High Throughput Computing for CERN's Large Hadron Collider

Havard Bjerke
CERN openlab
22 June 2008

CERN openlab

Overview of CERN

What is CERN?

- CERN is the world's largest particle physics centre
- Particle physics is about:
  - elementary particles, the constituents all matter in the Universe is made of
  - fundamental forces which hold matter together
- Particles physics requires:
  - special tools to create and study new particles
    - Accelerators
    - Particle Detectors
    - Powerful computers

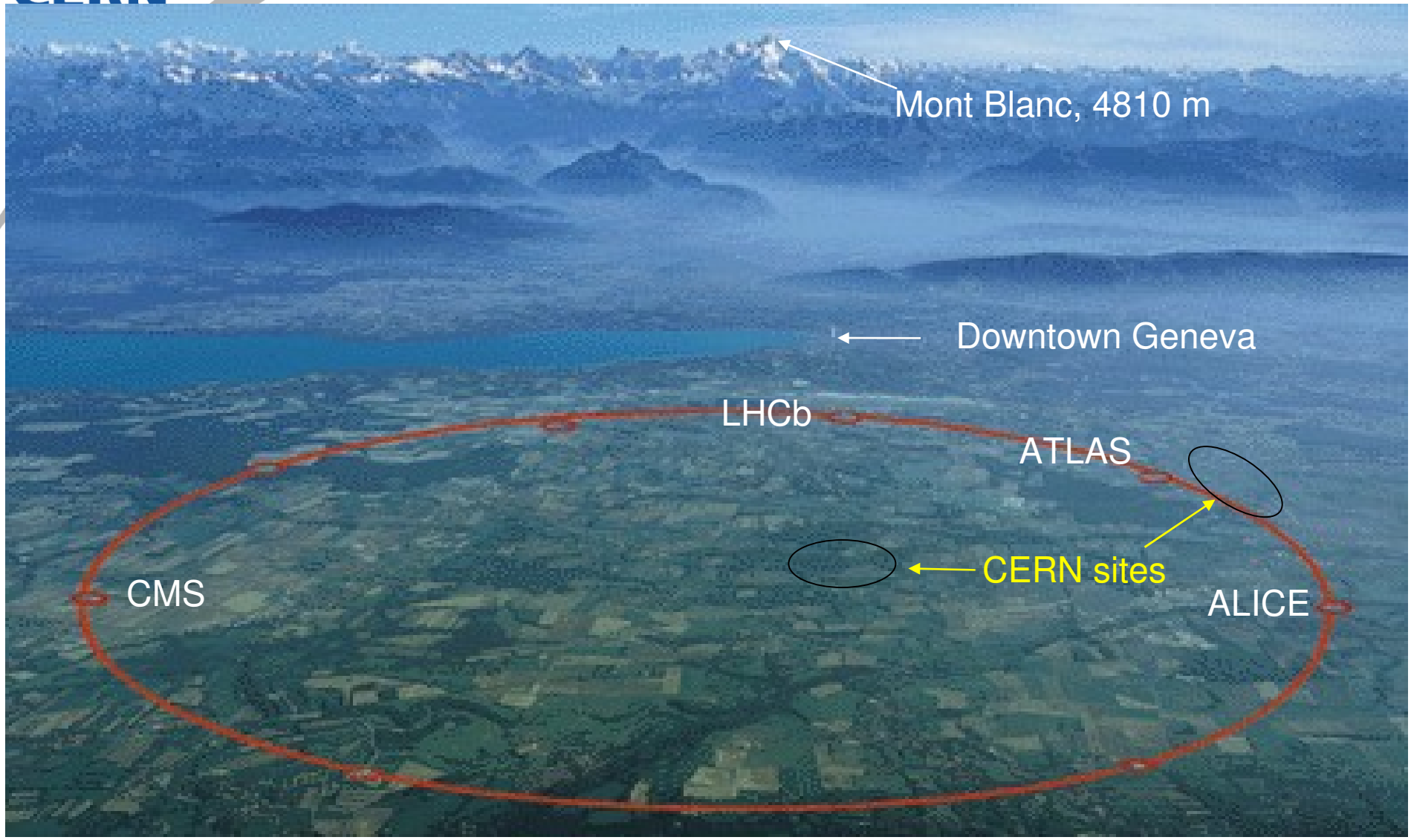
CERN is also:

-2500 staff (physicists, engineers, technicians, …)

- Some 6500 visiting scientists (half of the world's particle physicists)

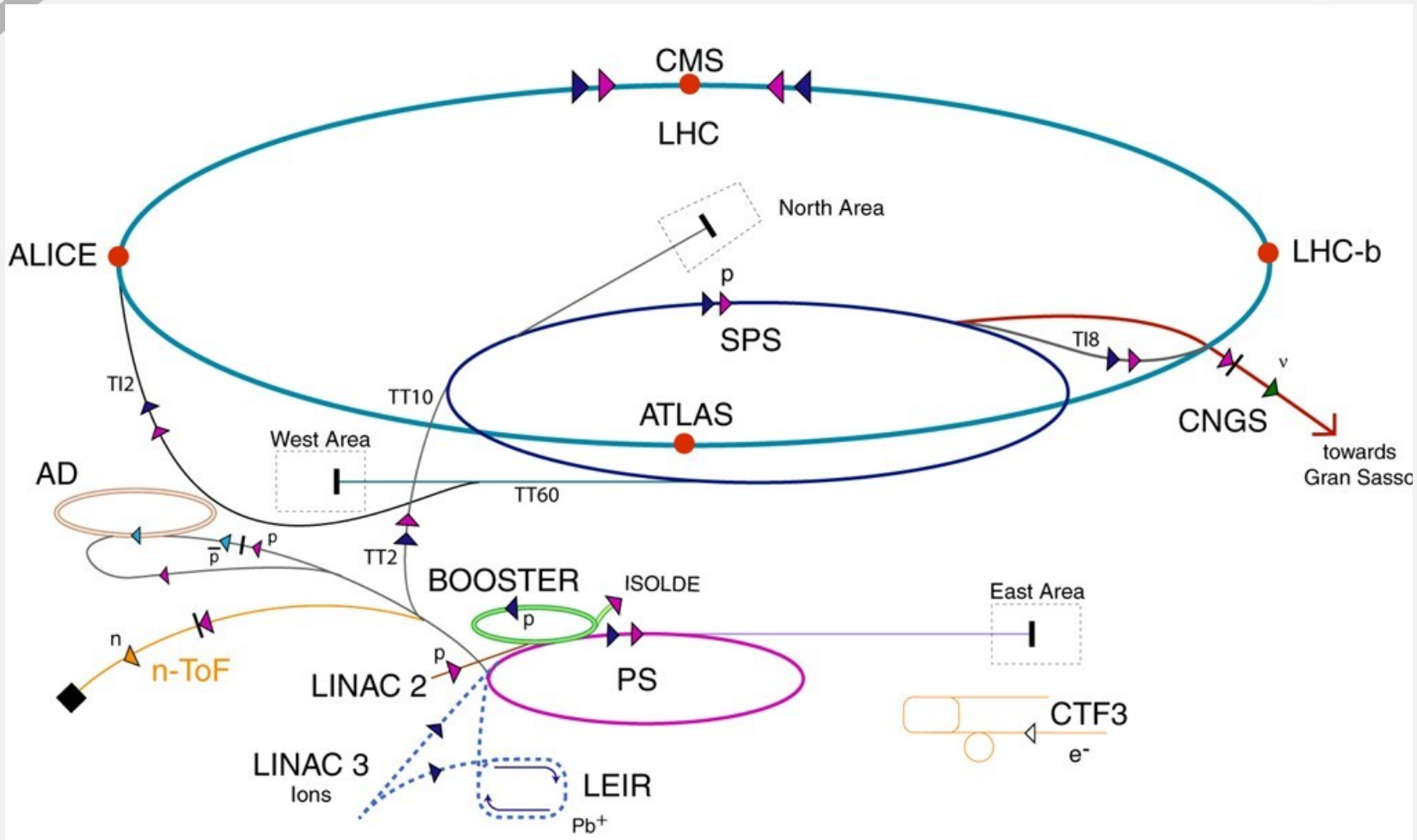
They come from 500 universities representing 80 nationalities.

