A Brief and Partial History of Computing in HEP

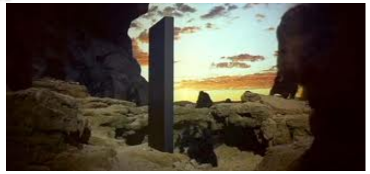
Roger Jones

Lancaster University

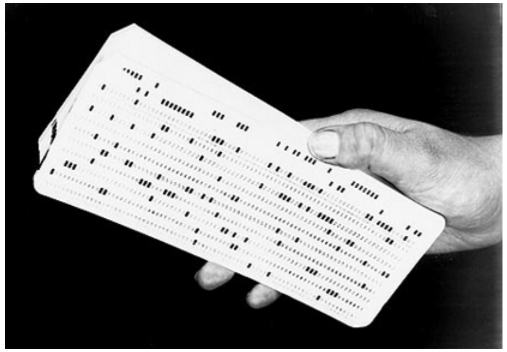
NormanFest November 2011

An Introduction and Apology

- In the original planning, the older history would have been covered by Chris Jones, who sadly cannot join us
- As a result, the view will be more slanted to my shorter and LHC view than might have been – not from the Dawn of Time.



  - It's long enough though – from an age (1983) when punched cards were still being used - though only for writing notes!
  - Not entirely inappropriate – ATLAS computing is one of the bigger contributions Norman has made
  - And LHC computing planning started soon after



- Thanks to those who, knowingly or otherwise, contributed to this talk
  - Notable names include Jim Shank, David Rouseau, Alistair Hart of Cray, Roger Goff of Dell, Neil Geddes and Norman himself

The World in 1983 & Now

- The interface on the desk was a VT100, if you were lucky
  - More likely you had a line-mode editor and anything vaguely graphical used something specialized like a Tektronix terminal

- Memorex full-screen terminals seemed like the Space Age and Megatek 3-d displays for UA1 etc seemed like magic

- Now we can do all that, and store a full LEP experiment dataset on a laptop!

Experiment Computing Evolution

- Perhaps more significantly, where computing is done kept changing
  - In the mid-1980s, ‘big’ experiments processed their data at the accelerator lab
  - This model lasted until into Run II of the Tevatron
  - Small experiments took their data away to be processed
    - On WA21, I had the whole data set on <10 round tapes, and did all the processing in Birmingham on an IBM mainframe and 370E emulators
- Experiment analysis was starting to be distributed, but usually to only a few large centres
- What changed?
  - The size and distribution of the collaborations
  - The volume of data
  - The politics of funding – people like to spend locally
    - Leverage from countries & from universities

HEP Distributed Computing Grew Through the 1990's

- **E.g. LEP**
  - Distributed Monte Carlo generation
  - Initially tape and then network data transfer
- **HERA**
  - Funnel – remote event processing
- **BaBar**
  - Planned a much more distributed model based on regional centres

Was the Message Getting Through?

5.6 Offline Computing Requirements

In the short term (i.e. in the preparation of the technical proposal) we will intensify our Monte Carlo studies in order to approach a final detector design. We estimate that the computing time required at CERN will be of the order of 500 000 hours (CERN units) of which 90% can be on the CSF farm with the remaining 10% on the IBM.

The m... by further M... ent, and a bu... uc- tion and analysis. We shall rely on the expertise at CERN in providing library packages and in coordinating the software development.

ATLAS LoI 1992
Offline budget ~5MCHF

In the long term, the requirements for data storage, computing power and computing infrastructure will exceed those of current experiments by several orders of magnitude. The data volume (raw data, DST, and Monte Carlo) is estimated to be several hundred Tbytes/year. A hierarchical system of data storage and access is essential.

The computing power needed for first-pass reconstruction at LHC is estimated to be three orders of magnitude larger than that available on currently installed online reconstruction farms. In 1999 it can be expected that the increase of computing power of processors will compensate for a good part of higher demand, such that a system of about 100 processors will fulfil the task.

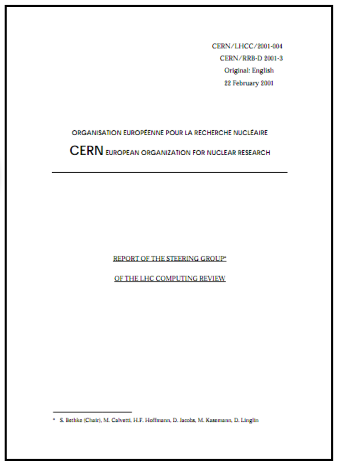
Most analysis and program development will be done on clusters of workstations. In an experiment with more than 1000 physicists, a few hundred workstations will be installed at CERN and in the other laboratories. The necessary infrastructure (file servers, high-speed networking, and manpower for system management) will have to be provided.

Table 9.2: *Preliminary cost estimates (in MCHF)*

Detector subsystem (baseline or baseline option)	Material cost mechanics or system	Material cost electronics	Total subsystem material cost
Inner detector			74
- SITV	15.0	10.0	
- GaAs	1.5	1.5	
- SIT	9.5	9.5	
- MSGC	2.0	10.0	

Slowly Things Changed

- **Much progress through the 1990's**
  - **Better understanding of the requirements**
  - **Better understanding of the opportunities**
  - **MONARC project**
- **Scale of the problem much bigger than originally thought**
  - **Scale of the requirements**
  - **Costs**
  - **Impact on National budgets**
- **→LHCC review in 2000**

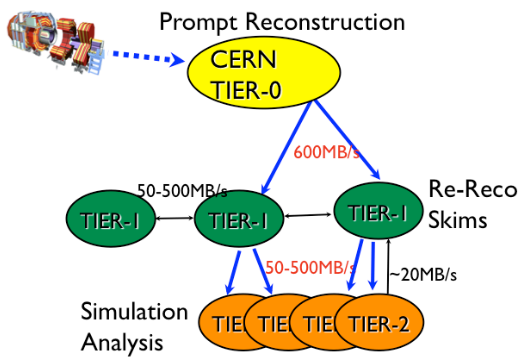


CERN/LHCC/2001-004
CERN/RRB-D 2001-3
Original: English
22 February 2001

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

REPORT OF THE STEERING GROUP*
OF THE LHC COMPUTING REVIEW

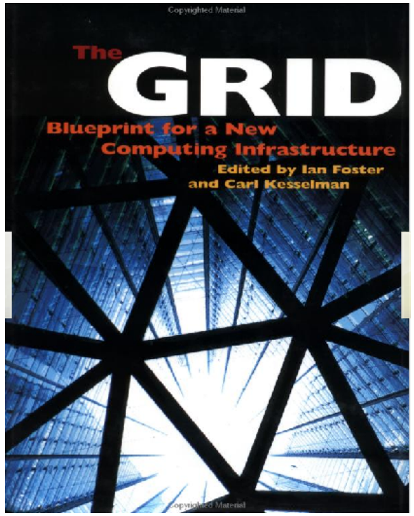
* S. Bethke (Chair), M. Calvetti, H.F. Hoffmann, D. Jacobs, M. Kasemann, D. Linglin



A Meme - The Grid

- **Timely vision from Foster and Kesselmann in 1998**

Highlighted a range of opportunities and technological developments in distributed computing, and how these could increasingly be exploited in large scale scientific



"*fundamentally change the way we think about – and use – computing*"

The LHC Computing Review

(Executive Summary - the LHC computing model)

1. review accepts scale of resource requirements of exp.s

2. recomm

10. hardware costs of initial set-up of LHC distributed computer centres (Tier-0 to -2): 240 MCHF
CERN-based Tier-0+1 centre: about 1/3 of total.

