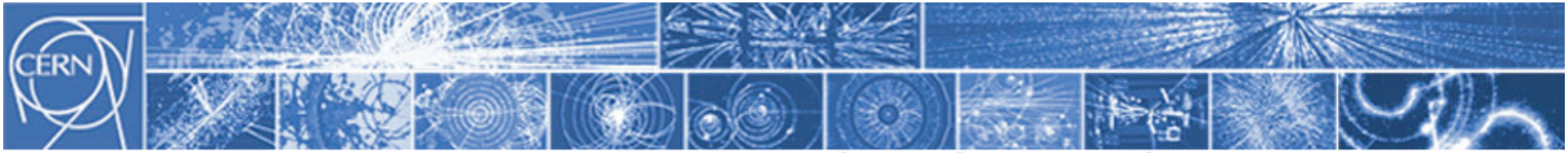


CERN
European Organization for Nuclear Research
Organisation Européenne pour la Recherche Nucléaire

The LHC Computing Challenge

Tony Cass
Leader, Fabric Infrastructure & Operations Group
Information Technology Department

15th November 2007

Outline

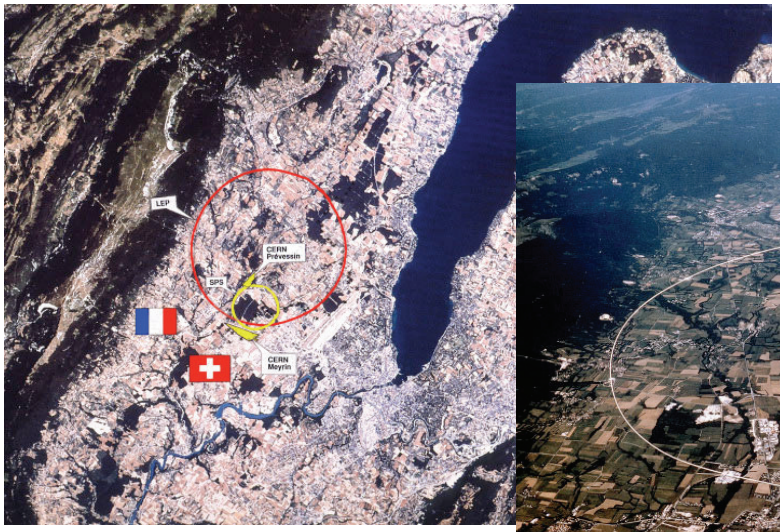
- Introduction to CERN and Experiments
- LHC Computing
- Challenges
- Summary/Conclusion

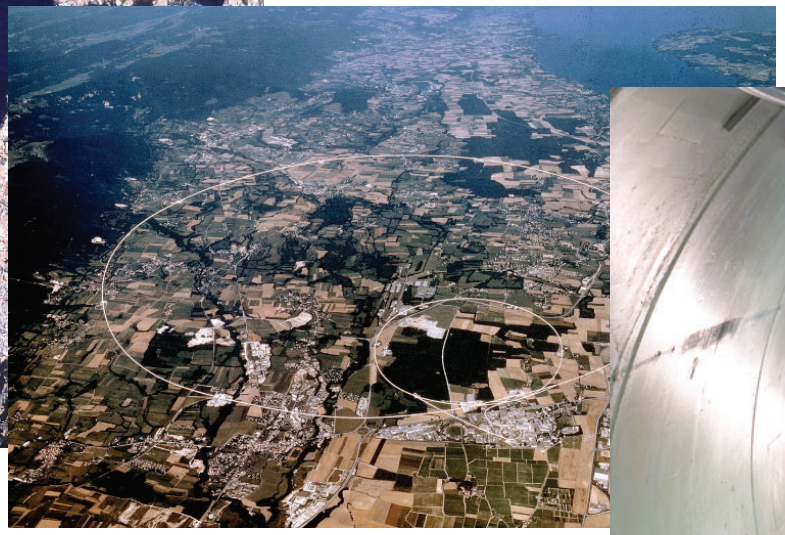
Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
- Summary/Conclusion

CERN

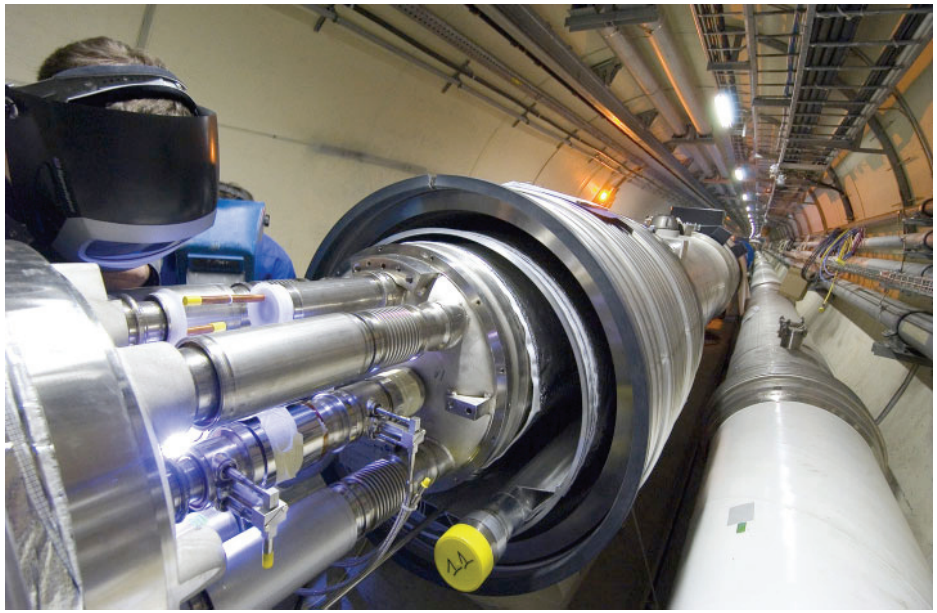
The fastest racetrack on the planet…

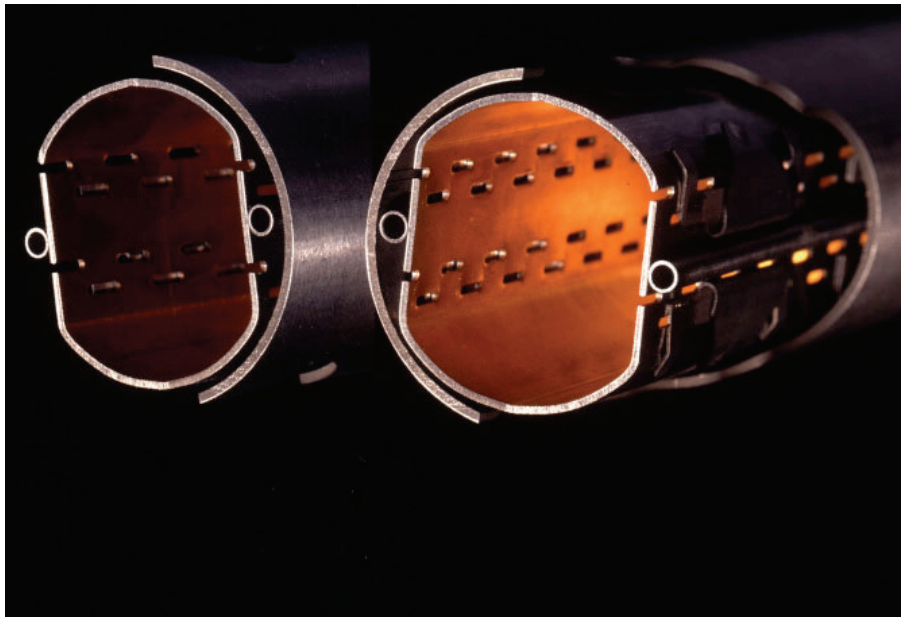




Trillions of protons will race around the 27km ring in opposite directions over 11,000 times a second, travelling at 99.999999991 per cent the speed of light.

The emptiest space in the solar system…



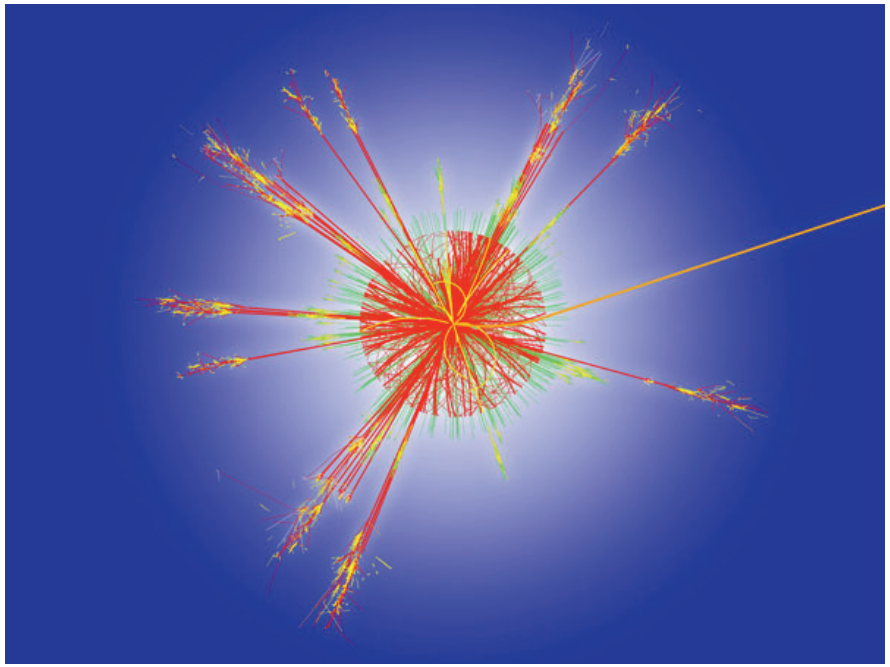


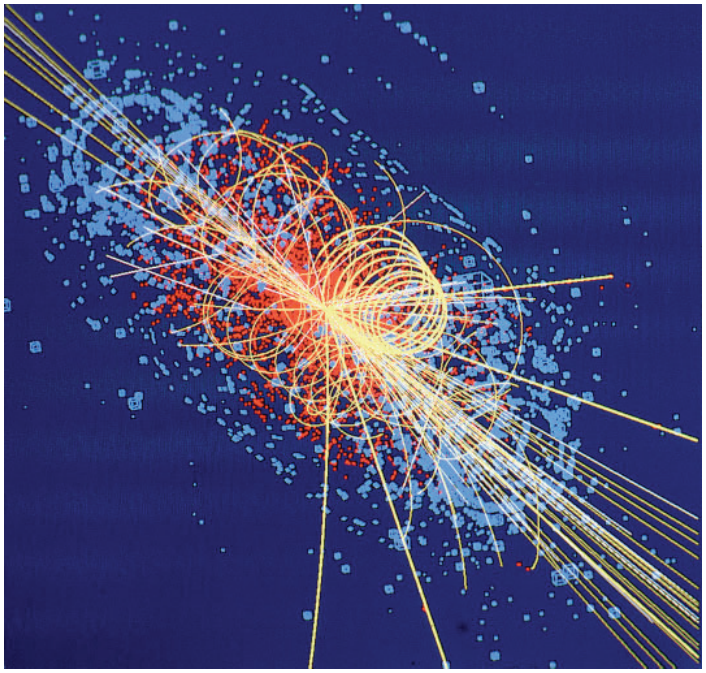
To accelerate protons to almost the speed of light requires a vacuum as empty as interplanetary space. There is 10 times more atmosphere on the moon than there will be in the LHC.

