

CERN openlab and Intel: Today and Tomorrow

> Sverre Jarp CERN openlab CTO 16 September 2008



LHC first beam 10 September 2008





# 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, ...)

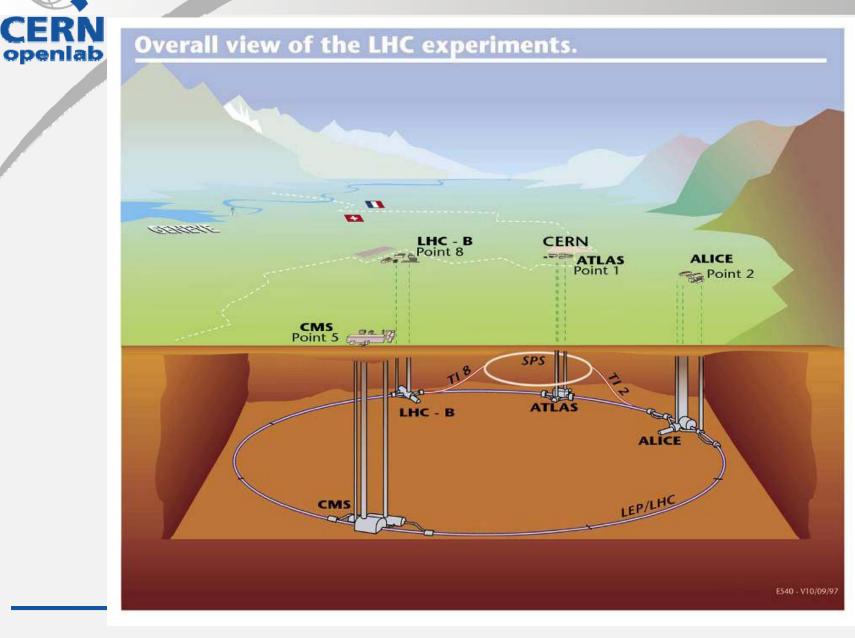
- Some 6500 visiting scientists (half of the world's particle physicists)

> They come from 500 universities representing 80 nationalities.



## The **CERN** Site CERN Mont Blanc, 4810 m Downtown Geneva LHCb ATLAS **CERN** sites CMS ALICE

### **CERN** underground



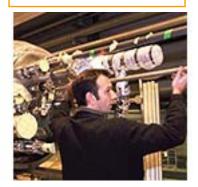
ER

## What is LHC?



- The Large Hadron Collider can collide beams of protons at an energy of 14 TeV (<u>inaugurated on 10 September!</u>)
- Using the latest super-conducting technologies, it operates at about – 271°C, just above the temperature of absolute zero.
- With its 27 km circumference, the accelerator is the <u>largest</u> superconducting installation in the world.

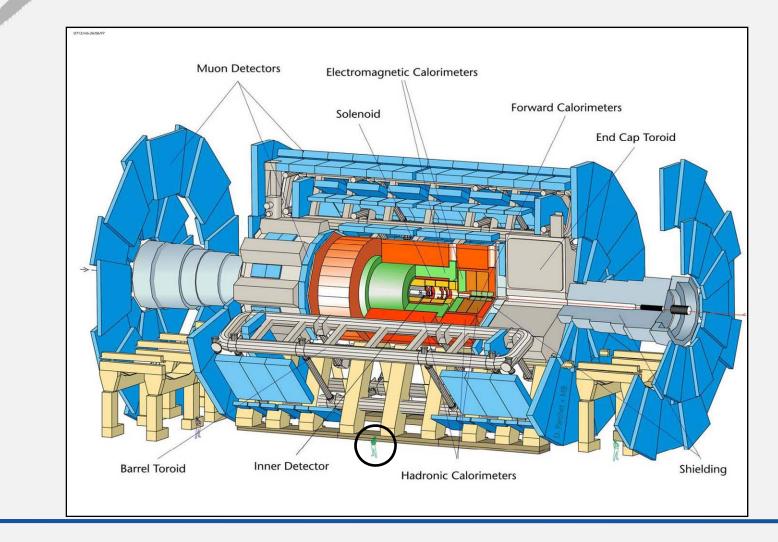
Four experiments, with detectors as 'big as cathedrals': ALICE ATLAS CMS LHCb

