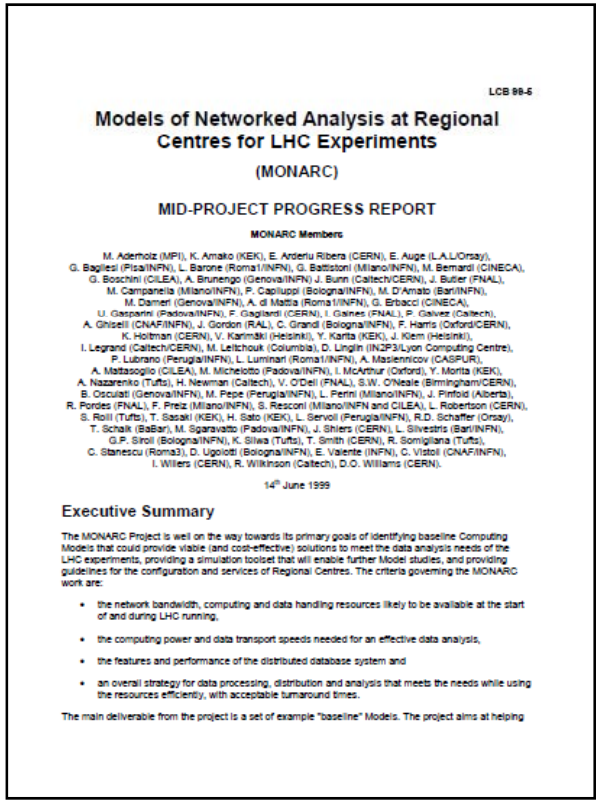
LHC Data Analysis will start on the Grid

What's Next?

Les Robertson - CHEP 2009, Prague - 23 March 2009

How did we get here?

- 1996 – funding hole
  - computing for physics was left out of the LHC budget approval process
- Clear that new funding would be needed and that most of it would have to come from outside the CERN budget
- 1998 – MONARC project
  - a distributed model
    - Integrate existing centres, department clusters, recognising that funding is easier if the equipment is installed at home
    - Devolution of control– local physics groups have more influence over how local resources are used, how the service evolves
  - a multi-Tier model
    - Enormous data volumes → looked after by a few (expensive) computing centres
    - Network costs favour regional data access
    - Simple model that HEP can develop and get into production ready for data in **2005**



LCB 98-6

Models of Networked Analysis at Regional Centres for LHC Experiments

(MONARC)

MID-PROJECT PROGRESS REPORT

MONARC Members

M. Aderholz (MPI), K. Amako (KEK), E. Arderiu Ribera (CERN), E. Auge (L.A.L/Orsay), G. Bagliesi (Pisa/INFN), L. Barone (Roma1/INFN), G. Battistoni (Milano/INFN), M. Bernardi (CINECA), G. Boschini (CILEA), A. Brunengo (Genova/INFN) J. Bunn (Caltech/CERN), J. Butler (FNAL), M. Campanella (Milano/INFN), P. Capiluppi (Bologna/INFN), M. D'Amato (Bari/INFN), M. Dameri (Genova/INFN), A. di Mattia (Roma1/INFN), G. Erbacci (CINECA), U. Gasparini (Padova/INFN), F. Gagliardi (CERN), I. Gaines (FNAL), P. Galvez (Caltech), A. Ghiselli (CNAF/INFN), J. Gordon (RAL), C. Grandi (Bologna/INFN), F. Harris (Oxford/CERN), K. Holtman (CERN), V. Karimäki (Helsinki), Y. Kamta (KEK), J. Klem (Helsinki), I. Legrand (Caltech/CERN), M. Leltchouk (Columbia), D. Linglin (IN2P3/Lyon Computing Centre), P. Lubrano (Perugia/INFN), L. Luminari (Roma1/INFN), A. Maslennicov (CASPUR), A. Maltasoglio (CILEA), M. Michelotto (Padova/INFN), I. McArthur (Oxford), Y. Morita (KEK), A. Nazarenko (Tufts), H. Newman (Caltech), V. O'Dell (FNAL), S.W. O'Neale (Birmingham/CERN), B. Osculati (Genova/INFN), M. Pepe (Perugia/INFN), L. Perini (Milano/INFN), J. Pinfold (Alberta), R. Pordes (FNAL), F. Prelz (Milano/INFN), S. Resconi (Milano/INFN and CILEA), L. Robertson (CERN), S. Rolli (Tufts), T. Sasaki (KEK), H. Sato (KEK), L. Servoli (Perugia/INFN), R.D. Schaffer (Orsay), T. Schalk (BaBar), M. Sgaravatto (Padova/INFN), J. Shiers (CERN), L. Silvestris (Bari/INFN), G.P. Siroli (Bologna/INFN), K. Sliwa (Tufts), T. Smith (CERN), R. Somigliana (Tufts), C. Stanescu (Roma3), D. Ugolotti (Bologna/INFN), E. Valente (INFN), C. Vistoli (CNAF/INFN), I. Willers (CERN), R. Wilkinson (Caltech), D.O. Williams (CERN).