One of the coldest places in the universe…

With an operating temperature of about -271 degrees Celsius, just 1.9 degrees above absolute zero, the LHC is colder than outer space.

The hottest spots in the galaxy…

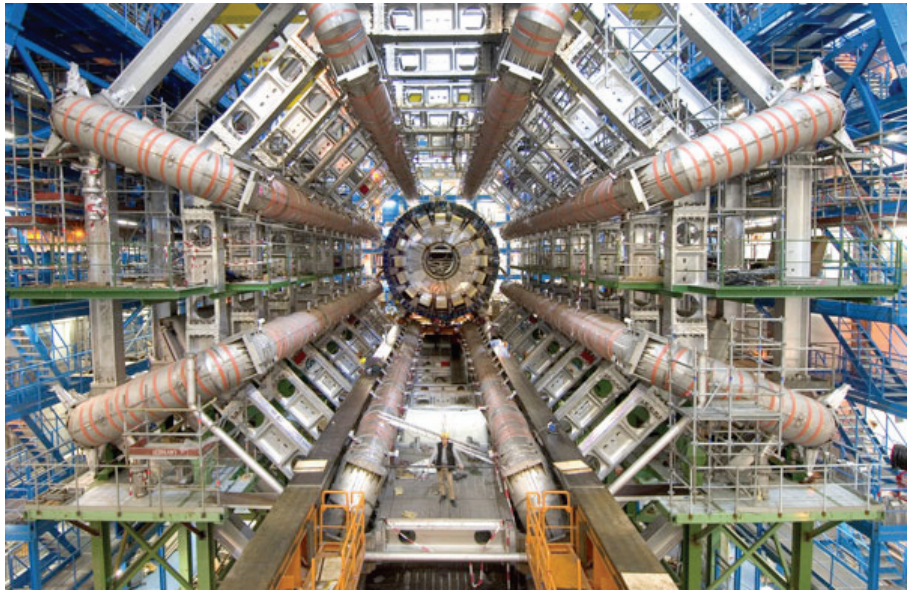




When two beams of protons collide, they will generate temperatures 1000 million times hotter than the heart of the sun, but in a minuscule space.

CERN

The biggest most sophisticated detectors ever built…



To sample and record the debris from up to 600 million proton collisions per second, scientists are building gargantuan devices that measure particles with micron precision.

CERN

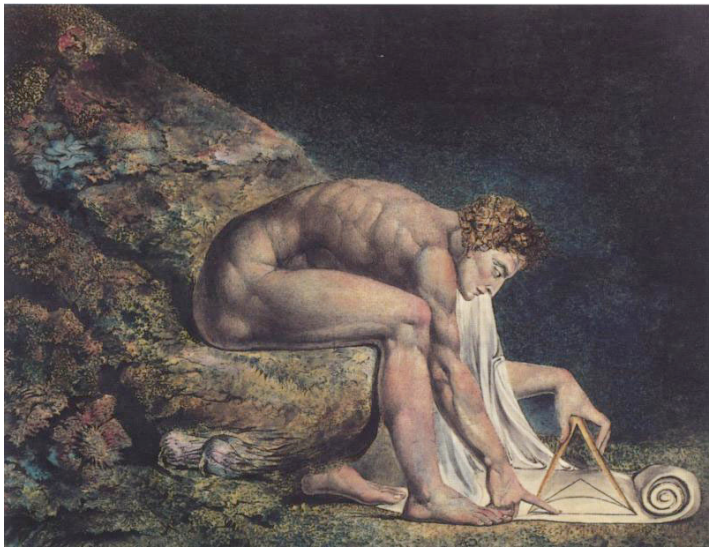
The most extensive computer system in the world…

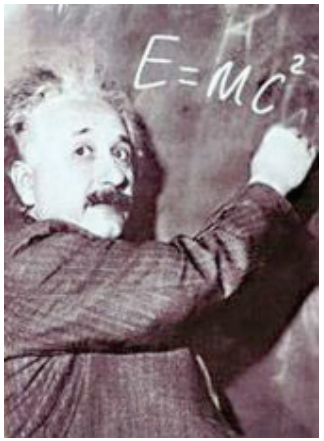
To analyse the data, tens of thousands of computers around the world are being harnessed in the Grid. The laboratory that gave the world the web, is now taking distributed computing a big step further.

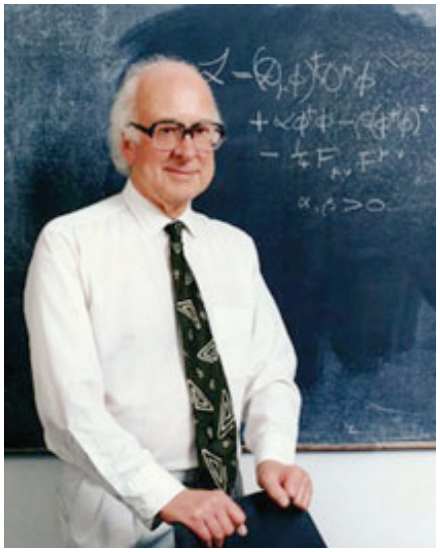
Why?

CERN

To push back the frontiers of knowledge…





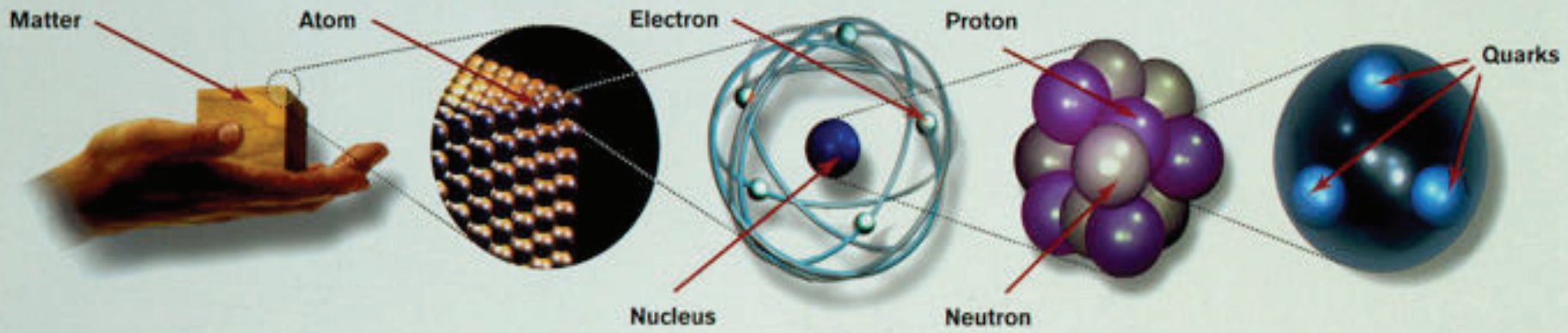


Newton's unfinished business… what is mass?

Science's little embarrassment… what is 96% of the Universe made of?

Nature's favouritism… why is there no more antimatter?

The secrets of the Big Bang… what was matter like within the first second of the Universe's life?



Matter particles
All ordinary particles belong to this group

These particles existed just after the Big Bang. Now they are found only in cosmic rays and accelerators

LEPTONS

FIRST FAMILY	**Electron** Responsible for electricity and chemical reactions; it has a charge of -1	**Electron neutrino** Particle with no electric charge, and possibly no mass; billions fly through your body every second
SECOND FAMILY	**Muon** A heavier relative of the electron; it lives for two-millionths of a second	**Muon neutrino** Created along with muons when some particles decay
THIRD FAMILY	**Tau** Heavier still; it is extremely unstable. It was discovered in 1975	**Tau neutrino** not yet discovered but believed to exist

QUARKS

Up Has an electric charge of plus two-thirds; protons contain two, neutrons contain one	**Down** Has an electric charge of minus one-third; protons contain one, neutrons contain two
Charm A heavier relative of the up; found in 1974	**Strange** A heavier relative of the down; found in 1964
Top Heavier still	**Bottom** Heavier still; measuring bottom quarks is an important test of electroweak theory

Force particles
These particles transmit the four fundamental forces of nature although gravitons have so far not been discovered

Gluons
Carriers of the strong force between quarks

Felt by: quarks

The explosive release of nuclear energy is the result of the strong force

Photons
Particles that make up light; they carry the electromagnetic force

Felt by: quarks and charged leptons

Electricity, magnetism and chemistry are all the results of electro-magnetic force

Intermediate vector bosons
Carriers of the weak force

Felt by: quarks and leptons

Some forms of radio-activity are the result of the weak force

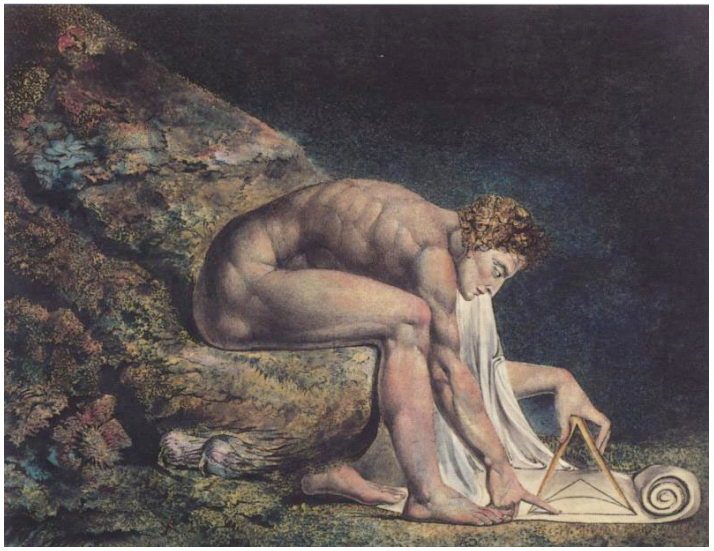
Gravitons
Carriers of gravity

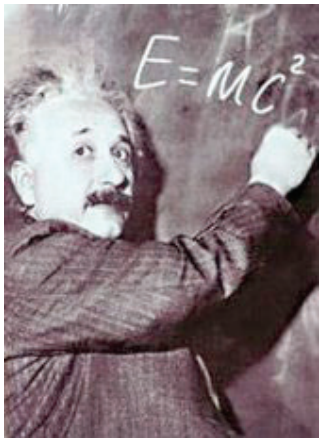
Felt by: all particles with mass

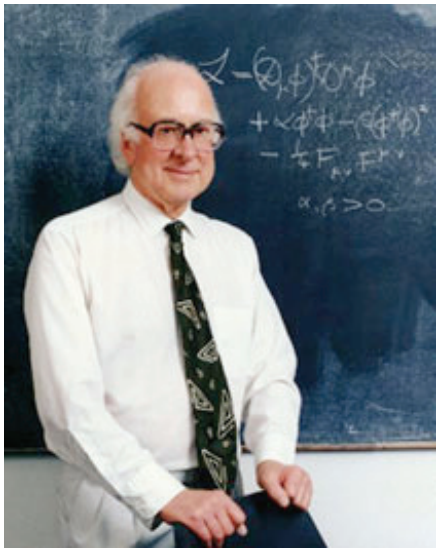
All the weight we experience is the result of the gravitational force

CERN

To push back the frontiers of knowledge…





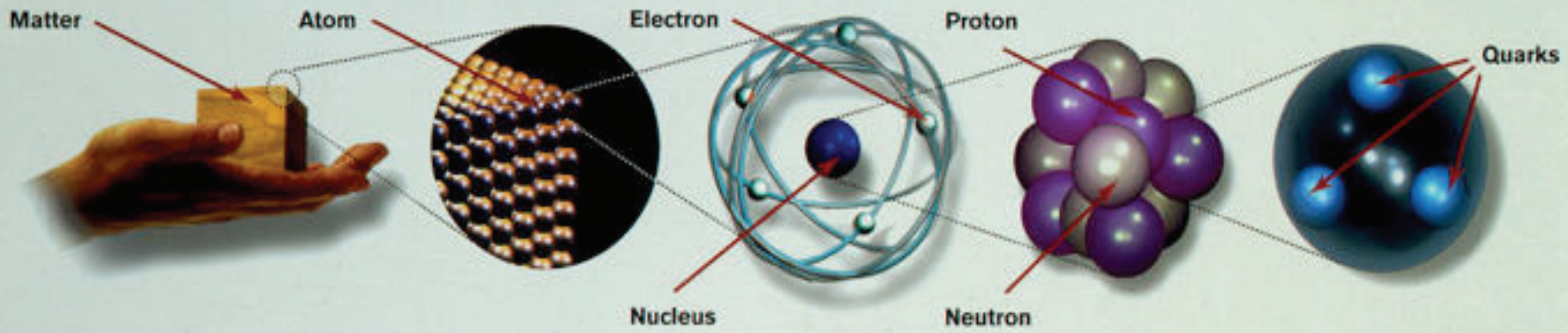


Newton's unfinished business… what is mass?

