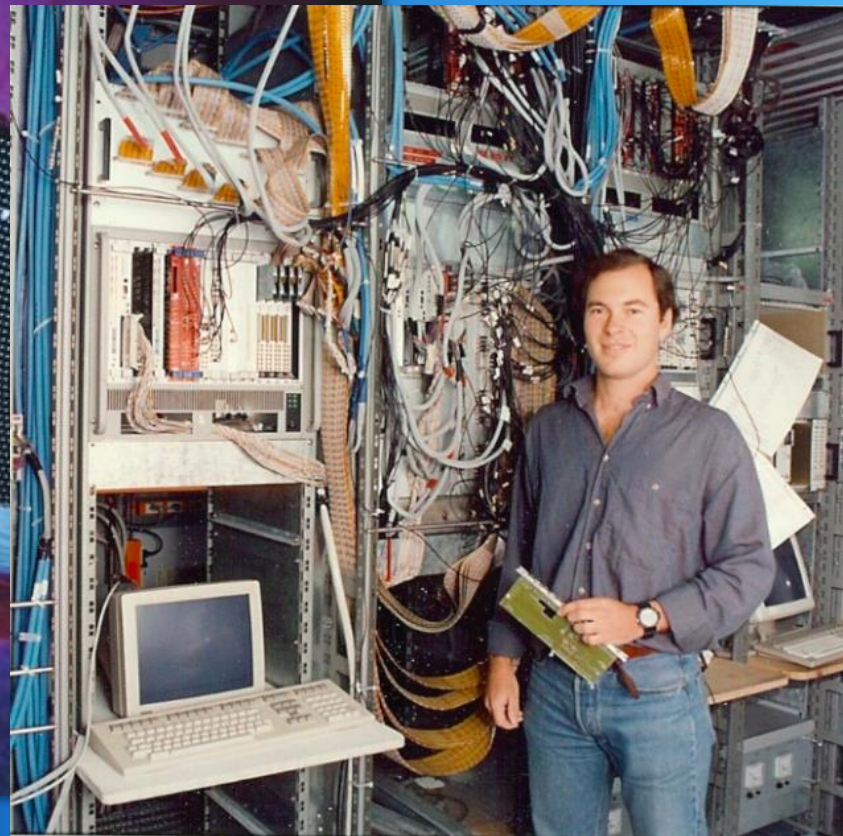


Computing throughout the Workflow



- ▶ More than half the CPU goes on simulation.
- ▶ Most of the rest is reconstruction.
- ▶ The remainder is analysis.

Where I learned to *weave*



Data Reduction / Analysis

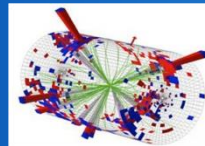
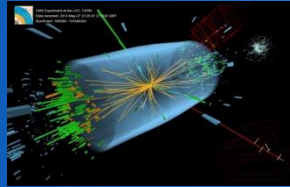
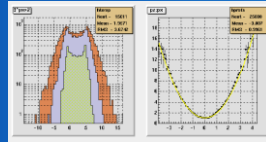
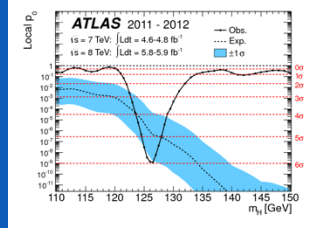
Publication