14th June 1999

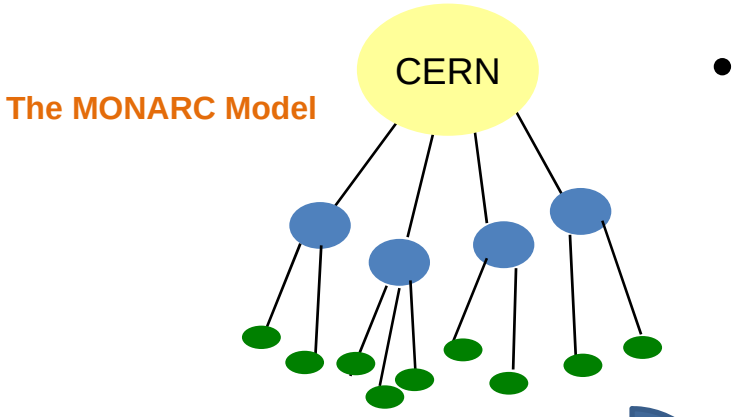
Executive Summary

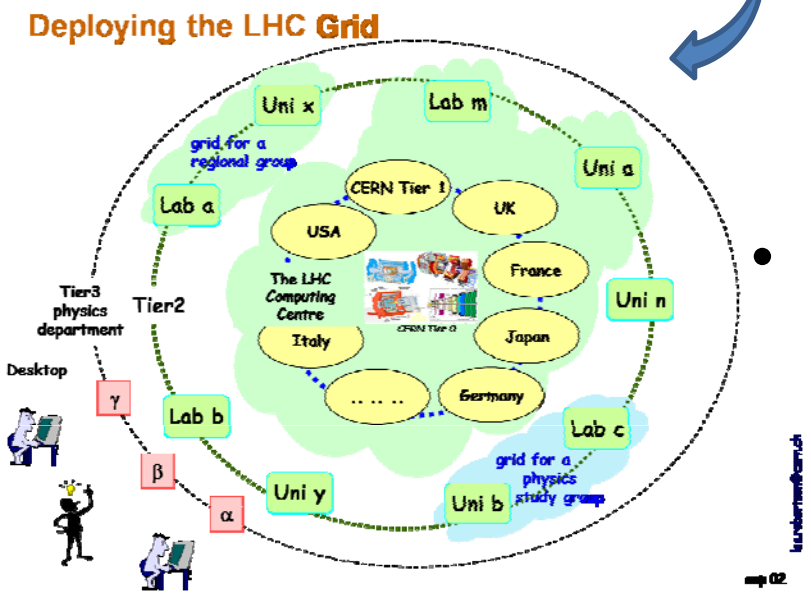
The MONARC Project is well on the way towards its primary goals of identifying baseline Computing Models that could provide viable (and cost-effective) solutions to meet the data analysis needs of the LHC experiments, providing a simulation toolset that will enable further Model studies, and providing guidelines for the configuration and services of Regional Centres. The criteria governing the MONARC work are:

- the network bandwidth, computing and data handling resources likely to be available at the start of and during LHC running,
- the computing power and data transport speeds needed for an effective data analysis,
- the features and performance of the distributed database system and
- an overall strategy for data processing, distribution and analysis that meets the needs while using the resources efficiently, with acceptable turnaround times.

The main deliverable from the project is a set of example "baseline" Models. The project aims at helping

Then came the GRID





- 1999 – Grid
  - More flexible → easier to use, adapt to the reality of data analysis
  - But – more complex to build and to manage
  - However – the basics were already there!
  - Prospect of a general science grid (memories of research networks)
  - With expectation of non-HEP funding
    - during development
    - and for long term operation
- Consensus on this approach emerged during CHEP 2000 in Padova

2000s - the Decade of the GRID

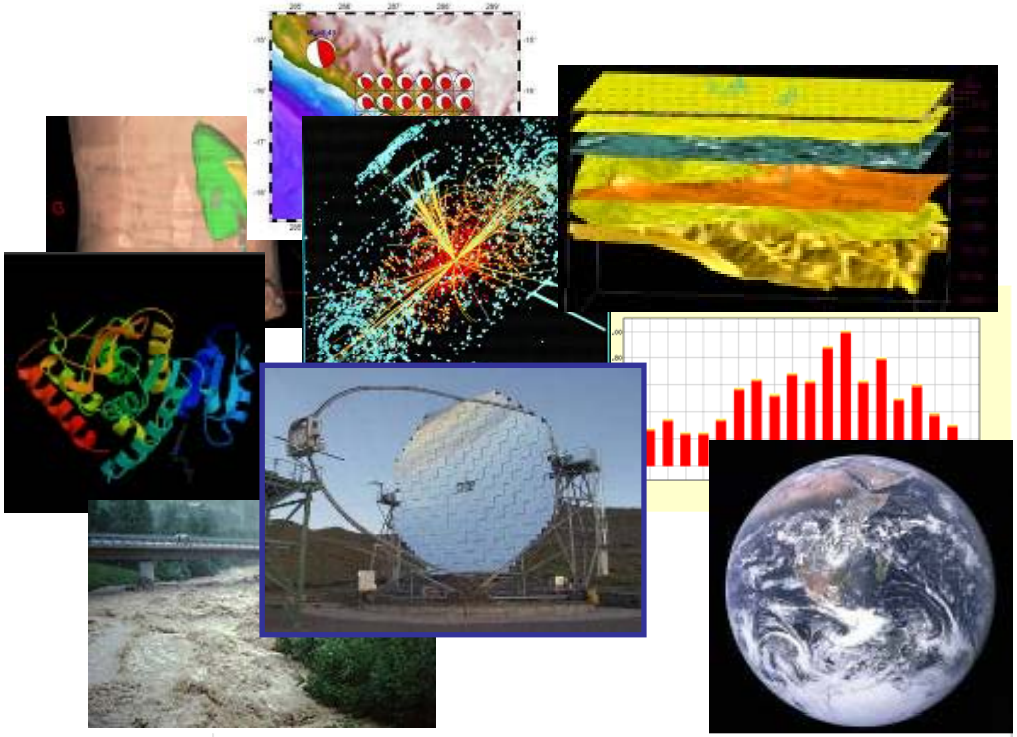
- E-Science in fashion
- Many grid/science/physics projects funded in Europe and US
  - Stimulated international collaboration – open to all LHC sites
  - WLCG could operate on top of these multi-science infrastructure grids – EGEE, OSG, ..