Science's little embarrassment… what is 96% of the Universe made of?

Nature's favouritism… why is there no more antimatter?

The secrets of the Big Bang… what was matter like within the first second of the Universe's life?



Matter particles
All ordinary particles belong to this group

These particles existed just after the Big Bang. Now they are found only in cosmic rays and accelerators

LEPTONS

FIRST FAMILY	**Electron** Responsible for electricity and chemical reactions; it has a charge of -1	**Electron neutrino** Particle with no electric charge, and possibly no mass; billions fly through your body every second
SECOND FAMILY	**Muon** A heavier relative of the electron; it lives for two-millionths of a second	**Muon neutrino** Created along with muons when some particles decay
THIRD FAMILY	**Tau** Heavier still; it is extremely unstable. It was discovered in 1975	**Tau neutrino** not yet discovered but believed to exist

QUARKS

Up Has an electric charge of plus two-thirds; protons contain two, neutrons contain one	**Down** Has an electric charge of minus one-third; protons contain one, neutrons contain two
Charm A heavier relative of the up; found in 1974	**Strange** A heavier relative of the down; found in 1964
Top Heavier still	**Bottom** Heavier still; measuring bottom quarks is an important test of electroweak theory

Force particles
These particles transmit the four fundamental forces of nature although gravitons have so far not been discovered

Gluons
Carriers of the strong force between quarks

Felt by: quarks

The explosive release of nuclear energy is the result of the strong force

Photons
Particles that make up light; they carry the electromagnetic force

Felt by: quarks and charged leptons

Electricity, magnetism and chemistry are all the results of electro-magnetic force

Intermediate vector bosons
Carriers of the weak force

Felt by: quarks and leptons

Some forms of radio-activity are the result of the weak force

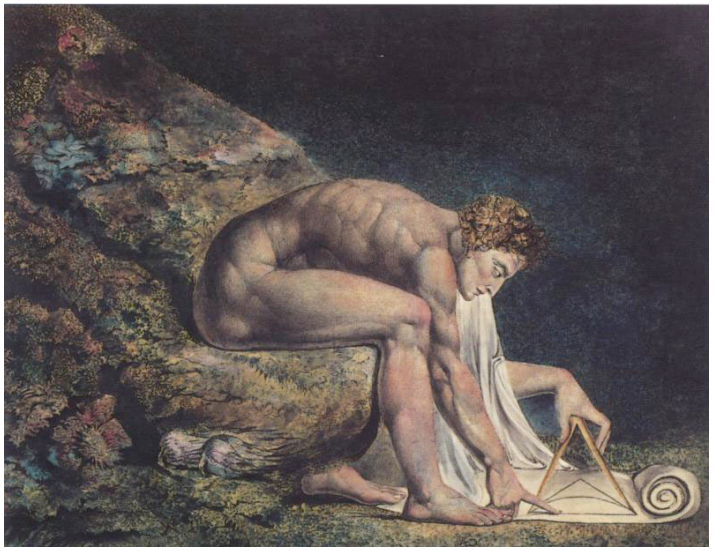
Gravitons
Carriers of gravity

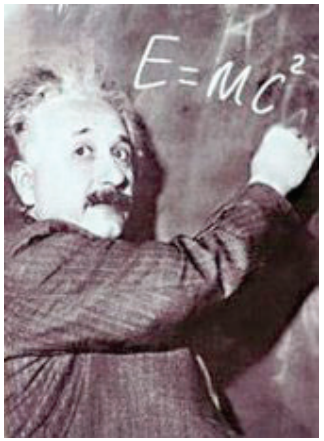
Felt by: all particles with mass

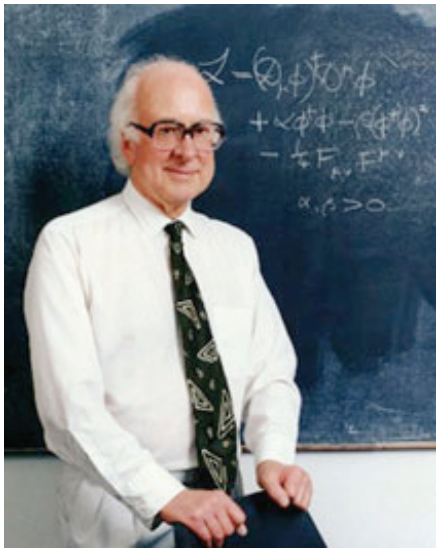
All the weight we experience is the result of the gravitational force

CERN

To push back the frontiers of knowledge…







Newton's unfinished business… what is mass?

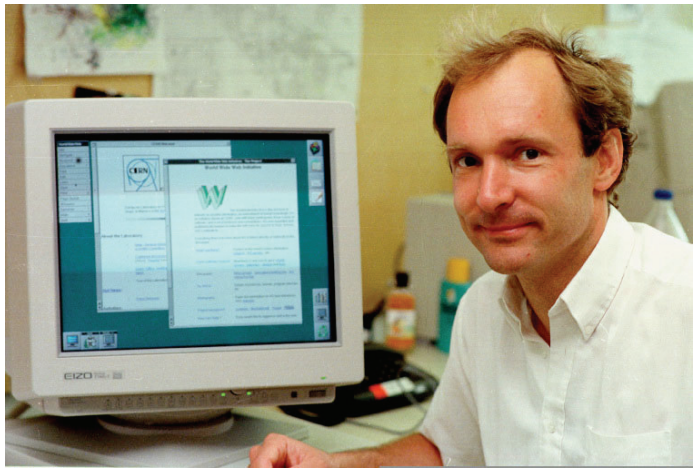
Science's little embarrassment… what is 96% of the Universe made of?

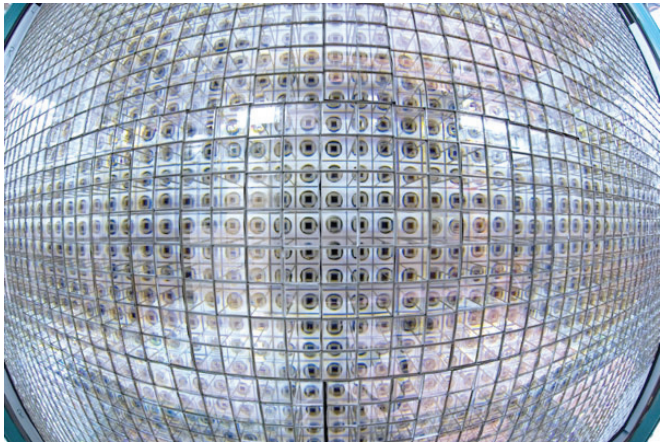
Nature's favouritism… why is there no more antimatter?

The secrets of the Big Bang… what was matter like within the first second of the Universe's life?

CERN

To develop new technologies…





Information technology - the Web and the Grid

Medicine - diagnosis and therapy

Security - scanning technologies for harbours and airports

Vacuum - new techniques for flat screen displays or solar energy devices

CERN

To unite people from different countries and cultures…

20 Member states

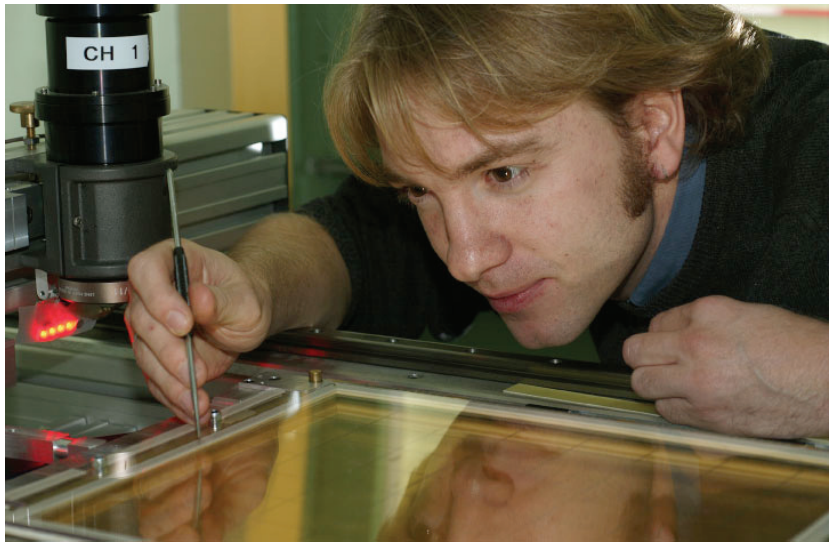
38 Countries with cooperation agreements

111 Nationalities

10000 People

CERN

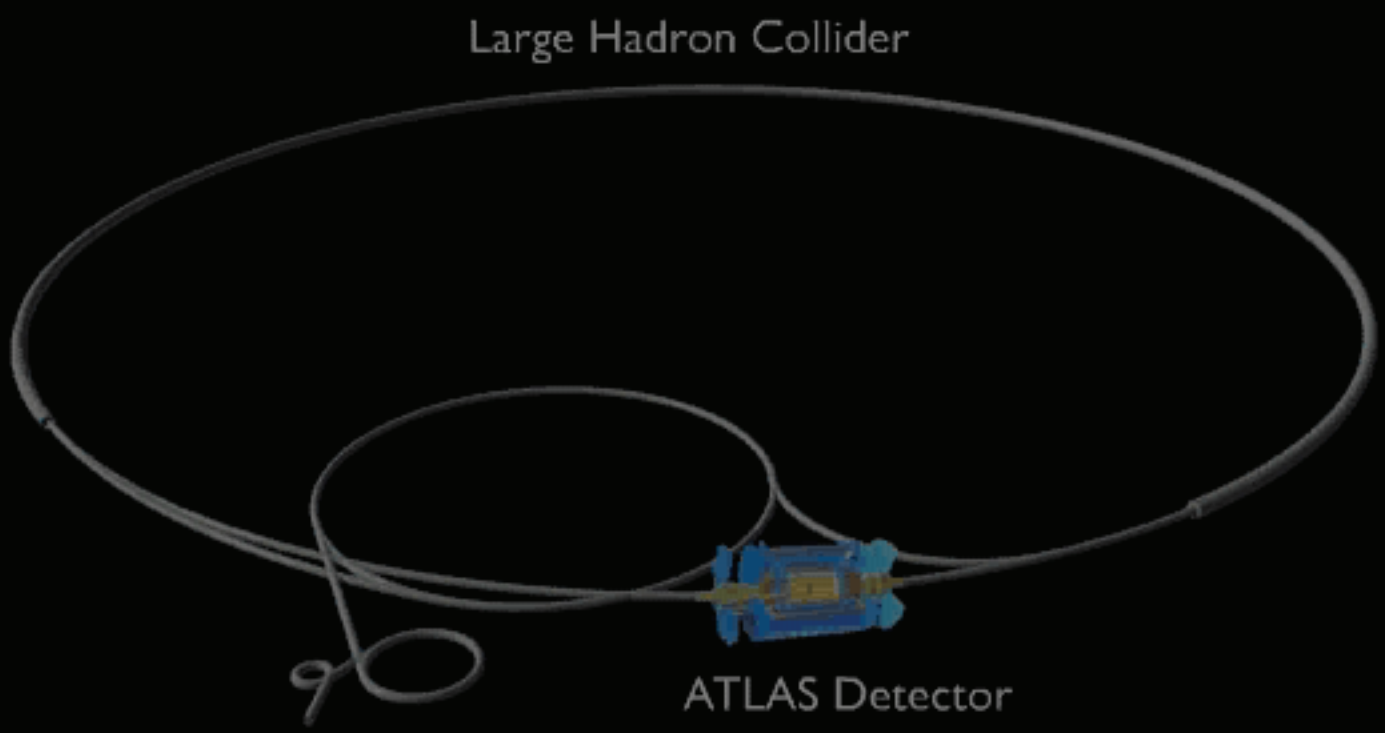
To train the scientists and engineers of tomorrow…