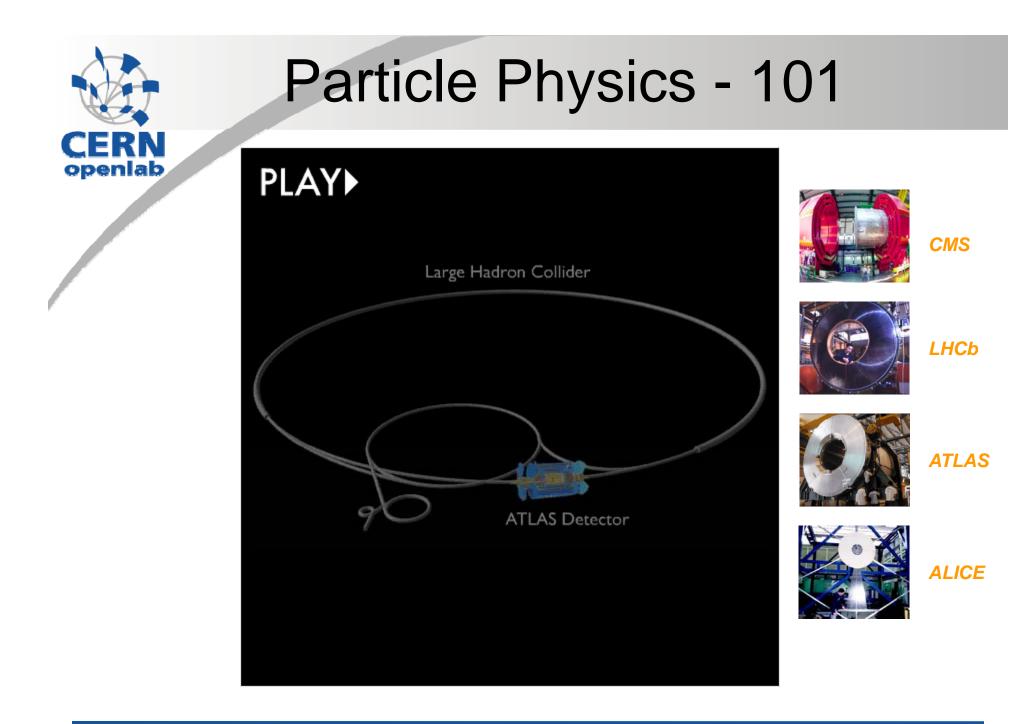




### ATLAS

### General purpose LHC detector – 7000 tons







# 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



### Computing at CERN today



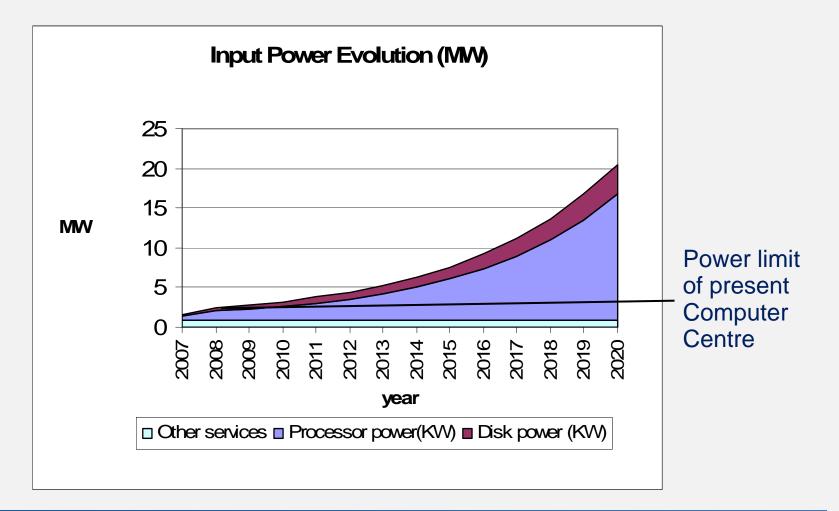
Nowhere near enough!