Reduced

Reconstructed

Raw

File Size



Files

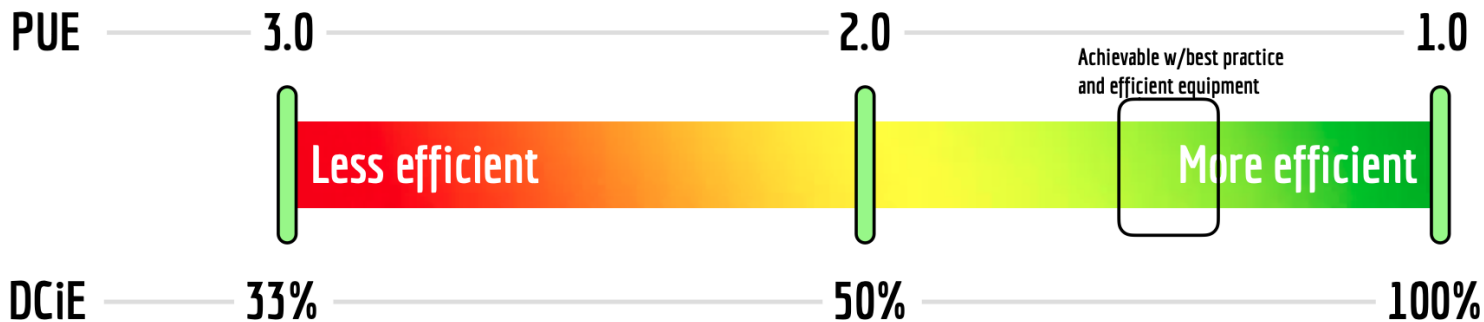
Researchers
T2s, T1s

Analysis Coordinators
T1s

Production Managers
T0, T1s

Green IT

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

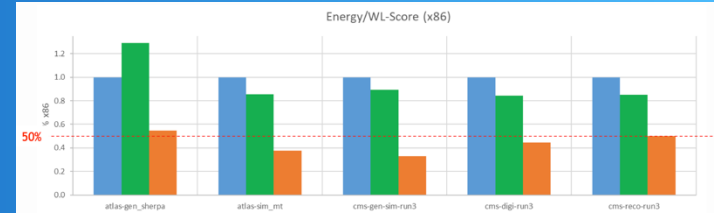


$$\text{DCiE} = \text{Data Center Infr. Efficiency} = \frac{1}{\text{PUE}}$$

Reducing IT's Energy Footprint: 3 lines of action

The Hardware

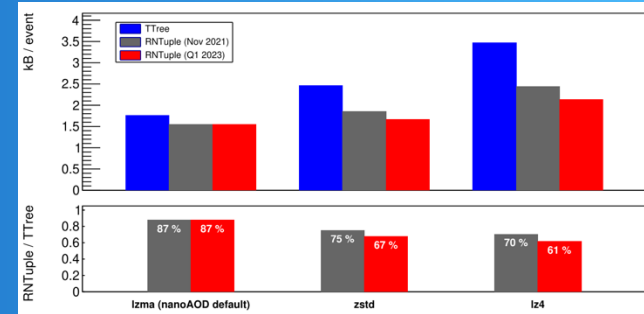
- Procuring power efficient hardware
- Extending hardware lifetime
- CPU virtualisation, Disk Server densification, Tape evolution



The Software

- Improving software efficiency; focused C++ training courses
- Innovating computing models
- Using accelerators to improve efficiency of Generation & Simulation

ATLAS and CMS workloads: on ARM of x86



ROOT foundation layer: less CPU, less disk hungry

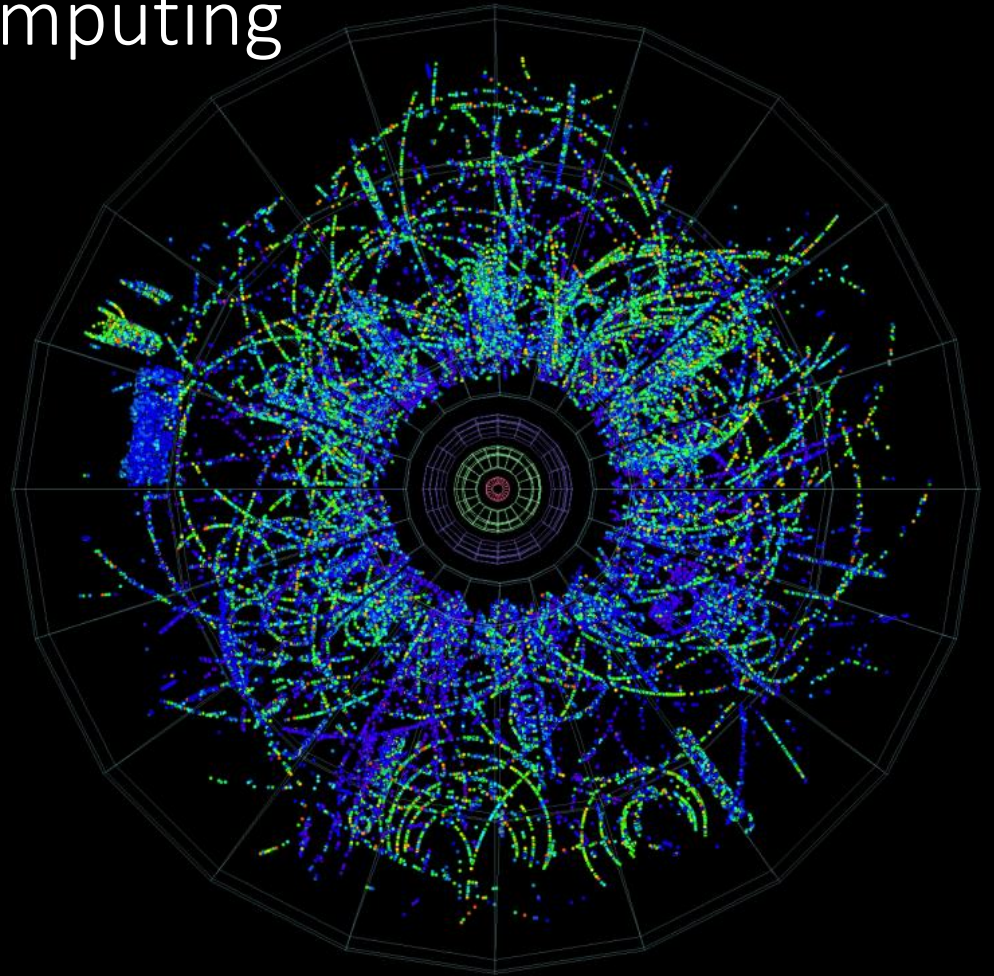
The Data Centre

- New data centre with efficient cooling and heat recovery
- An optimized hardware life-cycle