The CERN Site



CERN's accelerators

The world's most complete accelerator complex



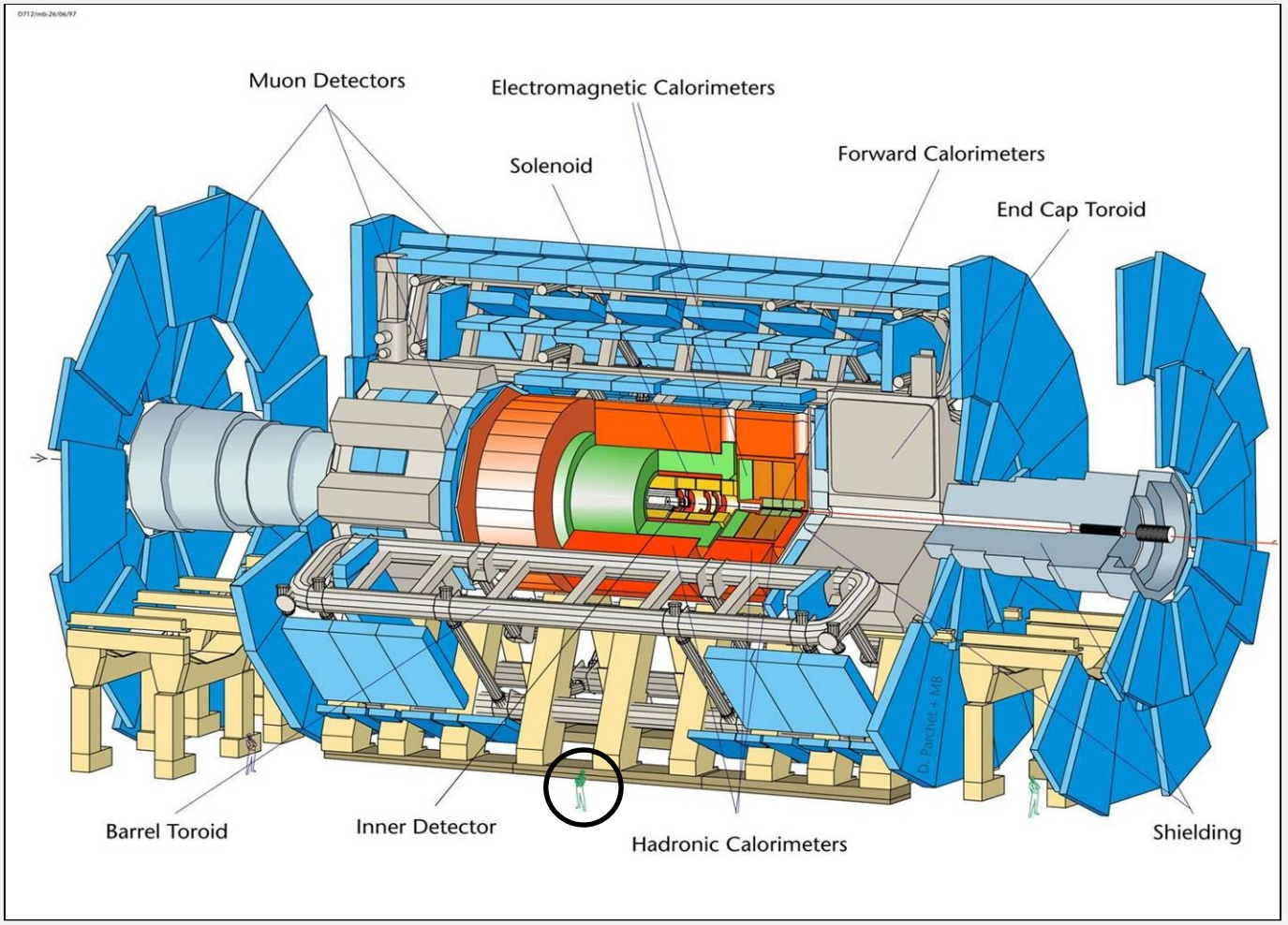
What is LHC?

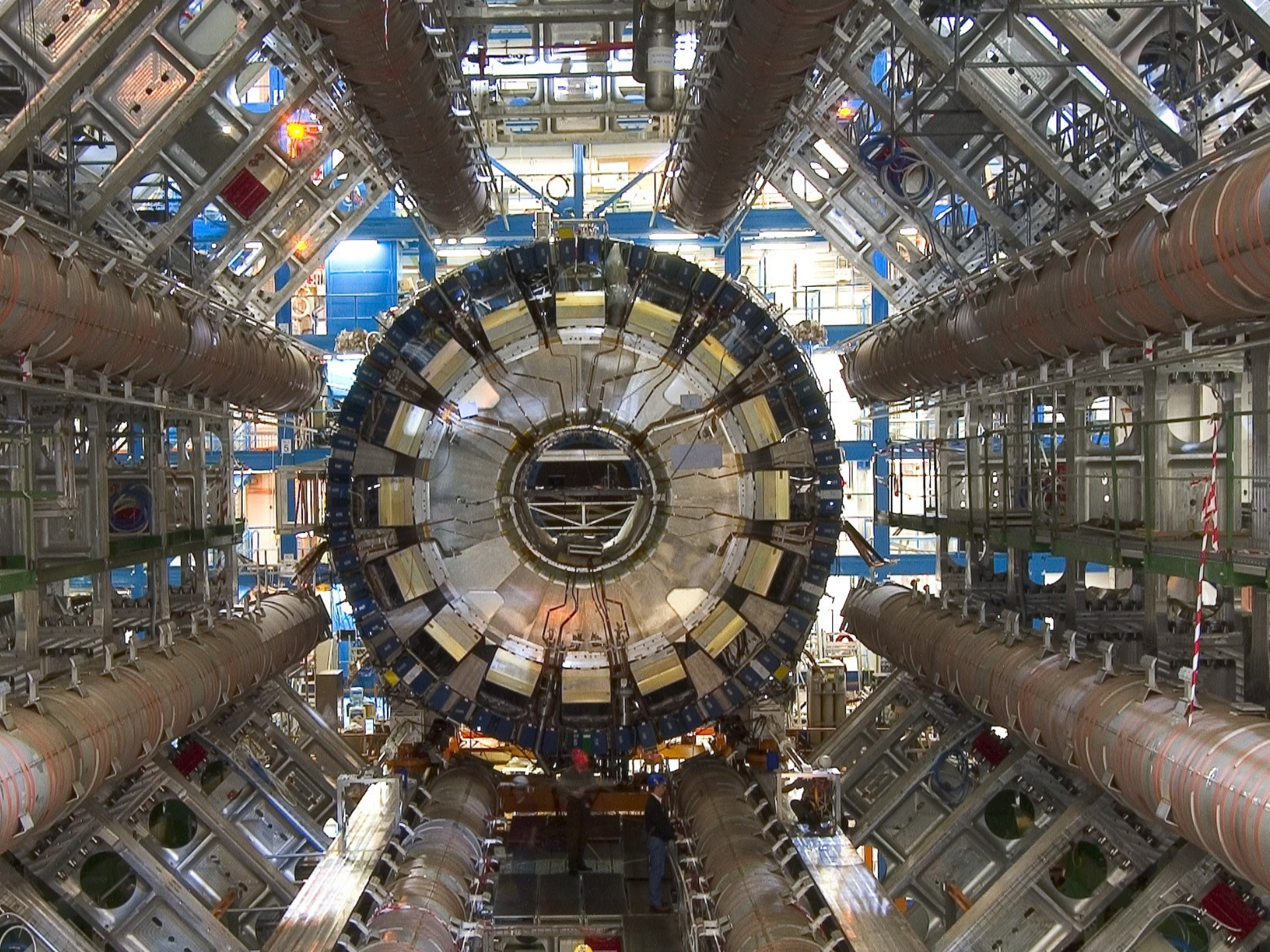
- The Large Hadron Collider will collide beams of protons at an energy of 14 TeV (in the summer of 2008)
- Using the latest super-conducting technologies, it will operate at about – 271°C, just above the temperature of absolute zero.
- With its 27 km circumference, the accelerator will be the largest superconducting installation in the world.

