



Introductions



Swiss-Japanese Seminar on Life science grids

***Miika Tuisku
Project manager
HIP-TEK***



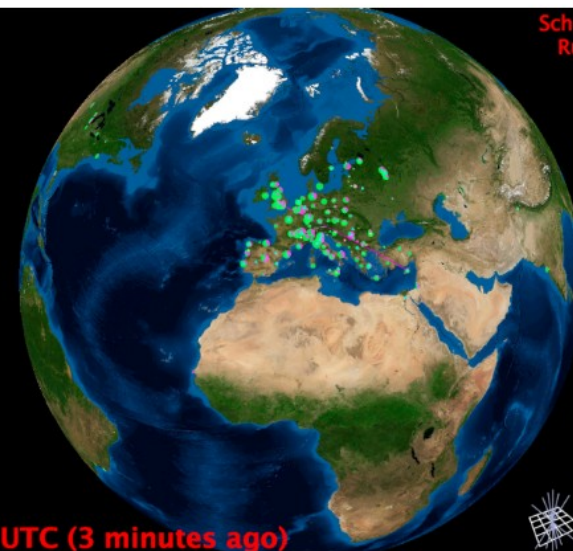
The CERN, LHC, Grid and HIP-TEK

Welcome to CERN! We're so glad you came.
While you're here, you can explore many questions:

- What is **CERN** and the **Large Hadron Collider** (LHC)?
- What is the **Worldwide LHC Computing Grid**?
- Who are the **HIP-TEK** guys, why they are here at CERN?



eGee
Enabling Grids
for E-science



10:13:27 UTC (3 minutes ago)



What is CERN?

CERN is the **world's largest particle physics centre**, and home to the LHC

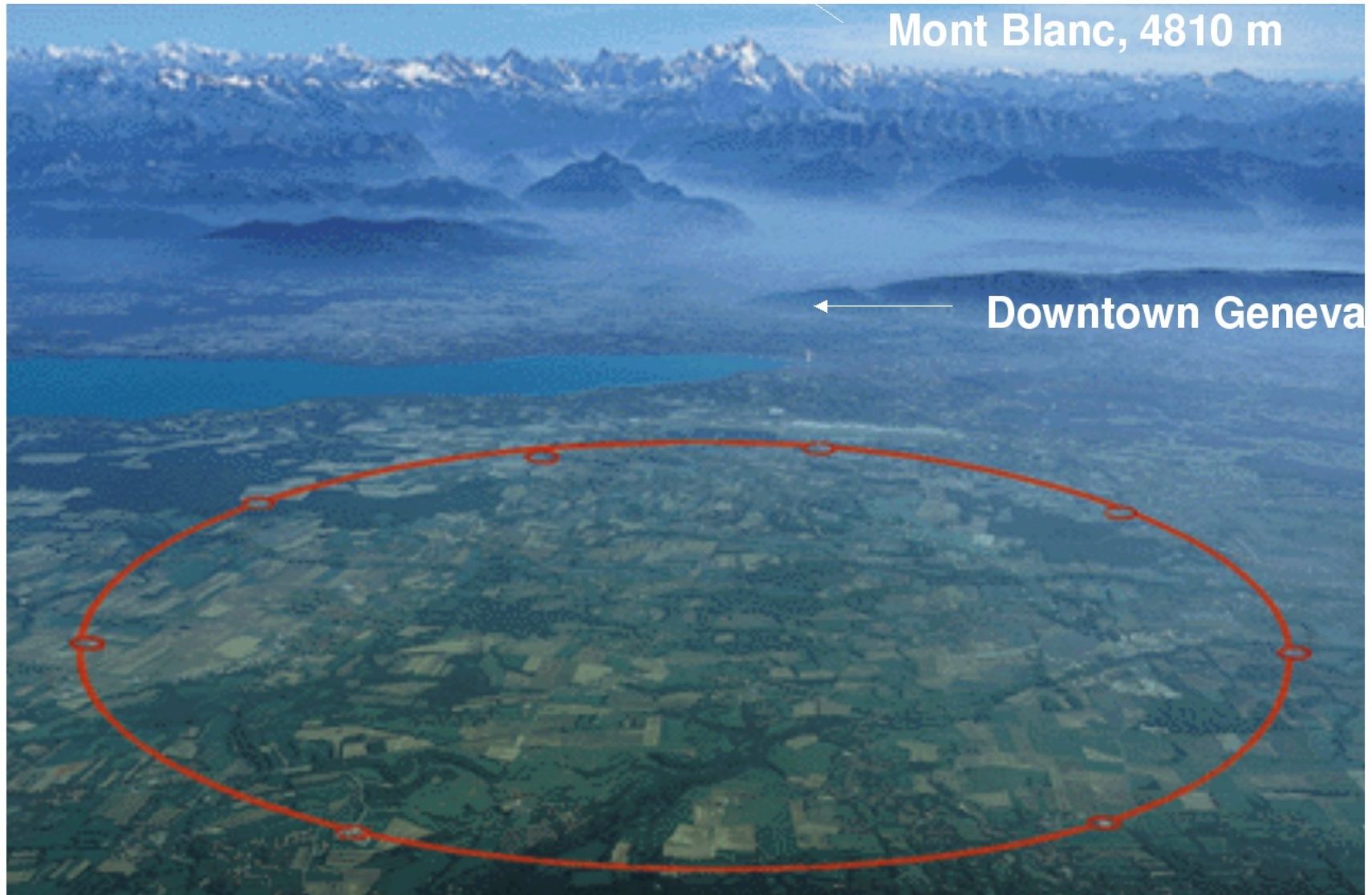
More than 2500 staff scientists (physicists, engineers, and more) work at CERN, with some 6500 scientists visiting at any time.

