



CERN, GRID and E-Science

Contents:

- **Introduction**
- **Computer intensive science**
- **Particle physics and the LHC**
- **The LHC data challenge**
- **LCG – the LHC Computing Grid**
- **The CERN IT Openlab**

Nils Høimyr, CERN IT

Includes presentation contents from Frédéric Hemmer, Ian Bird and the CERN IT Openlab



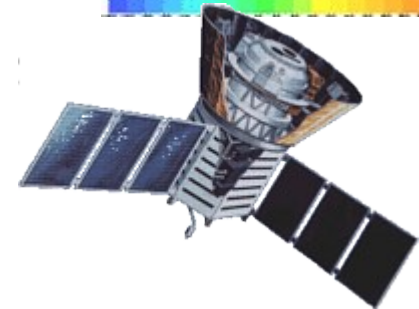
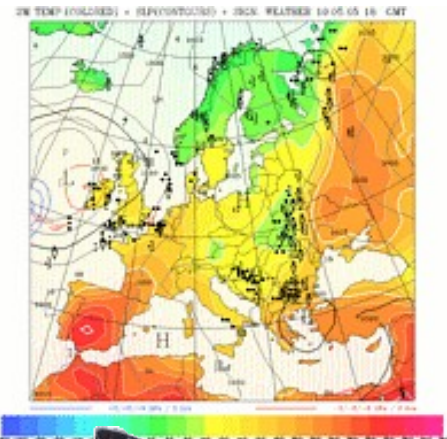
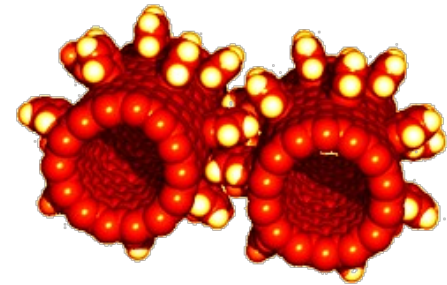
IT at CERN – more than the Grid

- **Physics computing – Grids (this talk!)**
- **Administrative information systems**
 - Financial and administrative management systems, e-business...
- **Desktop and office computing**
 - Windows, Linux and Web infrastructure for day to day use
- **Engineering applications and databases**
 - CAD/CAM/CAE (Autocad, Catia, Cadence, Ansys etc)
 - A number of technical information systems based on Oracle, MySQL
- **Controls systems**
 - Process control of accelerators, experiments and infrastructure
- **Networks and telecom**
 - European IP hub, security, voice over IP...

More information: <http://cern.ch/it>

Computing intensive science

- Science is becoming increasingly **digital** and needs to deal with increasing amounts of data
- **Simulations** get ever more detailed
 - Nanotechnology – design of new materials from the molecular scale
 - Modelling and predicting complex systems (weather forecasting, river floods, earthquake)
 - Decoding the human genome
- **Experimental Science** uses ever bigger **sensors** to make precise measurements
 - Compute a lot of statistics
 - Huge amounts of data
 - Serves user community around the world



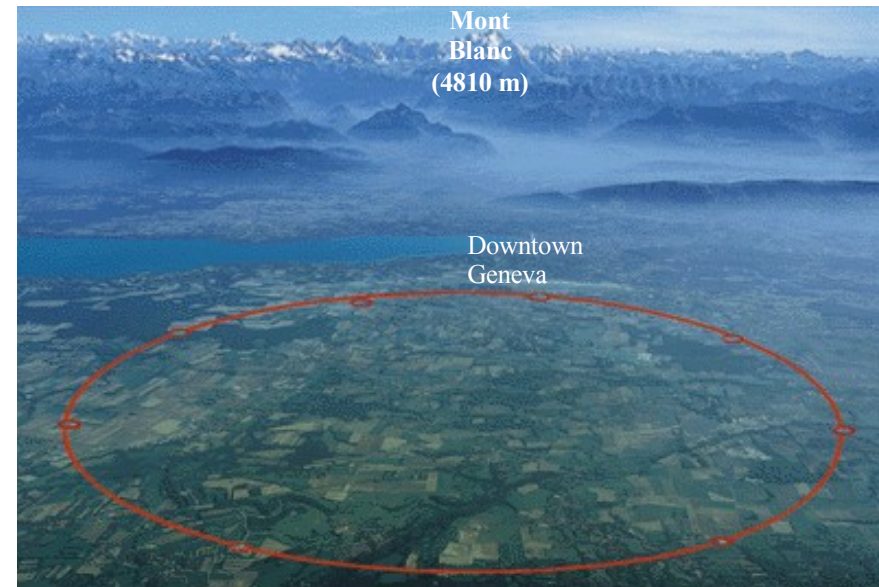


Particle Physics (I)