Four experiments, with detectors as 'big as cathedrals':
ALICE
ATLAS
CMS
LHCb

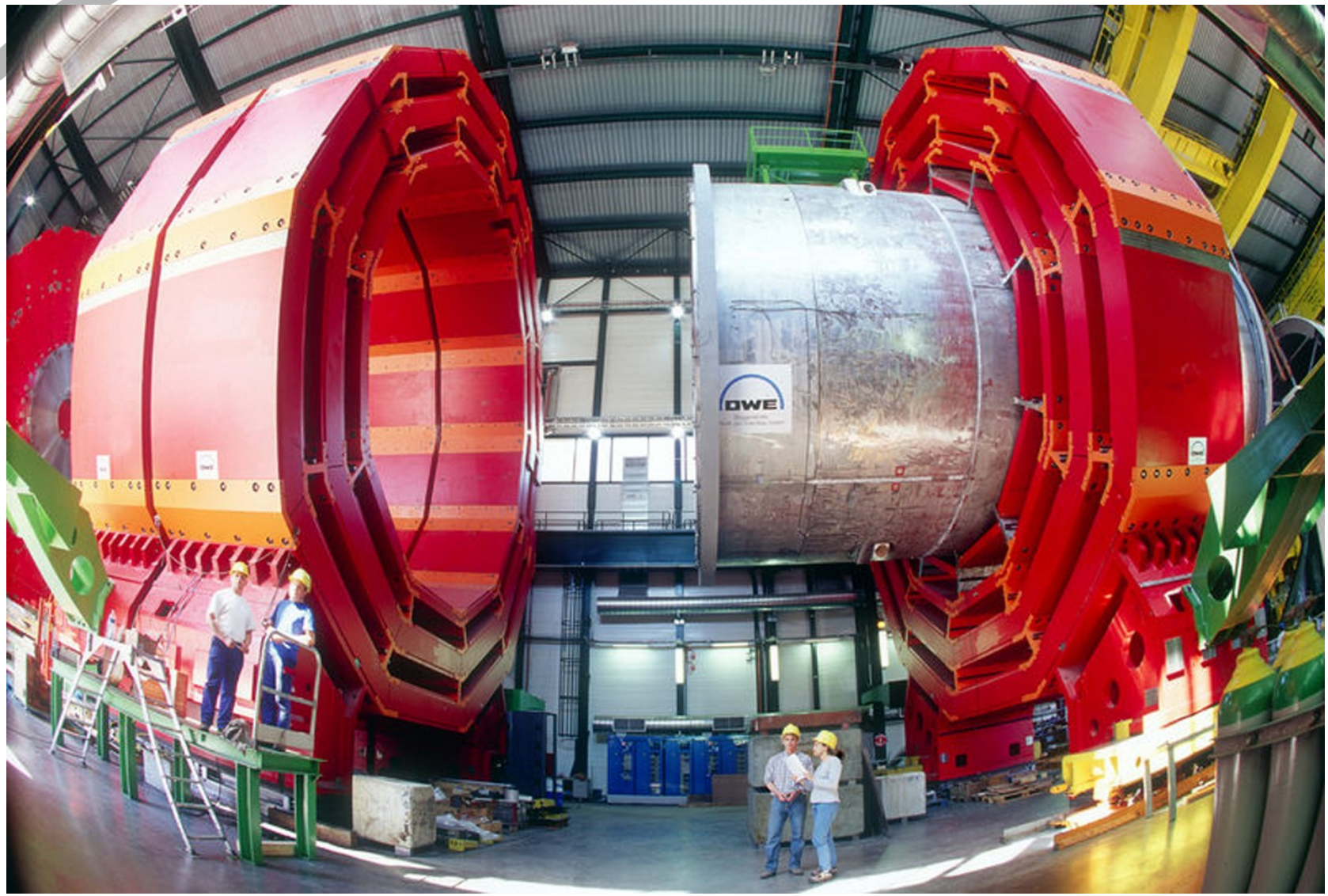
ATLAS

- General purpose LHC detector – 7000 tons





Compact Muon Solenoid
(CMS)



CERN openlab

Data management and computing

LHC data (simplified)

Per experiment:

- 40 million beam interactions per second
- After filtering, 100 collisions of interest per second
- A Megabyte of digitized information for each collision = recording rate of 0.1 Gigabytes/sec
- 1 billion collisions recorded = 1 Petabyte/year

1 Megabyte (1MB)
A digital photo

1 Gigabyte (1GB)
= 1000MB
A DVD movie

1 Terabyte (1TB)
= 1000GB
World annual book production

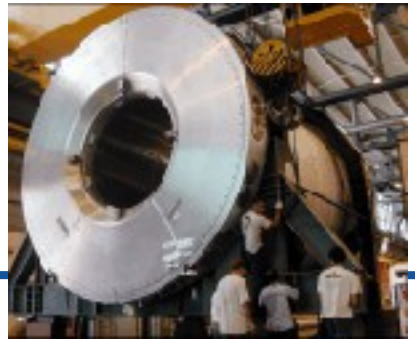
1 Petabyte (1PB)
= 1000TB
The annual production by one LHC experiment

1 Exabyte (1EB)
= 1000 PB
World annual information production

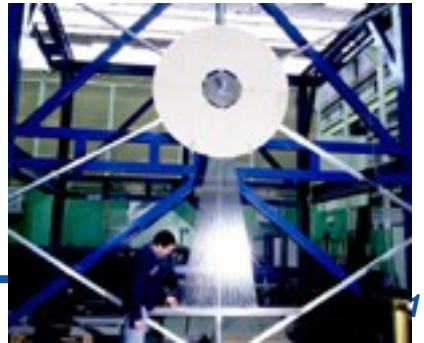
CMS

LHCb

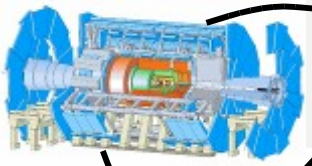
ATLAS



ALICE



Data Handling and Computation for Physics Analysis



selection & reconstruction

Reconstruction

event summary data

10%

processed data

100%

raw data

event reprocessing

batch physics analysis

Analysis

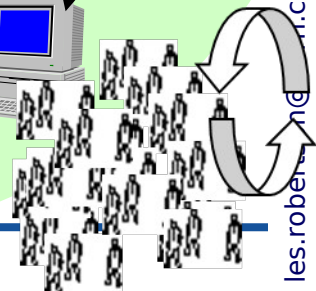
1%

analysis objects
(extracted by physics topic)

event simulation

Simulation

interactive physics analysis

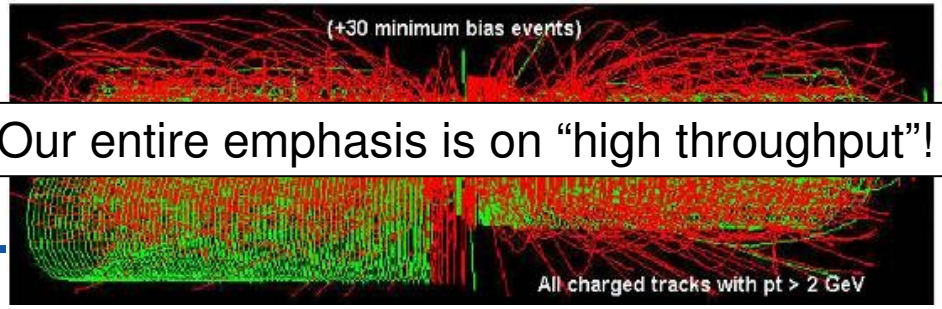


les.robertson@cern.ch

High Energy Physics Computing Characteristics

- Independent events (collisions of particles)
  - trivial (read: pleasant) parallel processing
- Bulk of the data is read-only
  - versions rather than updates
- Meta-data in databases linking to “flat” files
- Compute power scales with SPECint (not SPECfp)
  - But good floating-point (20% of total) is important!
- Very large aggregate requirements:
  - computation, data, input/output
- Chaotic workload –
  - research environment - physics extracted by iterative analysis, collaborating groups of physicists

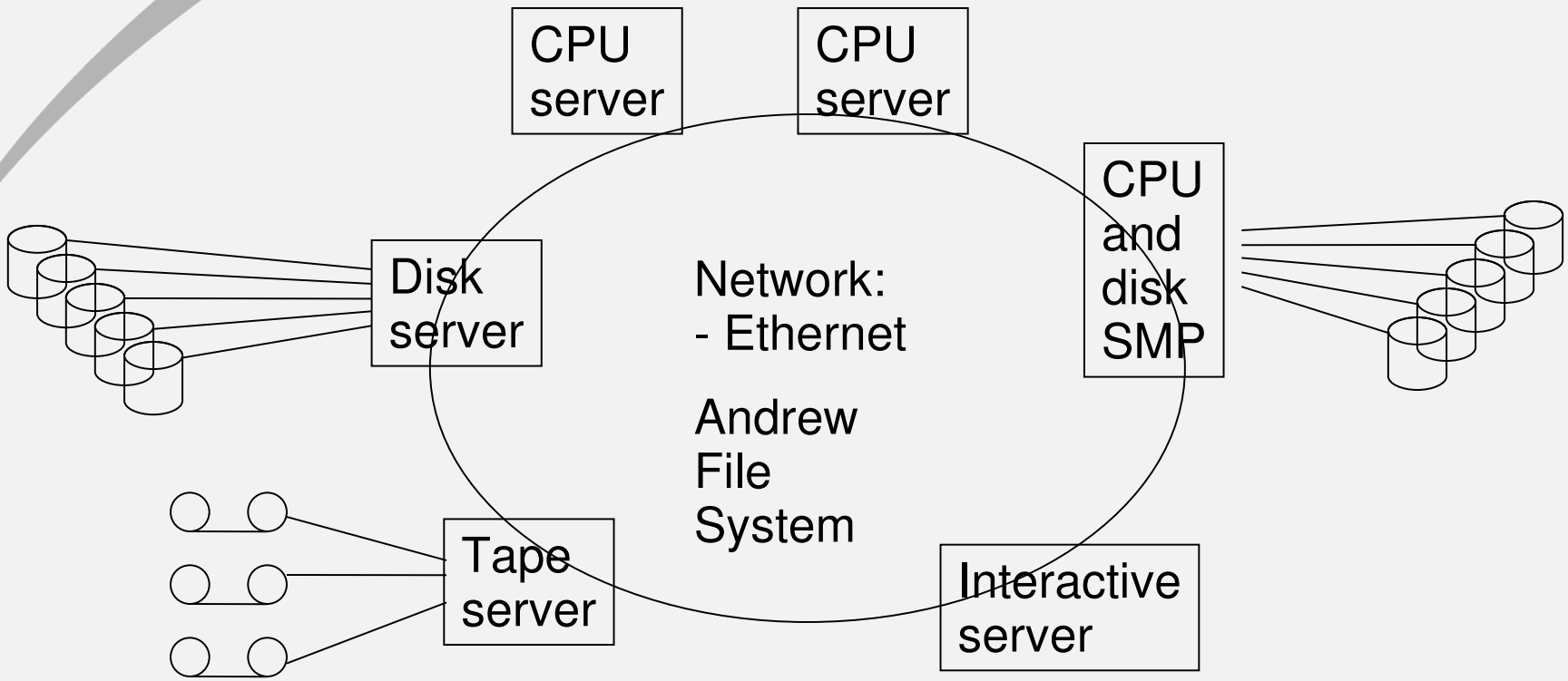
  → Unpredictable → unlimited demand



Our entire emphasis is on “high throughput”!