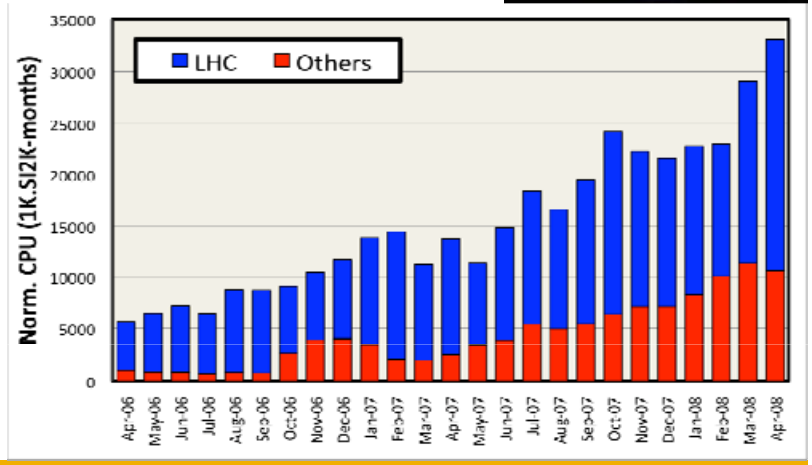
  → de-facto standards
  - Significant non-HEP funding was made available to LHC groups and centres – supporting operation, tools and middleware, application development and adaptation to grids
- Significant industrial interest came ….

  …. added confusion .... and went
- But many other sciences and also some industries have ported applications to the HEP style of Grid

EGEE Achievements - Applications

- **>260 VOs from several scientific domains**
  - Astronomy & Astrophysics
  - Civil Protection
  - Computational Chemistry
  - Comp. Fluid Dynamics
  - Computer Science/Tools
  - Condensed Matter Physics
  - Earth Sciences
  - Fusion
  - High Energy Physics
  - Life Sciences
- **Further applications under evaluation**

