Computing at CERN
→Hannah Short,
Computer Security

Hello!

My name's Hannah

CERN Fellow in Computer Security, Brit/Irish, Scientific Computing Enthusiast!

hannah.short@cern.ch
@hannah.short08

+ 10 years! =

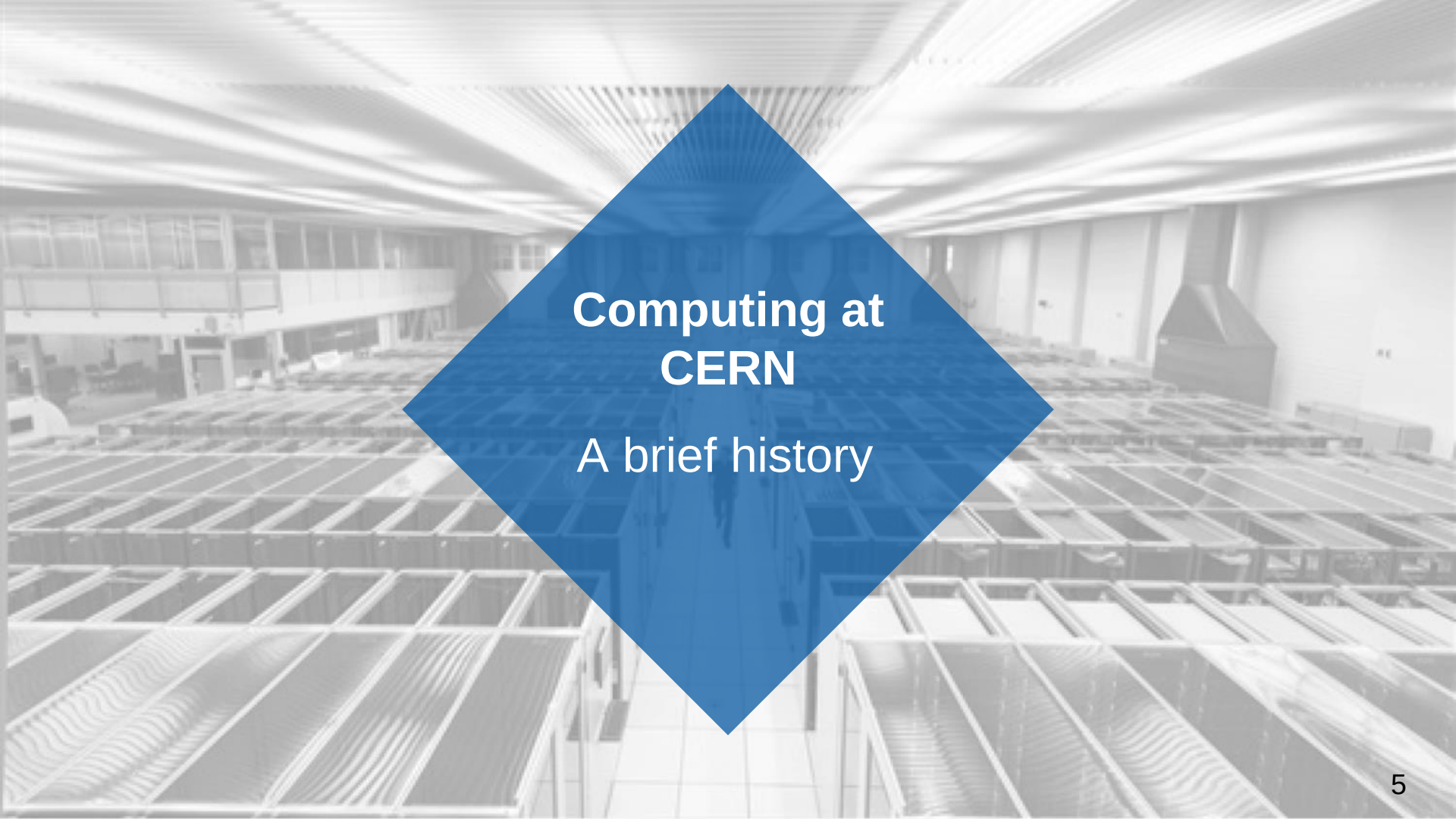
What will I talk about?

A brief history of Computing at CERN

The IT department (a small selection!)

- WLCG
- Innovation
- Collaborative tools

Born and bred at CERN



Computing at CERN

A brief history

computing, *n.*

Pronunciation:

Brit. /kəmˈpjuːtɪŋ/ , U.S. /kəmˈpjudɪŋ/

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.*

Credit to Tim Smith

An Early “Computer”

➜Wim Klein

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

➜Not the first CERN Computer! Two female computers were already working with mechanical calculators