SHIFT architecture

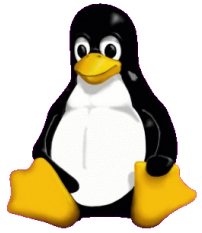
(Scalable Heterogeneous Integrated Facility)



In 2001 SHIFT won the 21st Century Achievement Award issued by Computerworld

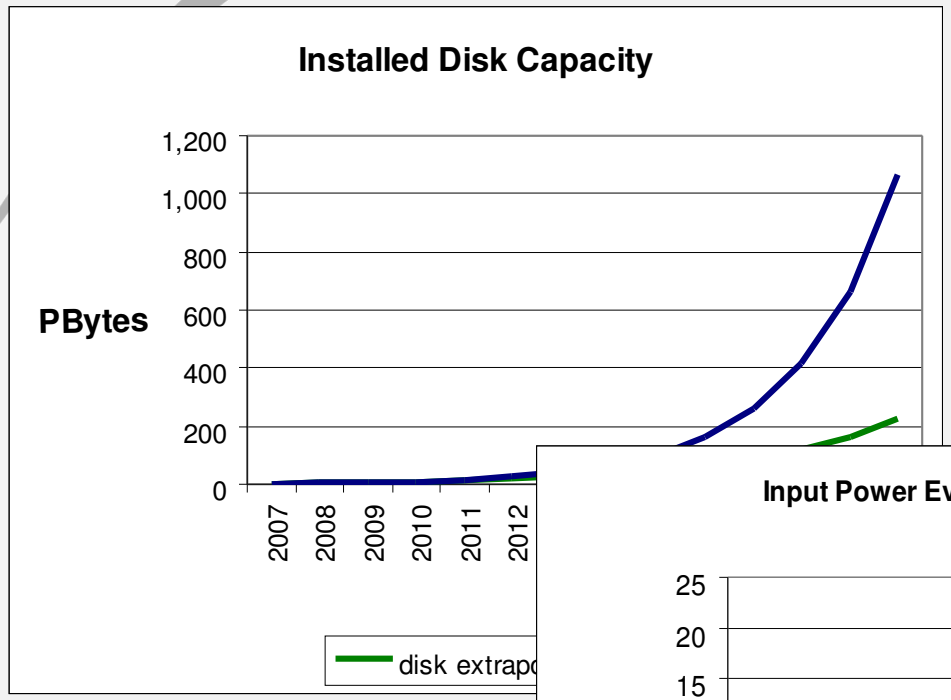
Computing at CERN today

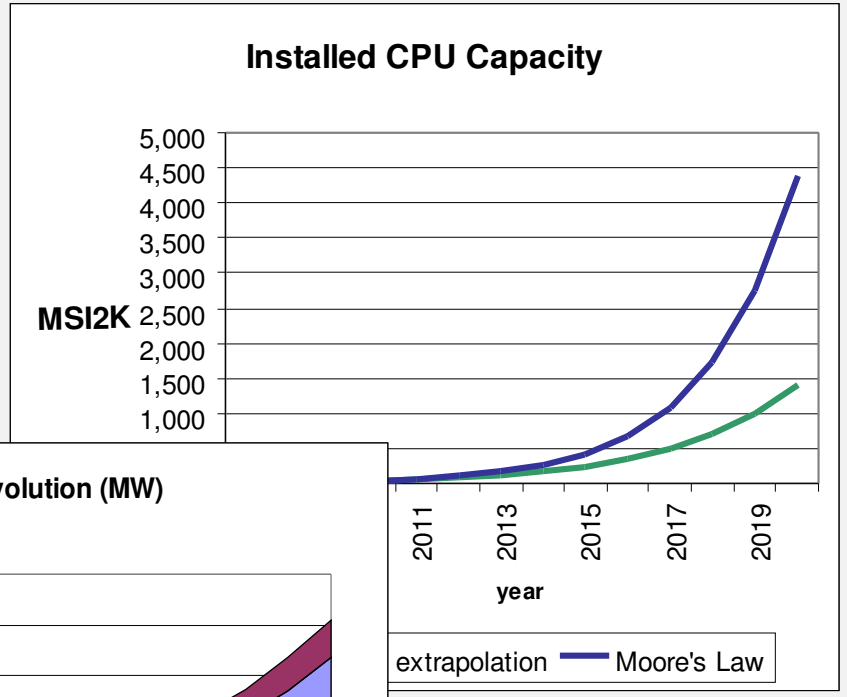
- High-throughput computing based on reliable “commodity” technology
- About 3000 dual-socket PC servers running Linux
- Petabytes of data on tape and on disk

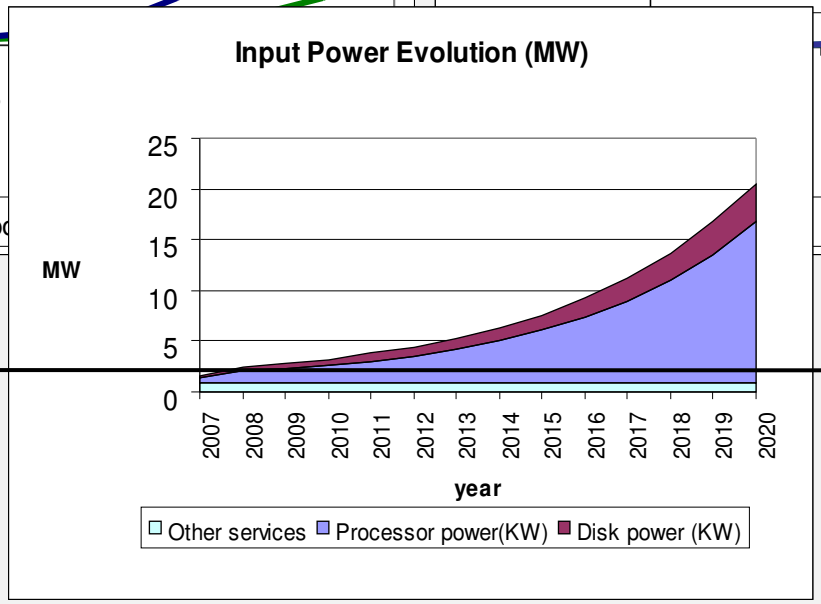


LHC computing capacity development

Development of computing capacity with a constant budget, given the increased cost for power and cooling