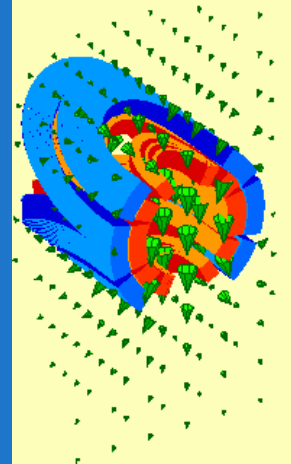
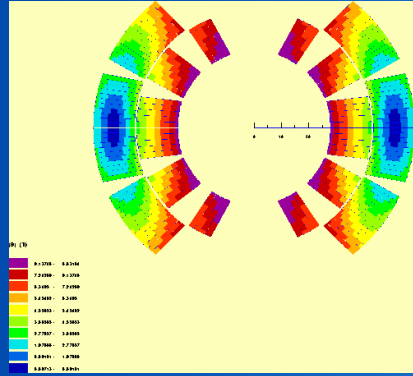
Computing

- Technical



Design

- Magnet Design



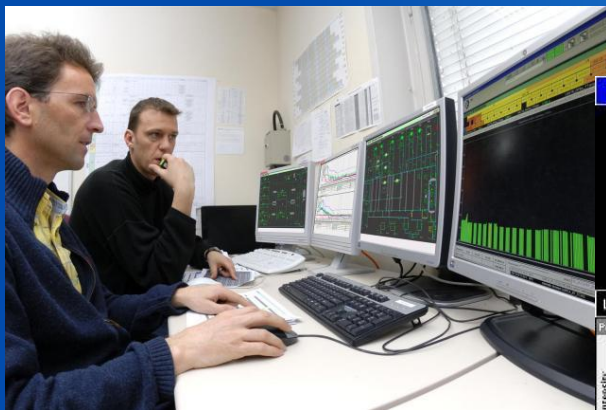
- Volunteer Computing

- LHC@home
- SixTrack

- Simulates particles accelerating around the 27 km LHC to find their orbit stability