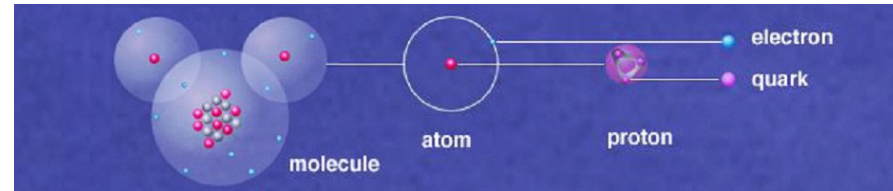
- **CERN: the world's largest particle physics laboratory**
- **Particle physics requires special tools to create and study new particles: accelerators and detectors**

- **Large Hadron Collider (LHC):**
 - most powerful instrument ever built to investigate elementary particles
 - four experiments: ALICE, ATLAS, CMS, LHCb
 - 27 km circumference tunnel
 - First beam 10th September 2008



- **Physicists smash particles into each other to:**

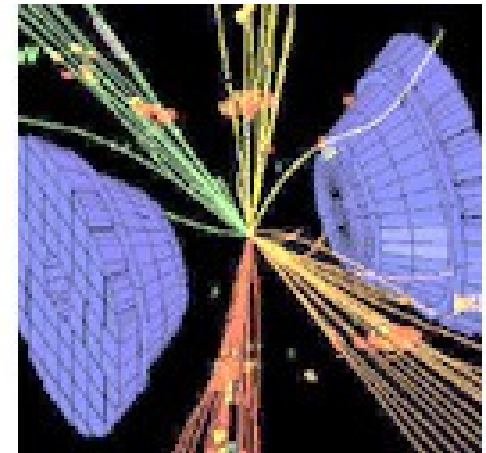
- identify their components
- create new particles
- reveal the nature of the interactions between them
- create an environment similar to the one present at the origin of our Universe



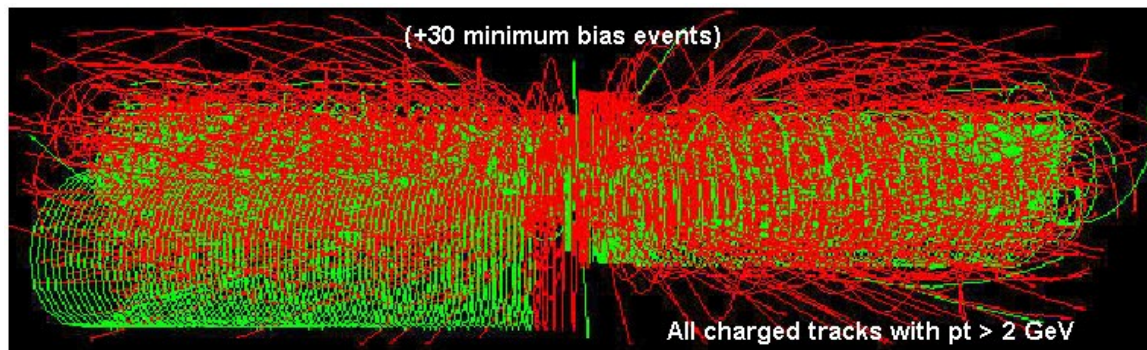
- **A particle collision = an event**

- need to count, trace and characterize all the particles produced and fully reconstruct the process

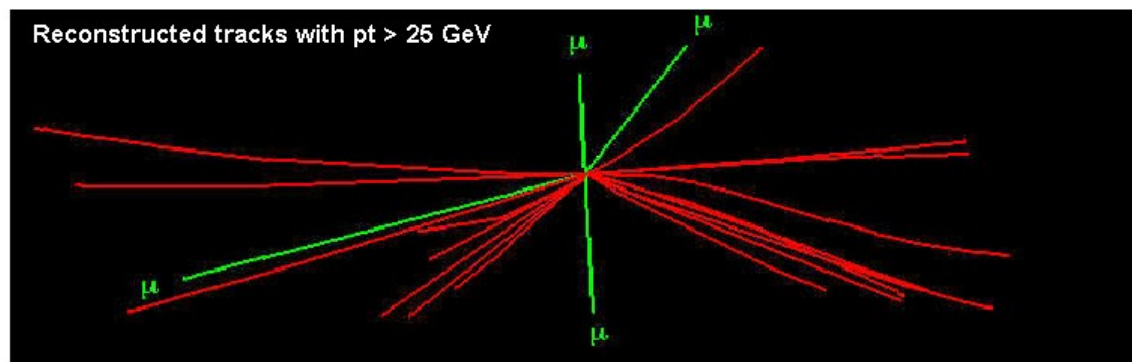
- **Among all tracks, the presence of “special shapes” is the sign for the occurrence of interesting interactions.**