Significant uncertainties due to performance of LHC, detectors, triggers, backgrounds ...

workstations

3. GRID technology to be used (efficient resource usage, rapid turnaround)

4. need well-supported Research Networking of 1.5–3 Gbps (for each experiment), at affordable costs, by 2006.

Politicians and commerce saw the potential … and money appeared

Equity Research

Global

Global Technology

29 June 2000

New Economy

Forget the Web, Make Way for the Grid

Keith Westhead
(44) 20 7545 6489
keith.westhead@db.com

Chris Mortenson
(1) 212 469 8340
chris.mortenson@db.com

Jim Moore
(1) 415 477 4269
jim.moore@db.com

Andrea Williams Rice
(1) 415 617 3343
andrea.w.rice@db.com

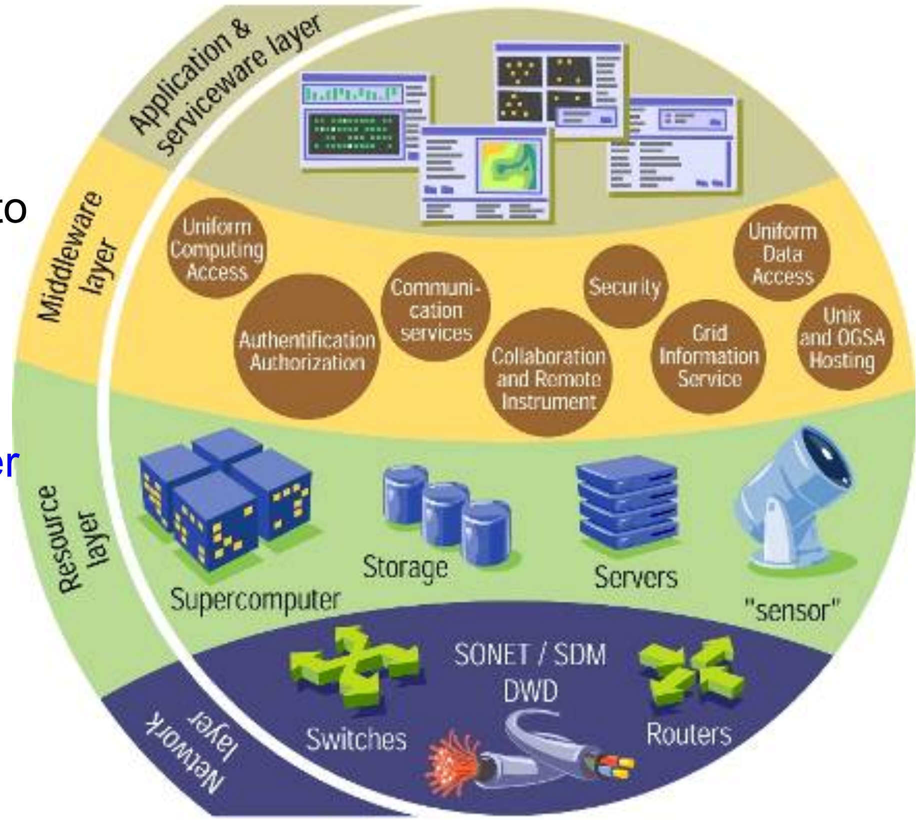
In conjunction with the Deutsche Bank Global Technology Team.

Deutsche Bank

Plus ca change?
But saved our bacon

The Grid Vision

- Users not concerned with with location of their files or CPUs
  - Partially achieved, though largely within the experiments
- Users have seamless transparent access to global resources.
  - Seamless? No, but workable
- Integration crosses organizational boundaries.
  - Up to a point – the UK does this better than most
- Coordinated resources not subject to centralized control
  - Hmm – I’m not sure the experiment operations would agree
- Using standard, open, general purpose protocols and interfaces
  - Let’s not over-egg it….

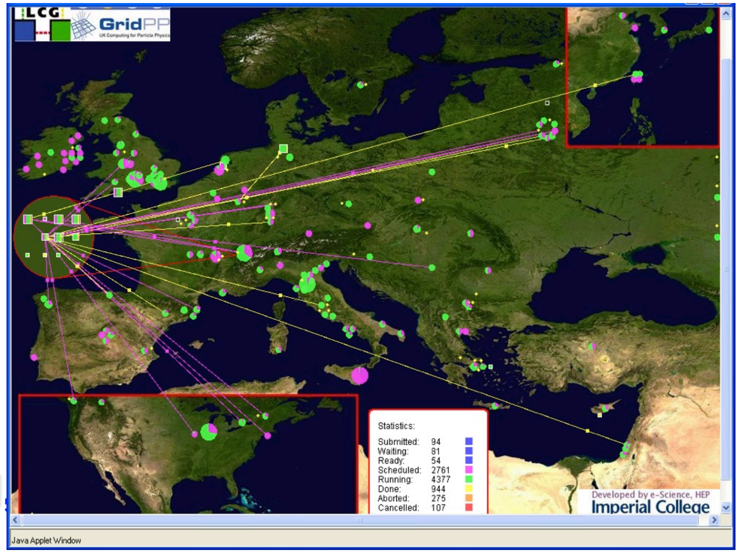


EU DataGrid/EGEE/EGI

- Flagship European grid infrastructure projects
- Now evolved into the European Grid infrastructure

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 Grid middleware
- Now evolved into European Grid Infrastructure
- 40 NGIs, 2 EIROs (CERN, EBI), 8 partners in Asia