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

1958, The Ferranti Mercury arrived

IBM 7090
DATA PROCESSING SYSTEM

C E R N C O M P U T E R N E W S L E T T E R

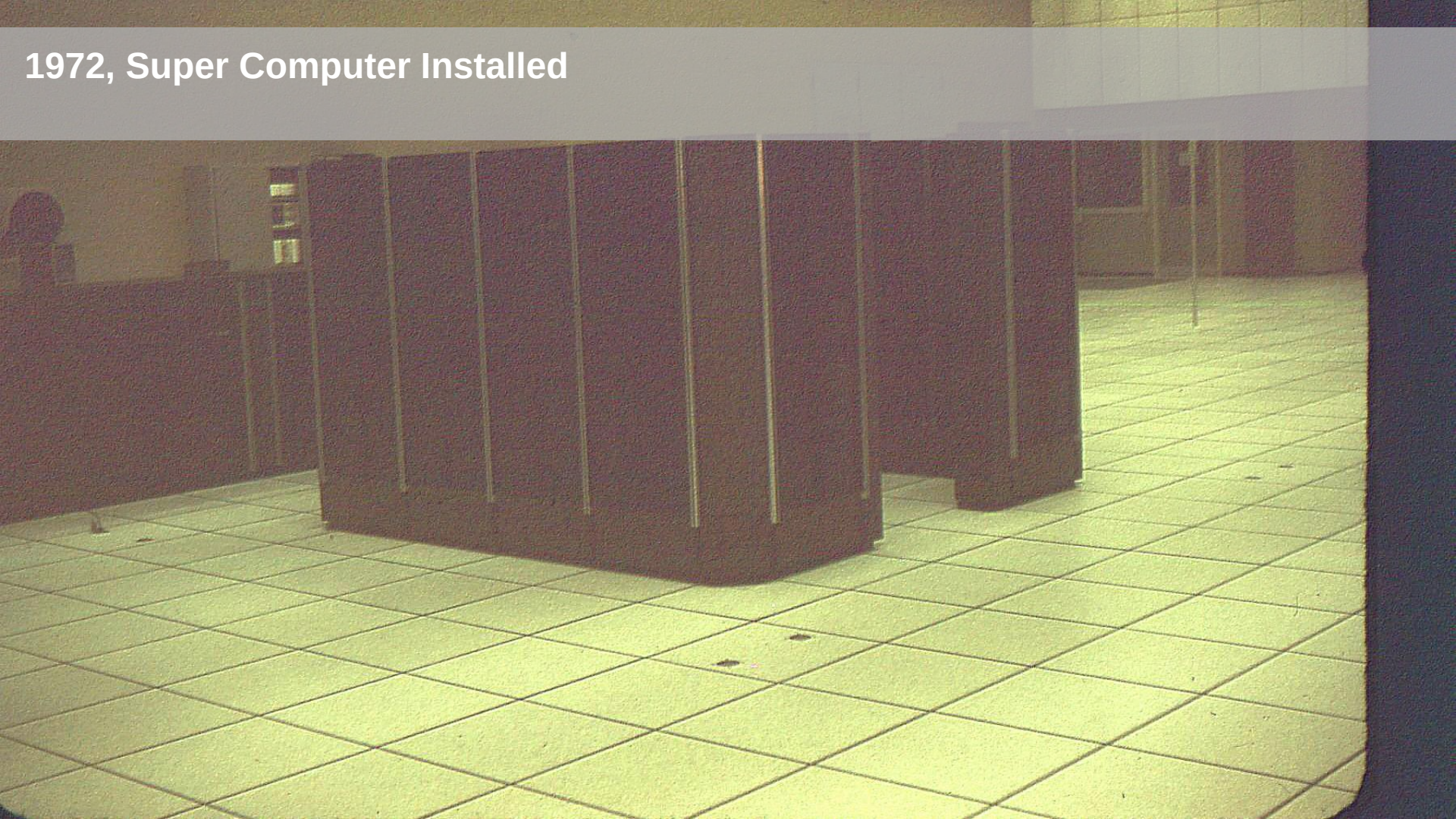
Number 1

15 February 1966

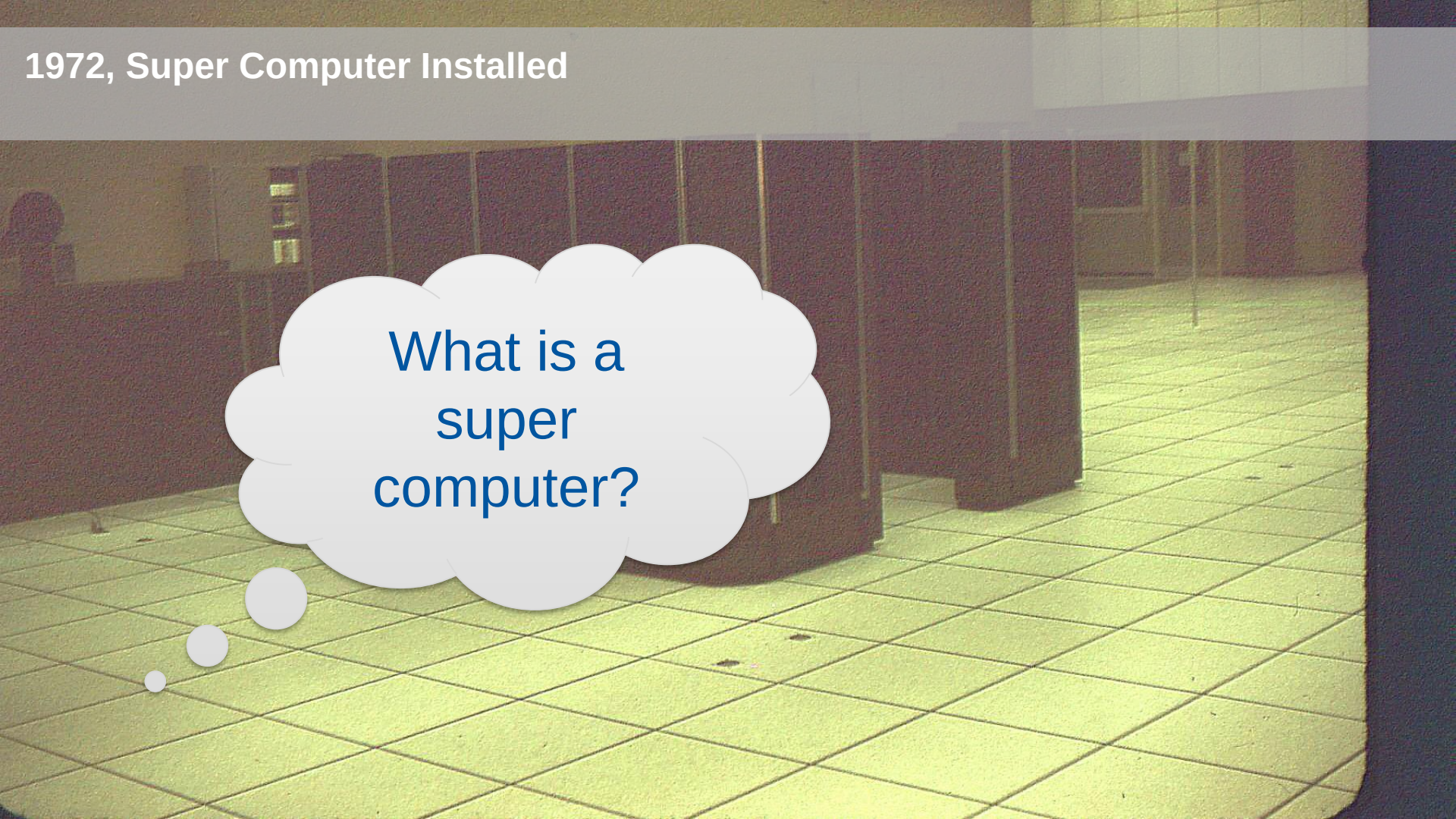
Introduction

As computing becomes a more and more widespread and complex activity in the laboratory, the need will increase for a means to have a wider general circulation of background information about different aspects of computing activities than is possible with the present system of Computer Notices. I therefore make no apology for introducing yet another circular which will find its way on to CERN desks. Rather I would express the hope that this newsletter will prove to be a useful source of general information on computer use and performance, programming developments and the requirements of different kinds of computer users, as well as on future plans for computers, programming and computer uses in the laboratory. The newsletter will be

1972, Super Computer Installed



1972, Super Computer Installed