Starting from
this event



Looking for
this “signature”



→ **Selectivity: 1 in 10^{13}**

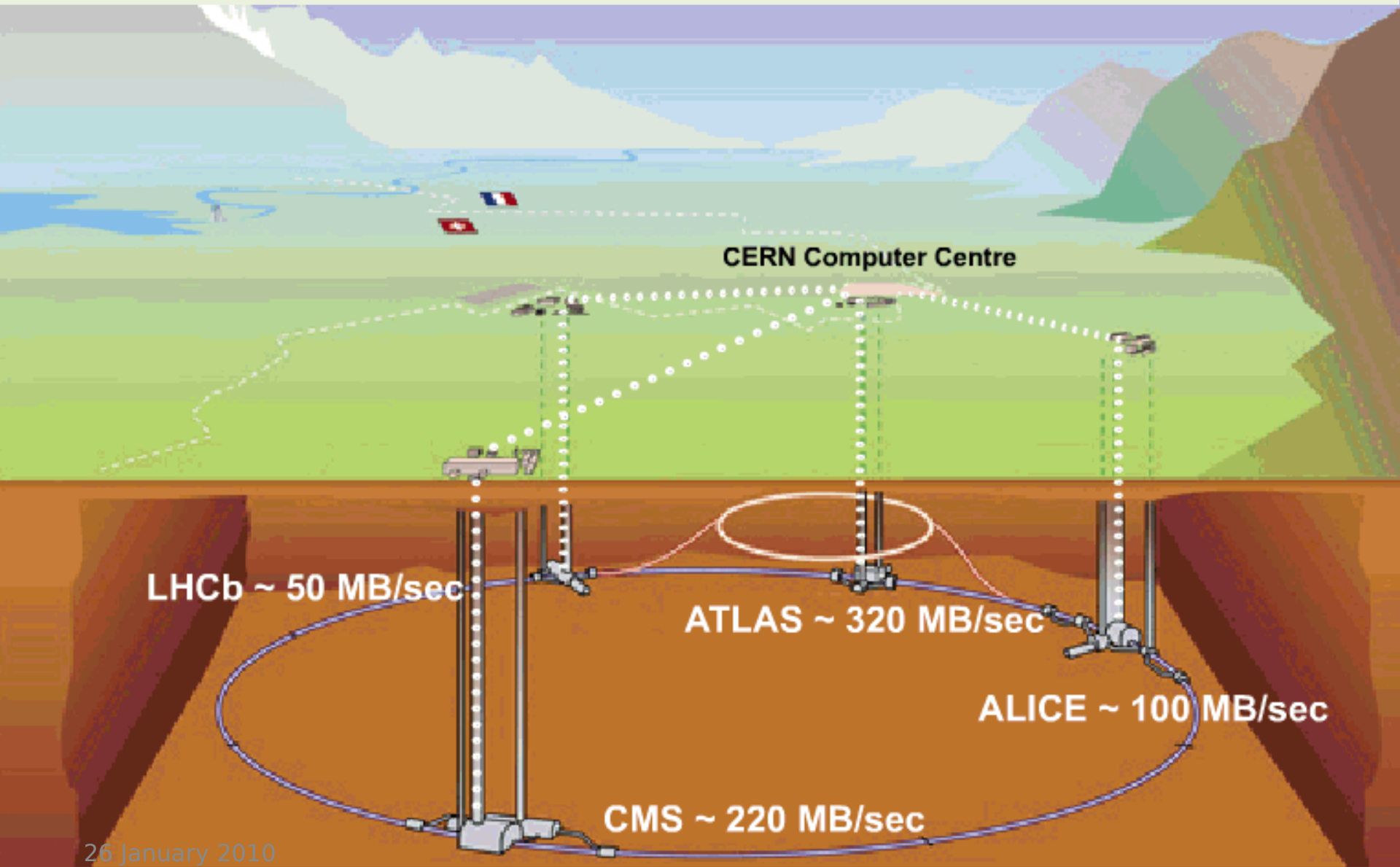
(Like looking for a needle in 20 million haystacks)