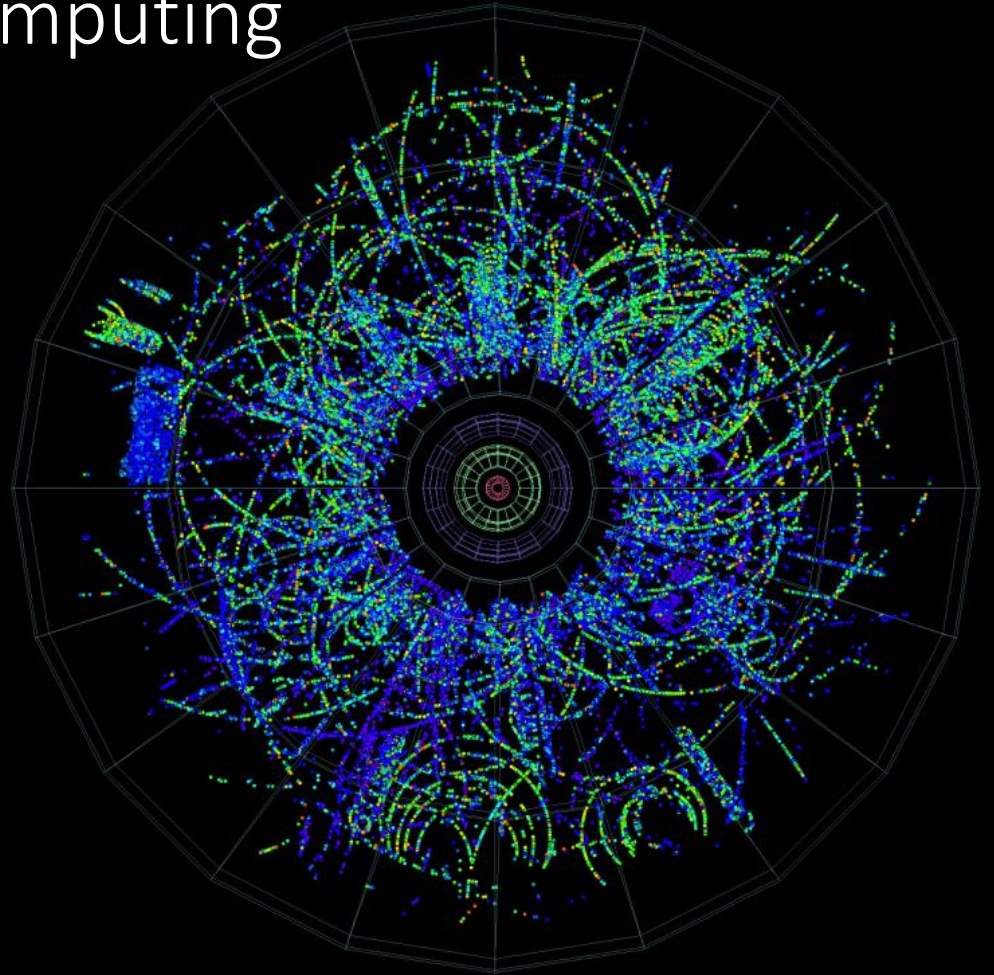


Operations

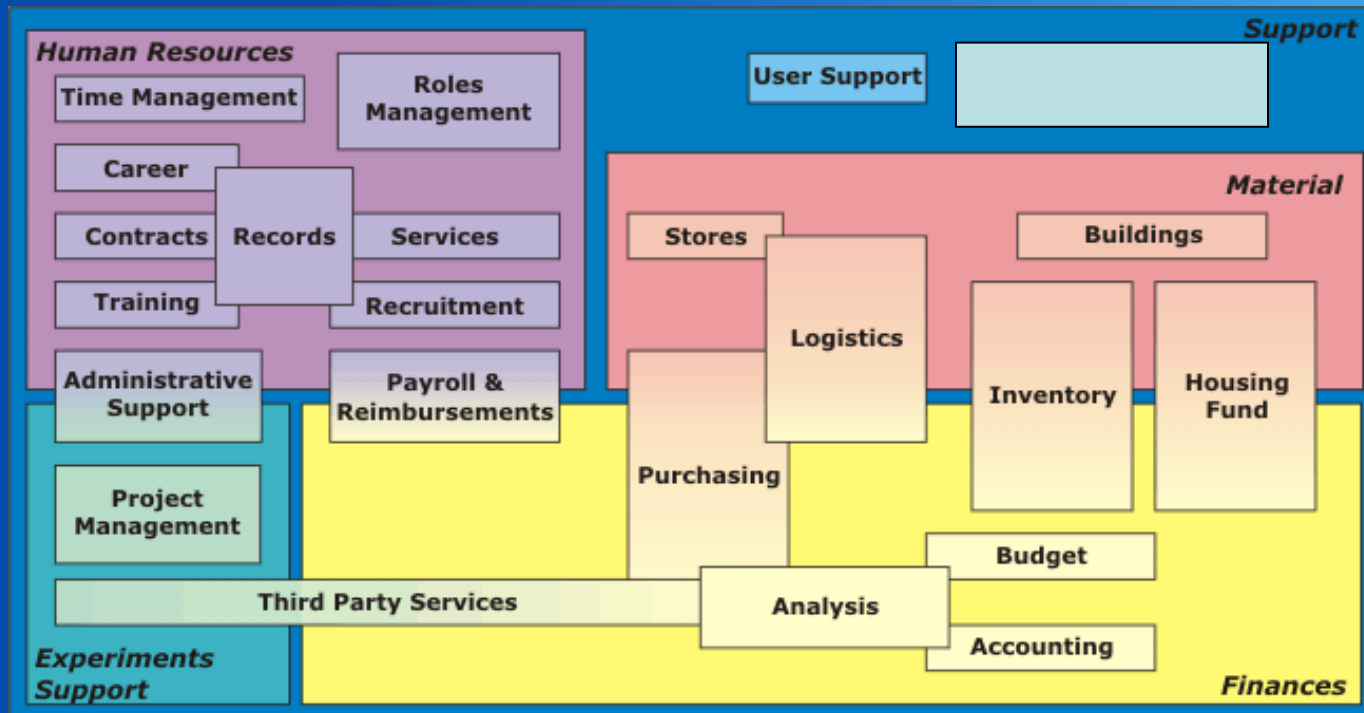


Computing

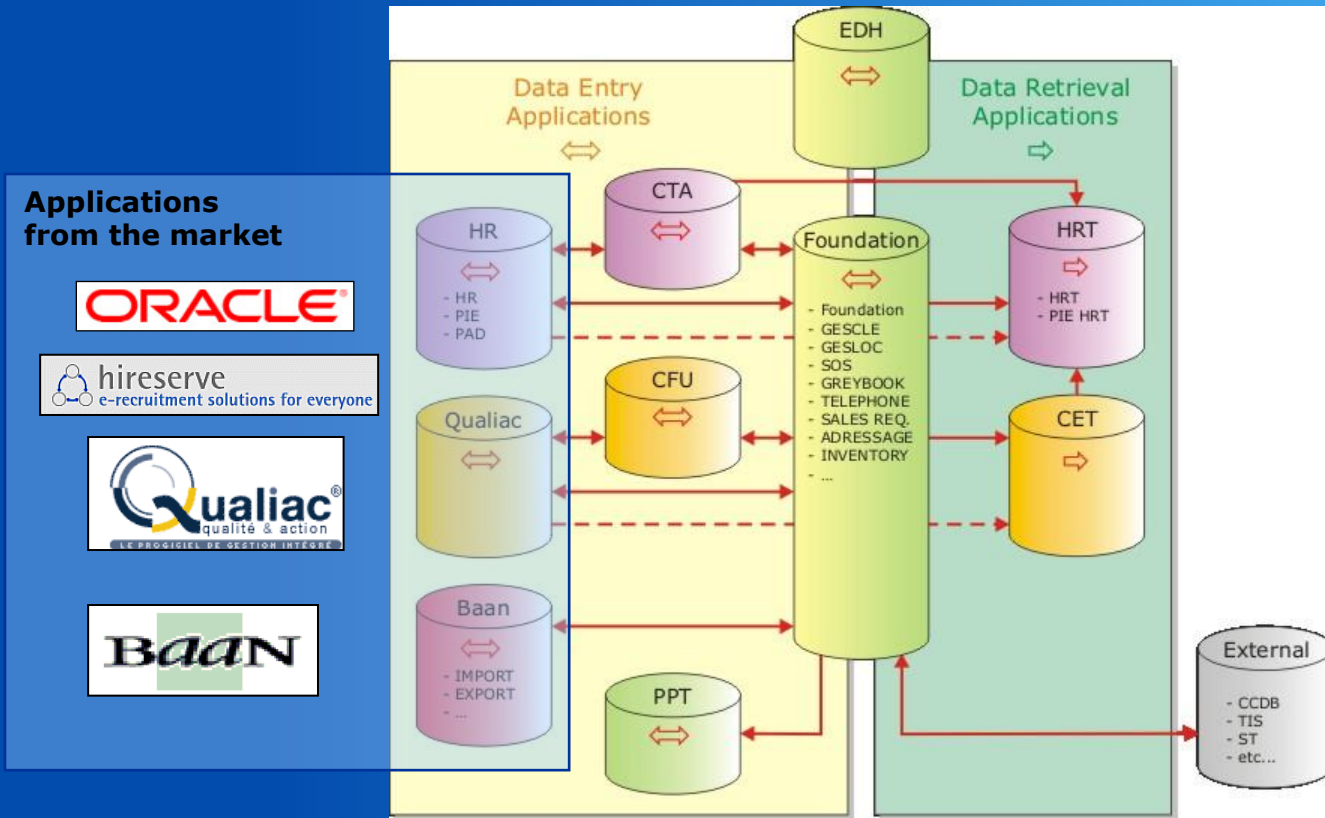
- Administrative
- Software



Administrative Information Services



Administrative Information Services



Conference, Meetings, Events...

700k events total

80k events in 2019

10k users /day

IT - Activities and Services Discussion Forum

Thursday 6 Apr 2017, 16:00 -> 17:20 Europe/Zurich
31-3-004 - IT Amphitheatre (CERN)
Emmanuel Ormancey (CERN)

Videoconference Rooms: IT_Activities_and_Services_Discussion_Forum

16:00 --> 17:20 Activity Reviews

- 16:00 DB-DBR: Working with Oracle 12c databases in the Oracle Cloud
Speaker: Eva Dafonte Perez (CERN)
- 16:05 CDA-DR: Enabling software citation and discovery workflows
Speaker: Lars Holm Nielsen (CERN)
- 16:10 CDA-DR: Migrating CDS content - from AFS to EOS aka from Invenio 1 to Invenio 3
Speaker: Ludmila Marian (CERN)
- 16:15 ST-TAB: New monitoring infrastructure for tape operations
Speaker: Julien Leduc (CERN)
- 16:20 CDA-IC: Videoconference and WebRTC development
Speaker: Mr. Marek Domaracky (CERN)

Room Booking

Monday 13 Aug 08:00-19:00

There are 15 rooms available for booking

- 6-2-024 - BE Auditorium Meyrin
- 13-2-005
- 14-4-010
- 14-4-022
- 18-3-008 - CLIC Meeting room
- 21-2-004 - IT Amphitheatre