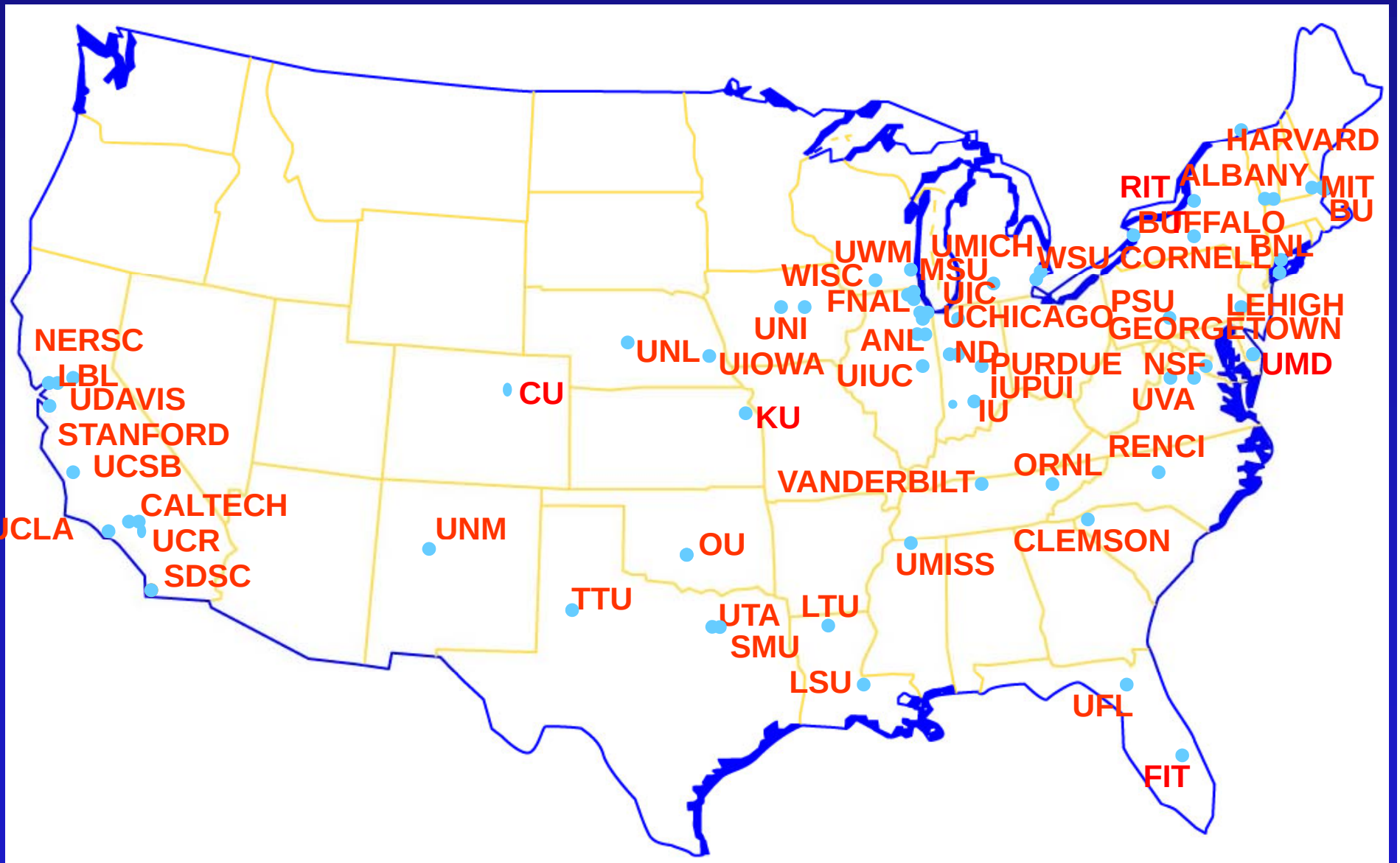


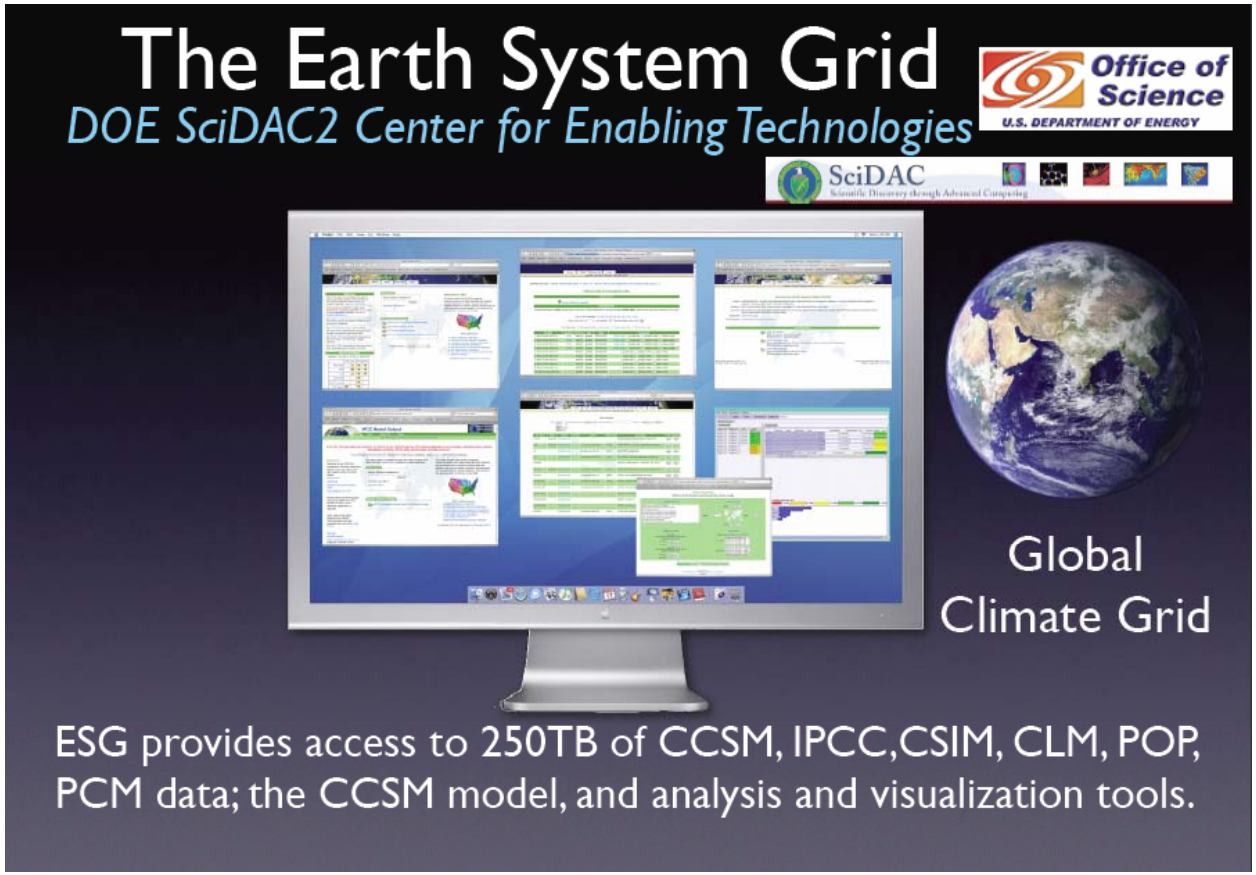


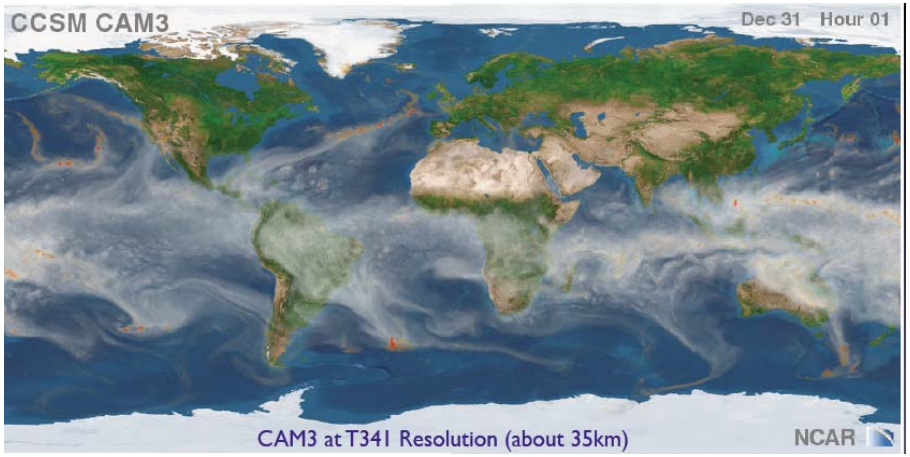
Applications have moved from testing to routine and daily usage

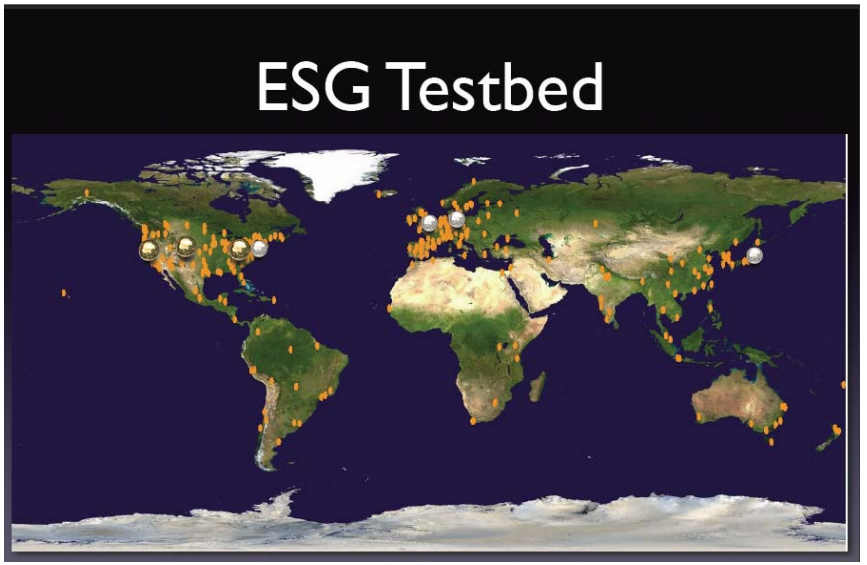
~80-95% efficiency

The Map of OSG Sites (in the US)









Collage of slides from presentation by Don Middleton, *National Center for Atmospheric Research, at the* OSG All Hands Meeting, Mar 09

Internal Grid experiences in Montepaschi Group

Server farm

- About 600 non-dedicated Windows and Linux servers
- We expect to have 1.000 central servers with Windows 2003 Server
- About 27.000 jobs every day (41.000 peak)

Accomplishments

- Current Accomplishments
  - Clickstream analysis (January 2005)
  - IAS 32-39 (August 2005)
  - Top Management Monitoring (January 2006)
  - Basel II (Algoritmics) (October 2005)
  - DOC1 (December 2006)
  - many other apps …. until today
- Next Steps
  - Host Batch downsizing
  - …….