View of the ATLAS detector (under construction)



Tier 0 at CERN: Acquisition, First pass processing

Storage & Distribution



- **Simulation**

- compute what the detector should have seen

- **Reconstruction**

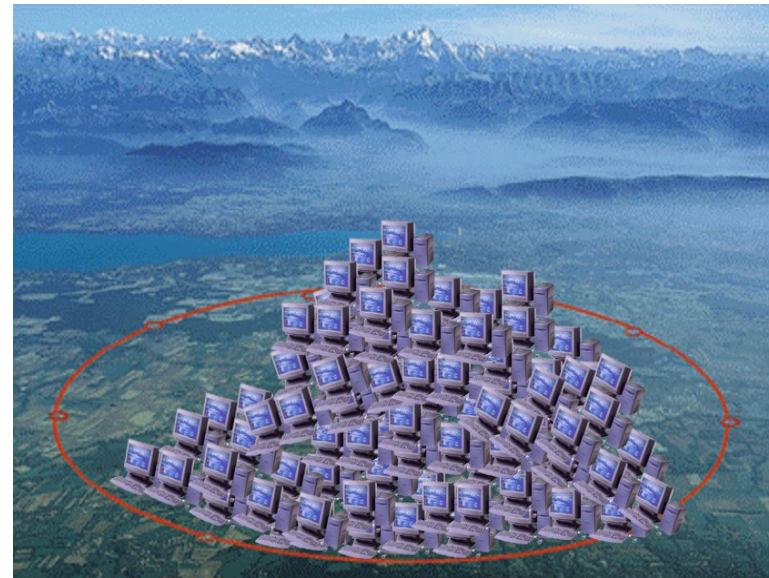
- transform signals from the detector to physical properties

(energies, charge of particles, ...)

- **Analysis**

- use complex algorithms to extract physics

- LHC data analysis requires a computing power equivalent to ~ 100,000 of today's fastest PC processors!





CERN Computing – Tier 0 in numbers

- Computing – CPU:
 - 8000 systems / 60k cores
 - Used for CPU servers, disk servers, general services
- Computing – disk:
 - 14 PB on 42.5k disk drives (+ planned 19 PB on 20k drives)
- Computing – tape:
 - 34 PB on 45k tape cartridges
 - 56k tape slots in robots, 160 tape drives
- Computer centre:
 - 2.9 MW usable power, + ~1.5 MW for cooling

Current status and numbers

Solution: the Grid

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

The **World Wide Web** provides seamless access to information that is stored in many millions of different geographical locations