- High-throughput computing based on reliable "commodity" technology
- About 3500 multi-socket multi-core PC servers running Linux
- More than 10 Petabytes of data on tape; 30% cached on disk



### **Expected Power Evolution**

Demand will grow continuously through the LHC era



## LHC Computing Grid

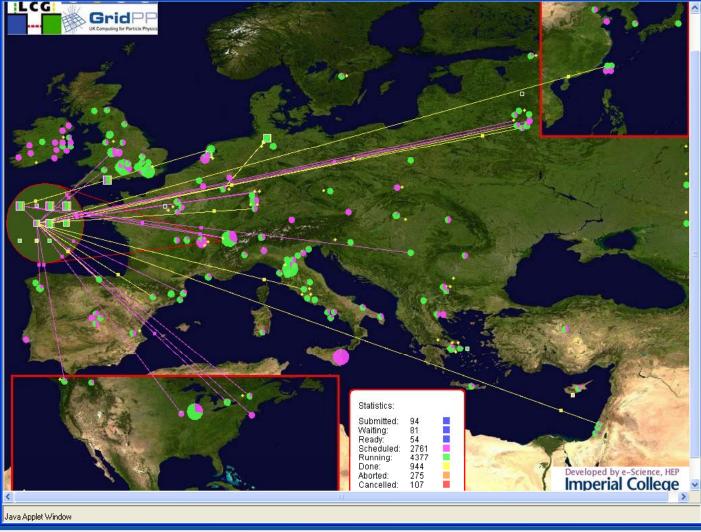
### Largest Grid service in the world !

• Almost 200 sites in 39 countries

openlab

• 100'000 IA processor cores (w/Linux)

• Tens of petabytes of storage

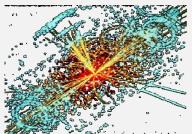


### CERN openiab

### 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  $(1x) \rightarrow RISC$  servers  $(30x) \rightarrow PC$  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!





# The CERN and Intel collaboration in openlab



### Platform Competence Centre

Related to today's Xeon- \_ 7400 announcement

### Intel-related activities:

- Thermal optimization
  - Servers and entire Computer Centre
- Virtualization
- Multi-core performance/throughput improvements
- 10 Gb networking
- Sneak peaks at new technologies

### Thermal management



- Joint white paper on Computer Centre efficiency
  - Based on issues with existing building
    - Just issued; see next slide
- Complementary project to understand thermal characteristics of each server component
  - Processors (frequencies and SKUs); Memory (type and size); Disks; I/O cards; Power supplies



### Project for new Computer Centre

- Understand all relevant issues (before starting)
- Aim at 2.5 + 2.5 MW



### Joint white-paper just issued

 Reducing Data Center
 Energy
 Consumption



### Reducing Data Center Energy Consumption

A summary of strategies used by CERN, the world's largest physics laboratory

White Paper Intel<sup>®</sup> Xeon<sup>®</sup> Processor Data Center Optimization To deploy massive new computing resources without exceeding the thermal limits of its 35-year-old data center, CERN is taking a comprehensive approach to improving energy efficiency. This paper outlines CERN's key strategies, including a move to the latest Intel<sup>®</sup> Xeon<sup>®</sup> processors that are helping the organization increase performance per Watt by a factor of five.



### Executive Summary

"Multi-core processors based on the Intel<sup>®</sup> Core" microarchitecture deliver about five times more compute power per Watt than singlecore processors based on the earlier Intel NetBurst<sup>®</sup> microarchitecture." - CERNI

Improving data center energy efficiency is becoming a fundamental requirement for most organizations, not only to contain operating costs, but also to support growth, extend the life of existing facilities, protect the environment, and address increasing regulatory requirements. Electricity costs are rising fast. Most businesses already spend about half as much for the electricity to power and cool their infrastructure as they do for the hardware itself, and this percentage can be expected to increase.