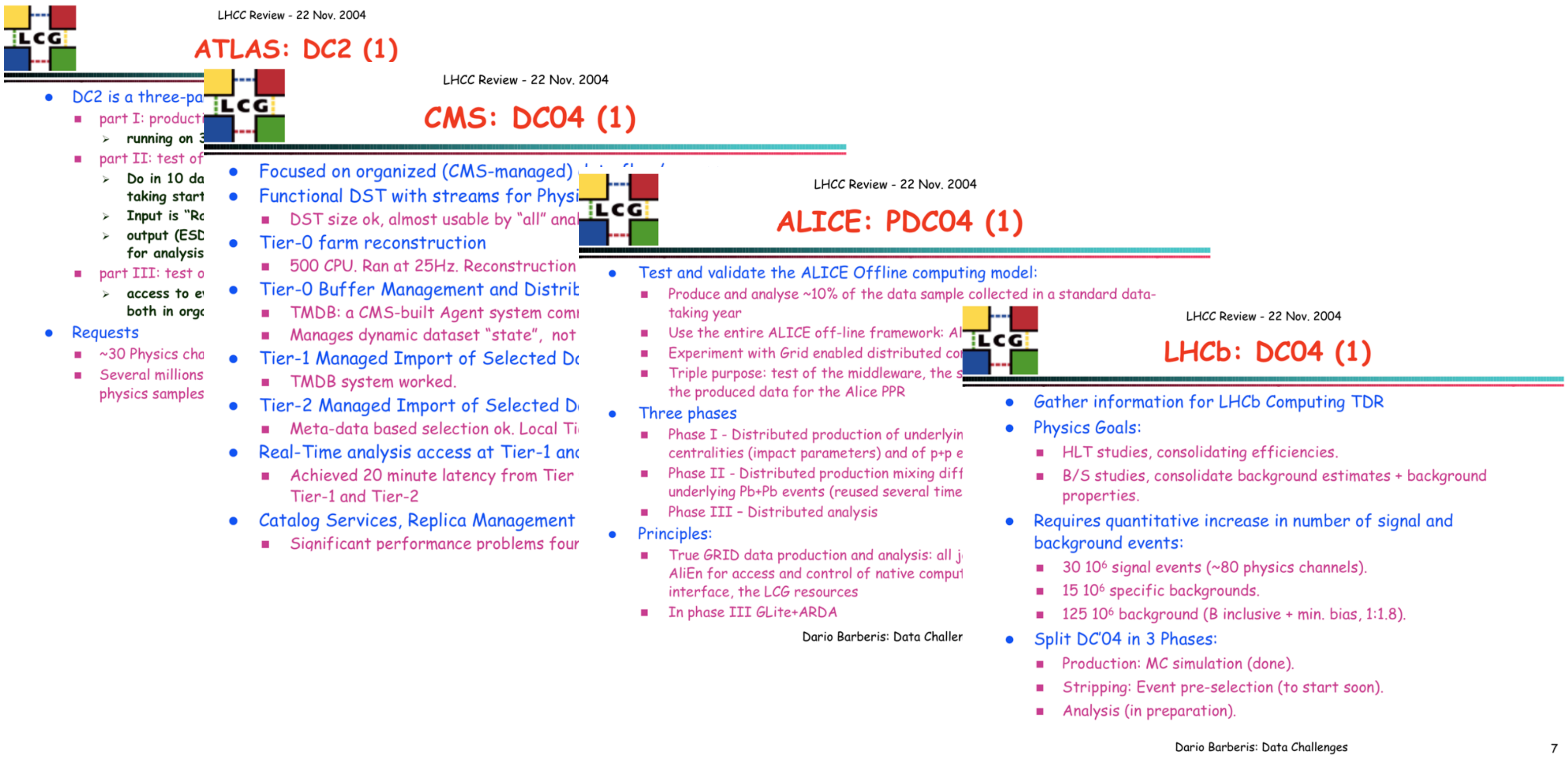


Worldwide LHC Computing Grid

- Union set of EGEE/EGI with US Open Science Grid, Nordic Grid Federation etc

2004 Data Challenges

- **To get a realistic design, the 4 experiments carried out major data challenges in 2004**



And a design that largely works emerged

LCG-TDR-001
CERN-LHCC-2005-024
20 June 2005

LHC Computing Grid

Technical Design Report

Version: 1.04
20 June 2005

The LCG TDR Editorial Board
Chair: J. Knobloch
Project Leader: L. Robertson

CONTENTS

Four years to make a production service

Independent Experiment Data Challenges

e.g. DC04 (ALICE, CMS, LHCb)/DC2 (ATLAS) in 2004 saw first full chain of computing models on grids

Service Challenges proposed in 2004

To demonstrate service aspects:

- -Data transfers for weeks on end
- -Data management
- -Scaling of job workloads
- -Security incidents ("fire drills")
- -Interoperability
- -Support processes

2004

SC1 Basic transfer rates

2005

SC2 Basic transfer rates

2006

SC3 Sustained rates, data management, service reliability

SC4 Nominal LHC rates, disk→tape tests, all Tier 1s, some Tier 2s

2007

2008

CCRC'08 Readiness challenge, all experiments, ~full computing models

2009

STEP'09 Scale challenge, all experiments, full computing models, tape recall + analysis

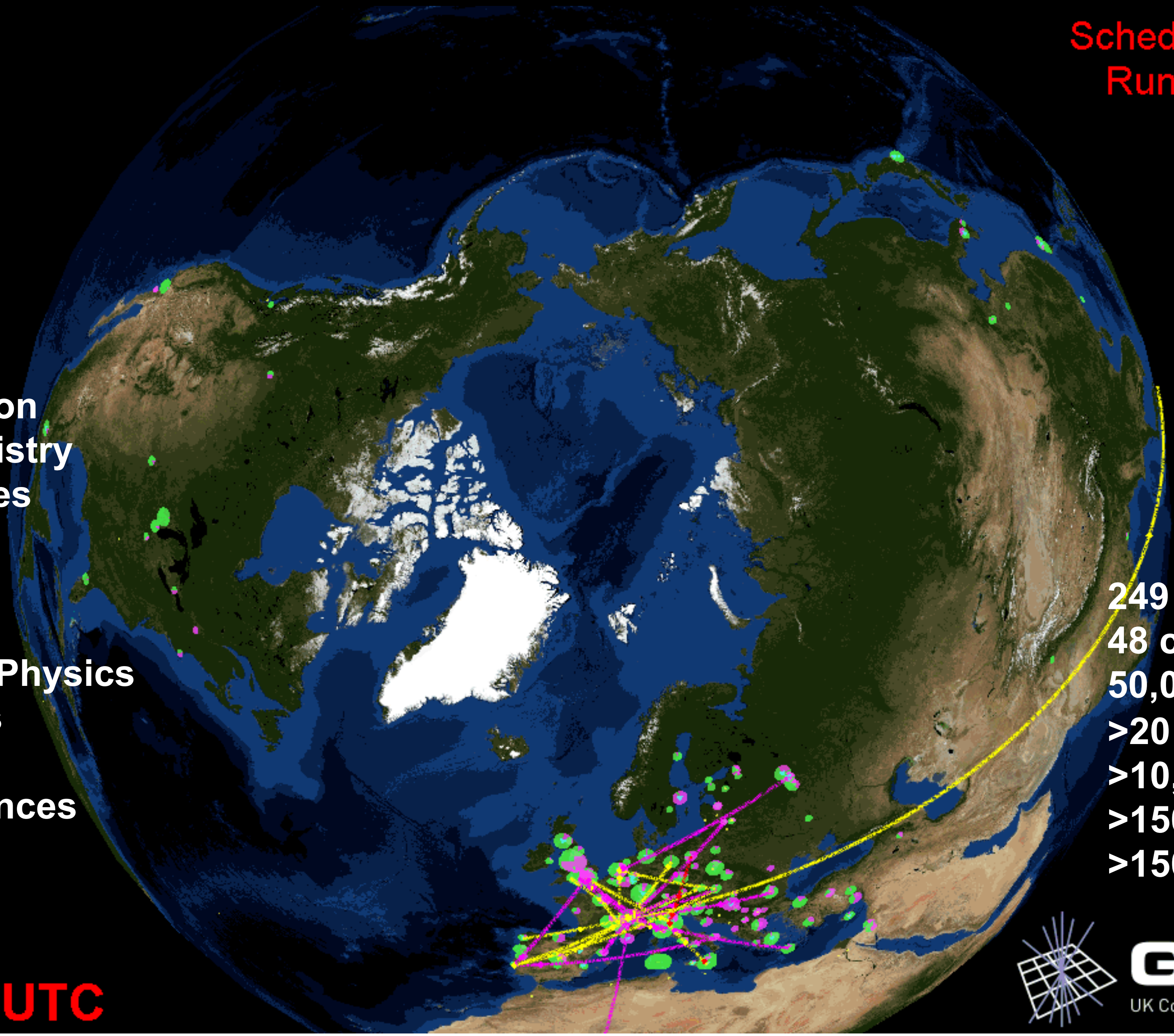
2010

- Focus on real and continuous production use of the service over several years (simulations since 2003, cosmic ray data, etc.)
- Data and Service challenges to exercise all aspects of the service – not just for data transfers, but workloads, support structures etc.