CERN brings together people from 500 universities representing 80 nationalities.





Where is the LHC?



Mont Blanc, 4810 m

Downtown Geneva



What is the LHC?

- With its **27-km circumference**, the LHC is the largest machine in the world. LHC stands for Large Hadron Collider.
- Particle physicists are using the LHC to collide protons at very high energies, aiming to learn more about the Universe,
- The LHC operates at about **-300°C** , just above absolute zero, and speeds particles to almost the speed of light.
- The LHC involves four experiments, with detectors as 'big as cathedrals': ALICE, ATLAS, CMS and LHCb

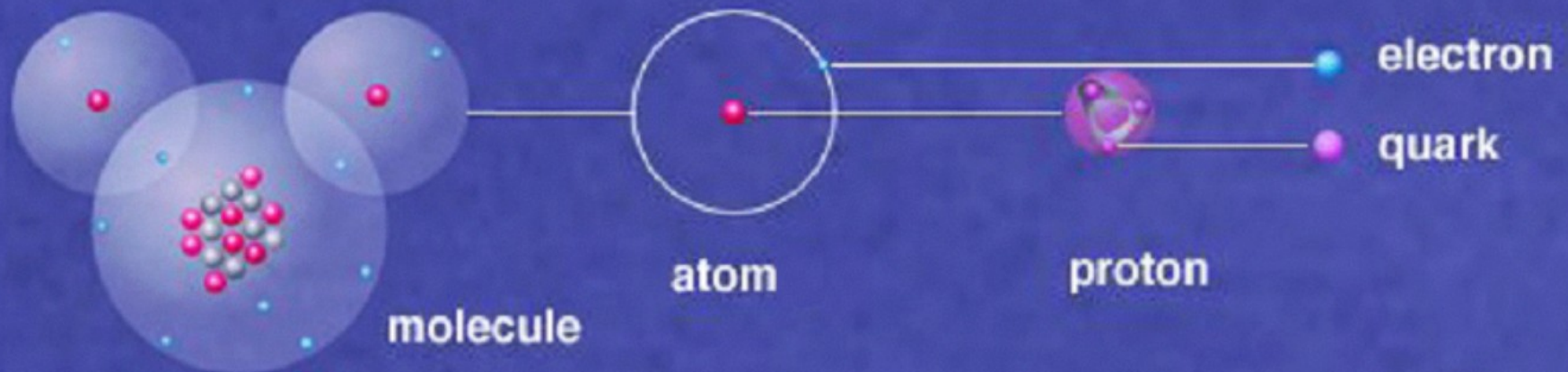




But what for?

To answer fundamental questions about the Universe:

- How did the Universe begin?
- What is the origin of mass?
- What is the nature of antimatter?
- Can we find the mysterious “God” particle: the missing Higgs boson?



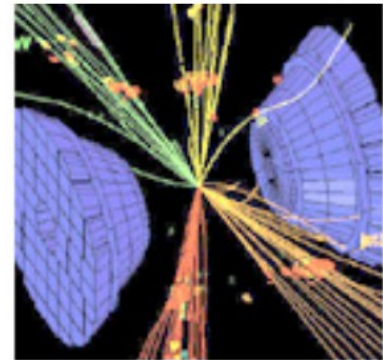


Finding the Higgs

One way to find the Higgs boson is to look at what happens to particles after they collide at high energies inside the LHC.

Physicists count, trace and characterize all the particles produced and **fully reconstruct the process**.

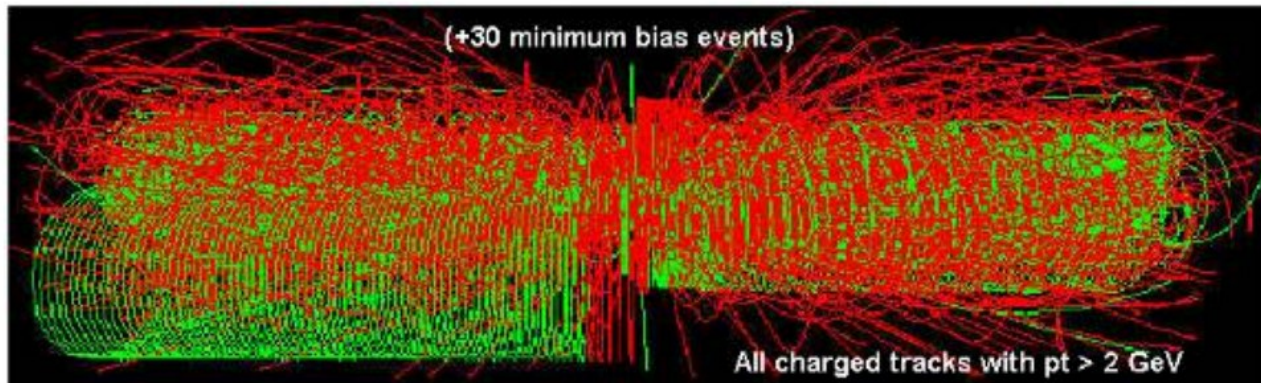
If they can find a characteristic decay pattern producing 4 “muon” particles, they know they’re on the trail of the Higgs!