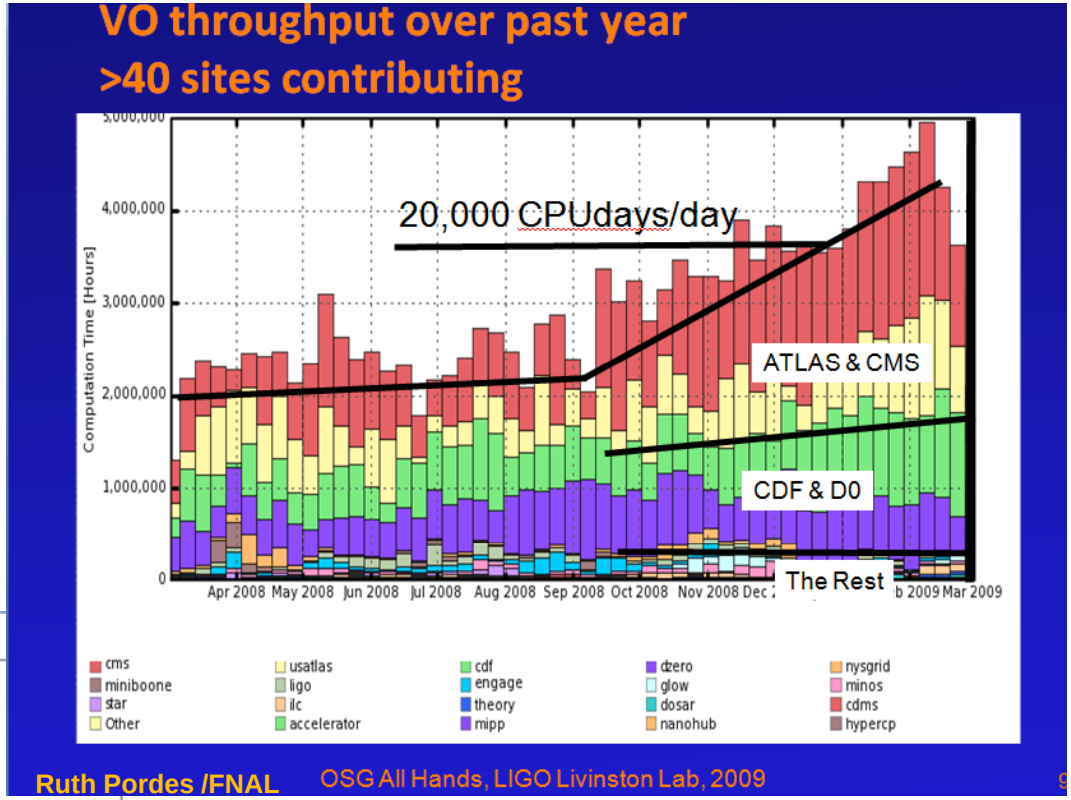
MONTE DEI PASCHI DI SIENA
BANCA DAL 1472

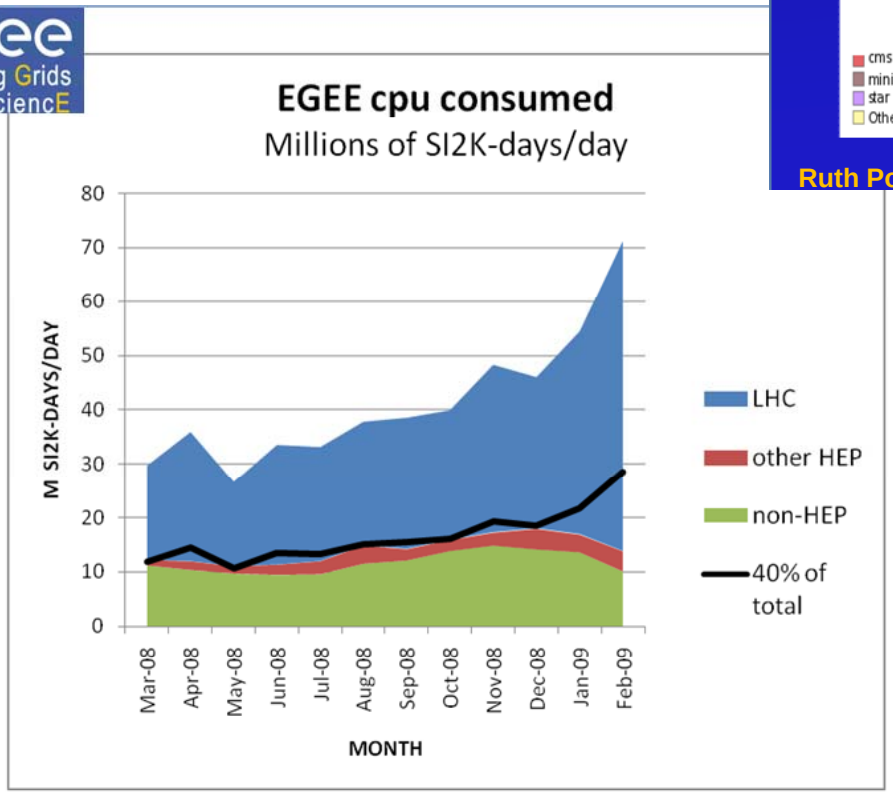
Collage of slides from a presentation by Piero Poccianti, Gruppo Motepaschi at the EGEE User Forum, Catania, March 2009

GRUPPOMONTEPASCHI

Lots of applications in production

Some potentially comparable to LHC in computation, storage, global reach





Source: EGEE accounting portal

But how many of them, other than HEP, really depend on a general science Grid?