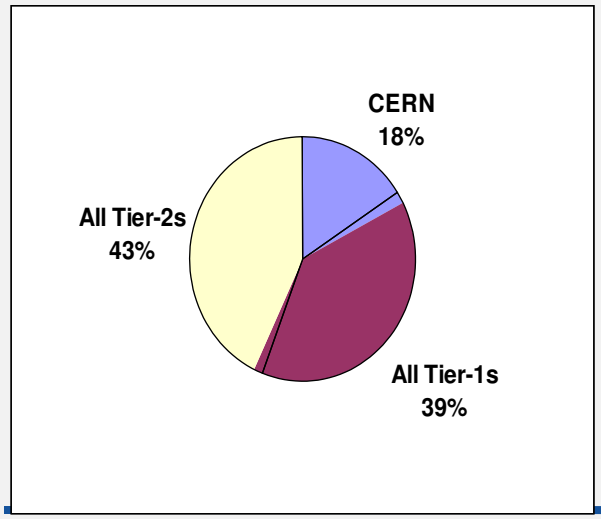


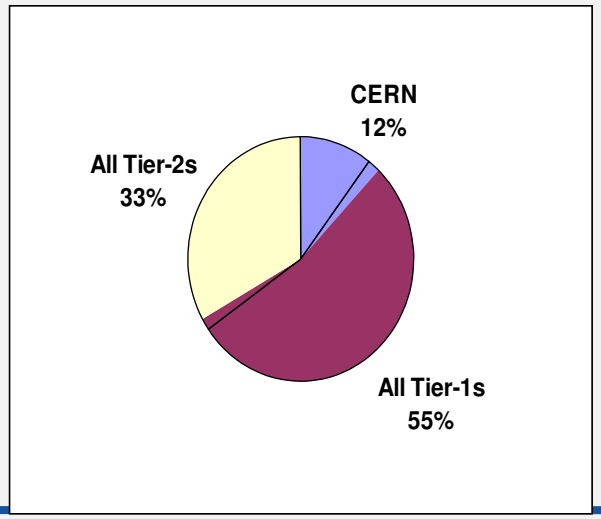


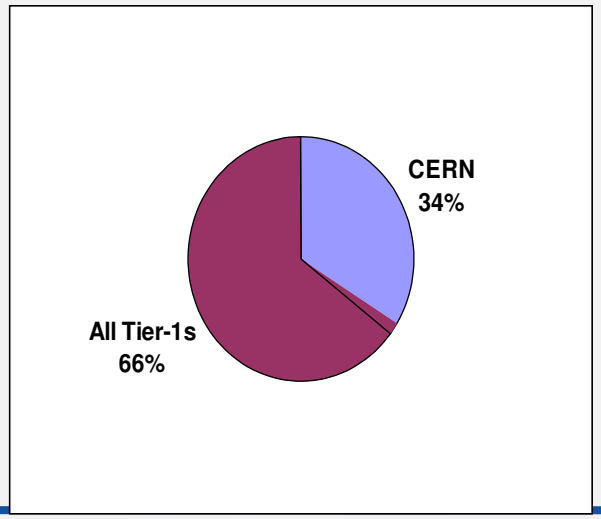
Power limit of present Computer Centre

Why do we need a Grid?

- The LHC Computing requirements are simply too huge:
  - Political resistance to putting everything at CERN
  - Impractical to build such a huge facility in one place
  - The users are in any case not necessarily at CERN
  - Modern wide-area networks have made distances shrink
    - But, latency still has to be kept in mind
- So, we are spreading the burden!



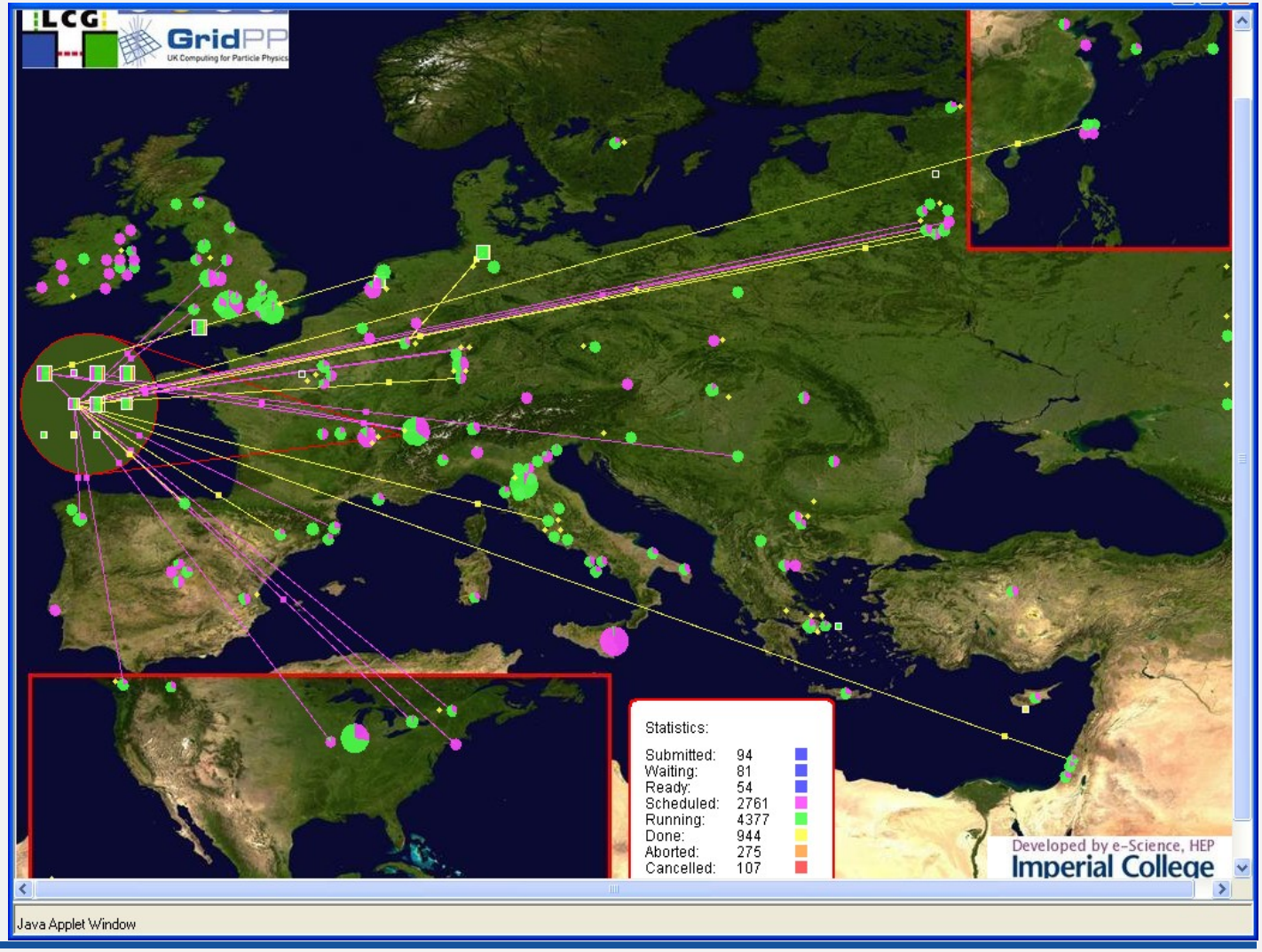




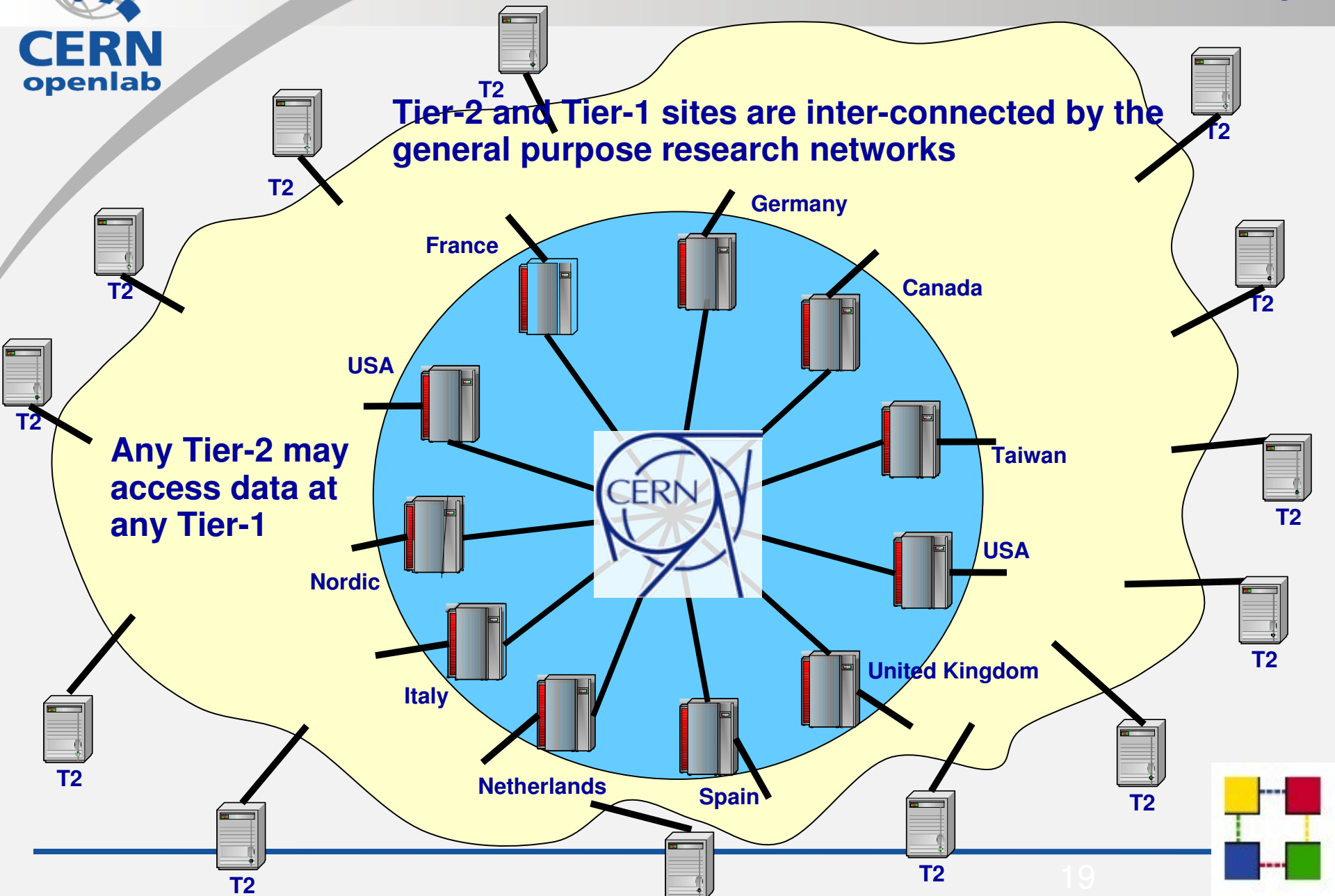
WLCG

- Largest Grid service in the world !