Tapes being sent up from B513 basement

7-1-83
Bob Marley live
SIDE
60 MIN
PER SIDE
SEL
SOUND TECH
120 MIN.

2016, Today's Data Centre

The IT Department

The IT Department

Department Infrastructure

WLCG

openlab

Security

EC Projects

Collaboration, Devices & Applications

Storage

Databases

Compute & Monitoring

Communication Systems

Computing Facilities

The IT Department

Department Infrastructure

WLCG

openlab

Security

E

We are responsible for Computer Security across CERN… pretty good idea of what is going on!

Collaboration, Devices & Applications

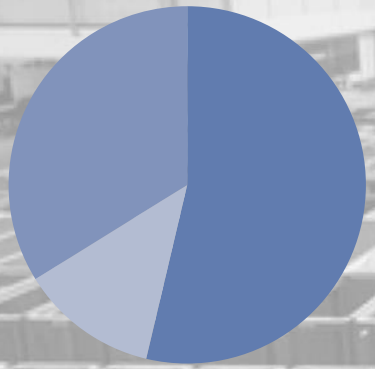
Storage

Compute & Monitoring

Communication Systems

Computing Facilities

Who are we?



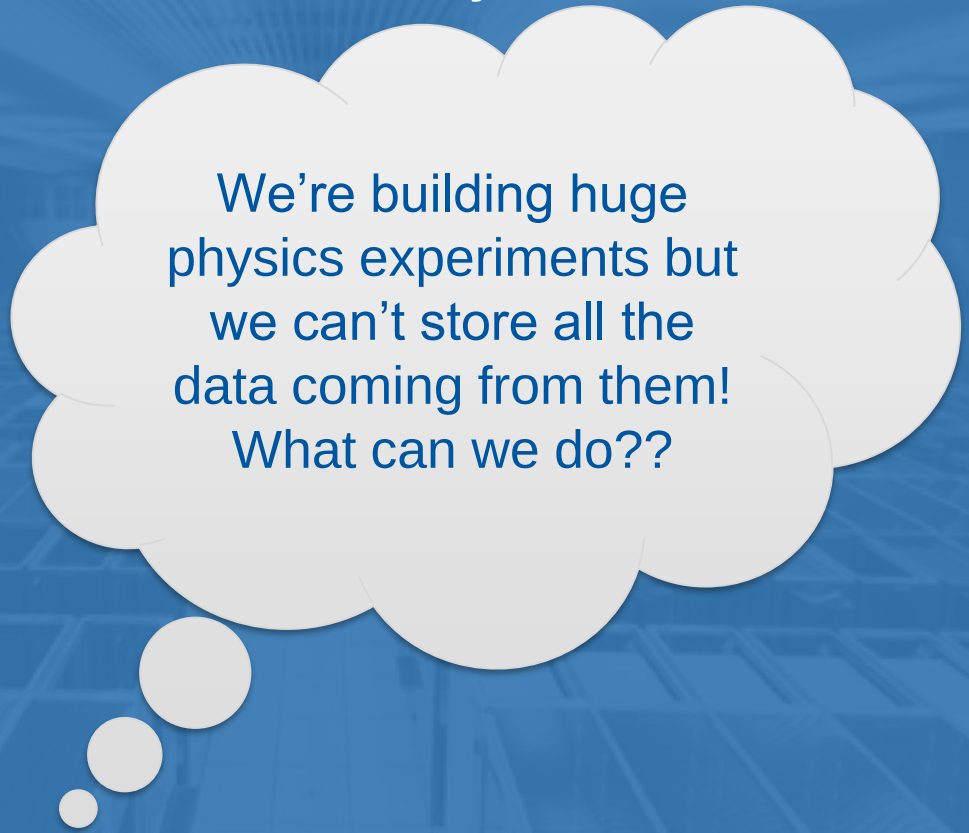
➔~400 people in IT
➔+ Many computing experts in other departments

https://cds.cern.ch/record/2265782/files/CERN-HR-STAFF-STAT-2016.pdf

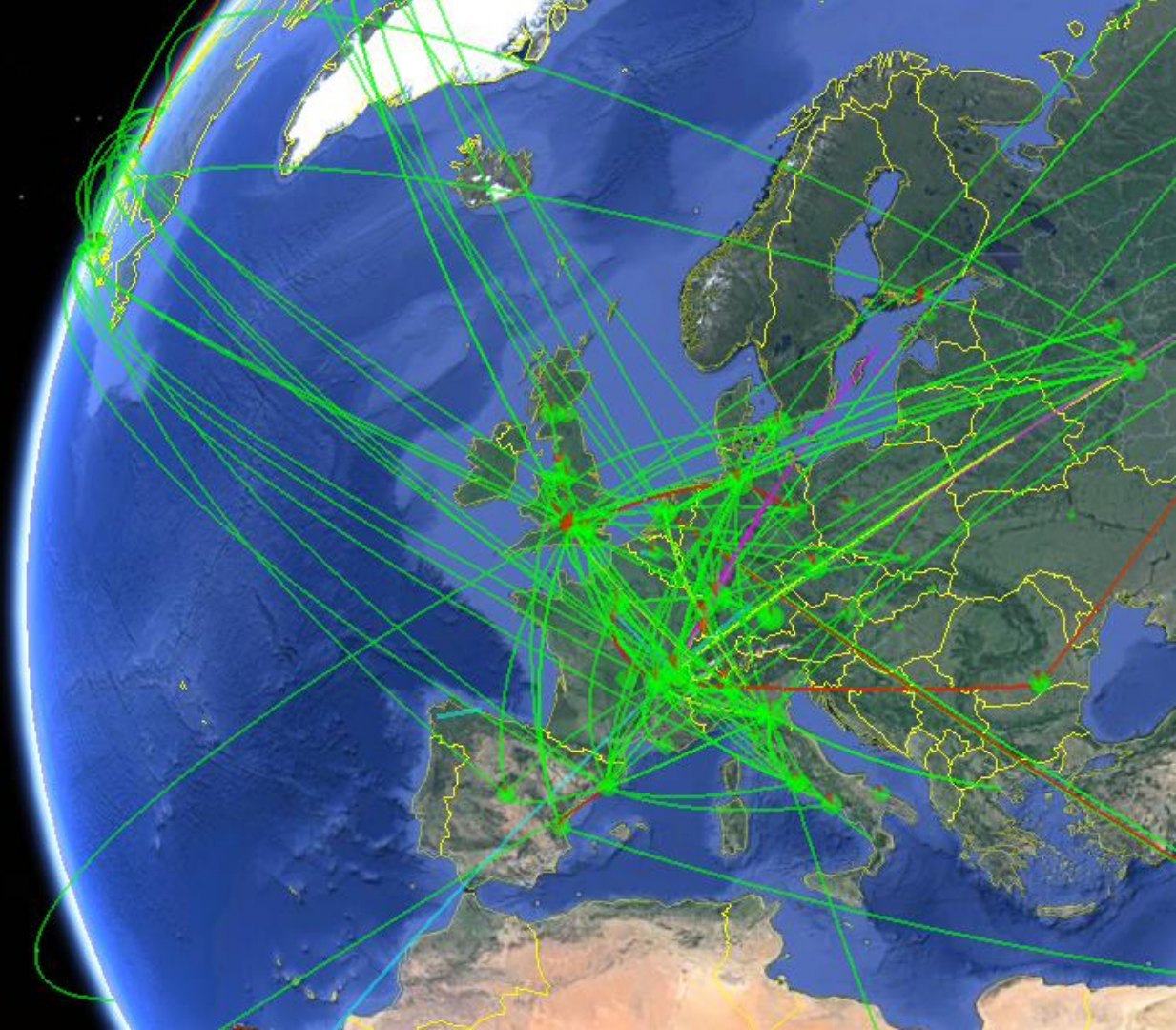
The role of the IT Department is to keep providing an excellent level of services while scaling up to the foreseen resource levels at an affordable cost