As LHC starts, data handling depends on a grid
But is the model of a general science grid sustainable?

- WLCG operates on top of multi-science grids
  - With short-term funding cycles <> incompatible with long-term services
  - Hard to find other sciences outside physics that ***depend*** on these grids
  - Proposal in Europe for a long-term infrastructure (EGI), but still some way from agreement and approval, and EGEE ends next April

    right in the middle of the first LHC run
  - Open Science Grid with a similar role to EGEE in US has 5-year funding from NSF and DoE through 2010
- Could be problematic – but ..
  - Tier-1 sites are still at the heart of EGEE and OSG operations
  - HEP institutes, collaborators and experiments are to a large extent responsible for the middleware
- So WLCG and LHC funding agencies can and **surely will** take on the necessary operational responsibility if EGEE and/or OSG close down

But MONARC was ten years ago!

While we waited for LHC to start in
~~2005~~ ~~2006~~ ~~2007~~ ~~2008~~ 2009

the world has surely changed!

1998 – Can you remember that far back?

- Clinton was President and the Lewinsky scandal broke in January
- DEC was bought by COMPAQ, which has since been absorbed by HP
- Amazon had been selling books online for only three years
- The Google name was not registered until September
- Intel introduced the Pentium II on 250 nm technology, with 7.5M transistors, and up to 450 MHz clock
  (2008 – Core 2 – 45 nm, 410M transistors, 3.2GHz)
- Top of the TOP 500 list was the ASCI Red at Sandia with 9,152 processors delivering 1.3TFLOPS
  (2008 – a 129,600 core system at Los Alamos delivering 1,105 TFLOPS)
- The European Research Network backbone was upgraded from 34 to 155 Mbps in December
- 64 kbps was an *excellent* home network connection
- WiFi prototypes were just appearing (IEEE 802.11-1997)
- and GPRS data services on GSM phones were yet to be launched

- **At the end of the 90s HEP was a major computer user**
  - Having broken out of the cost/performance constraints of mainframes and minis at the beginning of the nineties
  - Having led the way in international high speed networking
  - Having exploited the power of the Web it had conceived

  surpassed only by a few other large sciences, the military and the spooks

  …. and the 2005 needs of LHC looked gigantic
- But the past decade has seen business and industry also exploit PC clusters, mass market disks, the Internet, the Web,..

→ **the demand for computing power and storage has shot up** driven by **new applications**: search engines, web advertising, online commerce, digital libraries, interactive games, photo archives, social networking, ...

→ **HEP is now a *relatively* small player**

- Google had 36 data centers in 2008**
- WLCG had 134

***source: royal.pingdom.com*

BUT

- One Google Data Center is estimated to cost ~$600M
- An order of magnitude more than the new centre being planned at CERN

Google's data center at the Dalles on the Columbia river

- Microsoft's data center in Quincy, WA
  - 44K m^2 - 10 rooms each with up to 30K servers and storage for 6 trillion photos
- Yahoo, Amazon, IBM -- also building giant data centers

A look at the server area inside Microsoft's Quincy data center, courtesy of the BBC

- These major companies are expecting to build new markets for utility computing (clouds), software as a service (Google Apps), as well as absorbing the expansion of traditional computing services...

  ... by offering **very cost-effective computing** –
  - Economies of scale (hardware, management, operation)
  - Efficient resource scheduling for high utilisation
  - Tax-efficient locations
  - Cheap and green energy

The landscape has changed

Have we updated the maps?

Three issues:

- **Energy**
- **Virtualisation**
- **Clouds**
- **Networks and mobility**

Energy

- “Data centres consumed 1 per cent of the world’s electricity in 2005. By 2020 the carbon footprint of the computers that run the internet will be larger than that of air travel, a recent study by McKinsey and the Uptime Institute predicted.”

  Times Online – September 2008

- Even if the cost of oil is down at ~$50, if this growth rate really continues -
  power-efficient data centres and cheap renewable energy must be essential components of any infrastructure that is being planned today

☺ **The distributed (grid) model enables us to incorporate data centres wherever they may be located, and whoever is running them**

Virtualisation re-born

- A *virtual machine* implements a set of low-level functions
  - That can support a full-blown operating system such as Linux
  - And can be mapped on to a wide range of hardware
- The real operating system is reduced to being a virtual machine manager
- While the application is packaged together with the full function operating system and everything else it needs

→ The application decides which software configurations to support, when to upgrade

→ The computer centre is free to share processors between different applications,
and to install new hardware on its own timescale

→ No longer have to take global decision on versions of OS + middleware + ..+.. +.. freedom from the **dependency gridlock**

Hardware factors also favour virtualisation

- While Moore's law continues in its original form –
  the number of components on a chip grows exponentially
  - *chip makers are r*unning out of architectural ideas to use these components for faster execution
  - and trading off energy savings against performance

  → many more cores per processor

  → need to share processors between applications

- Intel and AMD both provide hardware support for efficient virtual machine execution
- And Linux is adapting to run efficiently within a virtual machine

No free lunches

- While the applications are now independent of each other they still need the functionality of Linux, software, middleware, and all the other external packages
  → integration, certification, test beds, ….
- Reduced scope for hackers? No more critical security upgrades?? Security no longer the responsibility of the computer centre???
- Can you really remove the *gridlock* and retain inter-operability?