Finding the Higgs

Starting from this event...

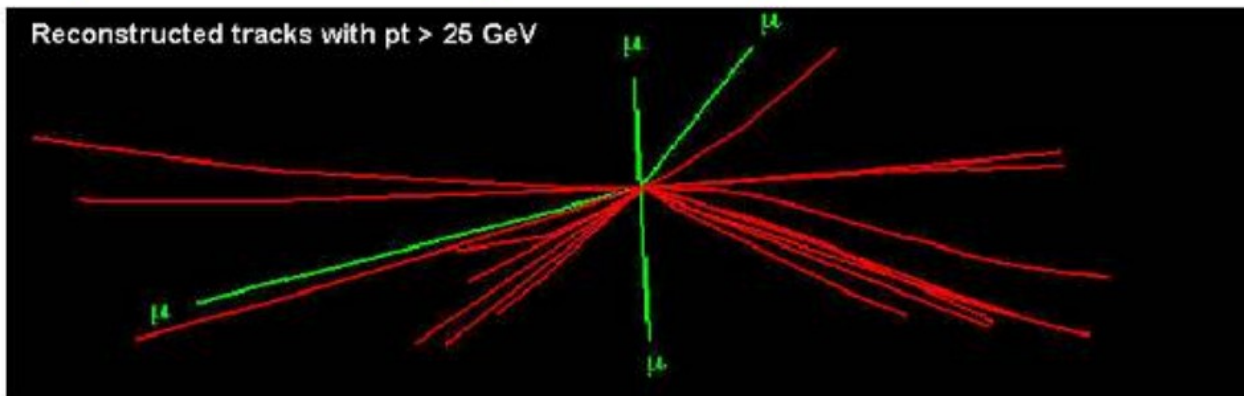


Selectivity: 1 in 10^{13}

Like looking for 1 person in a thousand world populations!

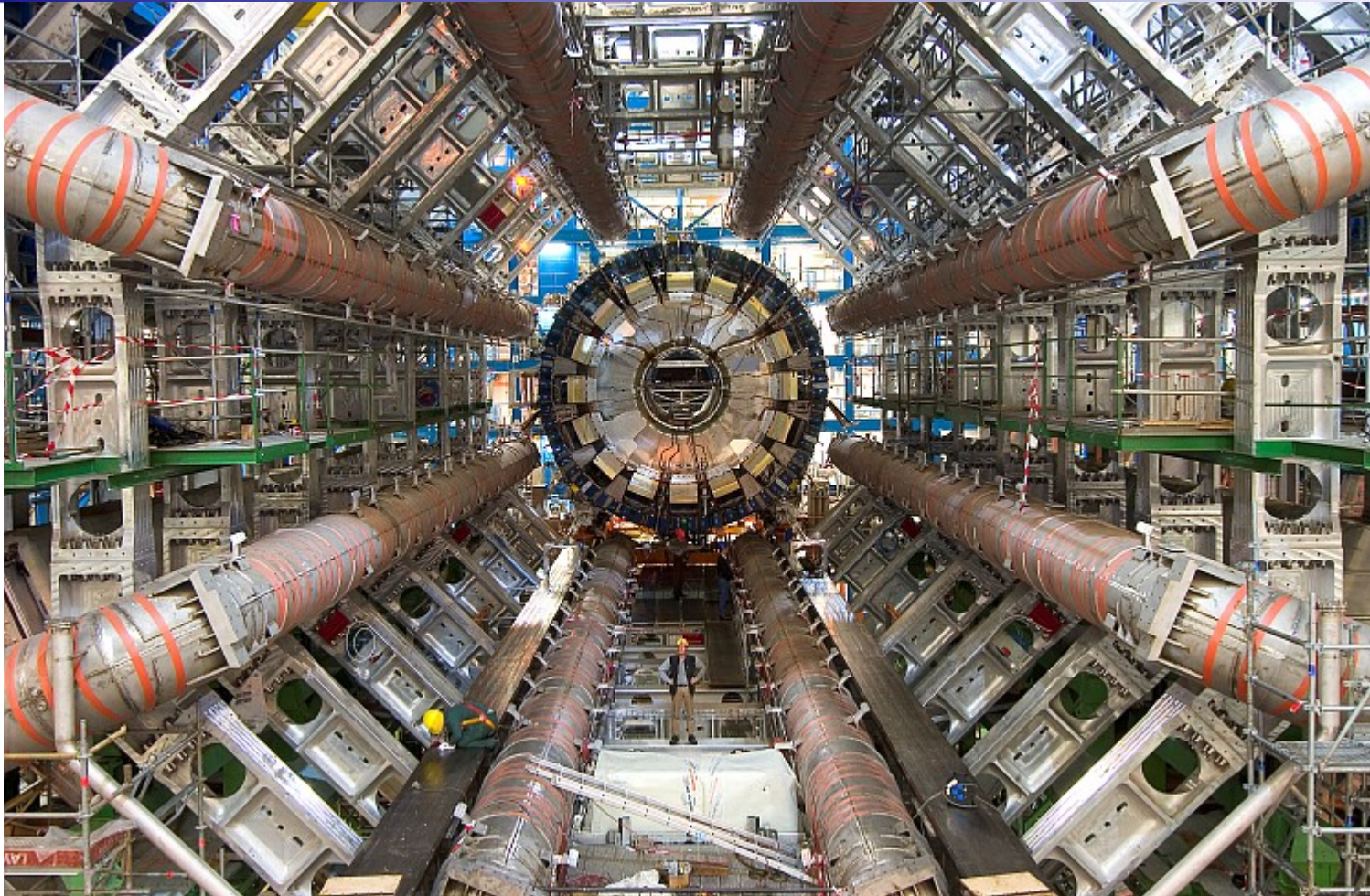
Or for a needle in 20 million haystacks!