- **Almost 150 sites in 35 countries**
- **100’000 IA-32 processor cores (w/Linux)**
- **Tens of petabytes of storage**

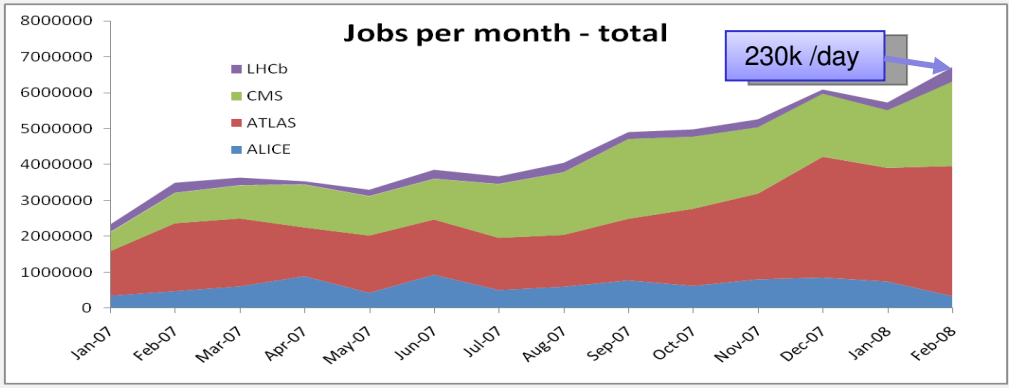


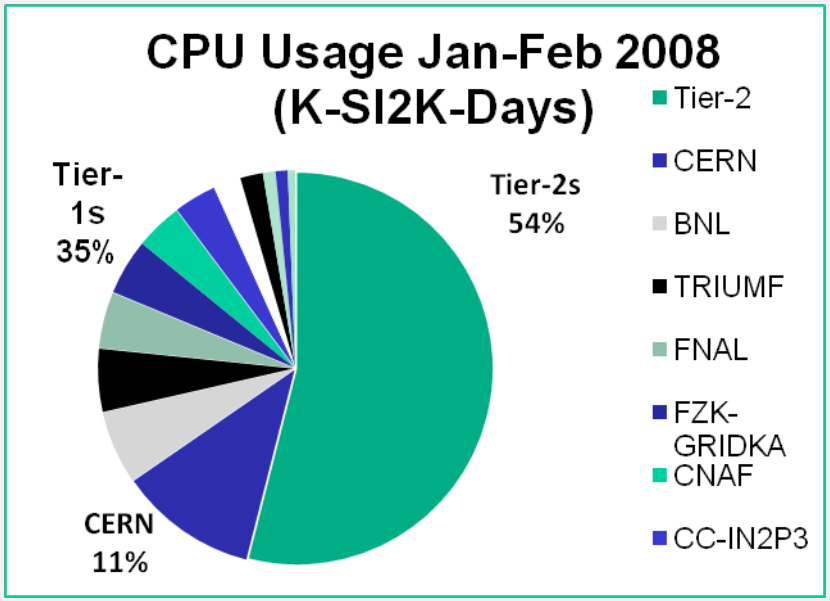
WLCG connectivity



WLCG Usage

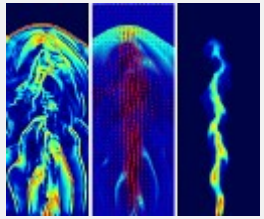
- WLCG ran ~ 44 million jobs in 2007 – workload has continued to increase
- Distribution of work across Tier0 / Tier1 / Tier 2 really illustrates the importance of the grid system
  - Tier 2 contribution is around 50%; > 85% is external to CERN
- Data distribution from CERN to Tier-1 sites
  - Latest test in February show that the data rates required for LHC start-up have been reached and can be sustained over long periods

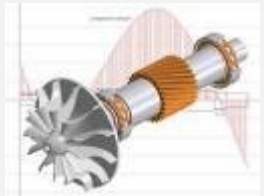




- EGEE
  - Started in April 2004, now in second phase with 91 partners in 32 countries
  - 3rd phrase (2008-2010) started 1st May 2008

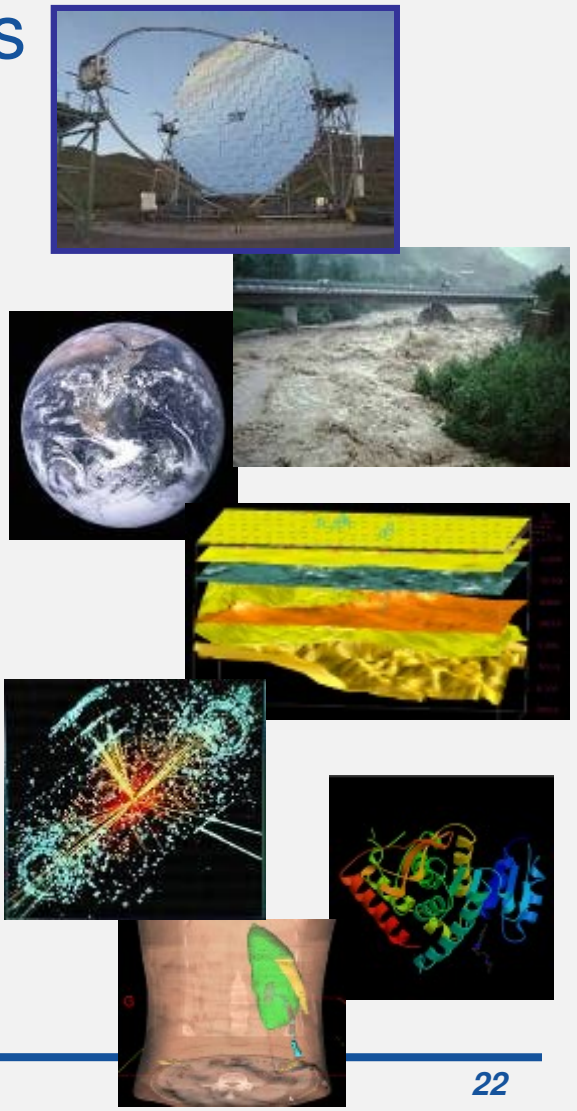
- Objectives
  - Large-scale, production-quality grid infrastructure for e-Science
  - Attracting new resources and users from industry as well as science
  - Maintain and further improve “gLite” Grid middleware





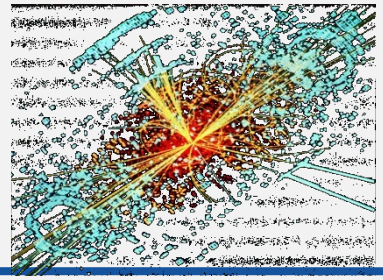
EGEE Application Domains

- More than 25 applications from an increasing number of domains
  - Astrophysics
  - Computational Chemistry
  - Earth Sciences
  - Financial Simulation
  - Fusion
  - Geophysics
  - High Energy Physics
  - Life Sciences
  - Multimedia
  - Material Sciences
  - …..



Background to the CERN openlab

- Information Technology has ALWAYS moved at an incredible pace
- During the LEP era (1989 – 2001) CERN changed its computing infrastructure twice:
  - Mainframes (1x) → RISC servers (30x) → PC servers (1000x)
- In openlab, we collaborate to harness the advantages of a continuous set of innovations for improving scientific computing, such as:
  - 10 Gigabit networks, 64-bit computing, Virtualization
  - Performance improvements (Moore's law): HW and SW
  - Many-core throughput increase, Thermal optimization
- We work with a long-term perspective:
  - LHC will operate until at least 2020!