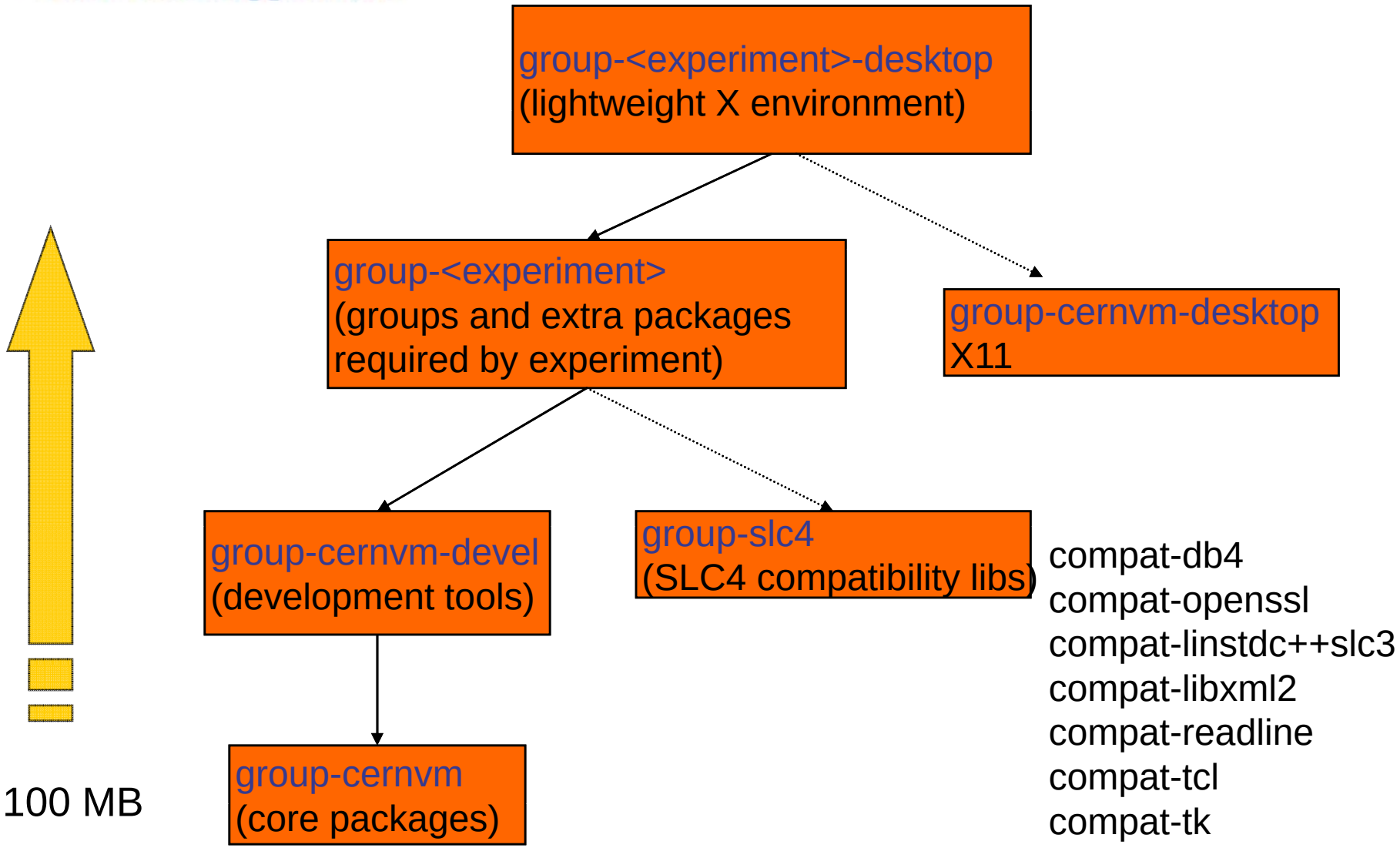
  … and maintainability?
  The reason CERN standardised on Linux+Intel (exactly 10 years ago) was to avoid the cost of supporting 7 or 8 flavours of Unix

But virtualisation also offers portability

CernVM Project

- Portable Analysis Environment using Virtualization Technology (WP9)
- Project goal:
  - Provide a complete, portable and easy to configure user environment for developing and running LHC data analysis locally and on the Grid independent of physical software and hardware platform (Linux, Windows, MacOS)
    - Decouple application lifecycle from evolution of system infrastructure
    - Reduce effort to install, maintain and keep up to date the experiment software
    - Lower the cost of software development by reducing the number of compiler-platform combinations
- Approved in 2007 (2+2 years) as R&D activity, started January 2008