Physicists are looking for this “signature”





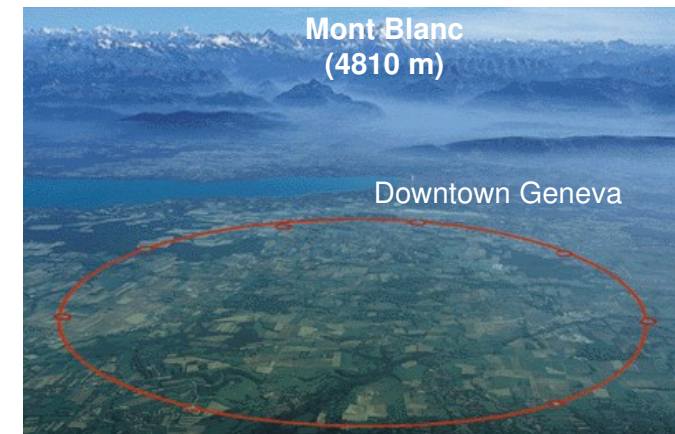
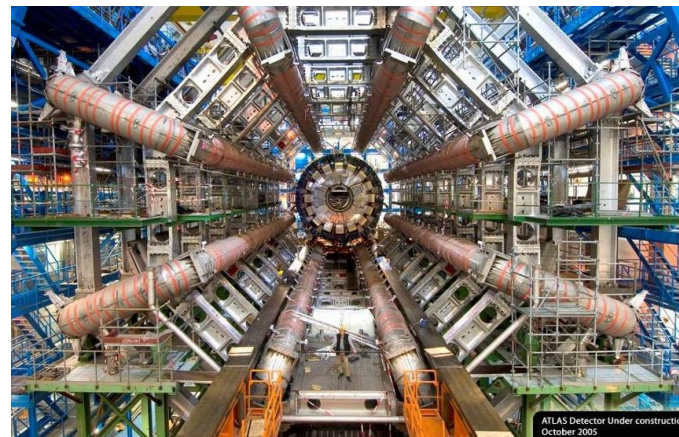
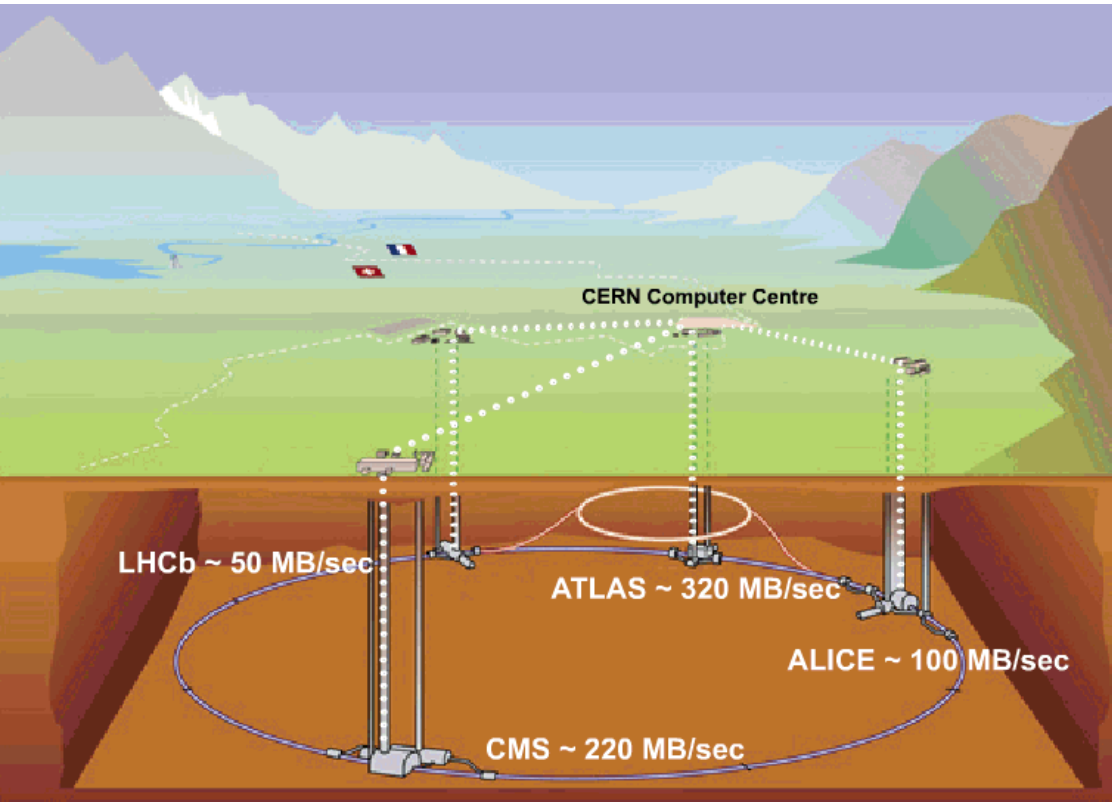
Once upon the time at CERN after WWW



7000 tons, 150 million sensors, 1 PB/s

Large Hadron Collider

- **27 km circumference tunnel**
- Due to start up in 2008**
- 40 Million Particle collisions per second**
 - Online filter reduces to a few 100 “good” events per second recorded on disk and magnetic tape at 100-1,000 MegaBytes/sec
 - ~15 PetaBytes per year for all four experiments
- Data analyzed by 100s of research groups world wide**



Data, data, data!!