Ready by: **2008** Sadly, the LHC was not

Enabling Grids for E-sciencE

Scheduled = 21539
Running = 25374



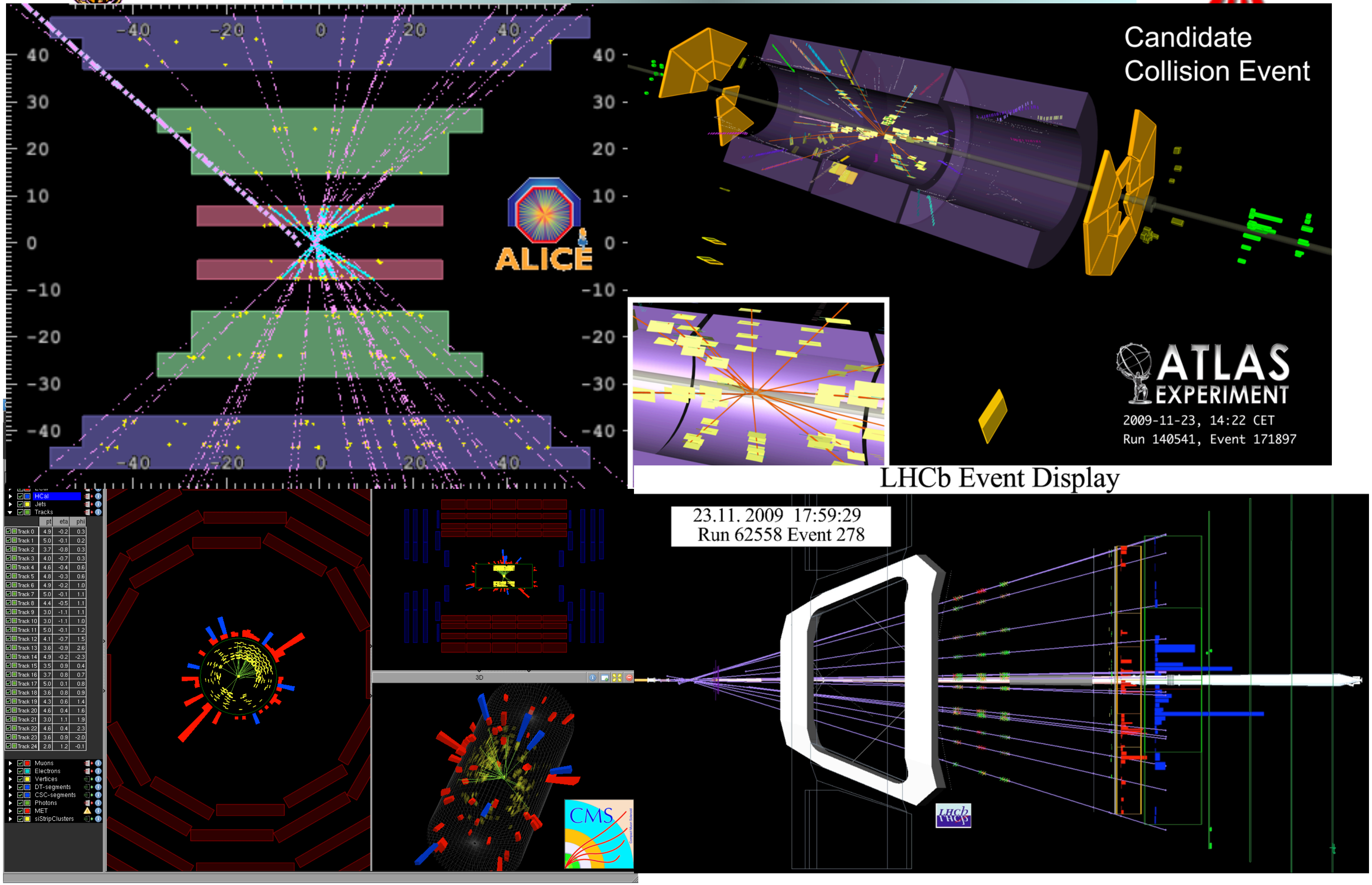
- Archeology
- Astronomy
- Astrophysics
- Civil Protection
- Comp. Chemistry
- Earth Sciences
- Finance
- Fusion
- Geophysics
- High Energy Physics
- Life Sciences
- Multimedia
- Material Sciences
- …