CernVM Components

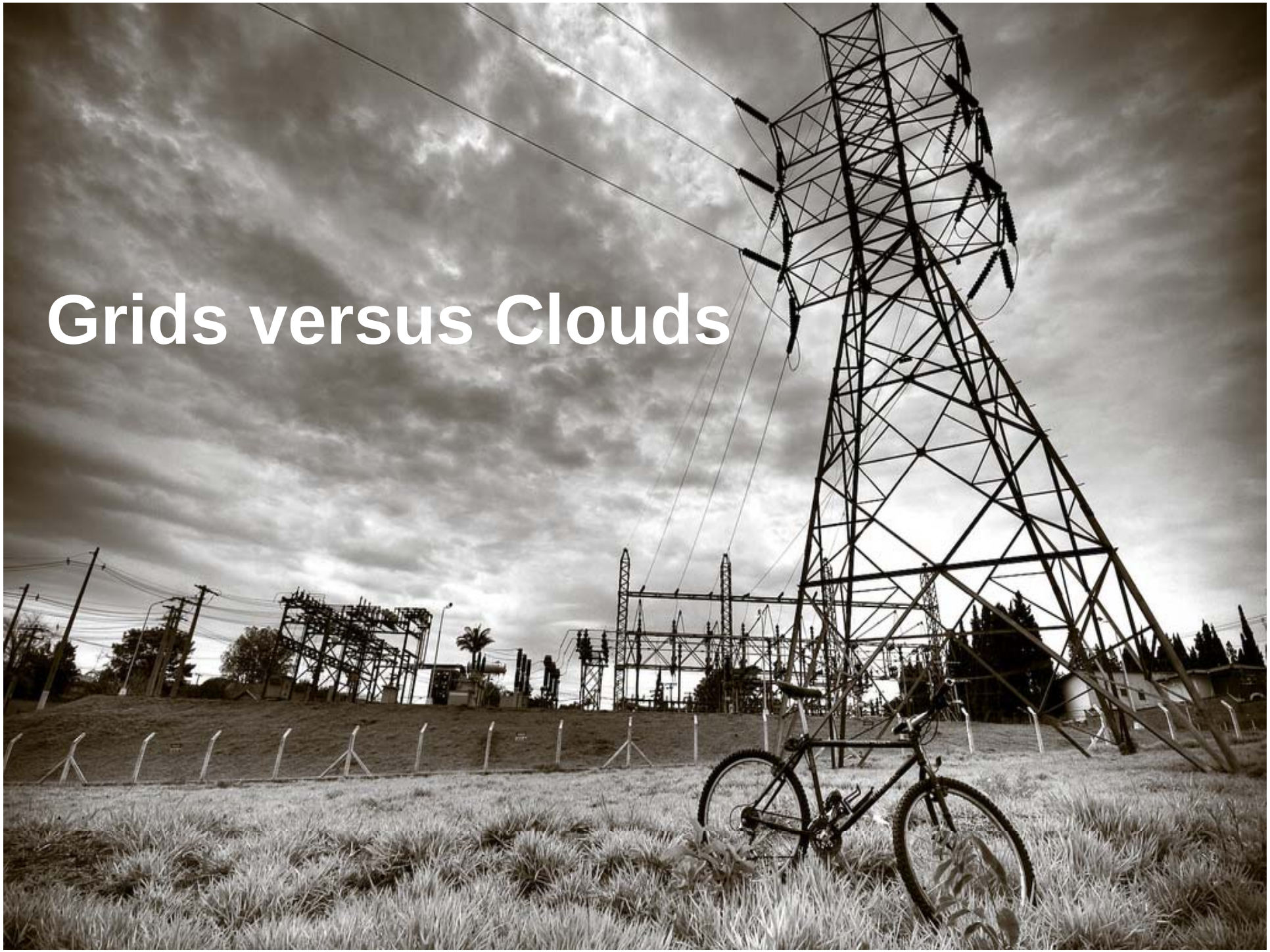


Virtualisation

- Maintenance of the application environment
- Efficient utilisation of farms of multi-core processors
- Portability of applications between operational environments

☺ If this technology had been practical ten years ago we would surely have built the grid on it

☺ The CernVM approach looks like a very good direction – start with end-user analysis, lightweight, platform-independent, maintainable model


Grids versus Clouds

Cloud v. Grid

- Clouds aim at efficient sharing of the hardware
  - low-level execution environment, Isolation between users
  - Operated as a homogeneous, single-management domain
  - Straight-forward i/o and storage
  - Expose only a high-level view of the environment - scheduling, data placement, performance issues are hidden from the application and the user
- Grids aim at collaboration
  - Add your resources to the community, but retain management control
  - Expose topology – location of storage, availability of resources
  - Choice of tools to hide the complexity from the user, and the application can write its own tools
- **Both need complex middleware to function**
  - Grids had a problem in trying to provide a universal high-functionality environment (OS, data management, ...), with intersecting collaborations and a naturally competitive environment
  - Clouds have an advantage in offering a simpler base environment, leaving much of the functionality to the application - where universal solutions are not necessary – and what they do have to provide can be decided within a single management hierarchy
- As the names suggest –

  the grids are transparent and the clouds are opaque

- These are early days – the market is not mature, the competition is not clear, the standards have not evolved, and we are in a world financial crisis that may blow the clouds away
- But there are powerful forces behind them
  – HEP should be adapting its model to be ready to integrate such offers if they become competitive
- And also ***learning from the innovative technology*** that is appearing stimulated by the cloud market
  - to manage, operate and supply these giant facilities with data
  - to package applications to use (heterogeneous) clouds

☺ **The evolution to virtualisation will ease the integration of clouds, grids and notebooks**

Grids and
Clouds

Networks ..

- The first MONARC report talked (1999) of –

  ".. the limited wide area network bandwidths that are likely to be available by the start of LHC data taking .."

  It was talking about 2005, but nevertheless high performance international networking has been one of the very good surprises
- HEP deserves a lot of credit for providing challenging requirements from several generations of global experiments ..

  .. and for taking the lead in operating its own networks ..

  .. and then integrating them with the research network infrastructure as that developed

☺ **We profit enormously from this infrastructure provided as a general science service, but based today on excellent commercial offerings in a very competitive market**

😐 **Although HEP still has work to do in some parts of the LHC world**

.. and Mobility

- ADSL at 20 Mbps, WiFi/WiMAX/3G
  - We are close to having good bandwidth data connections almost everywhere we go
  - And we already have a powerful high capacity computer in the backpack
- **This is where end-user analysis is going to be done**
- The physicist's notebook must be integrated with the experiment environment, the physics data, and the grid resources
- Without burdening the notebook or its user
- The grid environment is too complex to be extended to the notebook
- Ganga does a good job of bridging these environments

☺ **The approach of cernVM looks like the right direction for analysis, enabling the end-user to cache the data she needs and extend her environment on to the grid, or cloud, or ...**

Summary

- Grids are all about sharing.
  - groups distributed around the world can pool their computing resources
  - large centres and small centres can all contribute
  - users everywhere can get equal access to data and
- Grids are also flexible
  - place the computing facilities in the most effective and efficient places
  - exploiting funding wherever it is provided
- HEP and others have shown that
  - grids can support computational and storage resources on a massive scale
  - that can be operated around the clock
  - running hundreds of thousands of jobs every day
- The grid model has stimulated high energy physics to organise its computing
  - in a widely distributed way
  - building a collaboration involving directly a large fraction of the LHC members and their institutes
- This will be the workhorse for production data handling for many years and as such must be maintained and developed through the first waves of data taking

But – the landscape has changed dramatically over the past decade

- The Web, the Internet, powerful PCs, broadband to the home, …
  - have stimulated the development of new applications that generate a massive demand for computing remote from the user
  - …. that is being met by giant, efficient facilities deployed around the world
  - .... and creates a market for new technologies capable of operating on a scale equivalent to that of HEP
- Whether or not commercial clouds become cost-effective for HEP data handling is only a financial and funding-agency issue

BUT

- Exploiting the associated technologies is an obligation

Could there be a revolution here for physics analysis?