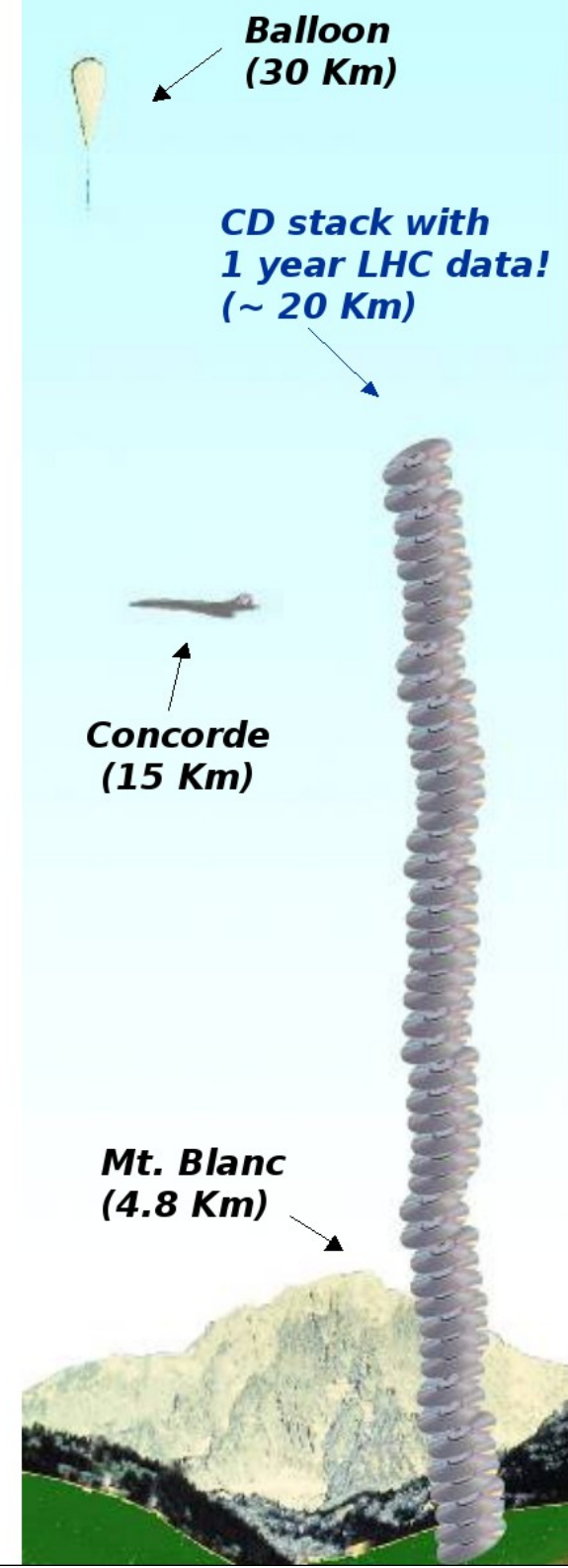
The LHC will produce around

15 Petabytes

of data every year!

That's about 20 million CDs each year!

Where will we store all of these data?





CERN Computer Centre

- High-throughput computing based on reliable “commodity” technology
- More than 5000 PCs with around 20,000 processor cores
- More than 8 petabytes (8 million Gigabytes) of disk storage and 18 petabytes of magnetic tape storage



**Nowhere
near
enough!**



Enter the GRID

GridCafé

LHC Computing Grid



Mission:

- Install a functioning grid to help the LHC experiments collect and analyse data coming from the detectors

Strategy:

- Integrate thousands of computers at hundreds of participating institutes worldwide into a global computing resource.

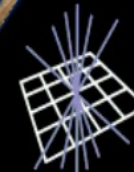
Results:

The Worldwide LCG launched in October 2008 with more than 100,000 processors from 140 institutions in 33 countries, producing a massive distributed supercomputer that will provide more than 7000 physicists around the world with near real-time access to LHC data, and the power to process it.

- **Expand/optimise existing EGEE infrastructure, include more resources and user communities**
- **Prepare migration from a project-based model to a sustainable federated infrastructure based on National Grid Initiatives**

~280 sites
48 countries
>139,000 CPU cores
>25 PetaBytes disk, >38PB tape
>13,000 users
>300,000 jobs/day