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

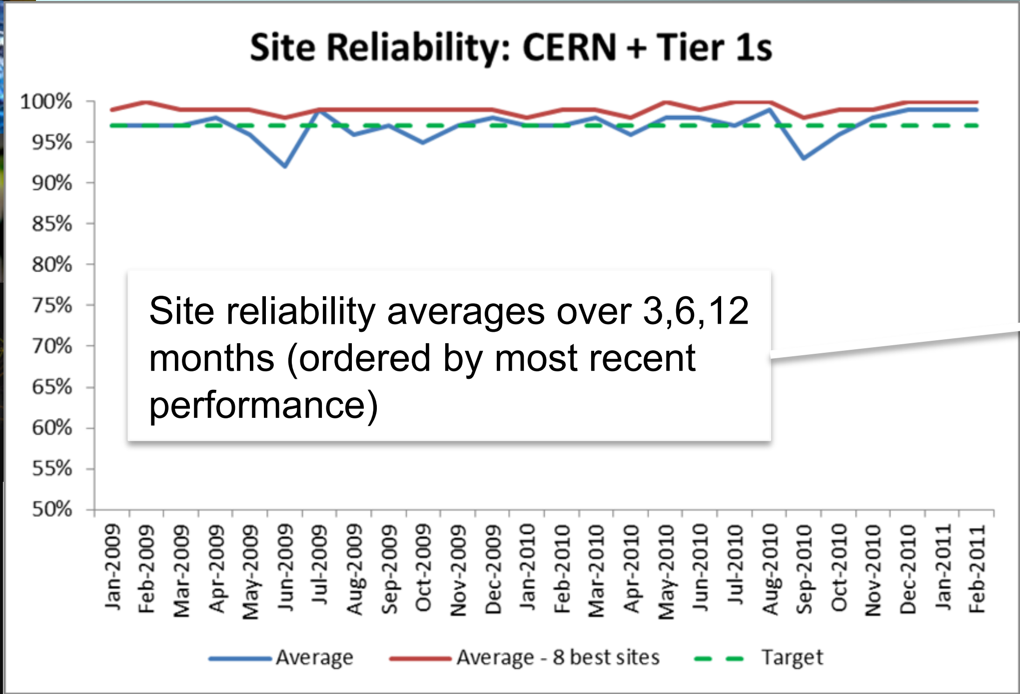
21:13:50 UTC

GridPP UK Computing for Particle Physics

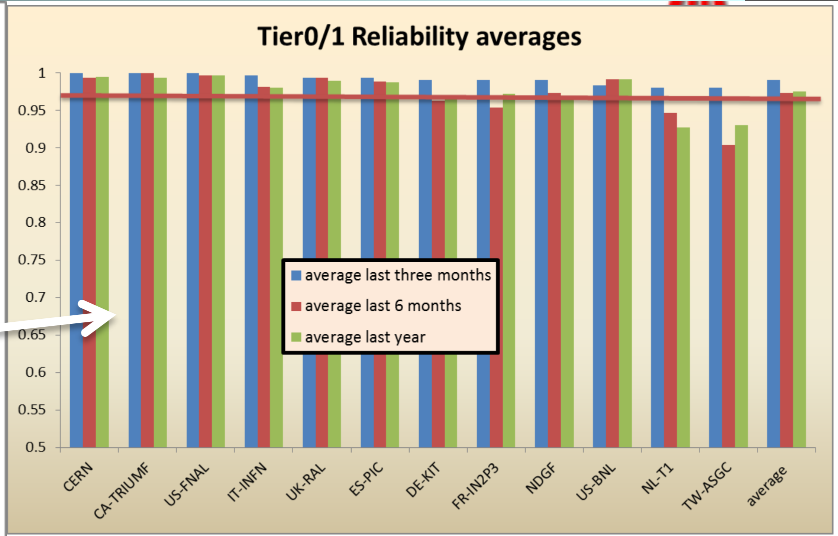
2009: First Physics Collisions

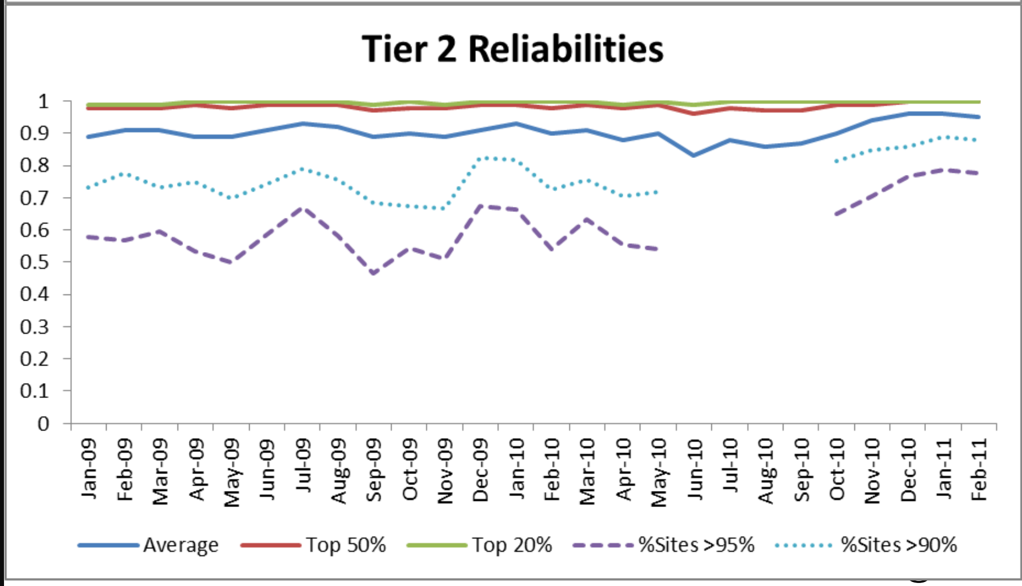


The system works reliably, despite user complaints!

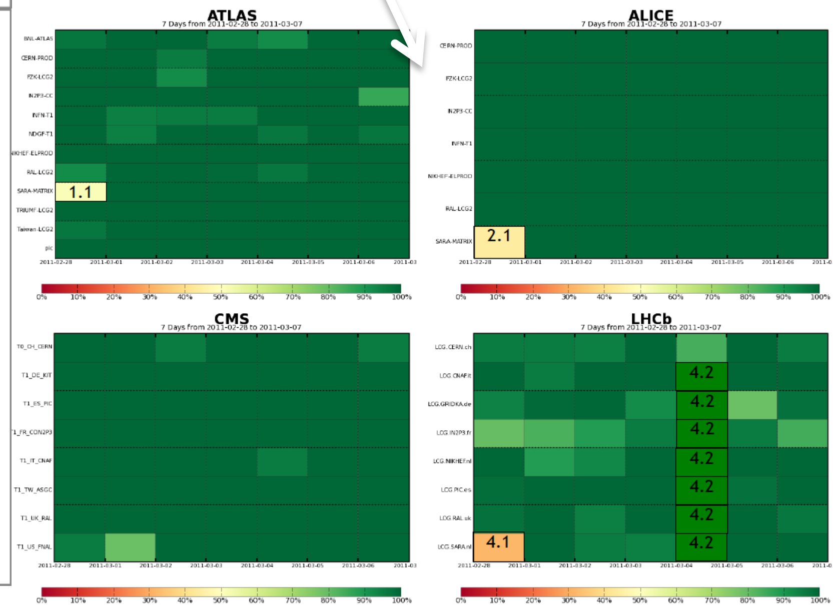


Site reliability averages over 3,6,12 months (ordered by most recent performance)





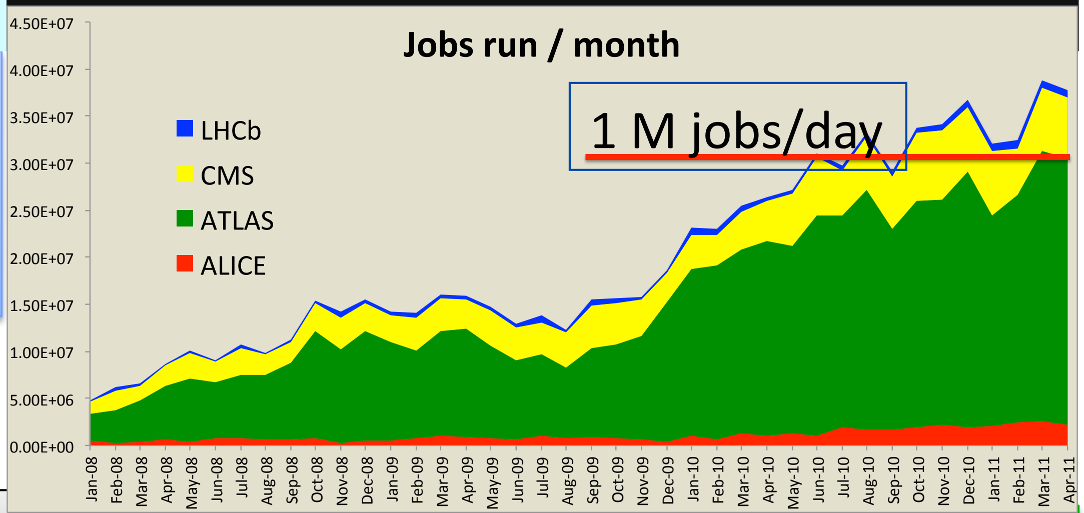
Typical site readiness now:



Grid Usage

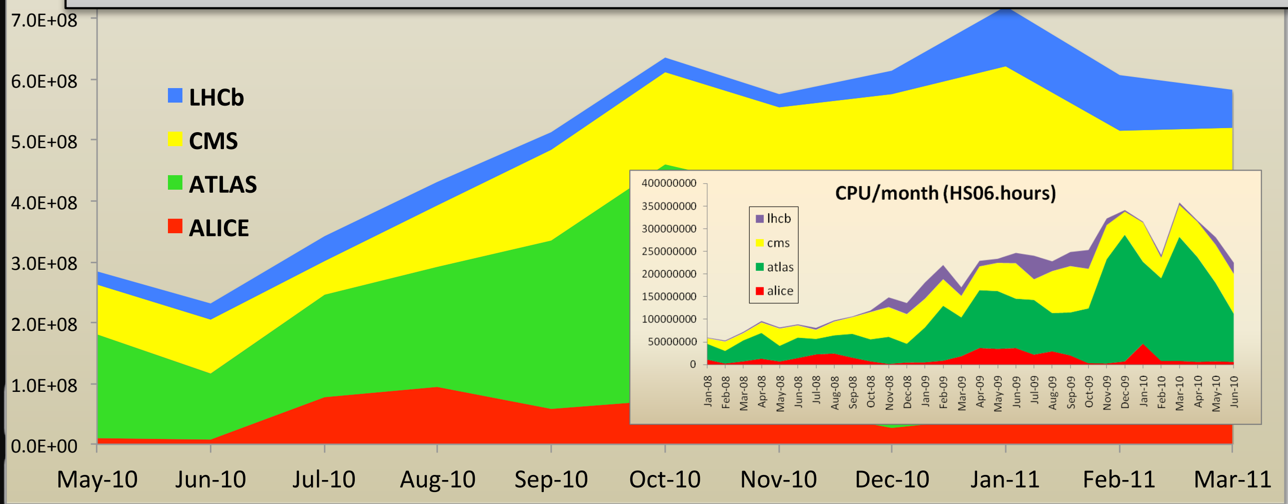
Use remains consistently high:

- >1 M jobs/day;
- ~150k CPU



At the end of 2010 we saw all Tier 1 and Tier 2 job slots being filled

CPU used at Tier 1s + Tier 2s (HS06.hrs/month) – over 12 months



CPU usage now >> double that of mid-2010 (inset shows build up over previous years)

As well as LHC data, large simulation productions always ongoing

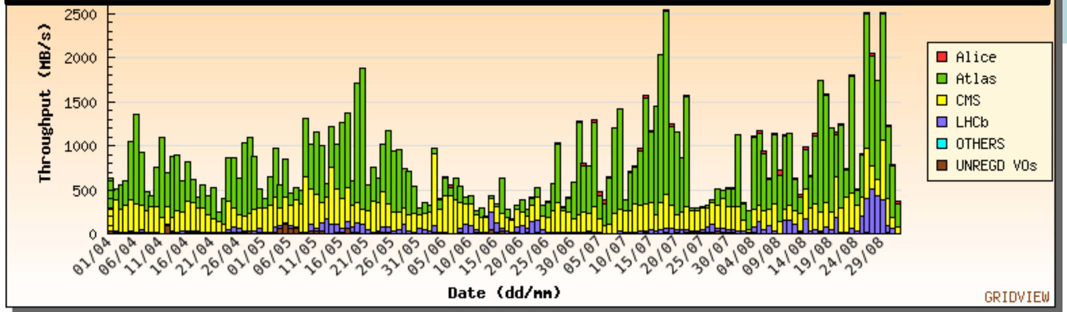
Large numbers of analysis users:

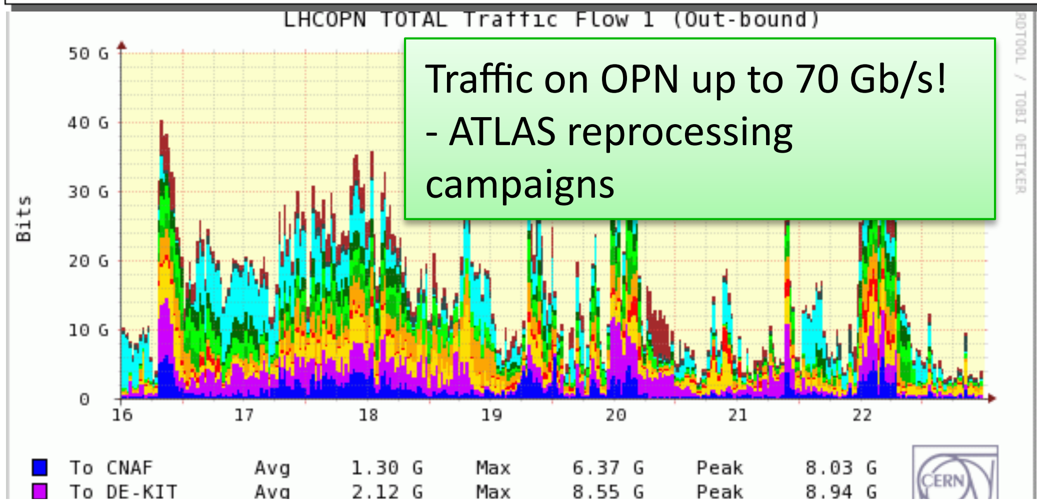
- ATLAS, CMS ~800
- LHCb,ALICE ~250

In 2010 WLCG delivered ~ 80-100 CPU-millennia!

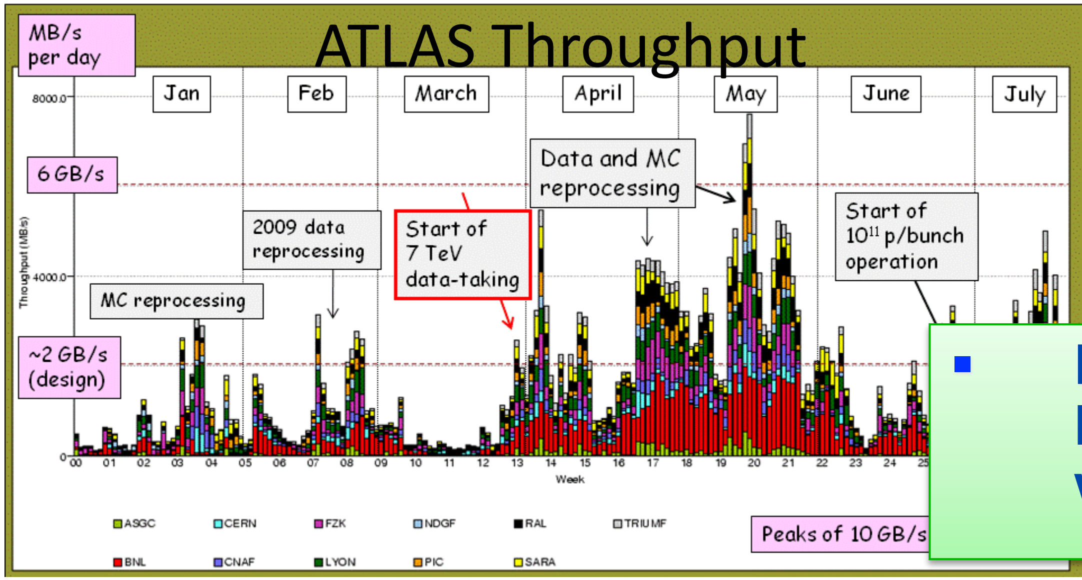
Data transfers

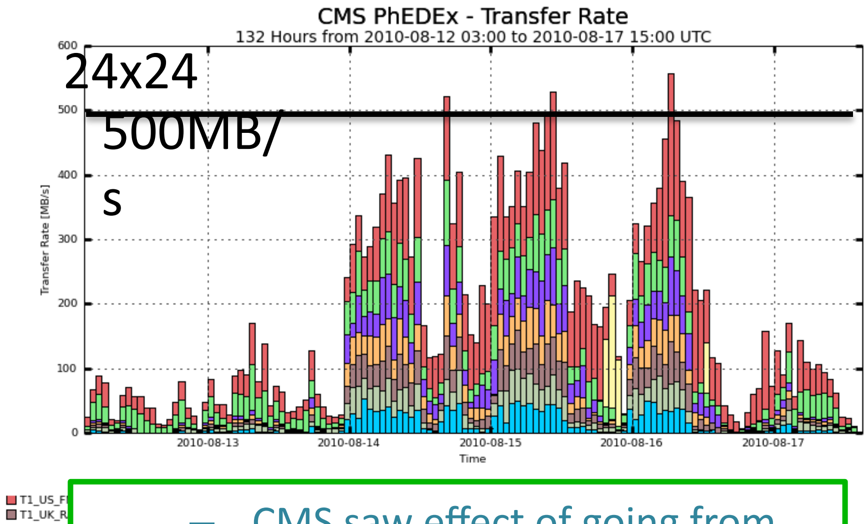
- Data transfers (mostly!) at rates anticipated