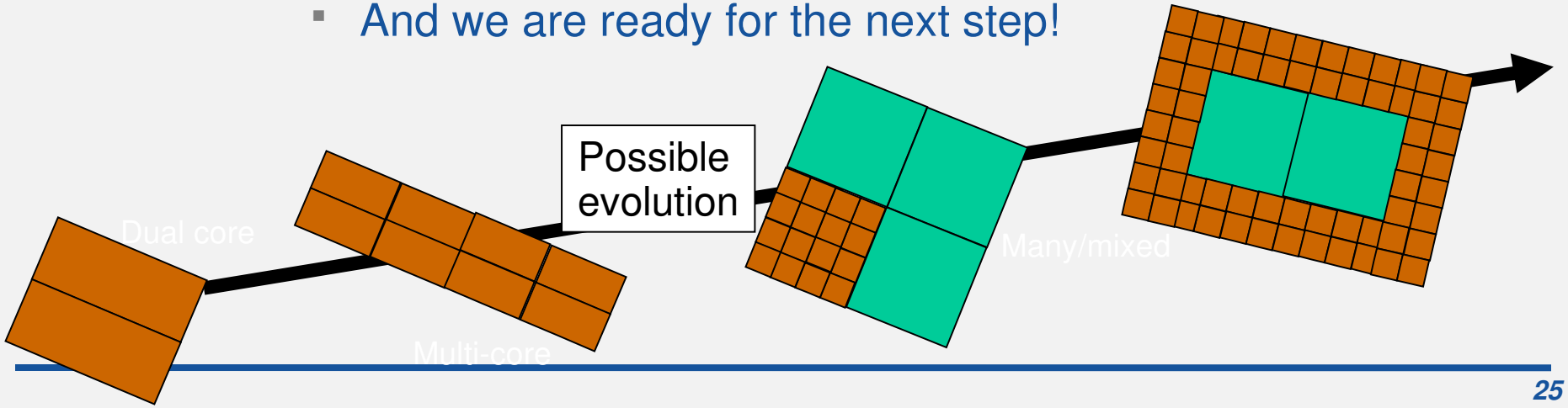


Platform Competence Centre

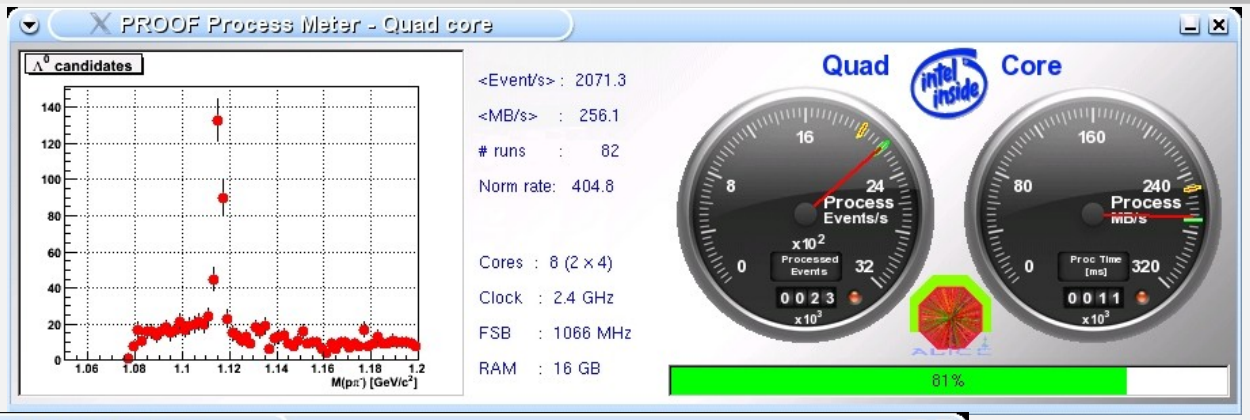
- Intel-related activities:
  - Performance/throughput improvements
    - Compiler improvement project
    - Tuning of physics applications
    - Performance Monitoring
    - Benchmarking w/SPEC and Oracle
  - Multithreading applications
  - TOP500 runs
  - Virtualization
  - Thermal optimization
    - Servers and entire Computer Centre
  - 10 Gb networking
  - New processors/new languages

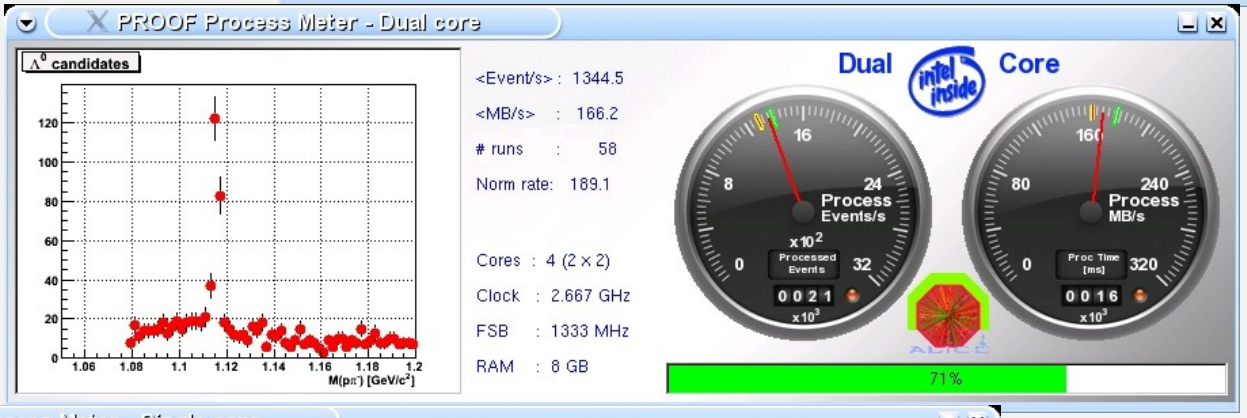
From Multi to Many

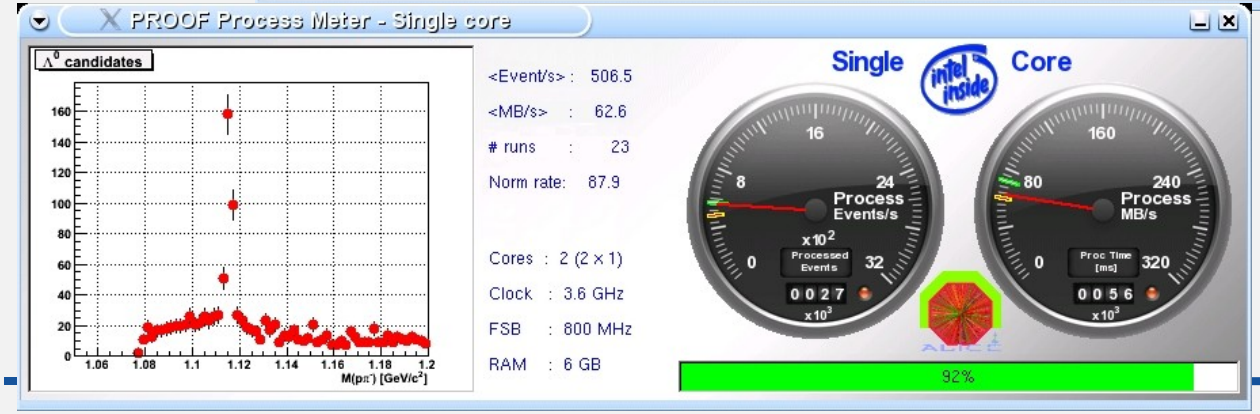
- Our “high throughput” computing model is ideally suited:
  - Independent processes can run on each core, provided that:
    - Main memory is added
    - Bandwidth to main memory remains reasonable
  - Testing, so far, has been very convincing
    - Single → Dual → Quad
- In November 2006, Intel announced their Quad-core processor in the CERN Globe
  - All our acquisitions are now QC (Clovertown/Harpertown)
  - And we are ready for the next step!