WLCG

WLCG… back in early 2000s



WLCG: The Worldwide LHC Computing Grid



The IT Department

Department Infrastructure

WLCG

openlab

Security

EC Projects

Collaboration, Devices & Applications

Storage

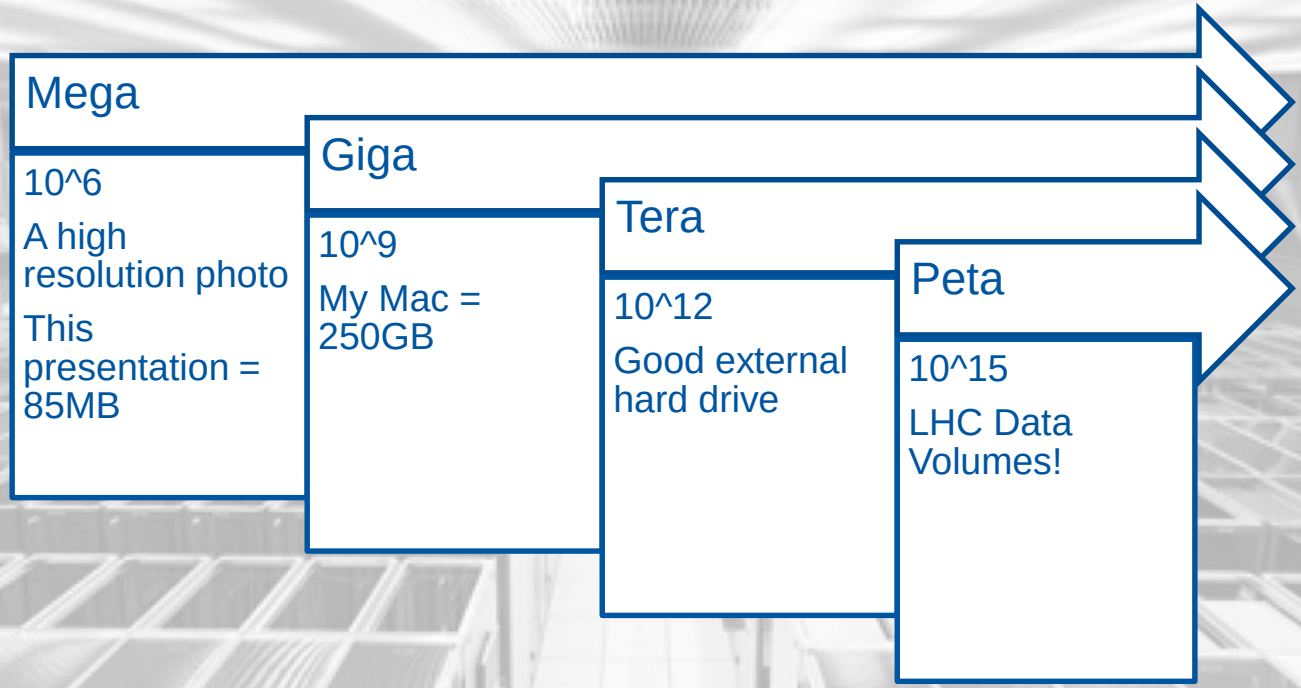
Databases

Compute & Monitoring

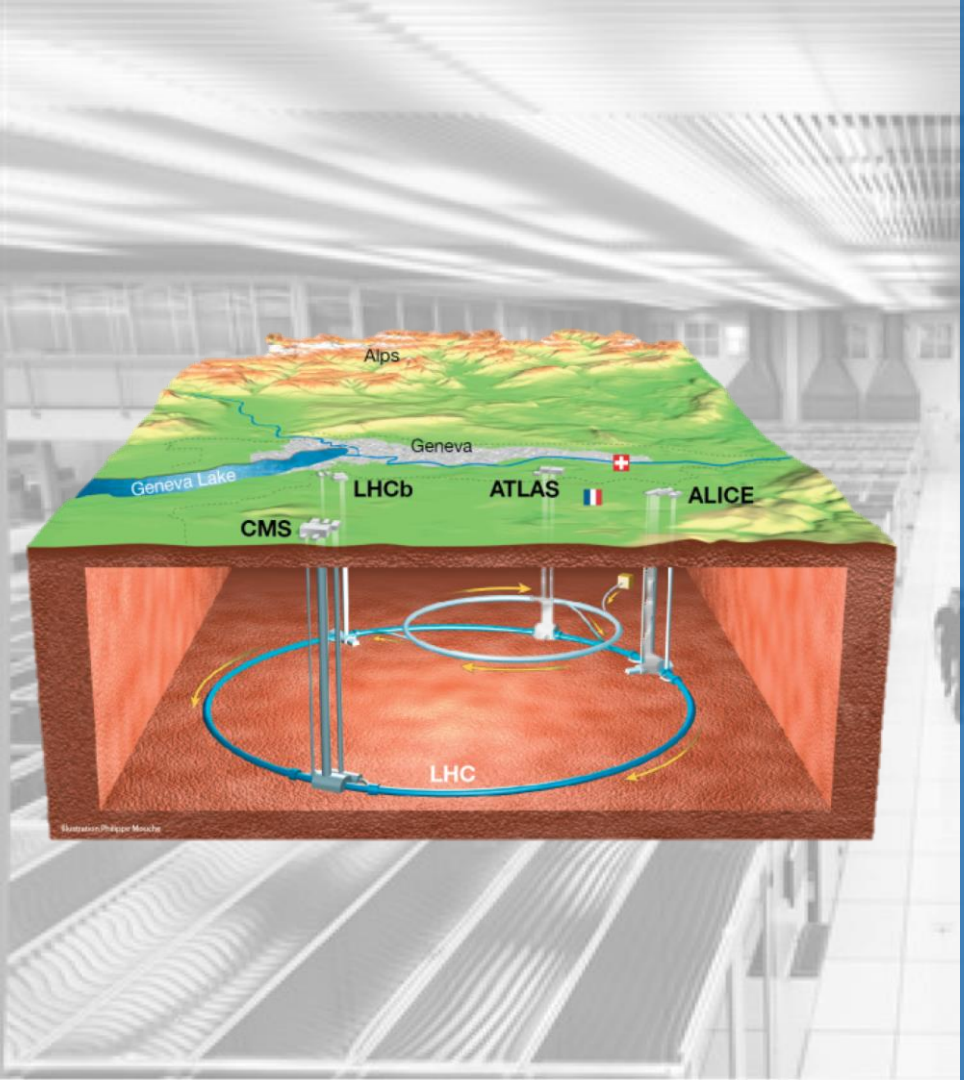
Communication Systems

Computing Facilities

Data Volumes

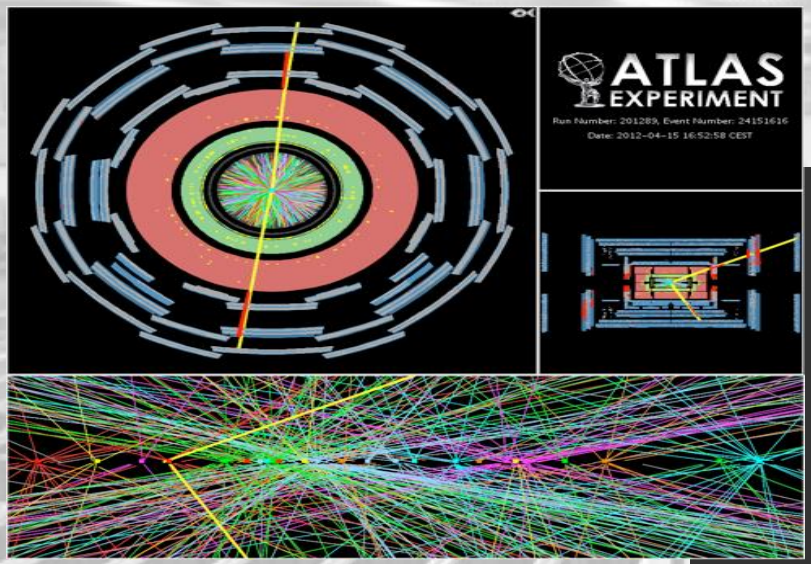


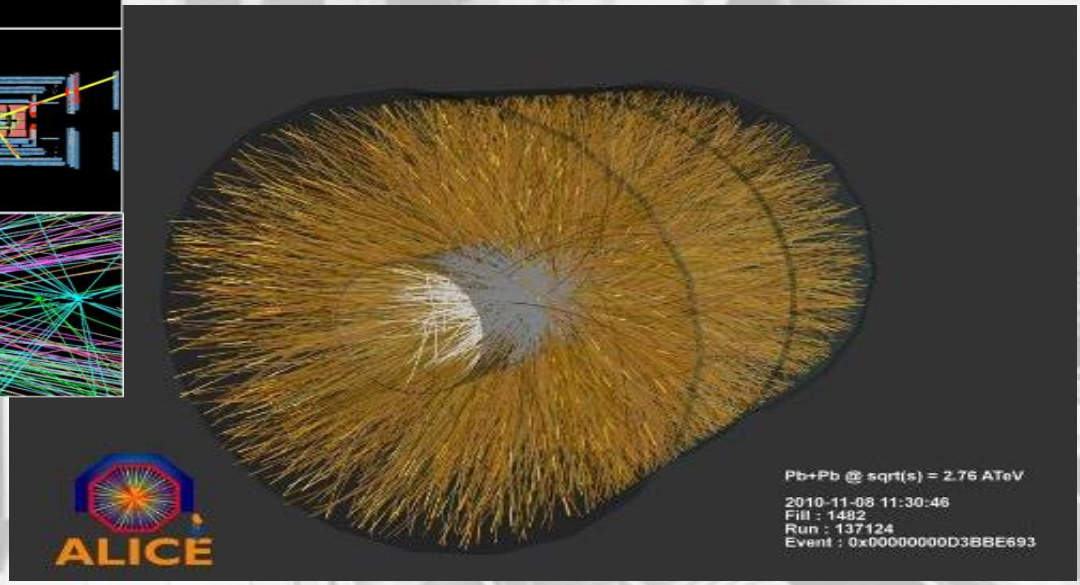
Where does the data come from?



➔Large Hadron Collider
➔(Mostly) 4 Experiments, ATLAS, ALICE, LHCb and CMS

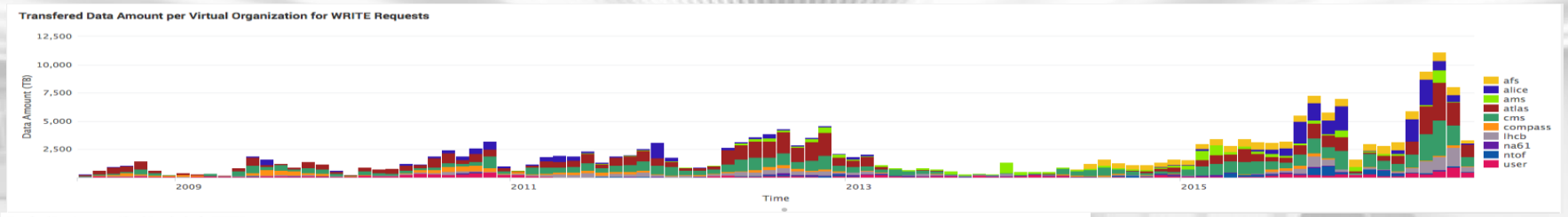
1PB/Second!

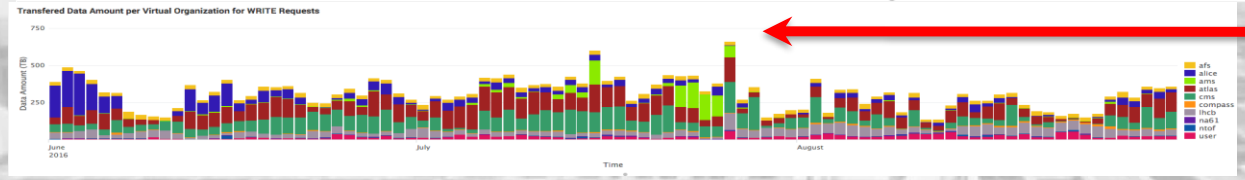




Credit to Lorena Lobato

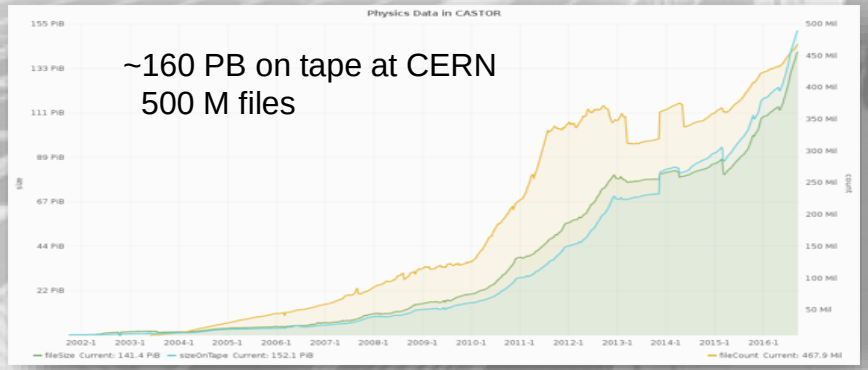
2016 LHC Data





June-Aug 2016 >500 TB / day

(Run 1 peak for HI was 220 TB)



LHC data – Continue to break records:
10.7 PB recorded in July 2016

Credit to Lorena Lobato

Breaking News!

Melissa Gaillard shared CERN Updates's post to the group: IT-Dep.
17 hrs ·

CERN Updates shared a link to the group: CERN Updates.
17 hrs ·

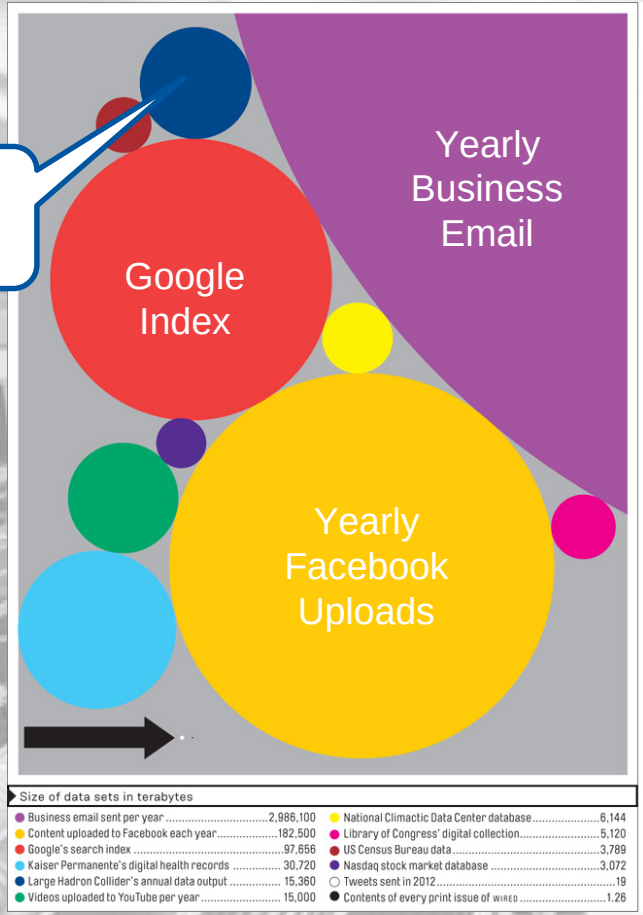
CERN Data Centre passes the 200-petabyte milestone | CERN

Where do these data come from? Particles collide in the Large Hadron Collider (LHC) detectors approximately 1 billion times per second, generating about...

HOME.CERN

Like Comment

How much data are we talking? (2012)



→2012, 15 PB
→2017 estimates 50 PB, equivalent to a 12km high stack of DVDs
→CERN can only provide 20%-30% storage and CPU