From mini-Einstein workshops for five to sixes, through to professional schools in physics, accelerator science and IT, CERN plays a valuable role in building enthusiasm for science and providing formal training..

“Compact” Detectors!

PLAY



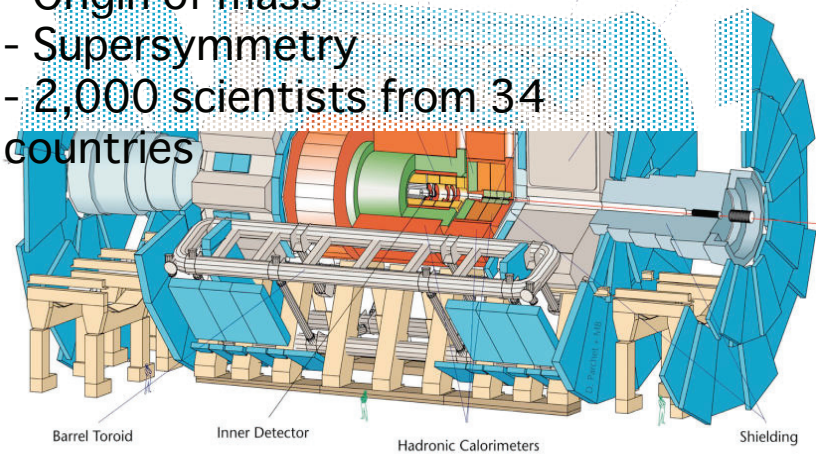
Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
- Summary/Conclusion

The Four LHC Experiments…

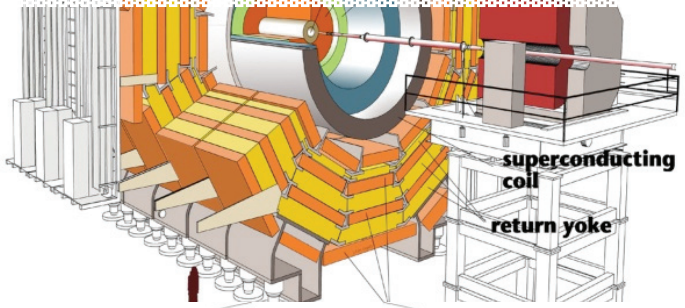
ATLAS

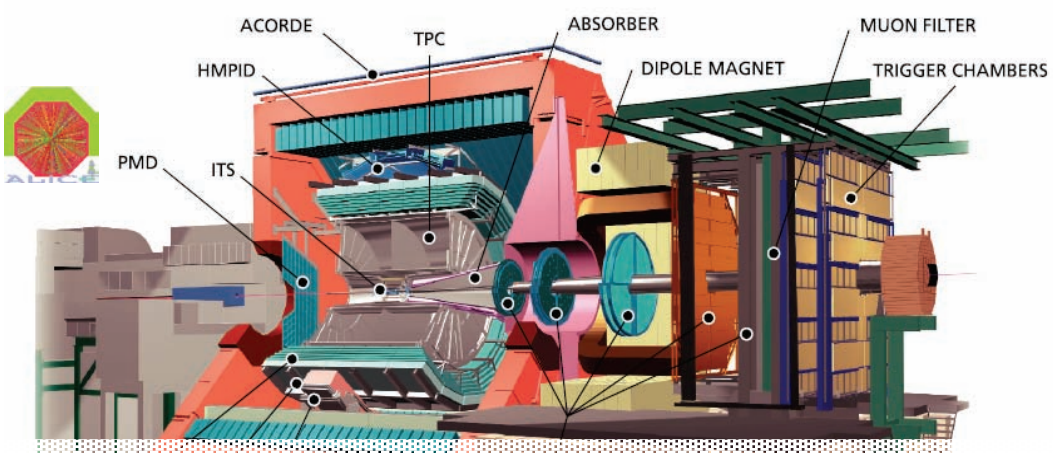
- General purpose
- Origin of mass
- Supersymmetry
- 2,000 scientists from 34 countries



CMS

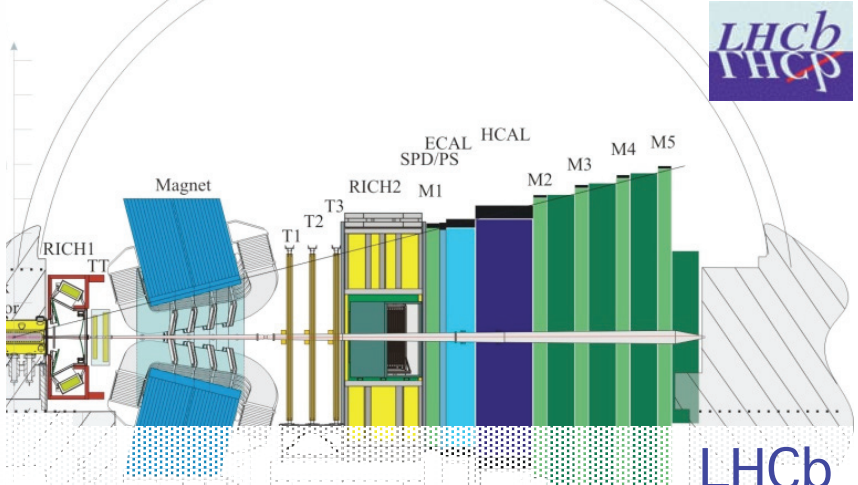
- General purpose
- Origin of mass
- Supersymmetry
-1,800 scientists from over 150 institutes





ALICE

- heavy ion collisions, to create quark-gluc plasmas
- 50,000 particles in each collision



LHCb

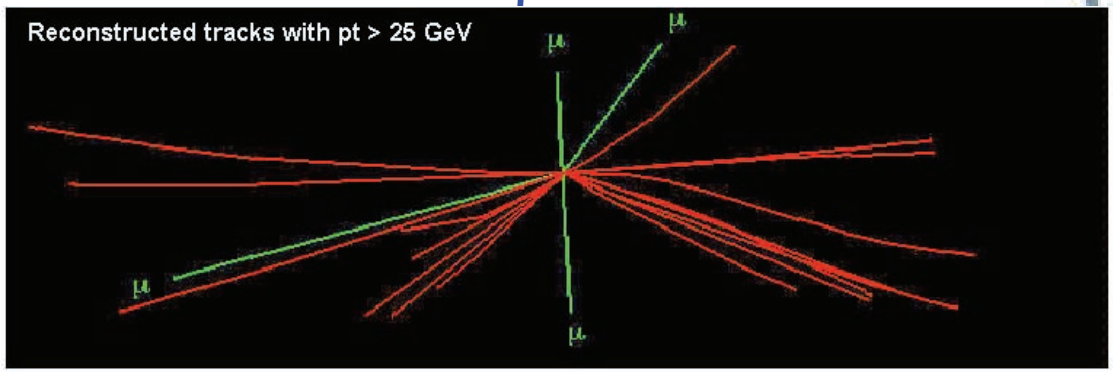
- to study the differences between matter and antimatter

… generate lots of data …

The accelerator generates 40 million particle collisions (events) every second at the centre of each of the four experiments' detectors

… generate lots of data …

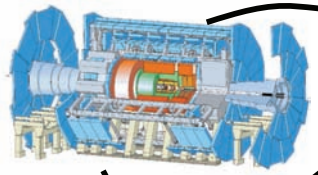
reduced by online computers to a few hundred "good" events per second.



Which are recorded on disk and magnetic tape at 100-1,000 MegaBytes/sec → ~15 PetaBytes per year

for all four experiments

Data Handling and Computation for Physics Analysis



detector

event filter (selection & reconstruction)

reconstruction

raw data

event reprocessing

event summary data

processed data

batch physics analysis

analysis

analysis objects (extracted by physics topic)

interactive physics analysis

event simulation

simulation

… leading to a high box count

Summary of Computing Resource Requirements

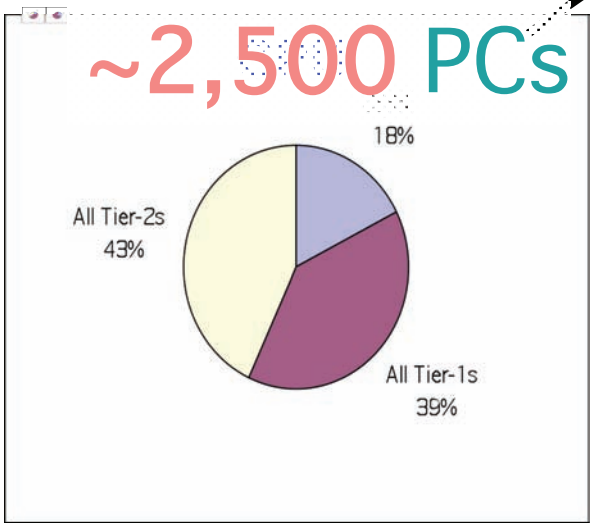
All experiments - 2008

From LCG TDR - June 2005

	CERN	All Tier-1s	All Tier-2s	Total
CPU (MSPECint2000s)	25	56	61	142
Disk (PetaBytes)	7	31	19	57
Tape (PetaBytes)	18	35		53