21:13:50 UTC



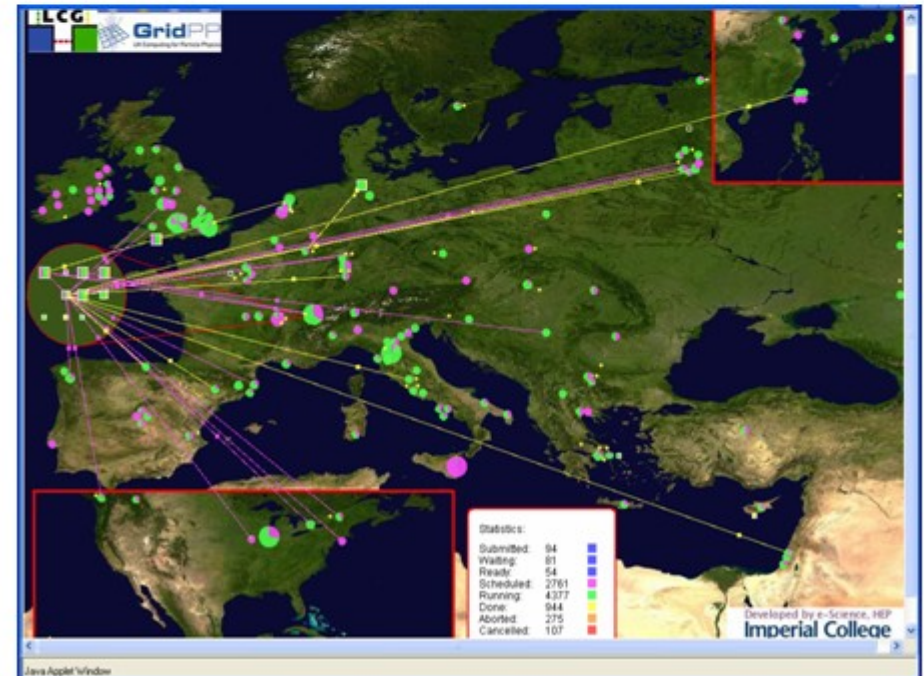
GridPP
UK Computing for Particle Physics



Flagship grid infrastructure project co-funded by the European Commission with 32M€. Now in 3rd phase with 126 partners in 34 countries

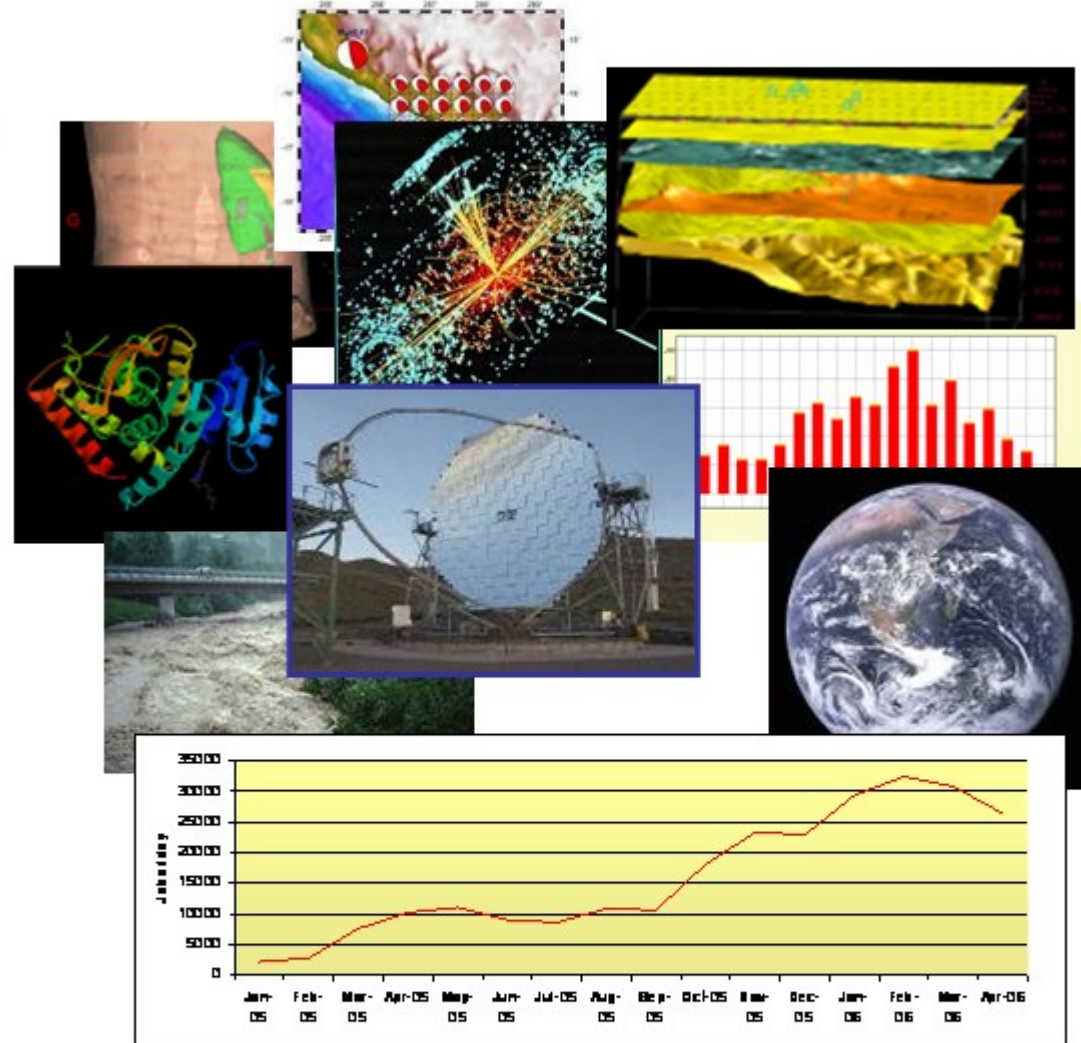
Objectives

- Large-scale, production-quality grid infrastructure for e-Science
- Attracting new resources and users from industry as well as science
- Maintain and further improve gLite Grid middleware



- **Many applications from a growing number of domains**

- Astrophysics
- Computational Chemistry
- Earth Sciences
- Financial Simulation
- Fusion
- Geophysics
- High Energy Physics
- Life Sciences
- Multimedia
- Material Sciences
- ...