https://www.wired.com/2013/04/bigdata/

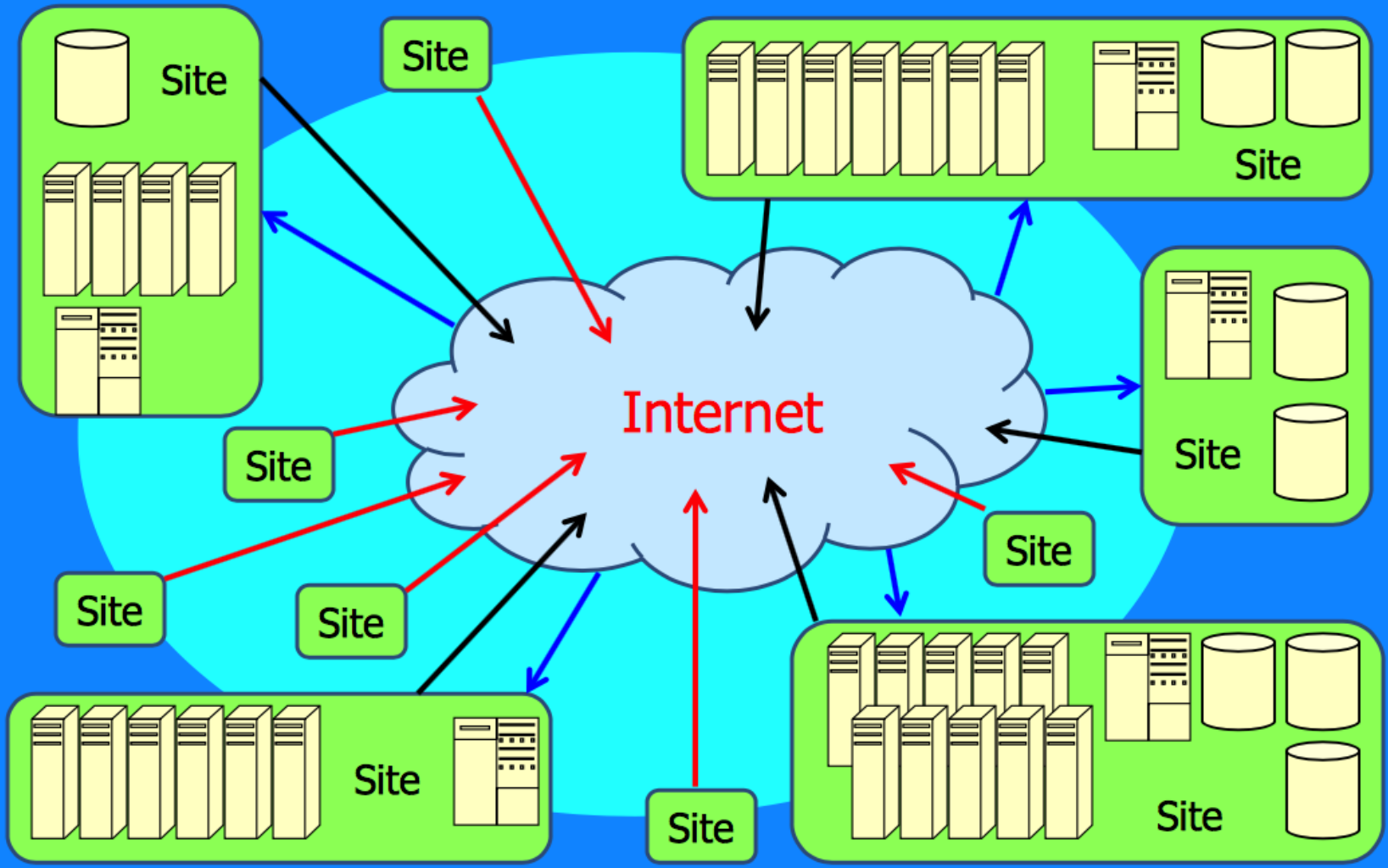
Distributed Computing

- ➜Collection of independent computers
- ➜Appear as a single system
- ➜Benefits
  - ✓ Continuous availability
  - ✓ Scalability

What is a grid?

➜*“A grid gives selected user communities uniform access to distributed resources with independent administrations”*
➜Like a power grid, you don’t need to know where the power comes from!
➜Don’t need to know where your computing is done
➜Don’t need to know where your data is stored

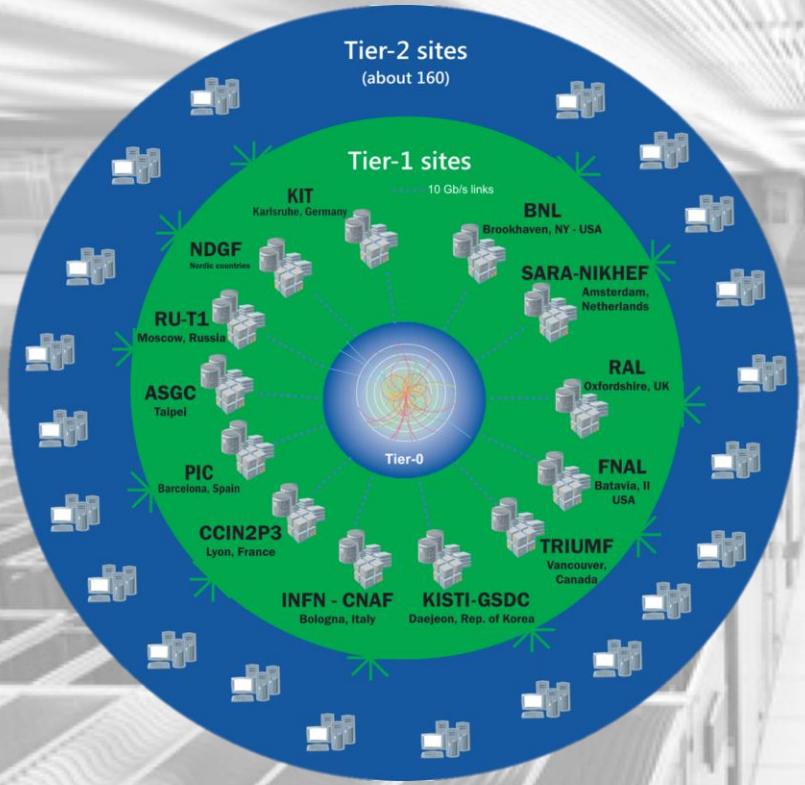
https://espace.cern.ch/cern-guides/Documents/WLCG-intro.pdf


Site
Site
Site
Site
Site
Site
Site
Site
Site
Site
Site
Internet

Isn’t this like a cloud?

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

Grid
➜Fixed size
➜Collaborative, run by community



Where is the data stored?

➔Tier 0
CERN & Budapest
➔Tier 1
13 large centres
24/7 Support
➔Tier 2
~160 smaller universities and institutes
➔Tier 3
Individuals, accessing the grid

CERN Meyrin Data Centre

http://goo.gl/maps/K5SoG

Credit to Lorena Lobato

Wigner Data Centre, Budapest

Credit to Lorena Lobato

3 100GB/s Connections

Credit to Lorena Lobato

Credit: Garwin Liu, Brian Nisbet

What has WLCG achieved?

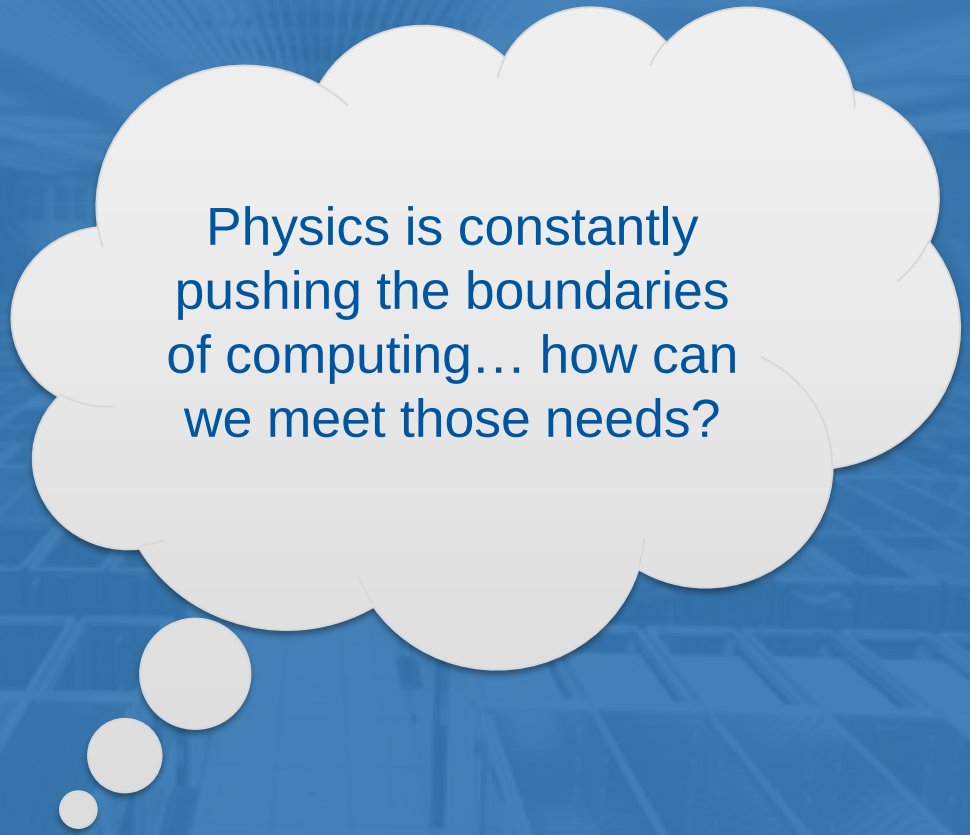
➜Unprecedented **speed and volume** of data processing

➜Analysis of **billions of collisions within weeks** to find the Higgs signal (Nobel Prize 2013)

➜Successful **collaboration** of diverse countries, organisations and people! 42 countries and 2 million jobs per day

Innovation

Collaborative tools



Innovation

→We need to be at the edge of commercial and academic developments!
→Openlab = Commercial Innovation
→EC Project = Research & Education Community Innovation

The IT Department

Department Infrastructure

WLCG

openlab

Security

EC Projects

Collaboration, Devices & Applications

Storage

Databases

Compute & Monitoring

Communication Systems

Computing Facilities

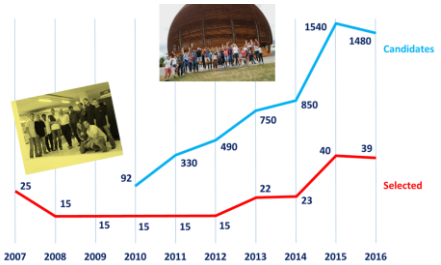
Openlab

➜*“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

Openlab R&D and Innovation

Seminars, training courses, academic training

CERN openlab Summer Students Programme

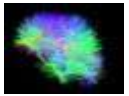


Joint R&D | Education | Management | Communication | Innovation and Knowledge Transfer