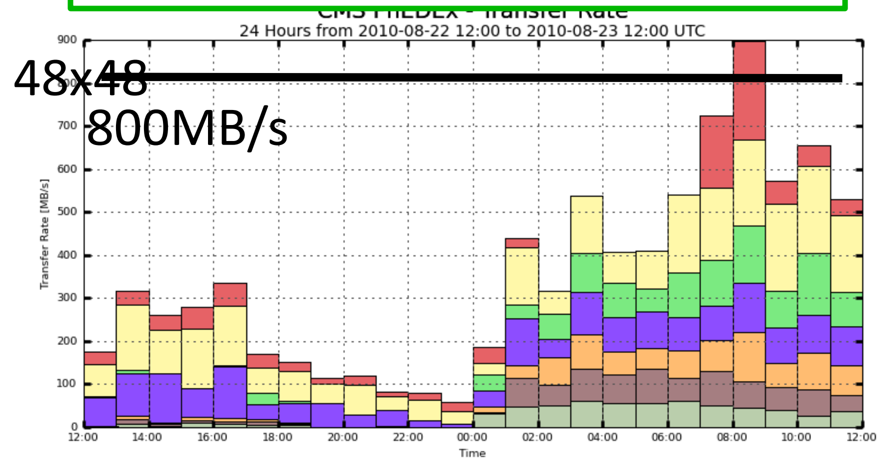


Traffic on OPN up to 70 Gb/s!
- ATLAS reprocessing campaigns





- CMS saw effect of going from 24x24 to 48x48 bunches



- **For all experiments: early data has been available for analysis within hours of data taking**

March 30, 2010

Rutherford Appleton Laboratory
UK Tier-1 official opening

The Tier 1 @ RAL in 2010

To/from CERN:
Up to 10Gb/s

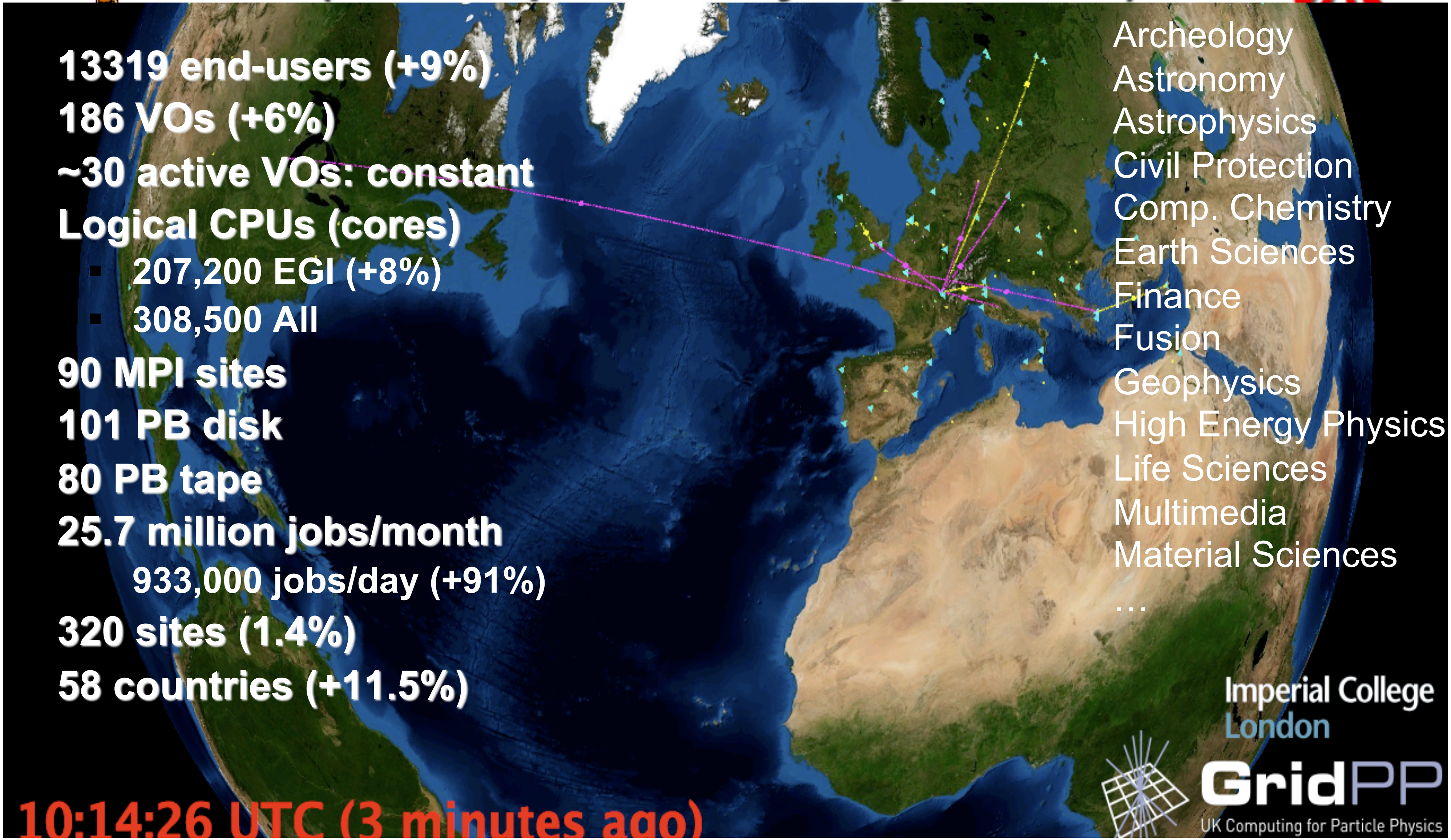
To/from SuperJanet5:
Up to 10Gb/s

- ~10000 cores
- ~8PB disk storage
- ~1.7PB tape storage
- ~450kW power

European Grid Infrastructure
(Status April 2011 – yearly increase)

- 13319 end-users (+9%)
- 186 VOs (+6%)
- ~30 active VOs: constant
- Logical CPUs (cores)
  - 207,200 EGI (+8%)
  - 308,500 All
- 90 MPI sites
- 101 PB disk
- 80 PB tape
- 25.7 million jobs/month
  933,000 jobs/day (+91%)
- 320 sites (1.4%)
- 58 countries (+11.5%)