newdle

Choose the time slots

Which will be presented as options to the participants

29 Jul 2019

Brandon Griffin

Kristian Turner

Payments

Payment methods

- PayPal
- Bank Transfer
- Credit Card

Exchange Intergration

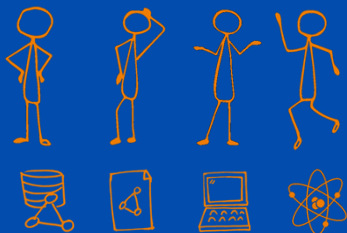
Pedro's test meeting

PDF conversion

Electronic payment Exchange Intergration PDF conversion

CERN Open Data

Education



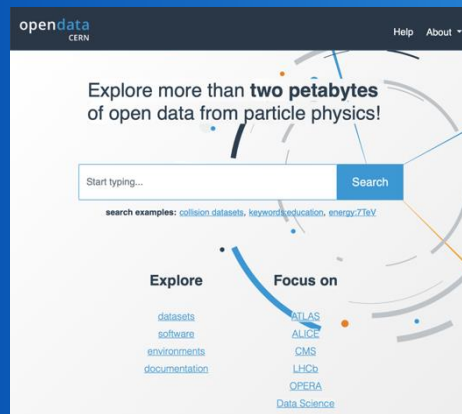
+



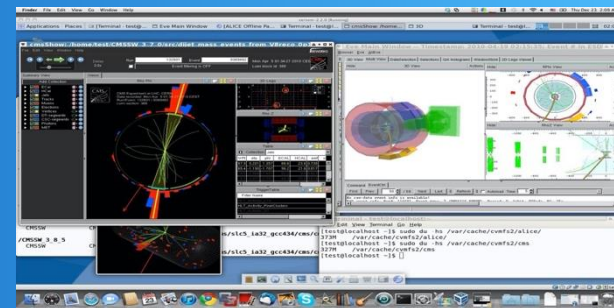
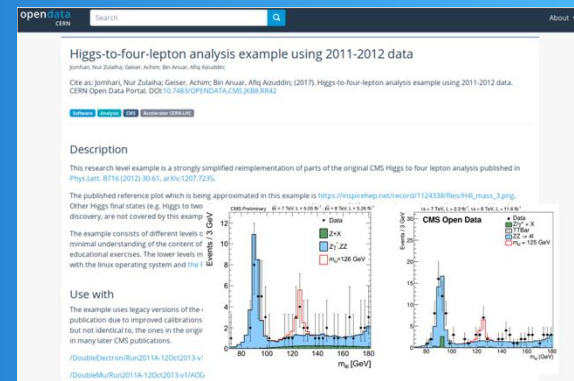
+



400 PB

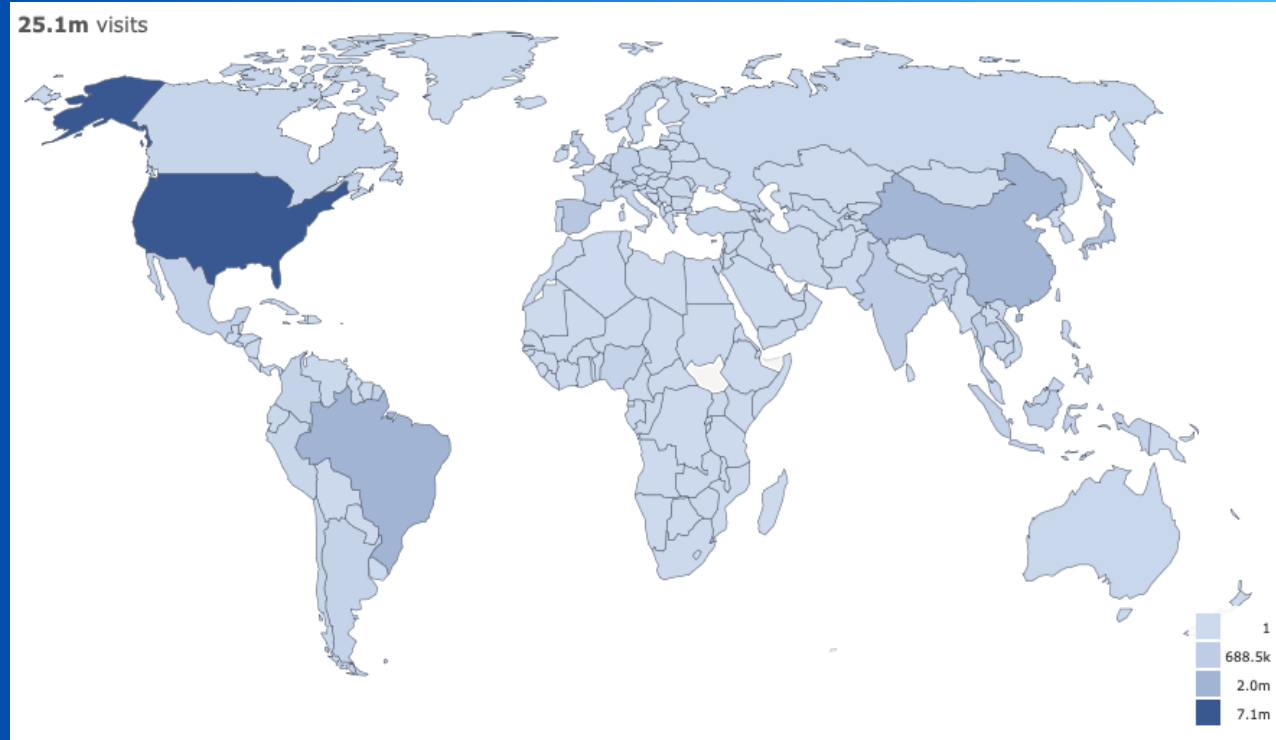
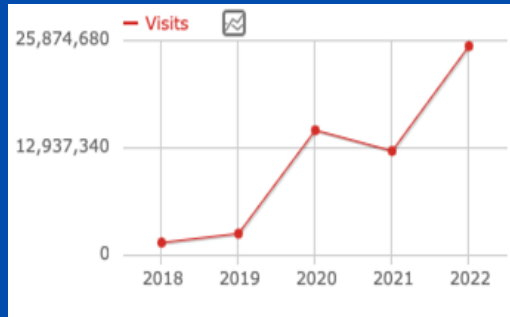