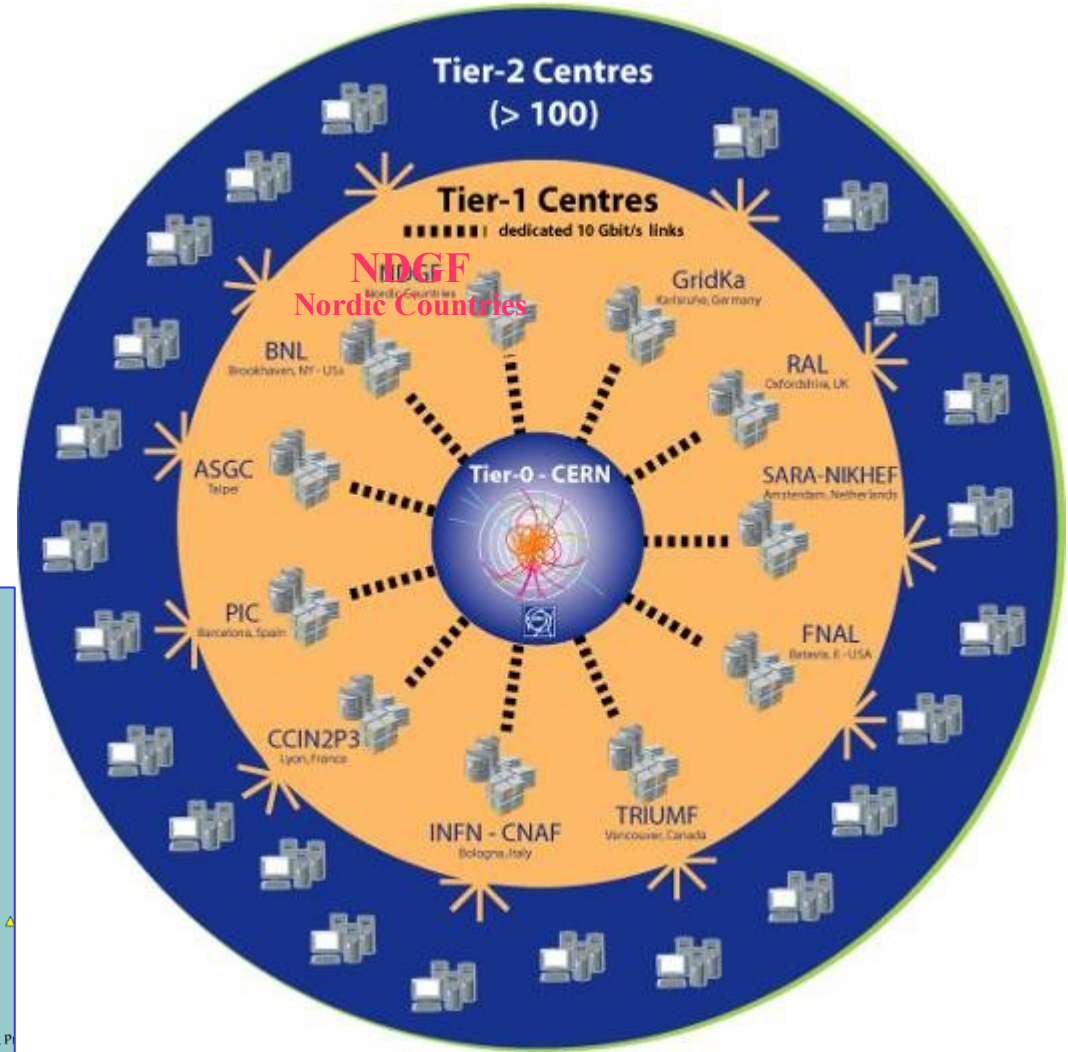
The **Grid** is an infrastructure that provides seamless access to computing power and data storage capacity distributed over the globe





LHC Computing Grid project (LCG)

- More than 200 computing centres
- 12 large centres for primary data management: CERN (Tier-0) and eleven Tier-1s
- 38 federations of smaller





WLCG Collaboration

- **The Collaboration**
 - 4 LHC experiments
 - ~200 computing centres
 - 12 large centres (Tier-0, Tier-1)
 - 38 federations of smaller “Tier-2” centres
 - ~35 countries
- **Memorandum of Understanding**
 - Agreed in October 2005
- **Resources**
 - Focuses on the needs of the four LHC experiments
 - Commits resources
 - each October for the coming year
 - 5-year forward look
 - Agrees on standards and procedures
- **Relies on EGEE and OSG (and other regional efforts)**



Open Science Grid

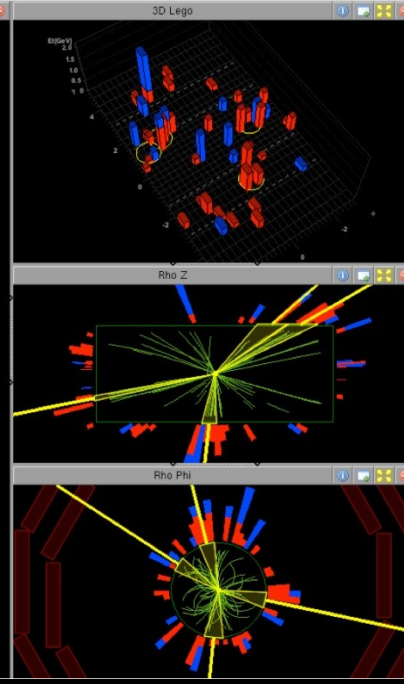
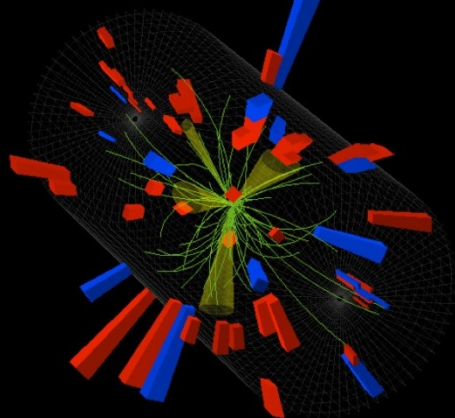
EGEE