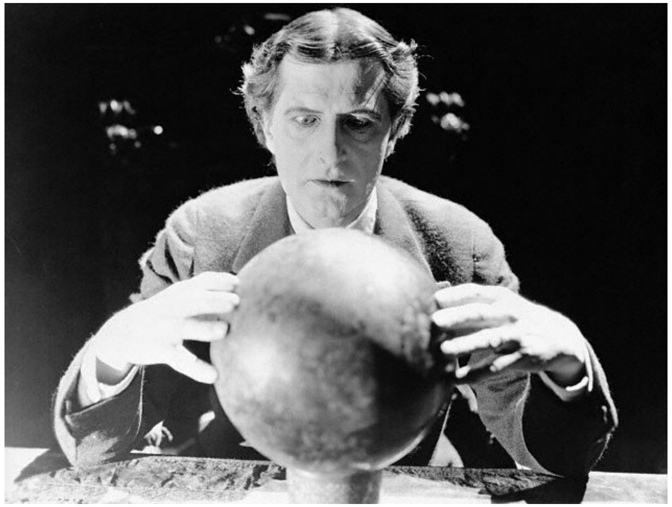
Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
…



Imperial College London
GridPP
UK Computing for Particle Physics

Projections - a Credo

- This requires a health warning, guided by the past
  - All projections get it wrong:
    - LEP resource projections were wrong by an order of magnitude!
    - And recall those early LHC planning papers
  - What can be done is driven by the technology, which is uncertain even 4 years out



The distributed future

- Grid➔Cloud
  - After moving away from big central facilities, the commercial clouds are trying to move things back
    - Their price models make no sense for our work
    - But the zeitgeist is with them
  - What we can do is use the technologies (virtualization) to realize the Grid objectives more fully
  - And create academic 'clouds' that have the right resonance at politician and University House level

The distributed future

- **Network has continued to improve in most territories**
  - **So we can place more data on demand and make better use of resources**
    - **But the number of user copies is driven by access requirements, not safe storage; so storage requirements may not drop**
    - **Can use this to move the burden of storage and make better use of CPU resources**

  - **The UK has particular issues – because we have many good, large facilities, and JANET works well, we do not fit well into the LHCONE view of the world.**
- **Architectures**
  - **Architectures are changing: co-processing, many-core (of which more in a bit)**
  - **The distributed systems need to evolve to handle this (e.g. whole-node scheduling rather than core-by-core)**

The distributed future (ii)

- **User demands**
  - **Triggers reject physics – increasing pressure to increase trigger rates, which increases the pressure on offline resources**
  - **And as datasets grow, user analysis rates look challenging**
  - **It would be a shame to force the users to throw away data if we can help it.**
- **Resources**
  - **We have a very large human operation to run the computing (which is still not acknowledged on a par with subdetectors). Is this sustainable?**
  - **No-one can have missed the financial and exchange rate issues – there are opportunities but big threats too**

But what about the software?

….Or what did you do in the Language Wars, daddy?

- Back in the 1980s, FORTRAN was the only realistic language for reconstruction, simulation and analysis
- In the 1990s, you could still make the joke…
- I don’t know what the language we will use in 20 years time will be, but it will be called FORTRAN
- But by the late 1990s, the joke sounded hollow
  - The cost of compilers
  - The training of students (the world did not want FORTRAN coders)
- And other arguments
  - Code reuse
  - Better code design
  - Better code readability

OO or nothing

- BaBar was probably the first to adopt
- Ah, but which OO language?
  - Eifel had a strong group in ATLAS.
    - (MOOSE group still dines out!)
  - C++ was the big commercial favorite
  - But the new semi-compiled, semi-interpreted JAVA was better for the purists
  - And then there were the FORTRAN95/2000 fans

- One of Norman's big headaches was to sort this out in ATLAS
  - And in true diplomatic style the answer was…
  - …C++ or Fortran or Java
  - (But C++ really)
  - ATLAS today: 4M lines C++
    - 1.4M lines Python
    - 0.1M F/F90, 0.1M Java

Concerns?

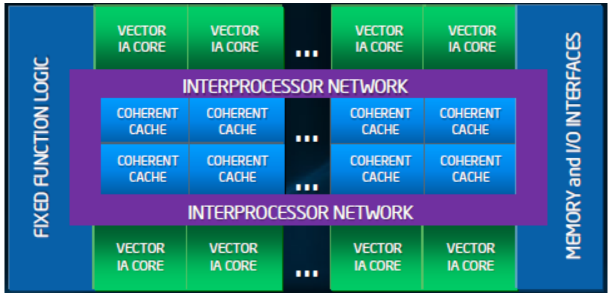
- Actually, we have put so much effort into the distributed computing, there is a worry the software is a poor relation

- ATLAS has
  - 2000 packages
  - 1000 developers over 3 years
  - But 305 developers in the last 6 months making 25 changes a day
  - High turn-over in developer community, and broad range of skill level
  - Rate of development does not decrease, many branches
  - Big challenges with integration – few integrators linking large developer and user communities

The Software Future

There are many challenges facing the LHC community

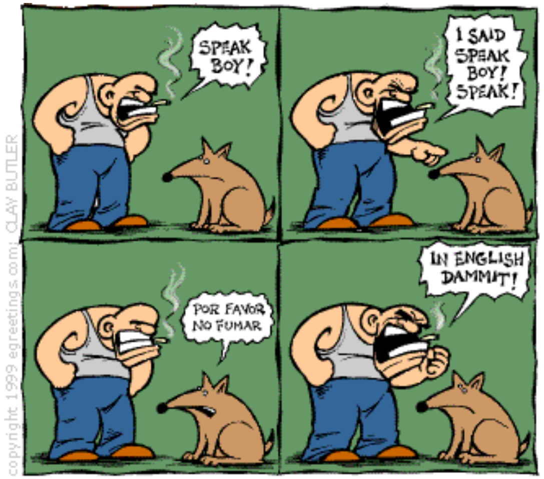
- **Increasing pile-up with increasing luminosity**
  - **Upgrades envisage ~200-500 interactions per read-out**
- **Architectural changes**
  - **From multi-core to many-core**
    - **Coarse-grained parallelism does not scale**
    - **Tricks to share memory between concurrent processes being used, but have limits at about 32 cores**
    - **Deeper parallelism required in our algorithms?**
  - **Co-processors (or more likely GPU onboard)**
    - **Efficient usage requires vector applications and local data to be identified and passed over**
  - **Compilers may give us a magic bullet, but work will be needed**



Intel MIC architecture:
Knights Ferry
Knights Corner

Back to language wars?

- CUDA developed for NViDia GPUs, but has general applicabaility
  - Developing rapidly
- OpenCL designed as an open standard, working for many-core and GPU
  - Slower development as a result?
- But we have a huge code-base and we cannot retrain a whole developer community
- Can we just hope for the compiler or libraries to abstract the accelerator?



Back to the Future?

An interesting observation from Ian Hart, Cray

- OpenMP is already established and allows relatively inexperienced users to get almost all the benefit you would get from an OpenCL/CUDA rewrite with the use of simple pragmas etc

- Supported as native in MIC architectures
- The catch:
  - You really want a nice, sequential language to do this…
  - Like Fortran!
- So, to be wicked, maybe the joke should have been:
- *“I don’t know what we call the language we use in 20 years time, but it will probably be a lot like Fortran”*

In Conclusion

At the end of 2010 LHCC made the following statement:

"The committee is impressed and pleased to see that the WLCG and the experiments' computing models are successfully processing and analysing the first LHC physics data without any major problems. This is the result of many years of careful planning and investment, and the hard work of a great many people. We send our congratulations to all the staff involved, at CERN and the many computer centres around the world."

And Norman was a big part of that success for ATLAS!