~2,500 PCs

Another ~1,500 boxes



Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
- Summary/Conclusion

Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

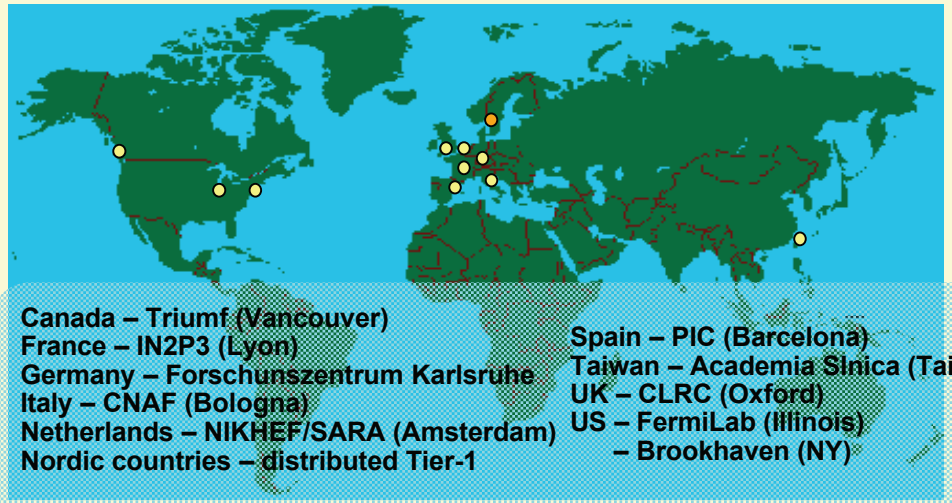
Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

Computing Service Hierarchy

Tier-0 – the accelerator centre

- Data acquisition & initial processing
- Long-term data curation
- Distribution of data Tier-1 centres



Tier-1 – "online" to the data acquisition process high availability

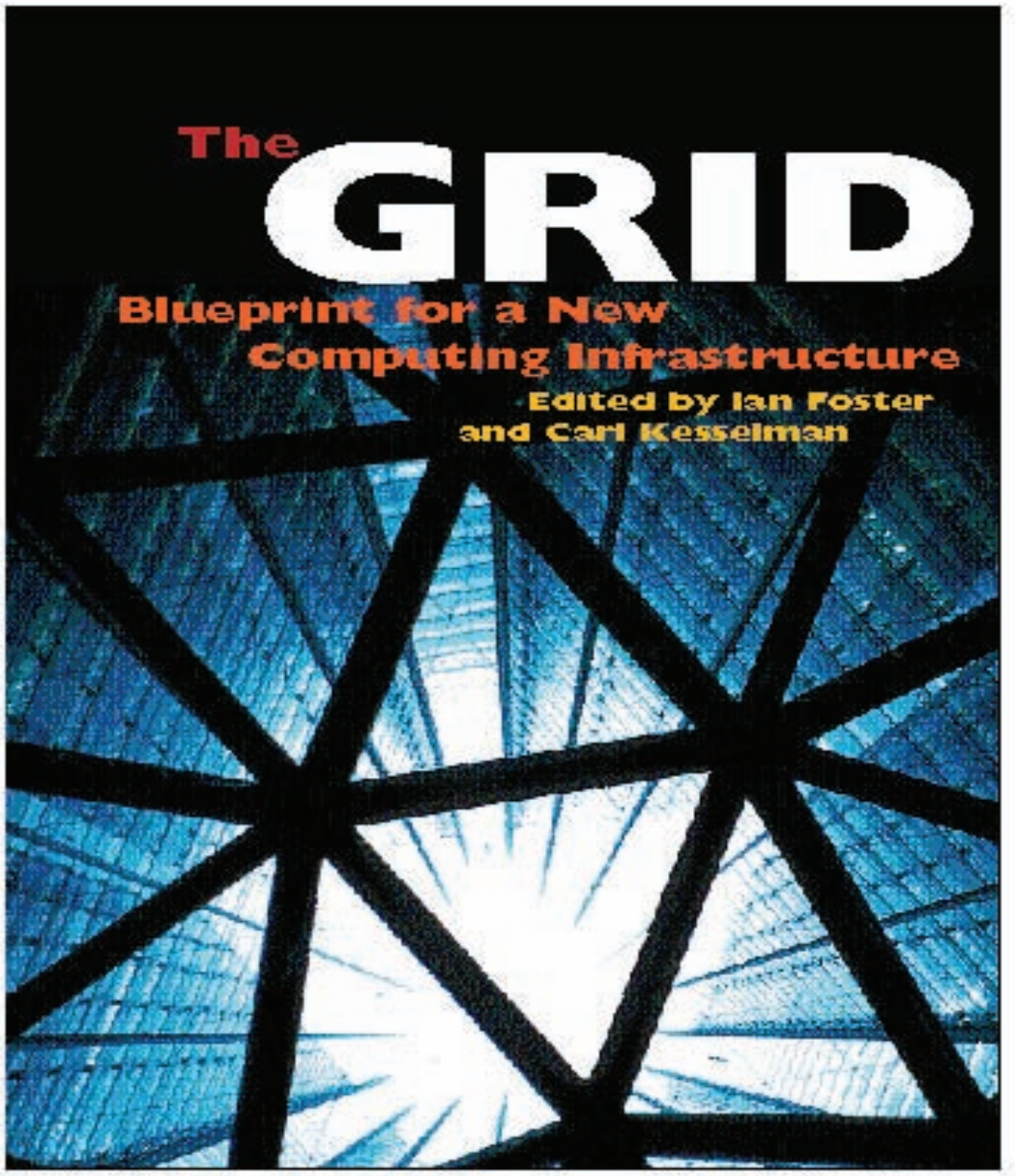
- Managed Mass Storage
- Data-heavy analysis
- National, regional support

Tier-2 – ~100 centres in ~40 countries

- Simulation
- End-user analysis – batch and interactive

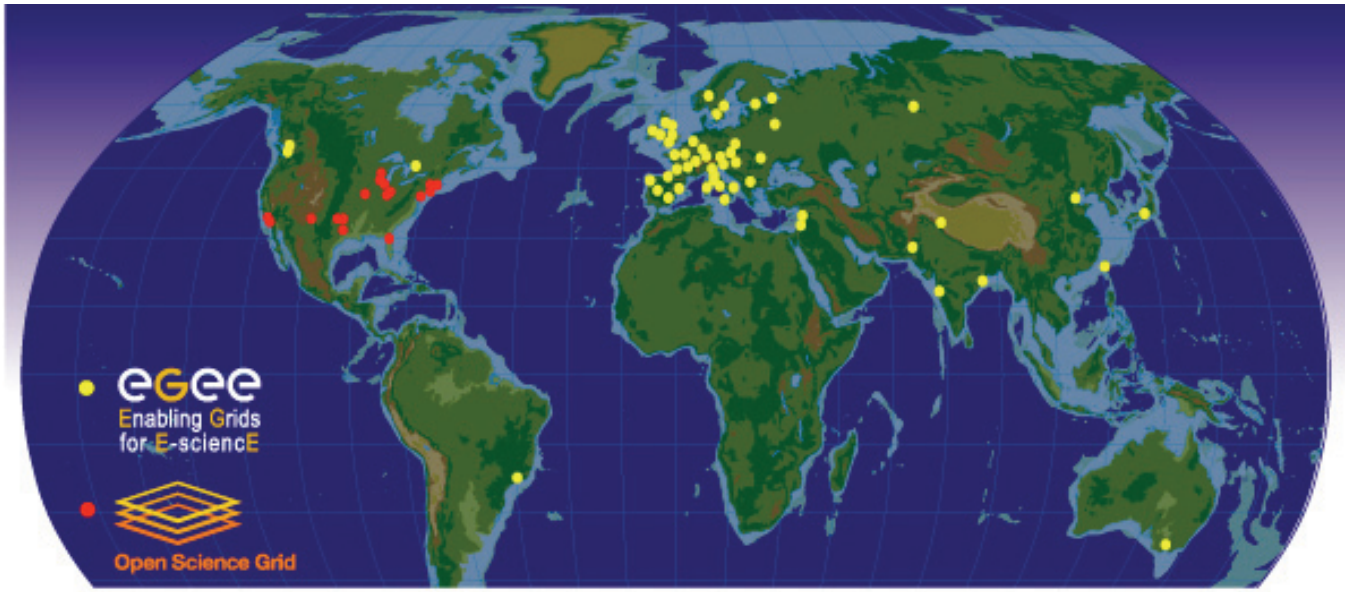
The Grid

- Timely Technology!
- Deploy to meet LHC computing needs.
- Challenges for the Worldwide LHC Computing Grid Project due to
  - worldwide nature
    - competing middleware…
  - newness of technology
    - competing middleware…
  - scale



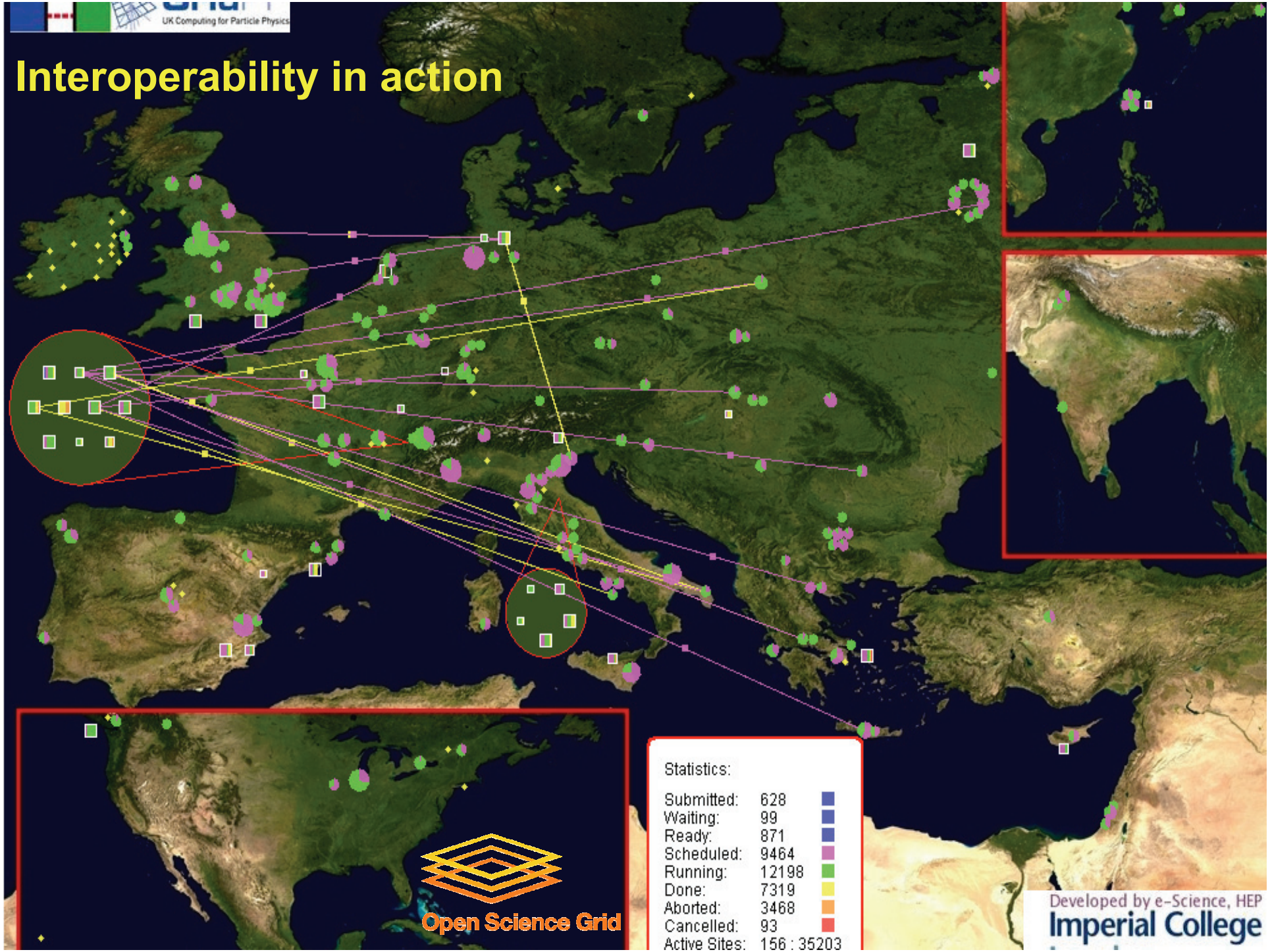
Interoperation between Grid Infrastructures

- WLCG service relies on three Grid infrastructures: EGEE, OSG and NorduGrid
- Interoperability required (and achieved) for
  - users (job submission)
  - administration (identity, monitoring, accounting, …)



A map of the worldwide LCG infrastructure operated by EGEE and OSG.