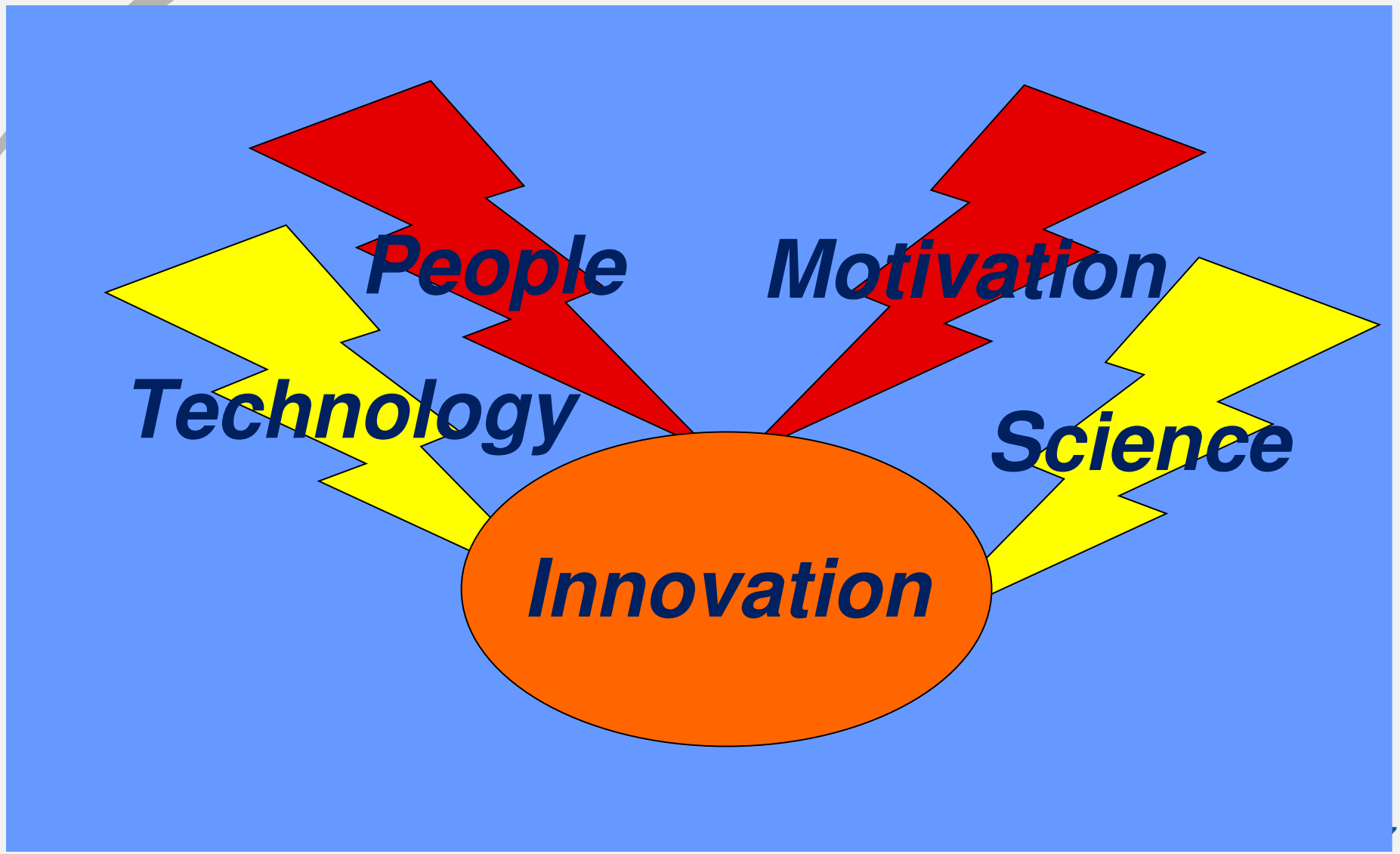
Multicore comparisons







Innovation in Computing

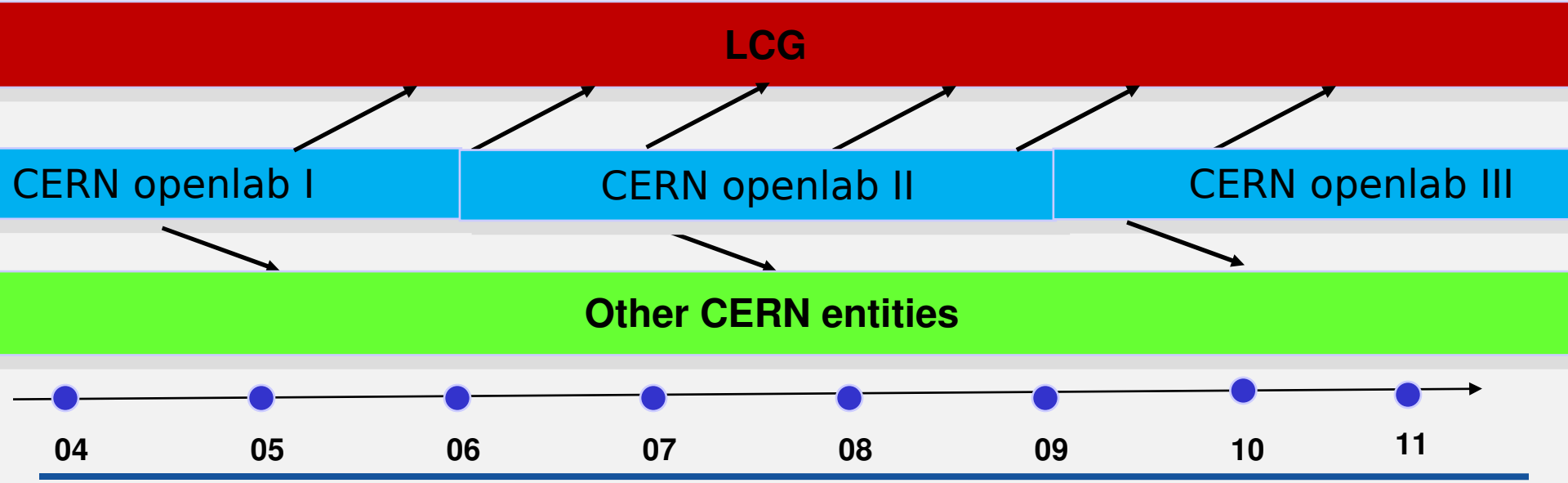


CERN openlab

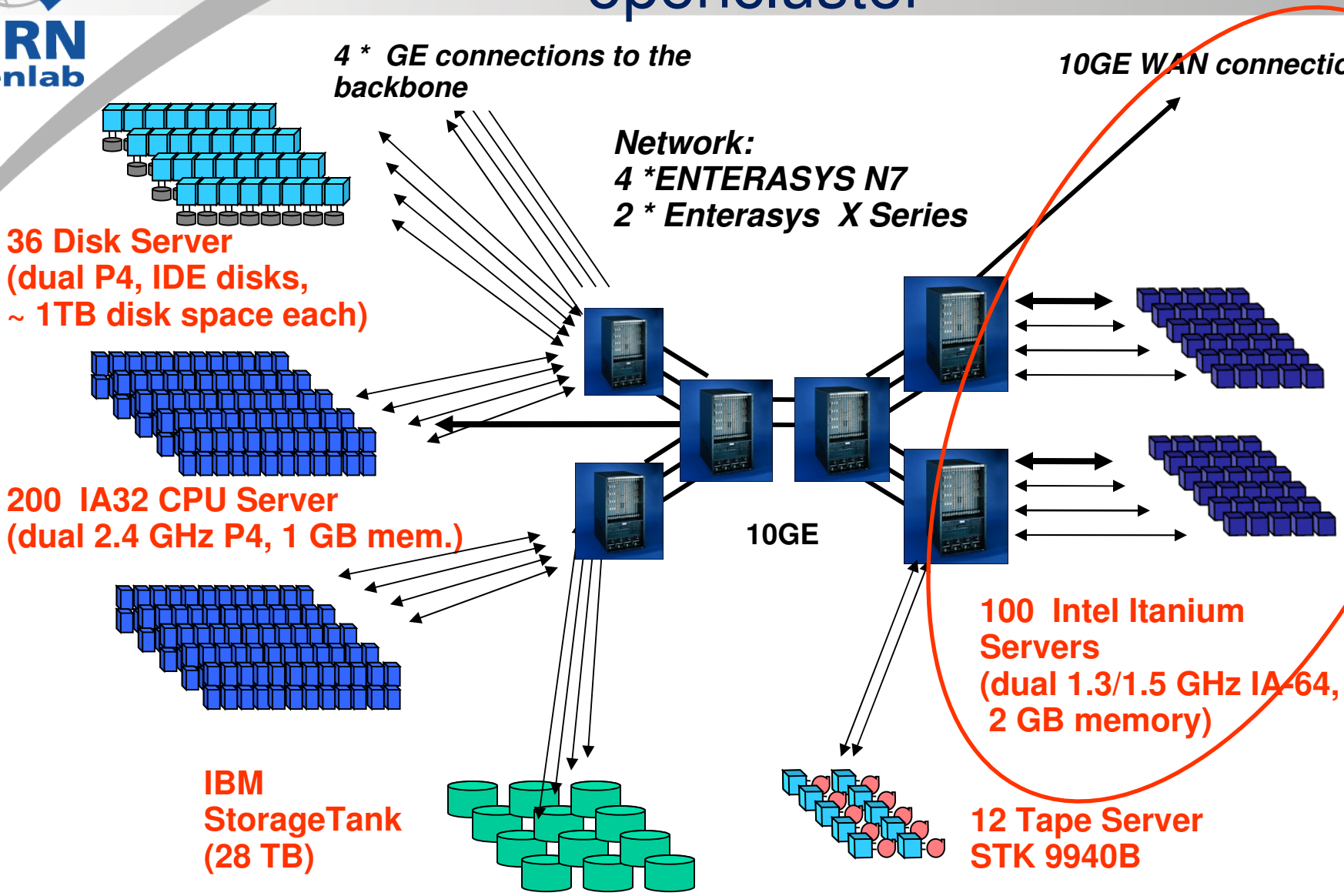
Backup

The CERN openlab

- Department's main R&D focus
- Framework for collaboration with industry
- Evaluation, integration, validation
  - **of cutting-edge technologies that can serve the LHC Computing Grid (LCG)**
- Sequence of 3-year agreements
  - **2003 – 2005: Phase I: the "opencluster" project**
  - **2006 – 2008: Phase II Multiple Competence Centres**



High performance test bed: the opencluster



Multi-threading activities

- Aim: Evangelize/teach parallel programming
- Two workshops arranged w/Intel in 2007
  - 1 day lectures, 1 day exercises
  - 5 lecturers (2 Intel + 3 CERN)
  - 45 participants
  - Next workshop: Late Spring 2008
- Licenses for the Intel Threading Tools (and other SW products) available
  - to all CERN users
- Advances in Geant4 parallelization experiment
- Recently also Minuit/ROOT