Applications have moved from testing to routine and daily usage ~80-90% efficiency



- **Business Partners**

- Funded

- **Business Associates**

- Collaborative

- **Business Forum & Task Force**

- Outreach

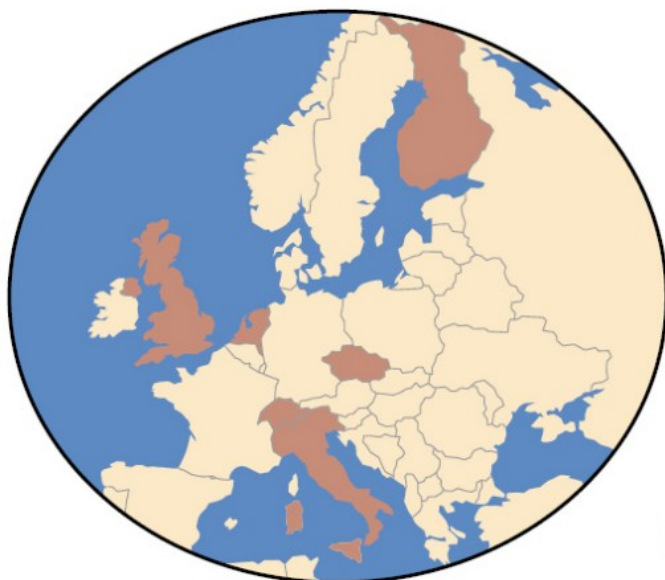
- **Industry Days**

- Events



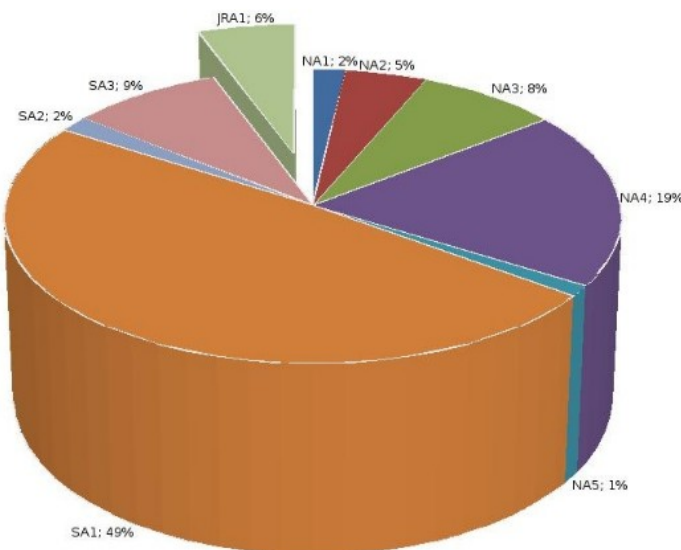
- **Avanade**
 - Bring expertise on .NET Technologies to fully exploit gLite interoperability with Windows machines
- **Excelian**
 - Providing high-level consultancy for the positioning gLite to business
- **Hitachi**
 - Plans to integrate their data transfer solution with the EGEE framework
- **2 New EBAs (approved in May)**
 - Constellation Technologies
 - Will offer additional gLite services and support to commercial customers
 - Linalis
 - Provide Industrial Grid training services
 - *Q13/EGEE-II surveys show market for standards-compliant, open source middleware with commercial support and training*





Country	Total PM planned at M24	Total FTE
CERN	78	3.3
Czech Republic	48	2.0
Finland	48	2.0
Italy	256	10.7
Netherlands	35	1.5
Switzerland	36	1.5
UK*	30	1.3
Total	531	22.1

*includes 18 PMs of voluntary effort



CERN openlab

Concept

- Partner/contributor sponsors latest hardware, software and brainware (young researchers)
- CERN provides experts, test and validation in Grid environment
- Partners: 500'000 €/ year, 3 years
- Contributors: 150'000 €, 1 year

Current Activities

- Platform competence centre
- Grid interoperability centre
- Security activities
- Joint events



www.cern.ch/openlab

PARTNERS



ORACLE®

CONTRIBUTORS



STONESOFT





EGEE as Grid



infrastructure, open issues

- Q1: How much does the EGEE Grid cost, I want to buy it?
 - A1: Can't be sold. Software is free to download and install on your own.
- Q2: How about doing business on EGEE Grid?
 - A2: No. EGEE Grid is built on not-for-profit academic IT infrastructures that prohibit commercial activities.
- Q3: Can I run my application "A" on the EGEE Grid?
 - A3: You must qualify an internal review process and to be assigned in an existing Virtual Organisation. Technically speaking your application must be adapted to Grid i.e. Gridified. Required libraries must be available in Grid.
- Q4: What kind of Service level is provided, e.g. Security?
 - A4: Best-effort, no guarantees given. Private data can be encrypted, yet it is not possible to avoid eventual distribution.