Enabling Grids for
E-science in Europe

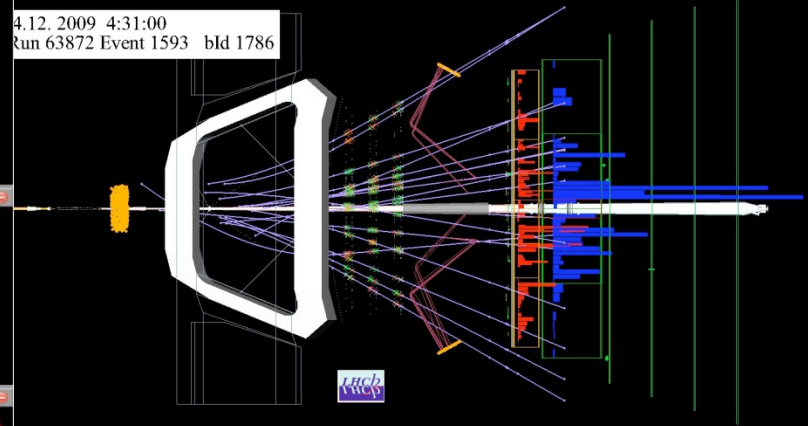




CMS Experiment at the LHC, CERN
 Date Recorded: 2009-12-14 05:41 CET
 Run/Event: 124120/16701049
 Candidate Multijet Event at 2.36 TeV