Interoperability in action



Feature

Seven Wonders of the IT World

The fastest supercomputer. The most intriguing data center. The constantly changing core at the heart of Linux. Take a tour of the most impressive and most unusual marvels of the IT world.

Leave a comment (65)

By C.G. Lynch

World's largest scientific grid computing project: The E-sciencE II (EGEE-II) project

Launched: September 2006, for use by scientists around the world.

A Google Earth view of European sites hooked into the EGEE grid computing project

Helps power: Large-scale scientific research projects in fields from geology to chemistry—for example, will analyze data from CERN's Large Hadron Collider, a particle accelerator being built to help investigate details around the Big Bang and related physics questions.

Amount of work it does: 98,000 jobs a day, more than 1 million per month.

Juggling ability: Runs about 30,000 jobs concurrently, on average.

RELATED SOLUTIONS

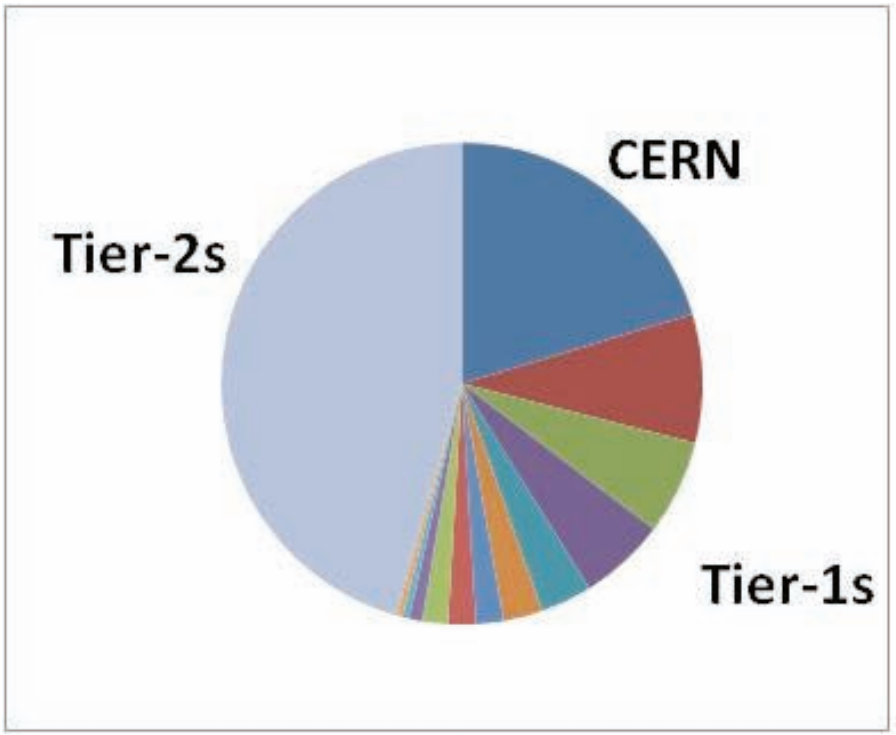
How-To »
- How to Lay Dead Technologies to Rest
- Ghostly Gear: Technology Tools for Paranormal Investigations

Research & Analysis »
- Lax Laptop Security Can Be Dangerous...and Expensive
- Reduce Information Technology Complexity, Costs with Consolidation

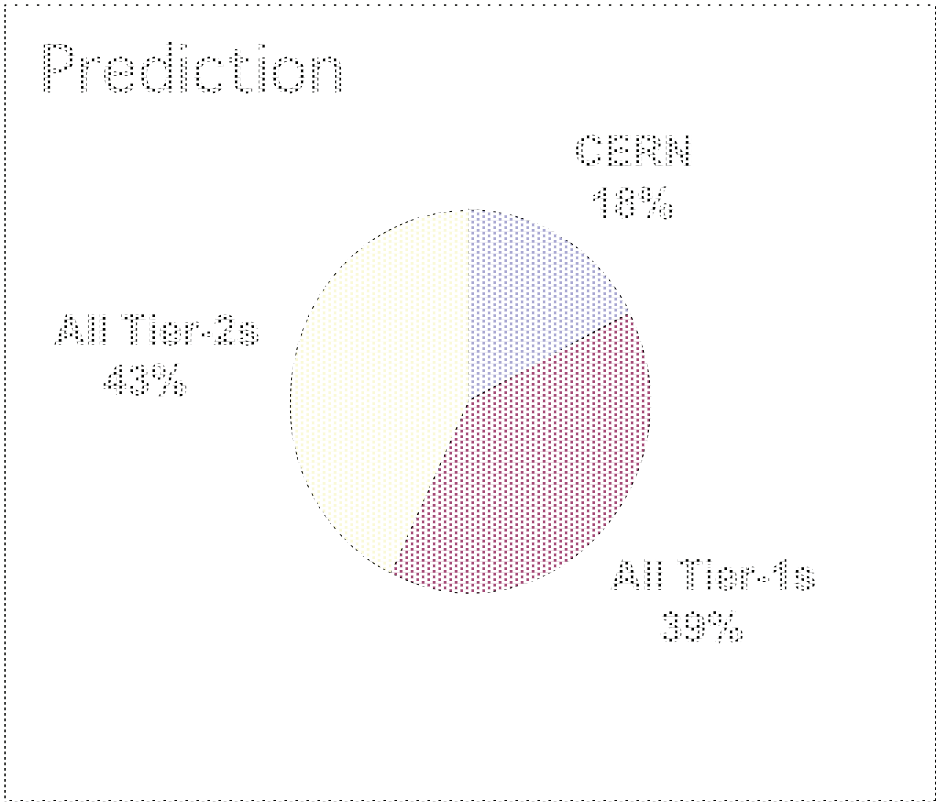
Advice & Opinion »
- Join the Conversation!
- Ask, Answer and Interact with the CIO Community

News »
- Nokia Makes the Ferrari of Phones...Literally
- Holiday Gift Guide 2007: Best Technology Bling

Accounted CPU Usage



80 Tier-2s 45%
11 Tier-1s 35%
CERN 20%



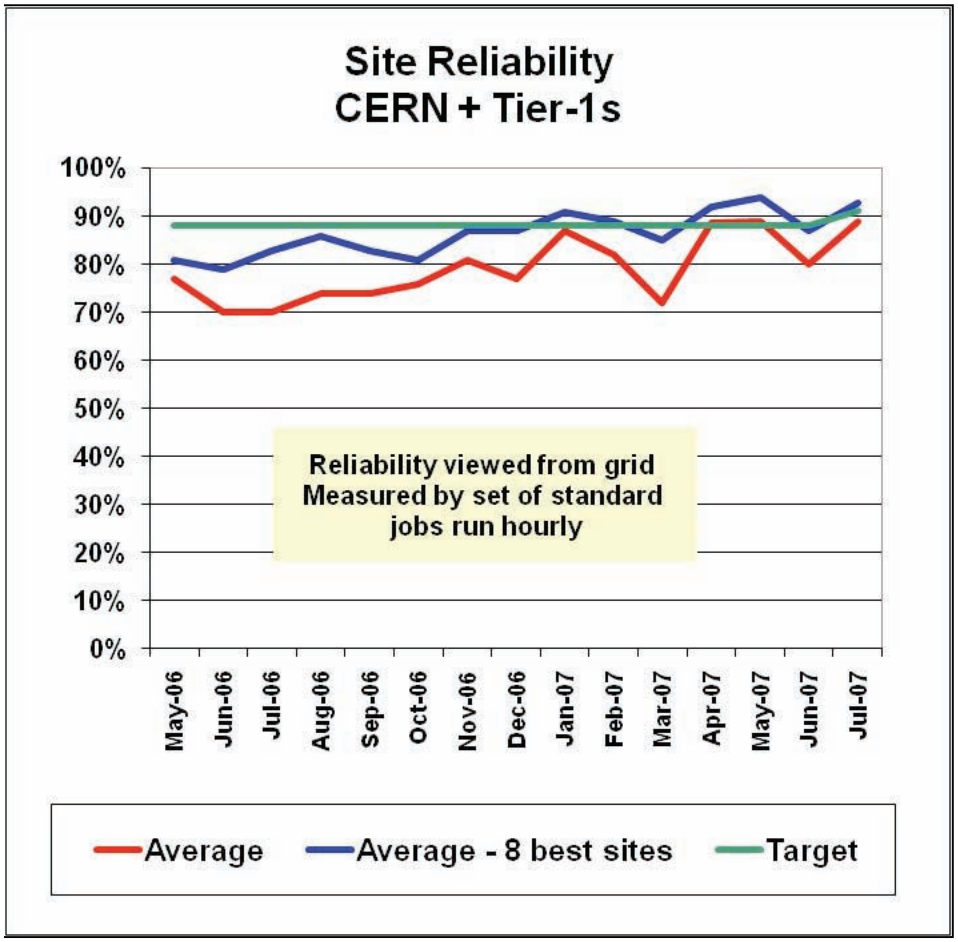
530M SI2K-days/month (CPU)

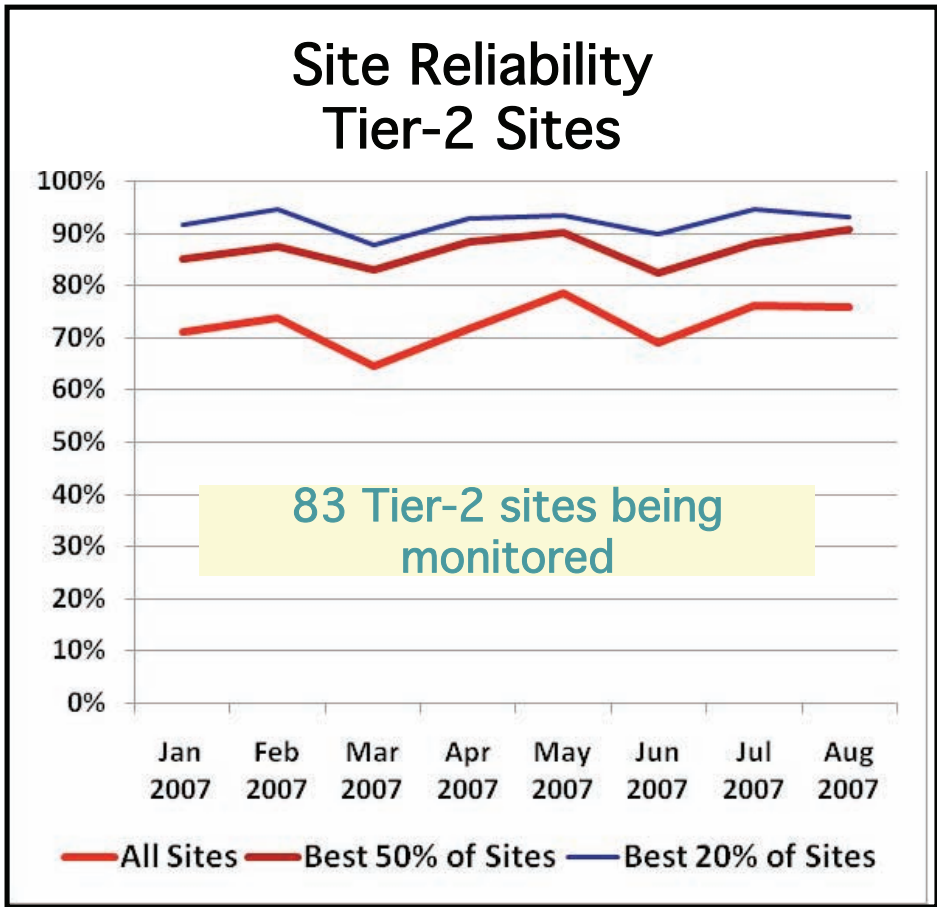
9 PB disk at CERN + Tier-1s

Remaining Challenges

- Creating a working Grid service across multiple infrastructure is clearly a success, but challenges remain
  - Reliability