Apart from eScience Grids

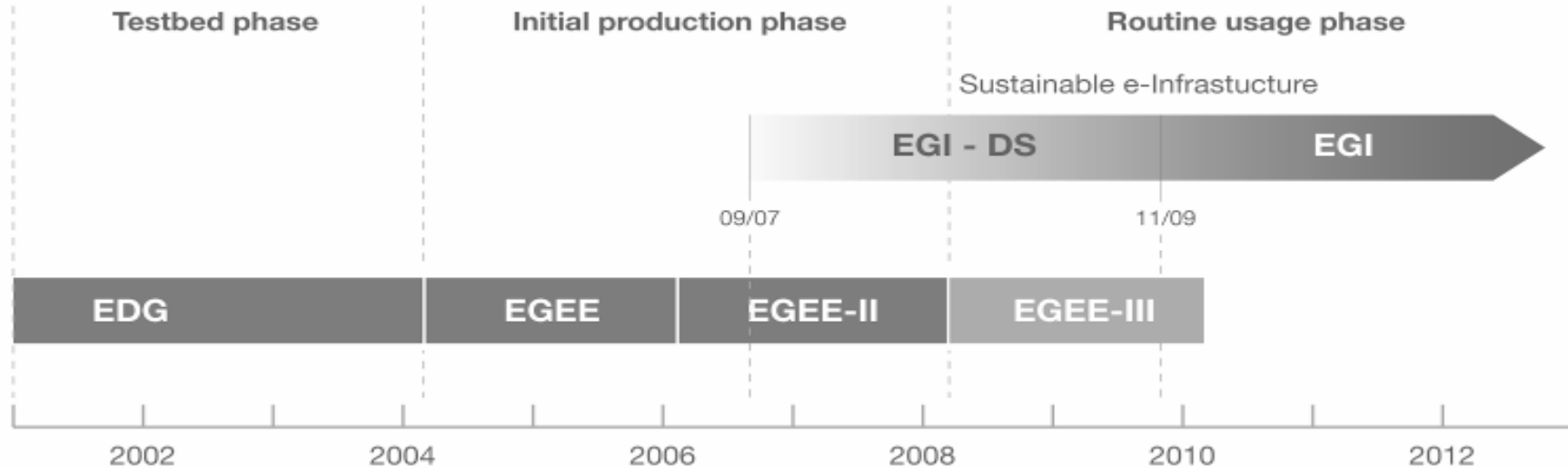


- **Enterprise Grids** have already proved to be commercially successful due to less critical organisational hurdles
 - Private data stays inside the company
- Enterprise Grids spread beyond the High-Performance Computing market
 - They deal with **clustering** and **virtualisation**
 - They share similarities with **Peer-to-Peer computing**
 - They can serve applications that are more typical to eBusinesses: Application hosting, Scaling out Web services: Google, eBay, Amazon ... (**Cloud Computing**)
- Still legacy licensing models are hindering Grid adoption
 - Applications must not be bound to dedicated bare-metal hardware

Apart from eScience Grids

- **Enterprise Grids** have already proved to be commercially successful due to less critical organisational hurdles
 - Private data stays inside the company
- Enterprise Grids spread beyond the High-Performance Computing market
 - They deal with **clustering** and **virtualisation**
 - They share similarities with **Peer-to-Peer computing**
 - They can serve applications that are more typical to eBusinesses: Application hosting, Scaling out Web services: Google, eBay, Amazon ... (**Cloud Computing**)
- Still legacy licensing models are hindering Grid adoption
 - Applications must not be bound to dedicated bare-metal hardware

What happens to Grids?



- European Grid Initiative (EGI) will provide sustainable European e-Science Services, similar to GEANT Research Networks in 2010
- EU commission has funded more than 100 M Euros for CERN led Grid projects since 2001; it's time for technology transfer.
- In addition to HIP-TEK see also <http://www.cern.ch/openlab> and <http://business.eu-egee.org>



The HIP-TEK group



Raison d'être



- Helsinki Institute of Physics (HIP)
 - Founded in 1996
 - Coalition between University of Helsinki and Helsinki University of Technology to unite efforts in experimental physics (www.hip.fi)
- Based on the presidential ruling (=law)
 - HIP is the official liaison between Finland and CERN
- Today
 - 3 more universities integrated
 - A national institute of physics
 - 120 people, budget €4.5M
 - 5 Programs



HIP-TEK



What the strategy says ...

- „The Technology Programme (HIP-TEK) serves to facilitate advanced technology transfer between international accelerator laboratories as CERN and Finnish industry and research institutes and to develop associated technology“
 - To develop technologies
 - To establish industrial collaboration
 - To collaborate through research networks
 - To initiate technology transfer



HIP-TEK, today



- Technology Programme (HIP-TEK)
 - Currently ~13 scientists (mainly computer science majors + students)
 - DataGrid project based at CERN, ClusterGrid project based in Finland
 - Two existing spin-offs from CERN collaboration: <http://www.kronodoc.com> and <http://www.emergence-tech.co.uk/>
- Grid research and development since 2001 together with CERN
 - Expertise covers Grid/Cluster computing, Grid/Web security software development and Server Based Computing (SBC)
- We aim to establish industrial project consortiums using EU, National R&D funding instruments
 - Produced degrees (Phd, MSc) and publications used as academic metrics, Active summer student programme used as main recruitment channel.
 - 32 MSc's, 3 PhD's produced + more than 70 summer students trained with >70 published articles and papers during past 10 years of operation.



The end



- My colleagues will tell more about our projects next ...
- For more information see <http://tek.hip.fi>
- Or contact: miika.tuisku@cern.ch
- Telephone: + 41 22 7676 195 (CERN)