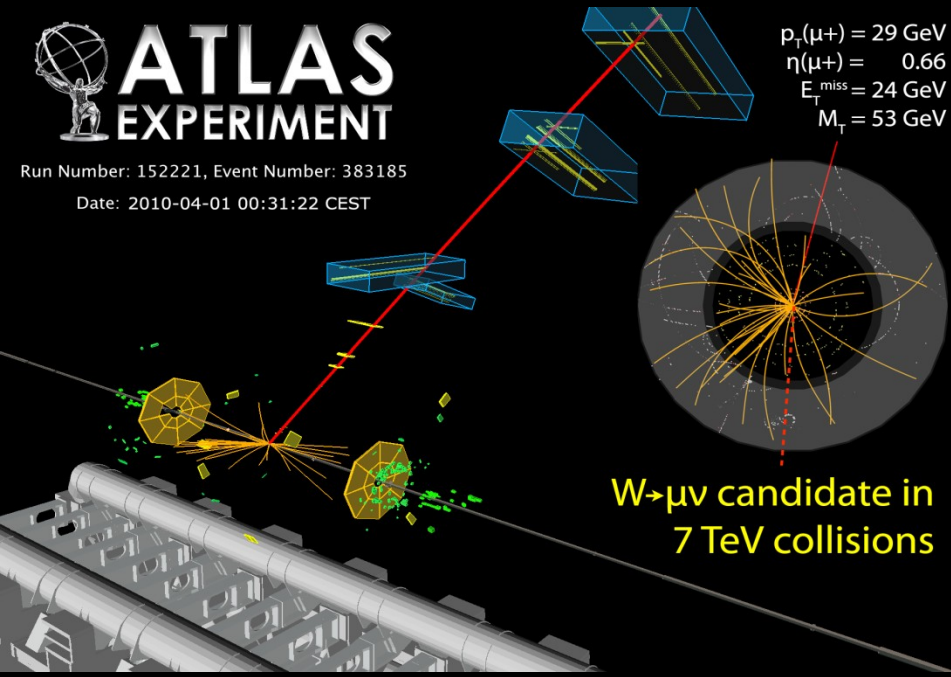


4.12.2009 4:31:00
 Run 63872 Event 1593 bld 1786



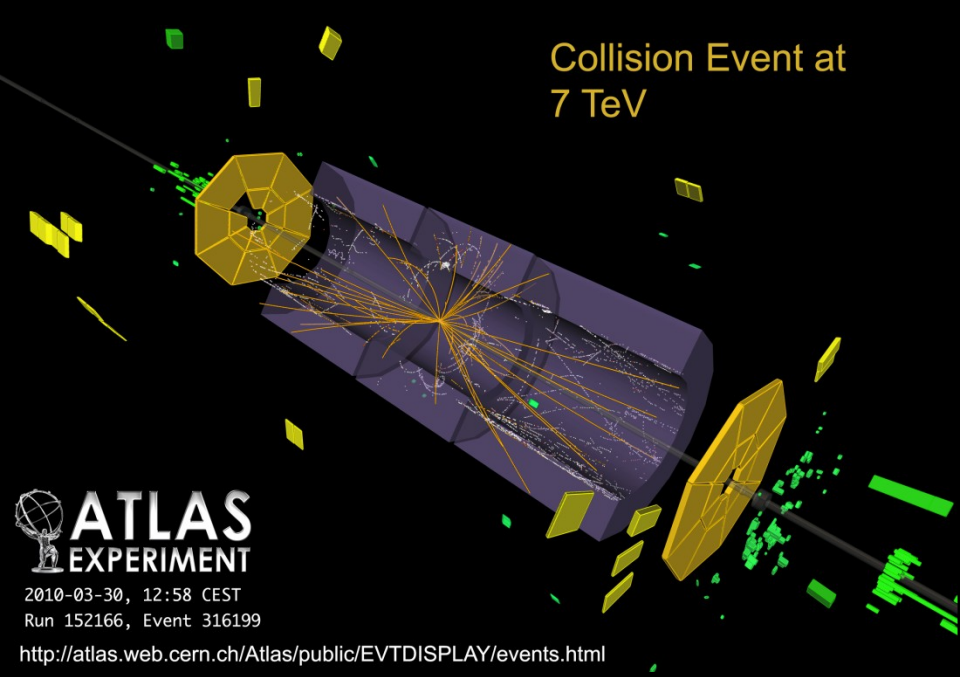
ATLAS EXPERIMENT

Run Number: 152221, Event Number: 383185
 Date: 2010-04-01 00:31:22 CEST



$p_{\mu^+} = 29 \text{ GeV}$
 $\eta(\mu^+) = 0.66$
 $E_{\text{miss}} = 24 \text{ GeV}$
 $M_{\text{T}} = 53 \text{ GeV}$

**$W \rightarrow \mu\nu$ candidate in
 7 TeV collisions**



**Collision Event at
 7 TeV**

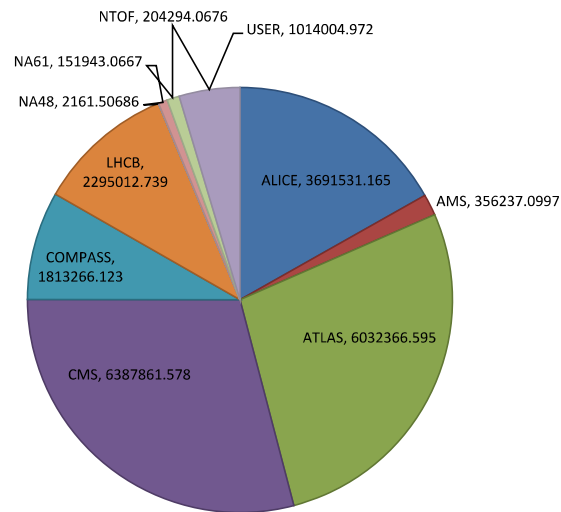
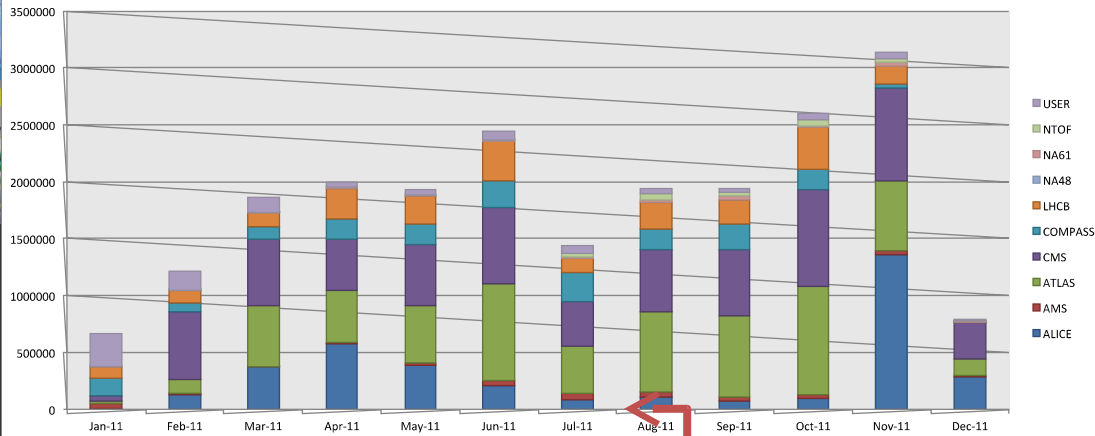
ATLAS EXPERIMENT