Reliability

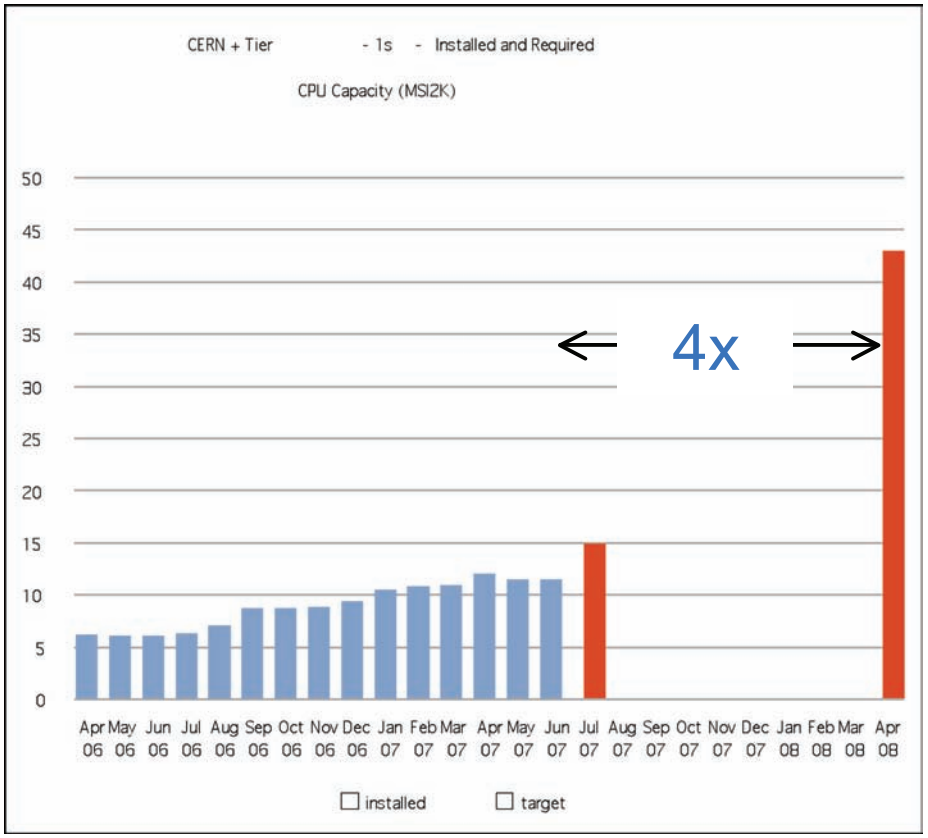


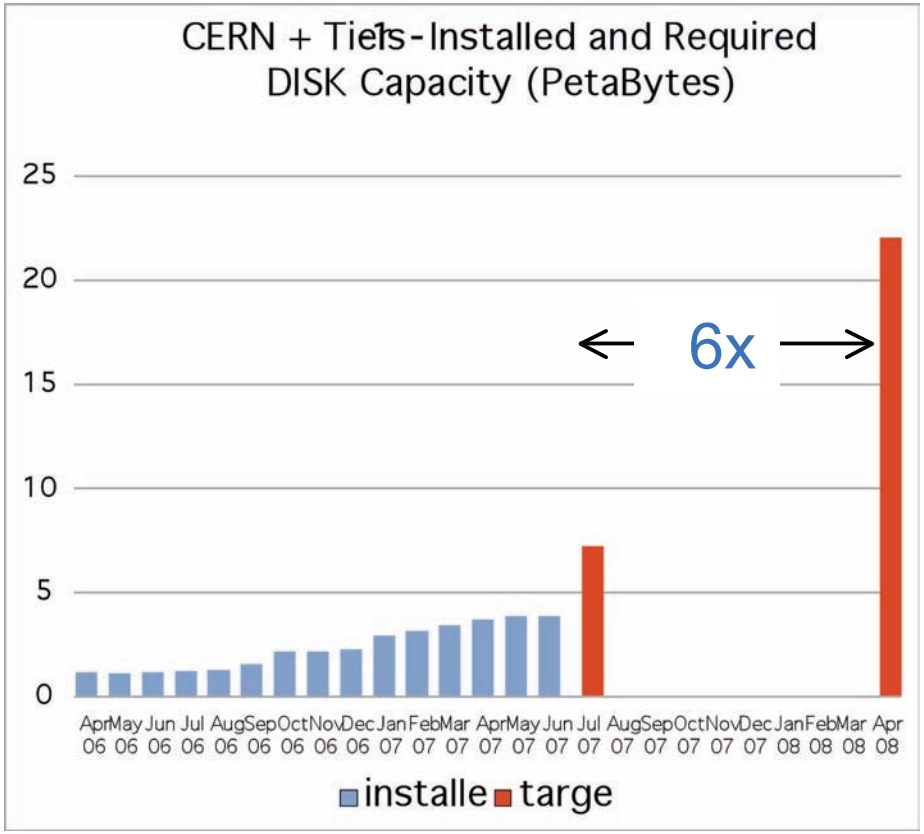


Remaining Challenges

- Creating a working Grid service across multiple infrastructure is clearly a success, but challenges remain
  - Reliability
  - Ramp-up

Steep ramp-up still needed before first physics run





Evolution of installed capacity from April 06 to June 07
Target capacity from MoU pledges for 2007 (due July07)
and 2008 (due April 08)

Remaining Challenges

- Creating a working Grid service across multiple infrastructure is clearly a success, but challenges remain
  - Reliability
  - Ramp-up
  - Collaboration
    - From computer centre empires to a federation
    - consensus rather than control
  - …

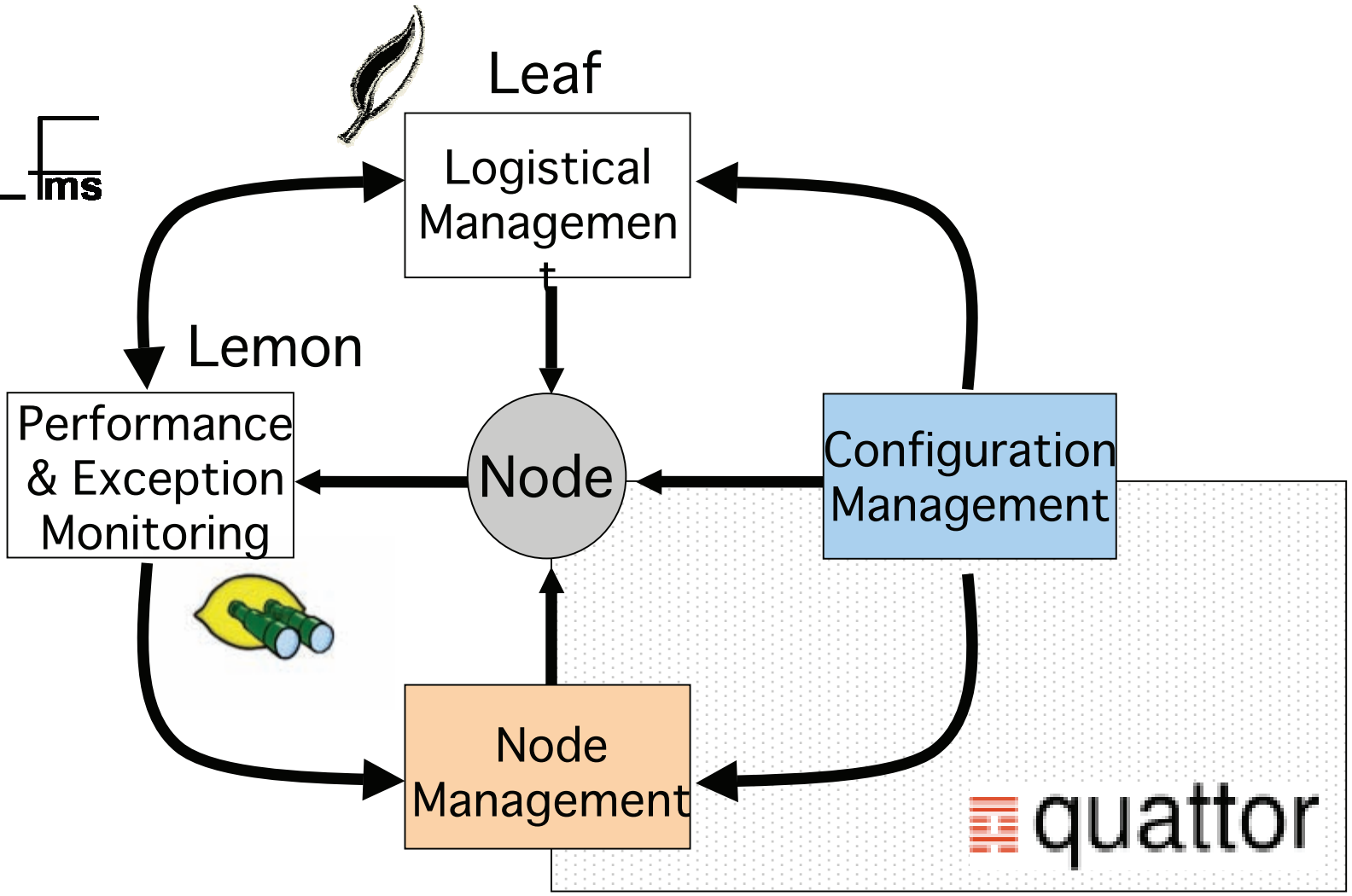
Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
    - Installation & Configuration
    - Monitoring
    - Workflow
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

ELFms Vision



Toolkit developed by CERN in collaboration with many HEP sites and as part of the European DataGrid Project.
See http://cern.ch/ELFms

Why a bespoke system?