<http://opendata.cern.ch>



Research

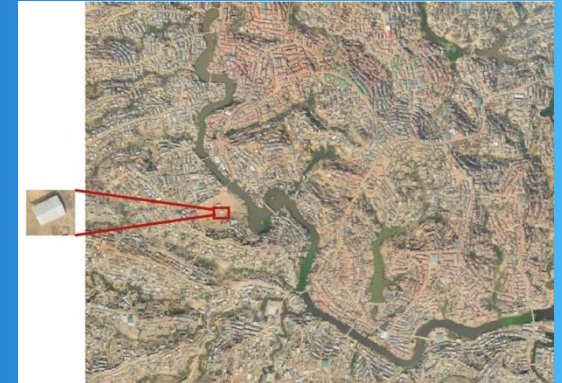
Zenodo: Open Science for All



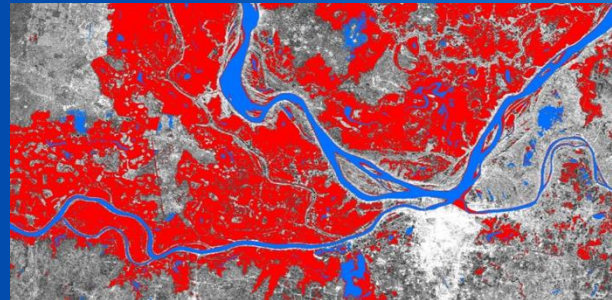
CERN as a host



UNOSAT
established at CERN in 2001,
based on IT infrastructure

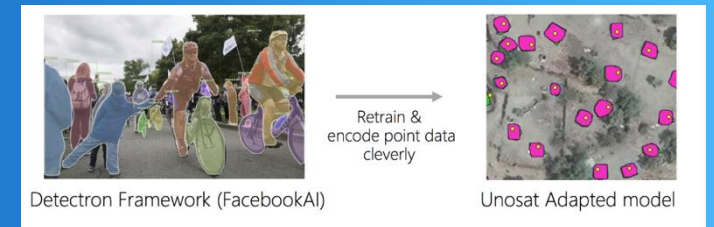


Flood detection



AI for Satellite
Imagery Analysis

Mapping shelters in refugee camps

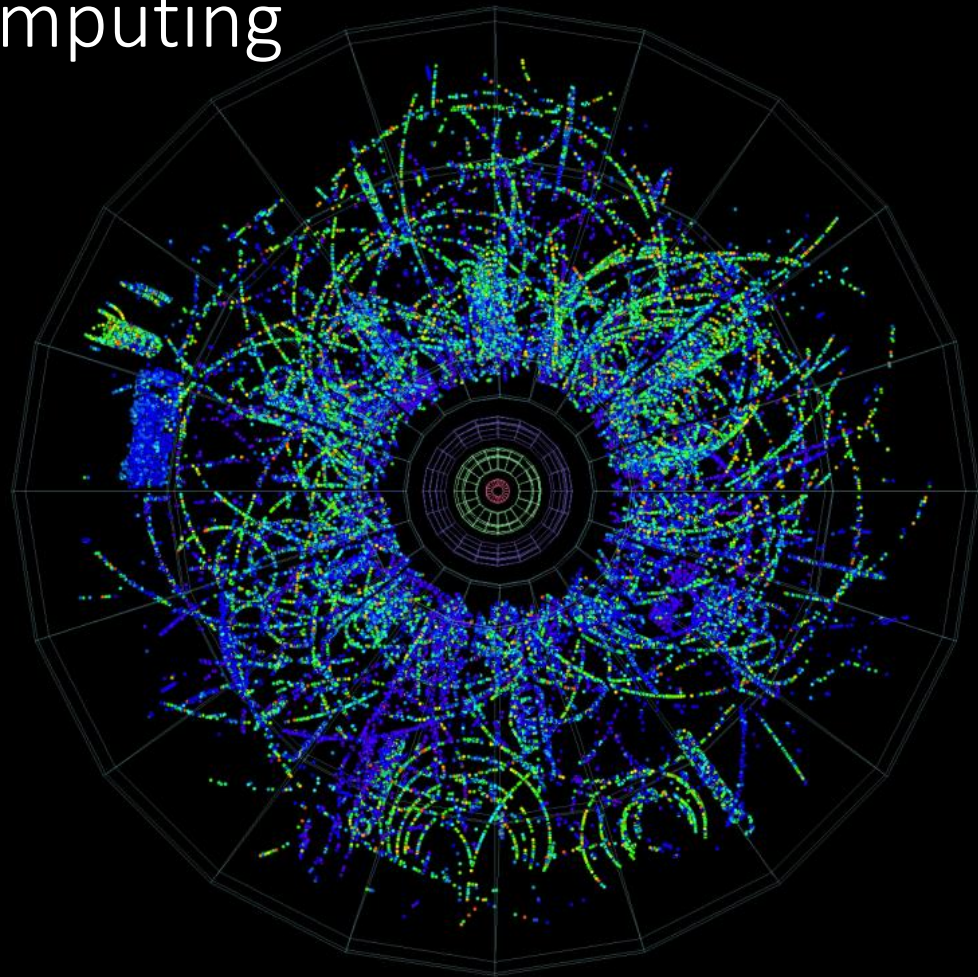


Detectron Framework (FacebookAI)

Unosat Adapted model

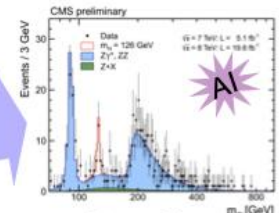
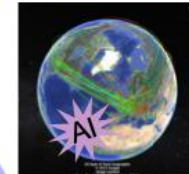
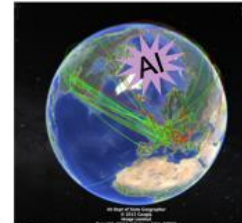
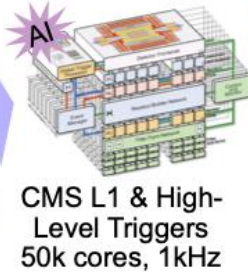
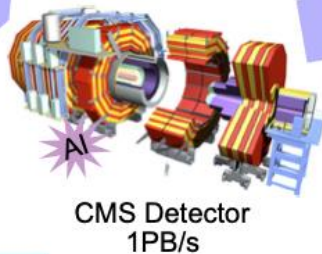
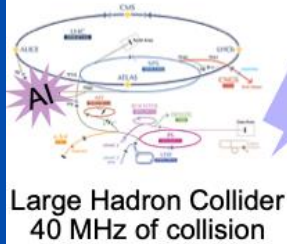
Computing

- Future

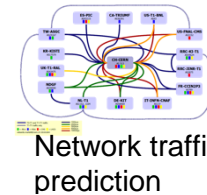


AI in HEP

Role of AI: accelerator control, data acquisition, event triggering, anomaly detection, new physics scouting, event reconstruction, event generation, detector simulation, LHC grid control, analytics, signal extraction, likelihood free inference, background rejection, new physics searches, ...



Rare Signal
Measurement
~1 out of 10^6



Thanks to J-R. Vlimant

Quantum Technology Initiative



- Assess the **areas of potential quantum advantage** in HEP applications (QML, classification, anomaly detection, tracking)
- Develop **common libraries of algorithms, methods, tools**; benchmark as technology evolves
- Collaborate to the development of shared, **hybrid classic-quantum infrastructures**

Computing & Algorithms



- Identify and develop techniques for **quantum simulation** in collider physics, QCD, cosmology within and beyond the SM
- Co-develop quantum computing and sensing approaches by providing **theoretical foundations** to the identifications of the areas of interest