- 200+ press cuttings
- 150,000 visits to our website
- 50+ events, visits, lectures
- 100+ presentations
- 50+ news articles, press releases, case studies

15 years CERN openlab

Innovation & Entrepreneurship



BioDynaMo
The Biology Dynamic Modeller

GeneROOT

Applications to cross-disciplinary research

European Commission projects

On-going

HELIX NEBULA THE SCIENCE CLOUD

EGI

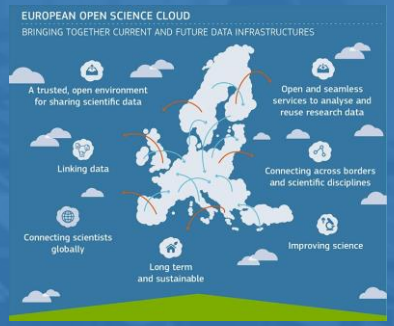
EUDAT

INDIGO DataCloud

AARC

OpenAIRE

Future



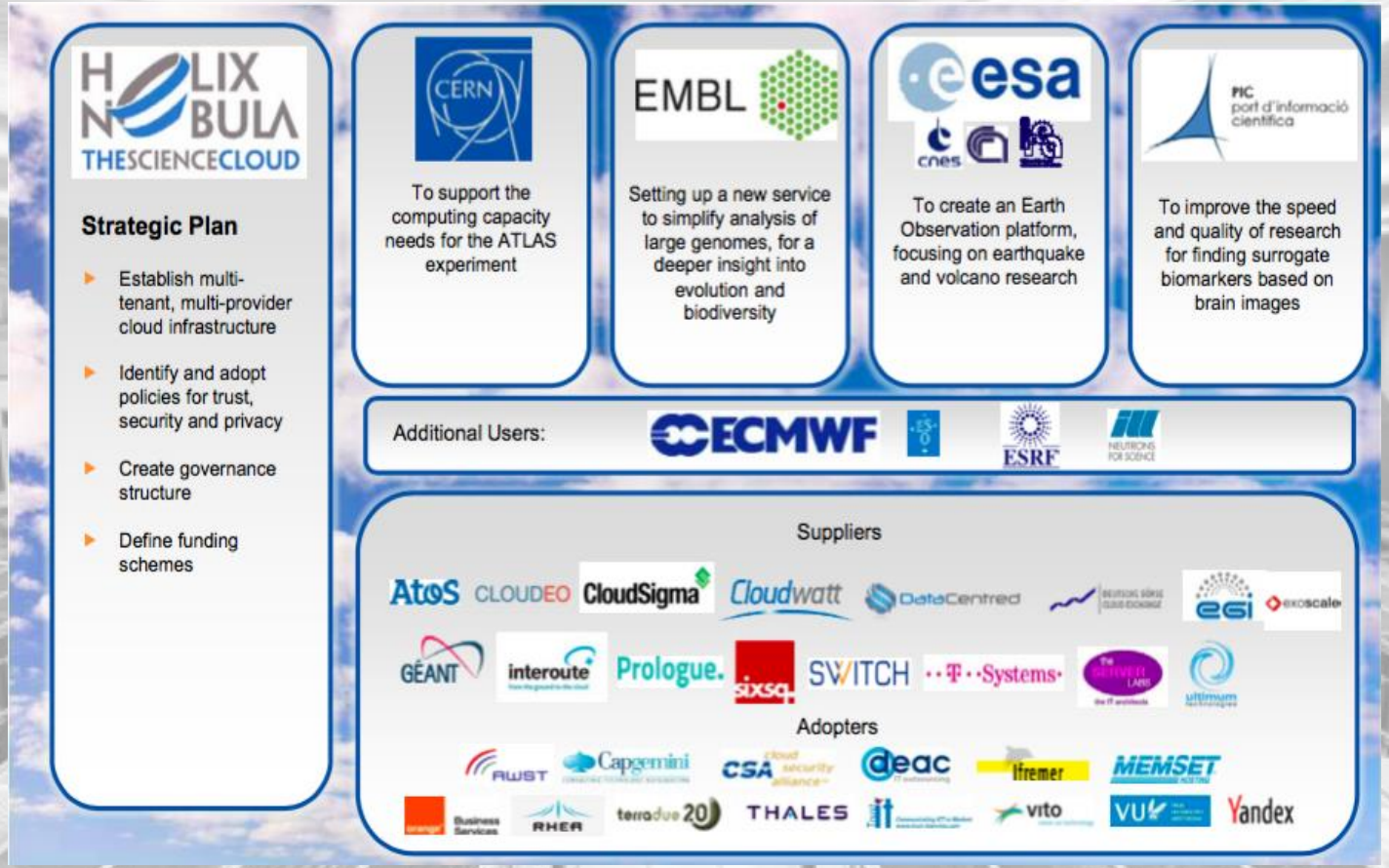
Bridging the gap between schools and universities through informal education

UP 2 U

https://up2university.eu

Challenges in EC Projects

Some of the many themes address:
➜On-Premise Vs. Public Clouds
➜Supporting the Long Tail of Science (LTOS)
➜Trust and collaboration



The HNSci Cloud public-private partnership

Open Science: Zenodo

Visitors from ~ all Countries
- Including
  - Antarctica
  - Vatican City
- 56% from Europe

57k Records
- 11k Software
- 3k Datasets

700 Communities
- Projects
- Institutes
- Subjects
- Conferences
- Publishers

→Infrastructure
Runs on 30 VMs in CERN Cloud

→Impact
Biggest issuer of DOIs for SW in world
Reference material for publications
F1000, Wiley, eLife, PLoS, Elsevier, Nature, etc
Recommended by EC and National programmes

→Supports LTOS & large groups

What's next?

Collaborative Tools

Collaborative tools

The IT Department

Department Infrastructure

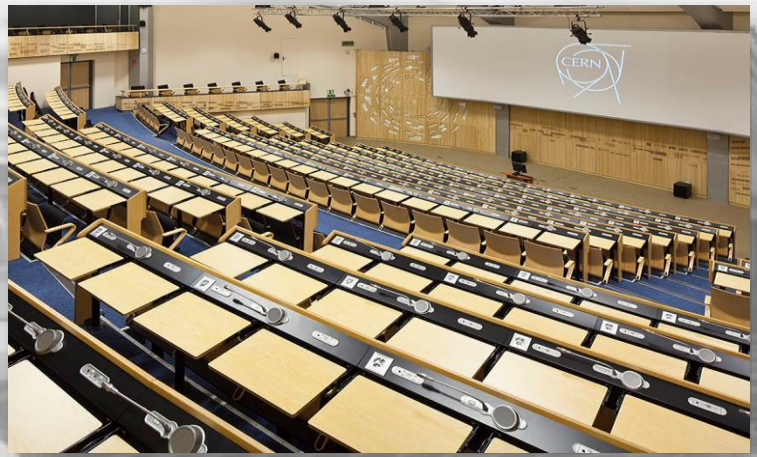
WLCG | openlab | Security | EC Projects

Collaboration, Devices & Applications

Storage | Databases

Compute & Monitoring

Communication Systems | Computing Facilities



Videoconference

→250 meeting rooms of all sizes on site

100 equipped for video conference

Legacy + VidyoPanorama

16 equipped for VC + Webcast

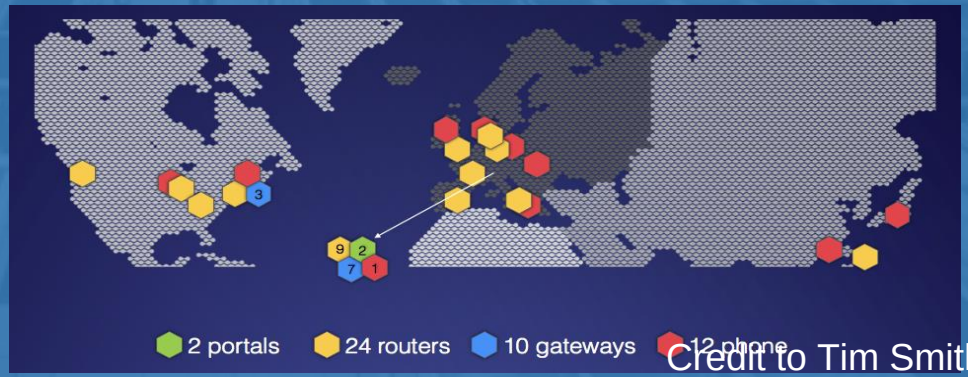
→500 legacy endpoints worldwide

Non centrally managed

Credit to Tim Smith

CERN Vidyo Worldwide Service Topology

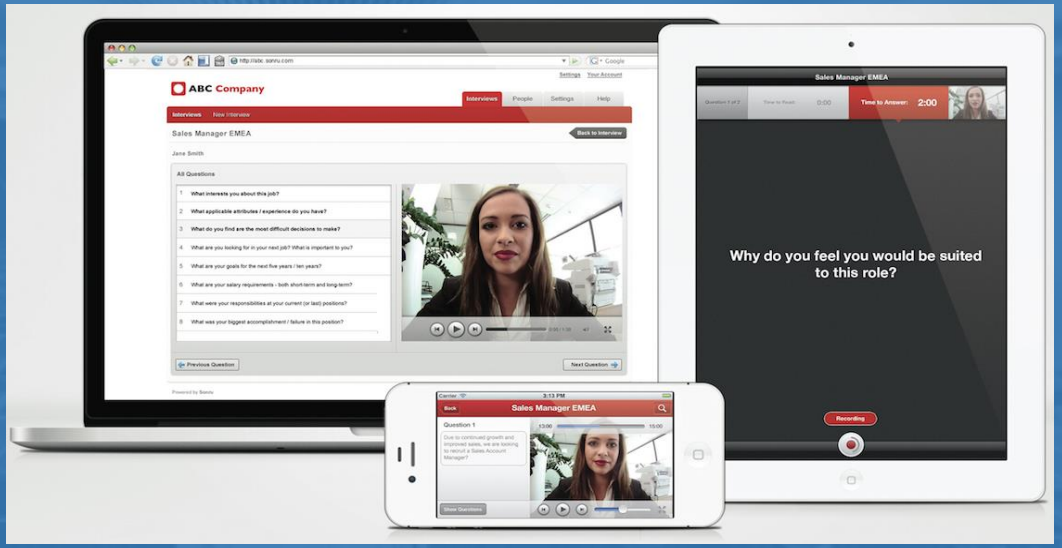
→8184 meetings/month
→941 simultaneous connections
→252 in one meeting
→50M minutes last year / 40k downloads