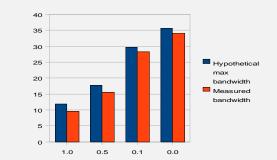


### Virtualization

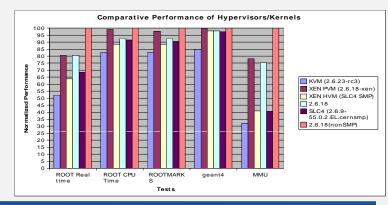
CERN

### Multiple aims:

- Server consolidation
- System testing (used in LCG)
- Improved flexibility and security
  - also in the grid
- Personalization of images
- Also: Development of complementary tools:
  - OSFarm
  - Content-Based Transfer mechanism
- Benchmarking



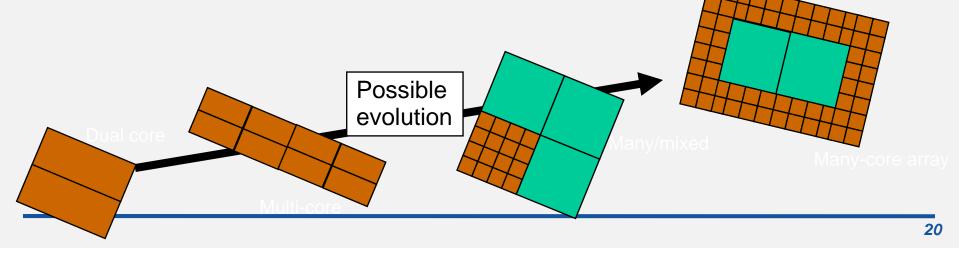
XS Len	Republicy About Log Statue Simple request Victual Appliances report Advanced request
software packages if need	ender OS images, and "virtual appliance" for use with Xen VMs. To create an image, enter a name for the image and select a "Class" a d. Click "Ceater image", and the image will be created and put in the <u>recontory</u> . If you check the "Doweload image upon creation" checks darwhen The image creation is firshed.
	e", the image will be named after the mdS checksum of the image configuration parameters. If an image with the exact same parameters existenced and can be downloaded immediately.
If you want to use wpet, th	an hara is an arannsia urt
	<pre>\create?name=&amp;download=on&amp;class=SLC4&amp;arch=386&amp;filetype=tar&amp;goup=core&amp;goup=base&amp;package=gite=EDI1*</pre>
"http://www.cem.ch/odfam	
"http://www.cem.ch/odfam	vcreate?name=&download=on&class=SLC4&arch=U86&lfletype= tarkgroup=core&group=base&package=gite=BDII*
"http://www.cem.ch/odfam	investel*name.adsemb.ad.enakciass-SLC44.achi-384.8f.letype=.tarlignop-corelign-up-baselipsckage-gite-EDI1* s for the image to be created.
"http://www.cem.ch/odfam	iventifehammadownitest-indicises-61,048ach-49683/irrype-tarligonup-constgorup-tasetpackage-gitre EDI** Stor the image to be created. Name
"http://www.cem.ch/odfam	iventerhammeddownbationakiasa6,044anh-0868/ingyes.tarkgosp.corekgosp.tasetpackage.gite EDI* s for the image to be created. Name
"http://www.cem.ch/odfam	Stortel hamme download and krises 61,04 karch 43 66 61 krypes, tarkgroup-corektgroup-basetpackage-gite EDI**  Storte image to be created.  Same  Synchronoux  Cases  SC-4



### Cores: From Multi to Many

CERN

- Our "high throughput" computing is ideally suited:
  - Independent processes can run on each core, provided that:
    - Main memory is added
    - Bandwidth to main memory remains reasonable
  - In openlab, we have had early access to multiple generations:
    - Woodcrest, Clovertown, Harpertown, Dunnington; Montecito
- Already in November 2006, we were proud to be part of Intel's movement to Quad core
  - All recent acquisitions have been Quad-core systems
  - And we are ready for the next step!