Simulation & Theory



- Develop and promote **expertise in quantum sensing** in low- and high-energy physics applications
- Develop quantum sensing approaches with emphasis on **low-energy particle physics measurements**
- Assess **novel technologies and materials** for HEP applications

Sensing, Metrology & Materials



- **Co-develop CERN technologies relevant to quantum infrastructures** (time synchron, frequency distribution, lasers)
- Contribute to the **deployment and validation of quantum infrastructures**
- Assess requirements and **impact of quantum communication on computing applications** (security, privacy)

Communications & Networks

Open Quantum Institute Pilot Phase

**Global and inclusive access to quantum computing
and the development of applications for the benefit
of humanity**



An initiative hosted by CERN, born at GESDA, supported by UBS

Take-aways



Fundamental science continues 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** fosters technological advancements that percolates to society



Want to Know More?

- Contact:
 - Tim.Smith@cern.ch
- More information:
 - IT Department: <http://information-technology.web.cern.ch>
 - The LHC Grid: <http://wlcg.web.cern.ch>
 - Google Street view in CC:
 - https://www.google.ch/maps/@46.232624,6.045747,3a,75y,162.48h,90t/data=!3m5!1e1!3m3!1sBU7JKhoaY_H9JVPFHcH8JA!2e0!3e5?hl=en
 - <http://lego-scavenger-hunt.web.cern.ch>
 - IT Archives: <https://it-archives.web.cern.ch>

Want to Follow More?

Social Media at CERN

<http://twitter.com/CERN>

http://twitter.com/CERN_FR

<http://facebook.com/cern>

<http://google.com/+CERN>

<http://youtube.com/CERN>

<http://linkedin.com/company/cern>

Thanks to ...

Hannah Short, Lorena Lobato Pardavila, Xavier Espinal
for their suggestions & contributions