Credit to Tim Smith

Credit: Steven Goldfarb

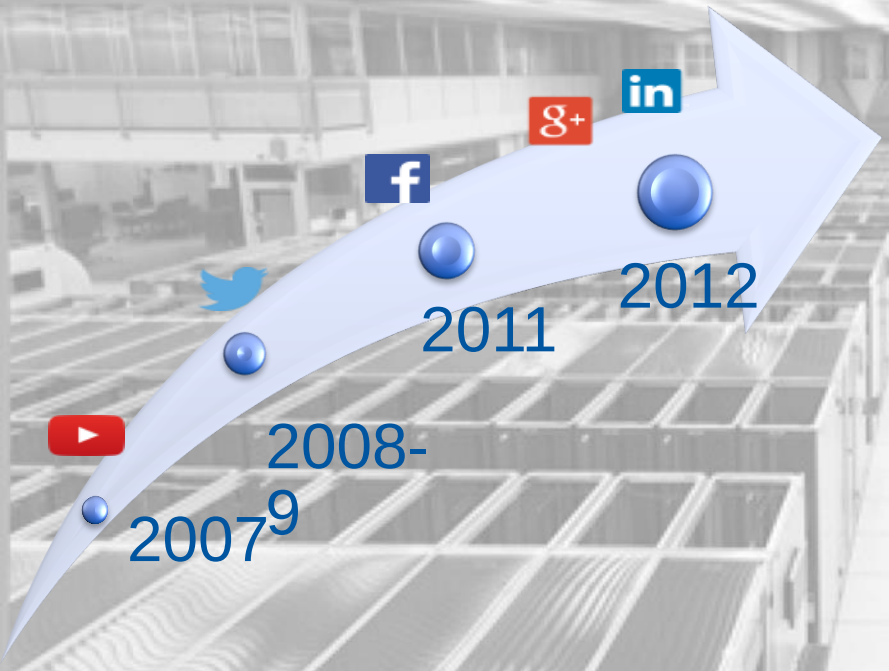
Recruitment



➜Asynchronous video screening
➜Cost savings in bringing people to interview
➜Multi-lingual – recruit from over 20 countries

Credit to Tim Smith

CERN's social media



975K
309K
92K
40K
20K
12K

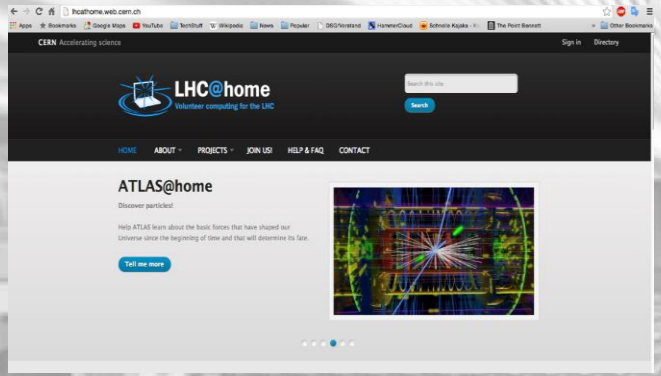
Credit to Tim Smith

Impact

Courtesy of Twiplomacy, November 2013

Credit to Tim Smith

Volunteer Computing



BOINC

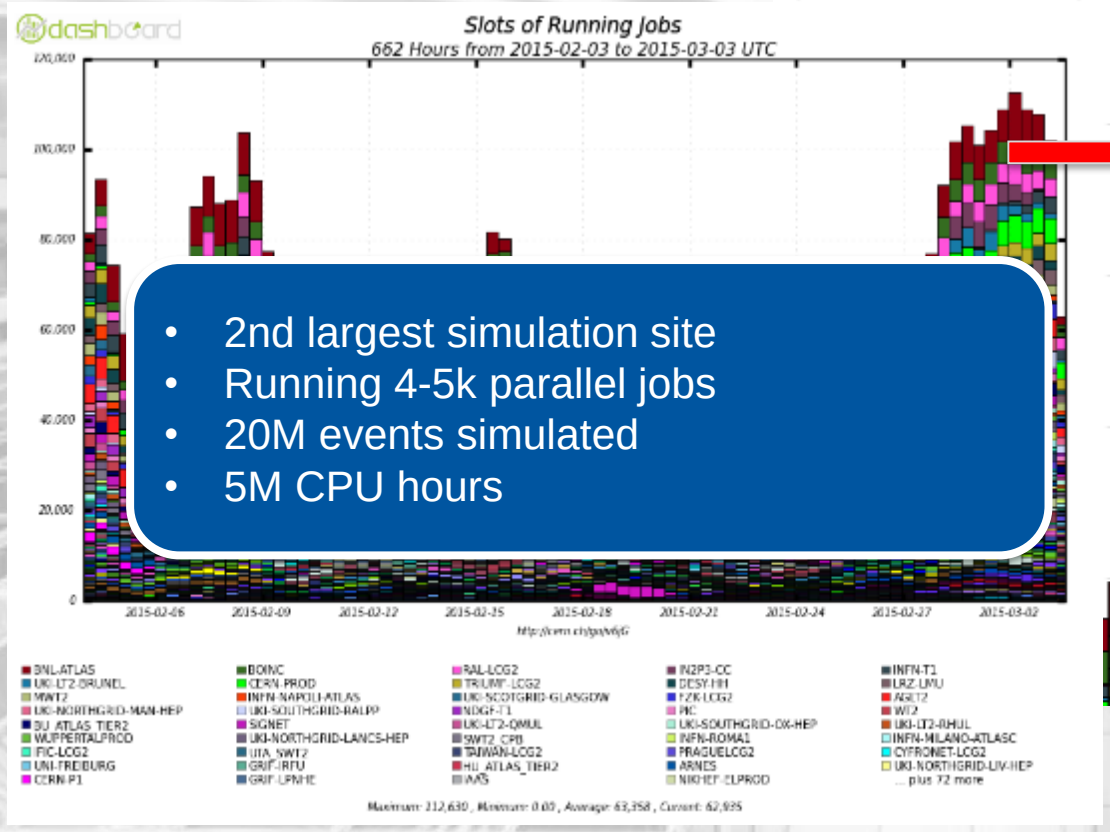
Scavenged resources
➜Volunteers (e.g. home PCs)
➜Institute desktops
➜Even mobile phones!

Unpredictable but significant resources
➜Target CPU bound simulations (not data intensive)
➜Over 50% of LHC compute is simulation!

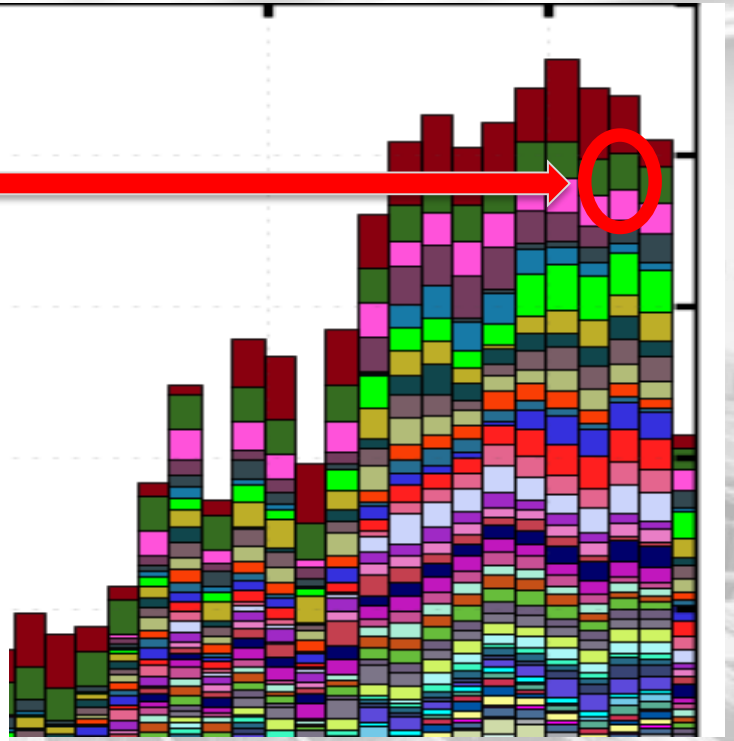
Outreach benefits, LHC@Home

http://lhcathome.web.cern.ch

Volunteer Computing



- 2nd largest simulation site
- Running 4-5k parallel jobs
- 20M events simulated
- 5M CPU hours