2010-03-30, 12:58 CEST
 Run 152166, Event 316199

<http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html>

WLCG in 2011

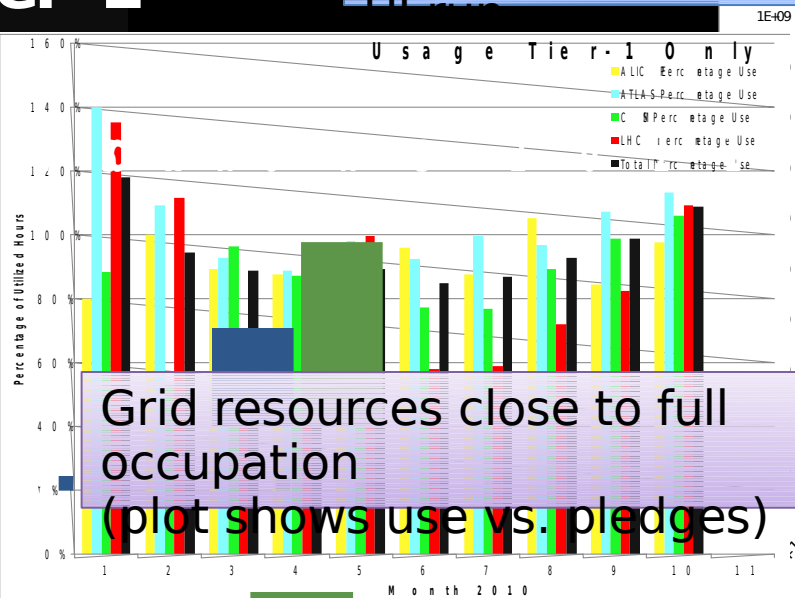
CASTOR data written, 01/01/2011 to 6/12/2011 (in GB)



22 PB data written in 2011
More than 6 GB/s to tape during LHC runs

CASTOR data written, 01/01/2011 to 6/12/2011 (in GB)

er-1



Grid resources close to full occupation
(plot shows use vs. pledges)

109 HEPSPEC-hours/month
(~150 k CPU continuous use)

LHCb
CMS
ATLAS
ALICE

CPU Use continues to grow... even during technical stops

WLCG - no technical stop for

Running jobs: 151768.0
Transfer rate: 4.72 GiB/sec

Activity on 3rd



WLCG
Worldwide LHC Computing Grid



© 2011 Europa Technologies
US Dept of State Geographer
© 2011 Google
© 2011 MapLink/Tele Atlas

1616 Google earth
Eye alt 18391.55 km



Impact of the LHC Computing Grid in Europe

EGEE

Enabling Grids
for E-science

Archeology

Astronomy

Astrophysics

Civil Protection

Comp. Chemistry

Earth Sciences

Finance

Fusion

Geophysics

High Energy Physics

Life Sciences

Multimedia

Material Sciences

...

- LCG has been the driving force for the European multi-science Grid EGEE (Enabling Grids for E-science)
- EGEE is now a global effort, and the largest Grid infrastructure worldwide
- Co-funded by the European Commission (Cost: ~170 M€ over 6 years, funded by EU ~100M€)
- EGEE already used for >100 applications, including...

Scheduled = 21539
Running = 25374

> 250 sites
48 countries
> 50,000 CPUs
> 20 PetaBytes
> 10,000 users
> 150 VOs
> 150,000 jobs/day



Grid



GRID vs Cloud

- **“Cloud computing” is gaining importance**
 - Web based solutions (http/https and RES)
 - Virtualization, upload machine images to remote sites
- **GRID has mainly a scientific user base**
 - Complex applications running across multiple sites, but works like a cluster batch system for the end user
 - Mainly suitable for parallel computing and massive data processing
- **Expect convergence in the future**
 - “Internal Cloud” at CERN - Ixcloud
 - CernVM – virtual machine running e.g. at Amazon
 - “Volunteer Cloud” - [LHC@home 2.0](#)



A second Tier 0 for CERN

- **Wigner Research Centre for Physics in Budapest will host the new facility**
 - Contract not yet signed
 - Stay tuned for more news
- **A second Tier 0 will allow for future redundancy of IT services, not only provide extra computing capacity**

- A science – industry partnership to drive R&D and innovation
- Started in 2002, now in phase 4

Motto: “you make it – we break it”

- Evaluates state-of-the-art technologies in a very complex environment and improves them
- Test in a research environment today what will be used in industry tomorrow
- Training:
 - openlab student programme
 - Topical seminars
 - CERN School of Computing



More information



www.cern.ch/openlab



www.gridcafe.org



www.cern.ch/lcg



www.nordugrid.org



www.eu-egi.org/



lhcat.home.cern.ch