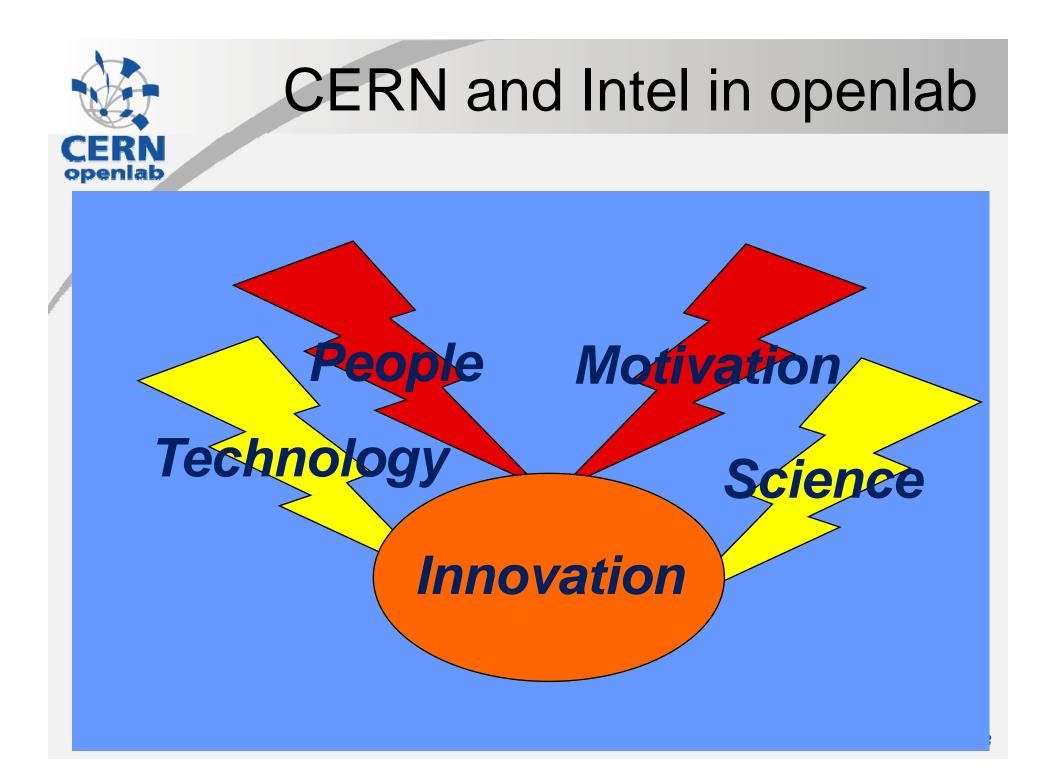
### **Multicore comparisons** 🕤 🔍 🗙 PROOF Process Meter - Quad core Λ<sup>0</sup> candidates Quad Core openlab <Event/s>: 2071.3 140 <MB/s> : 256.1 160 120 # runs : 82 100 Norm rate: 404.8 Process Process 80 Events/s x 10 60 Cores : 8 (2 × 4) 320 40 Clock : 2.4 GHz 0011 0023 20 : 1066 MHz FSB 1.1 1.12 1.14 1.16 1.18 1.2 1.08 RAM : 16 GB 1.06 81% M(pπ') [GeV/c<sup>2</sup>] 🕤 🔍 🗙 PROOF Process Meter - Dual core LX Λ<sup>0</sup> candidates Core intel " insid Dual <Event/s> : 1344.5 <MB/s> : 166.2 120 58 # runs 100 Norm rate: 189.1 24 Process Events/s 80 Process 60 x10<sup>2</sup> Cores : 4 (2 x 2) 320 32 40 Clock : 2.667 GHz 0021 0016 x 10<sup>3</sup> 20 FSB : 1333 MHz 0 1.16 1.18 1.2 Μ(pπ) [GeV/c<sup>2</sup>] 1.06 1.08 1.1 1.12 1.14 RAM : 8 GB 💿 🤍 🗙 PROOF Process Meter - Single core Single Core Λ<sup>0</sup> candidates <Event/s> : 506.5 160 <MB/s> : 62.6 140 F 23 # runs 120 Norm rate: 87.9 Process Proces 100 80 Cores : 2 (2 x 1) 60 Clock : 3.6 GHz 40 F 0027 0056 20 F FSB : 800 MHz 1.1 1.12 1.14 1.16 1.18 1. Μ(pπ<sup>-</sup>) [GeV/c<sup>2</sup>] 1.06 1.08 1.2 RAM : 6 GB

### Conclusions



 As a flagship scientific instrument, the Large Hadron Collider and the corresponding Computing Grid will be around for 15 years

- It will rely on continued innovation in Information Technology:
  - Thermal improvements
  - Multicore throughput improvements
  - Virtualization improvements
- Today's Xeon announcement is a great step in this direction!





# Backup