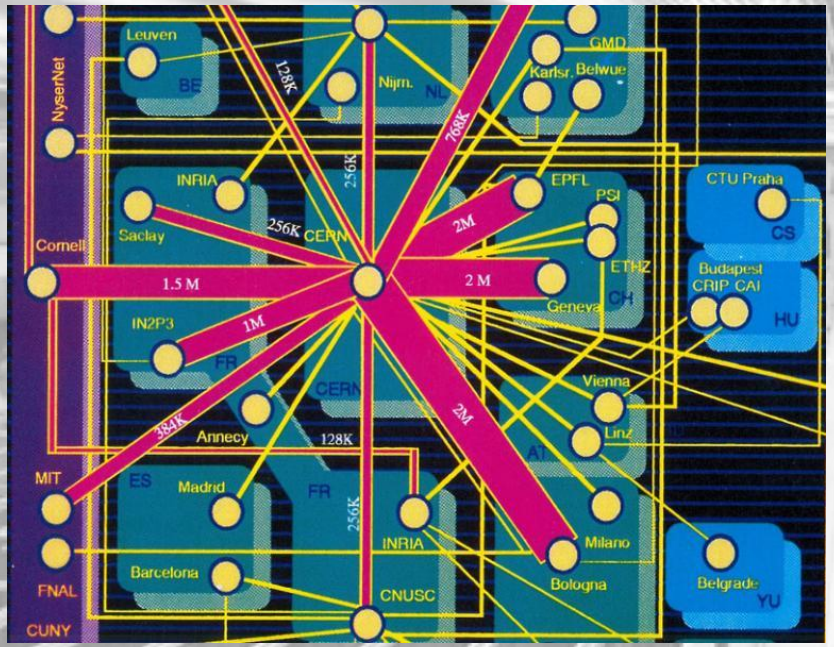


Born and bred at CERN
Technical Advances

CERN’s influence in computing

➜Several inventions directly from CERN
➜Several evolved at CERN
➜Direction of scientific computing strongly influenced and continues to be so!

The Internet

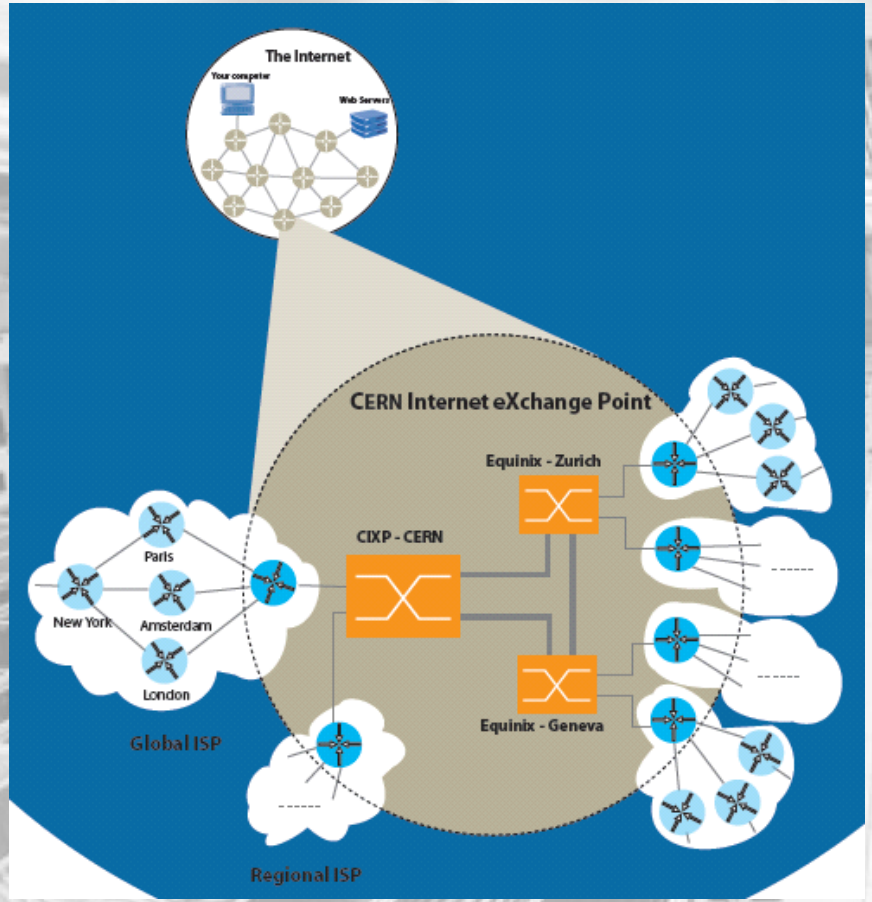


→CERN was one of the early European adopters of the internet

→1991 80% of internet capacity in Europe!

→CERN contributed to standardization

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



The Internet

→HEP centres set up links to enable data sharing
→Data exchange across the iron curtain
→1988 first data connection between China and scientific world – IHEP to CERN

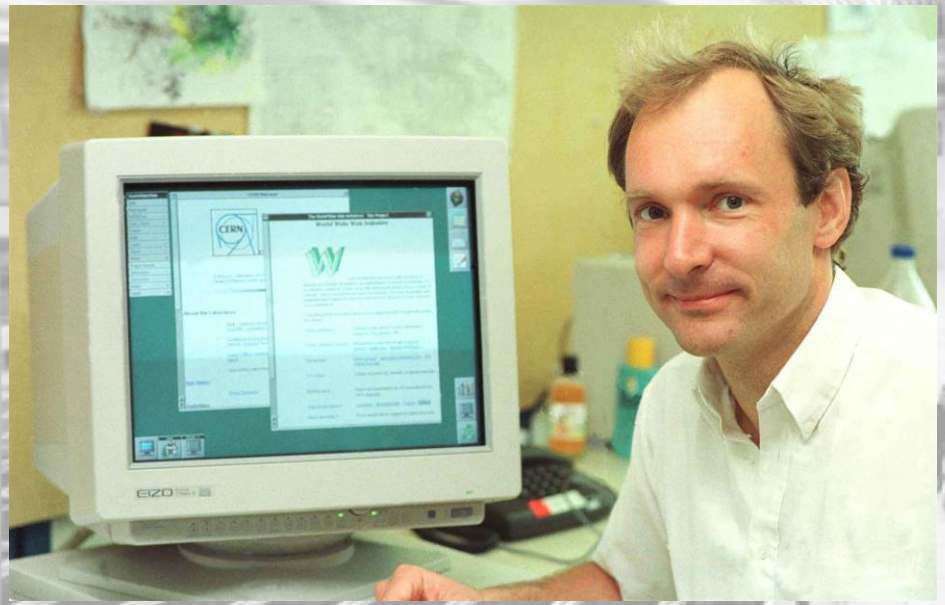
Credit to Tim Smith

Touch Screens

→Whilst not strictly an invention of the IT Department…
→Super Proton Synchrotron control system required complex controls
→Developed capacitive touch screen
→Based on open standards and moved into industry

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

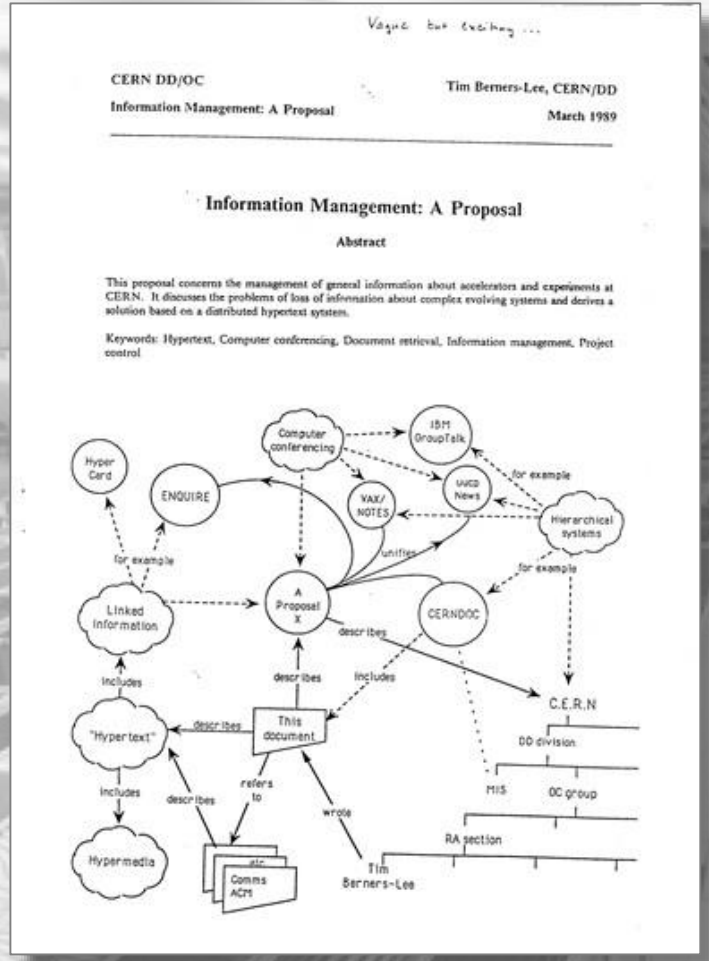
www



➔Tim Berners Lee found a solution to the information sharing backlog
➔In 1993, software was made public and quickly changed our lives!

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

www



Vague but exciting ...

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

- ➜“Vague but exciting”
- ➜Not hierarchical, or centrally controlled
- ➜Experts store locally, update independently
- ➜Community is distributed: remote access

Credit to Tim Smith

Thanks!

Any questions?

You can find me at:
@hannah.short08
hannah.short@cern.ch

CREDITS

- Lorena Lobato, Tim Smith and Melissa Gaillard for their content and ideas
- Presentation template by SlidesCarnival
- Photographs by Death to the Stock Photo (license)