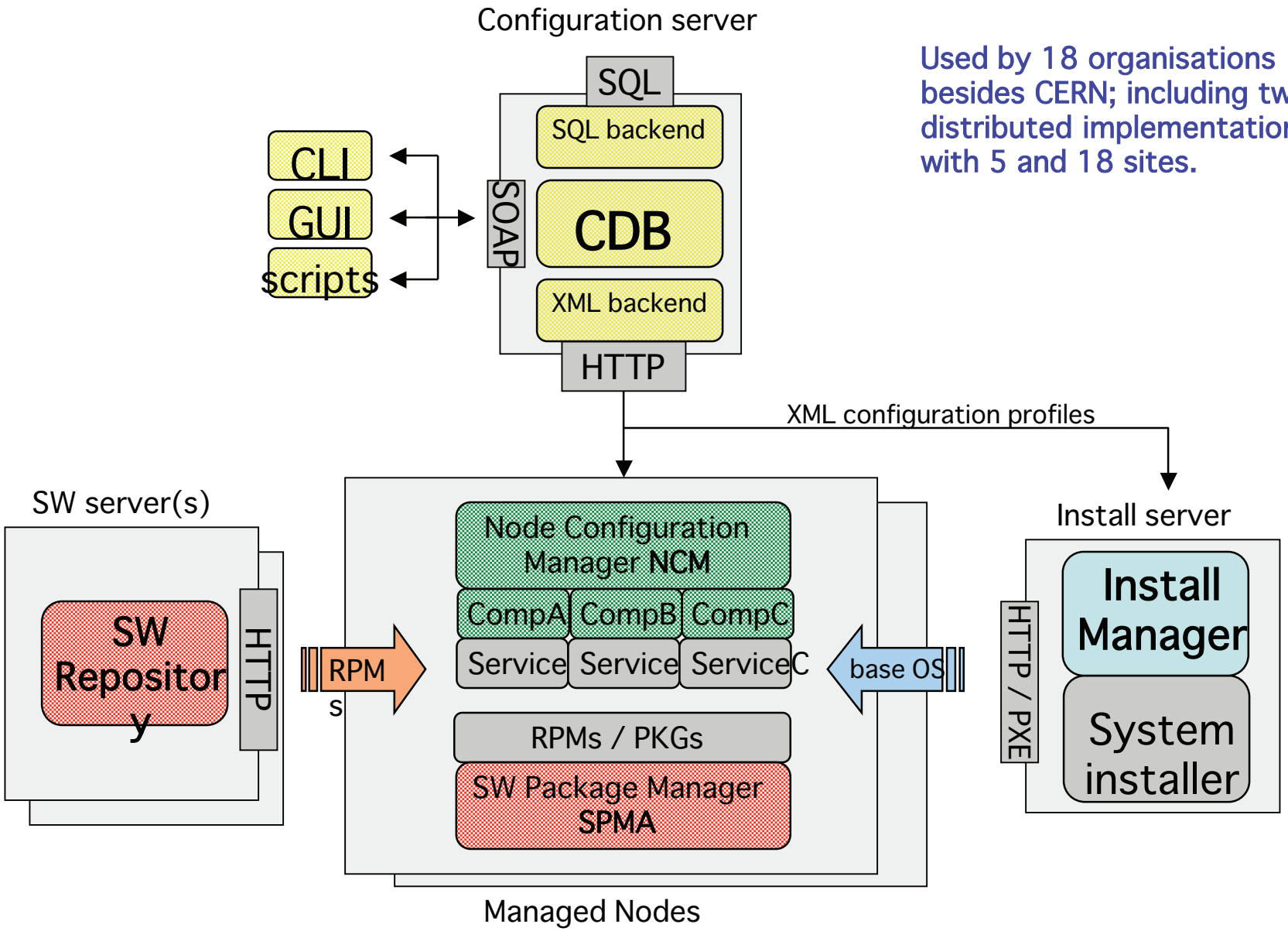
- Commercial Management Suites
  - (Full) Linux support rare (5+ years ago…)
  - Much work needed to deal with specialist HEP applications; insufficient reduction in staff costs to justify license fees.
- Open Source Systems
  - Many packages with interesting features, but none featuring all of items considered essential
    - Declarative, hierarchical configuration specification permitting validation, integrated software distribution and configuration management, separation of configuration data and code, feedback loop to avoid configuration drift, ability to update running systems

See EDG/WP4 report on current technology (http://cern.ch/hep-proj-grid-fabric/Tools/DataGrid-04-TED-0101-3_0.pdf) or "Framework for Managing Grid-enabled Large Scale Computing Fabrics" (http:/cern.ch/quattor/documentation/poznanski-phd.pdf) for reviews of various packages.

Outline

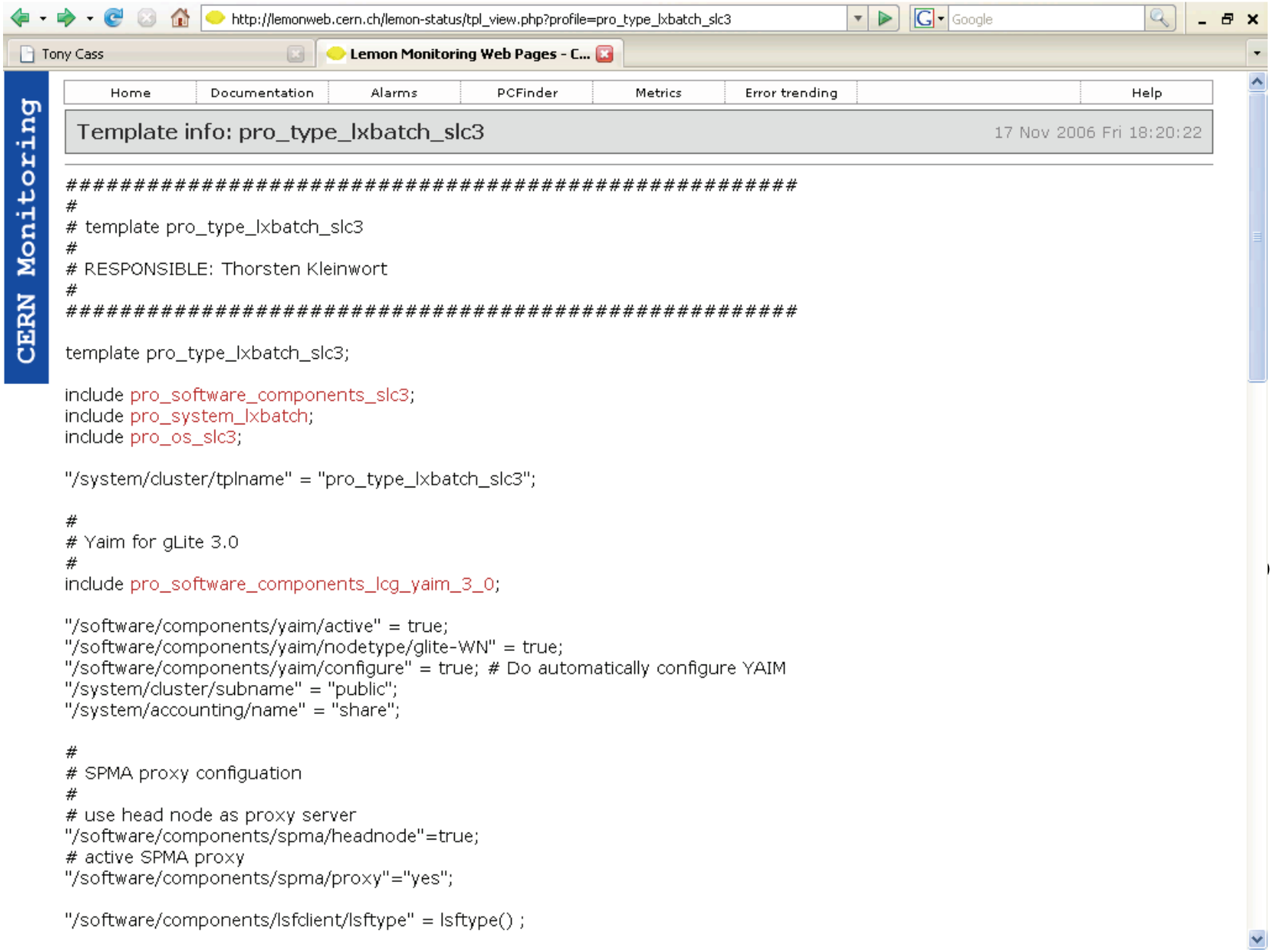
- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
    - Installation & Configuration
    - Monitoring
    - Workflow
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

Quattor

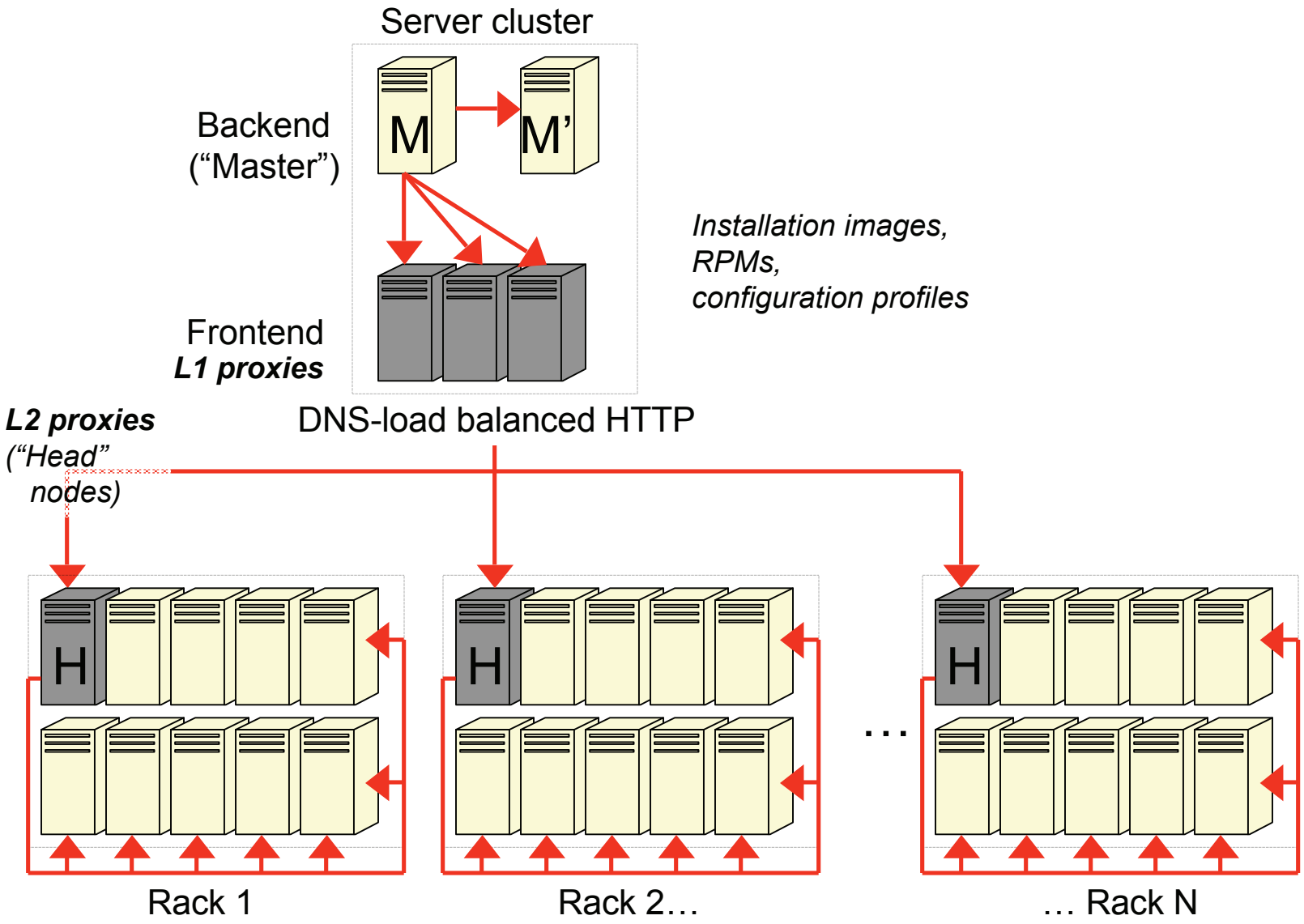


Used by 18 organisations besides CERN; including two distributed implementations with 5 and 18 sites.

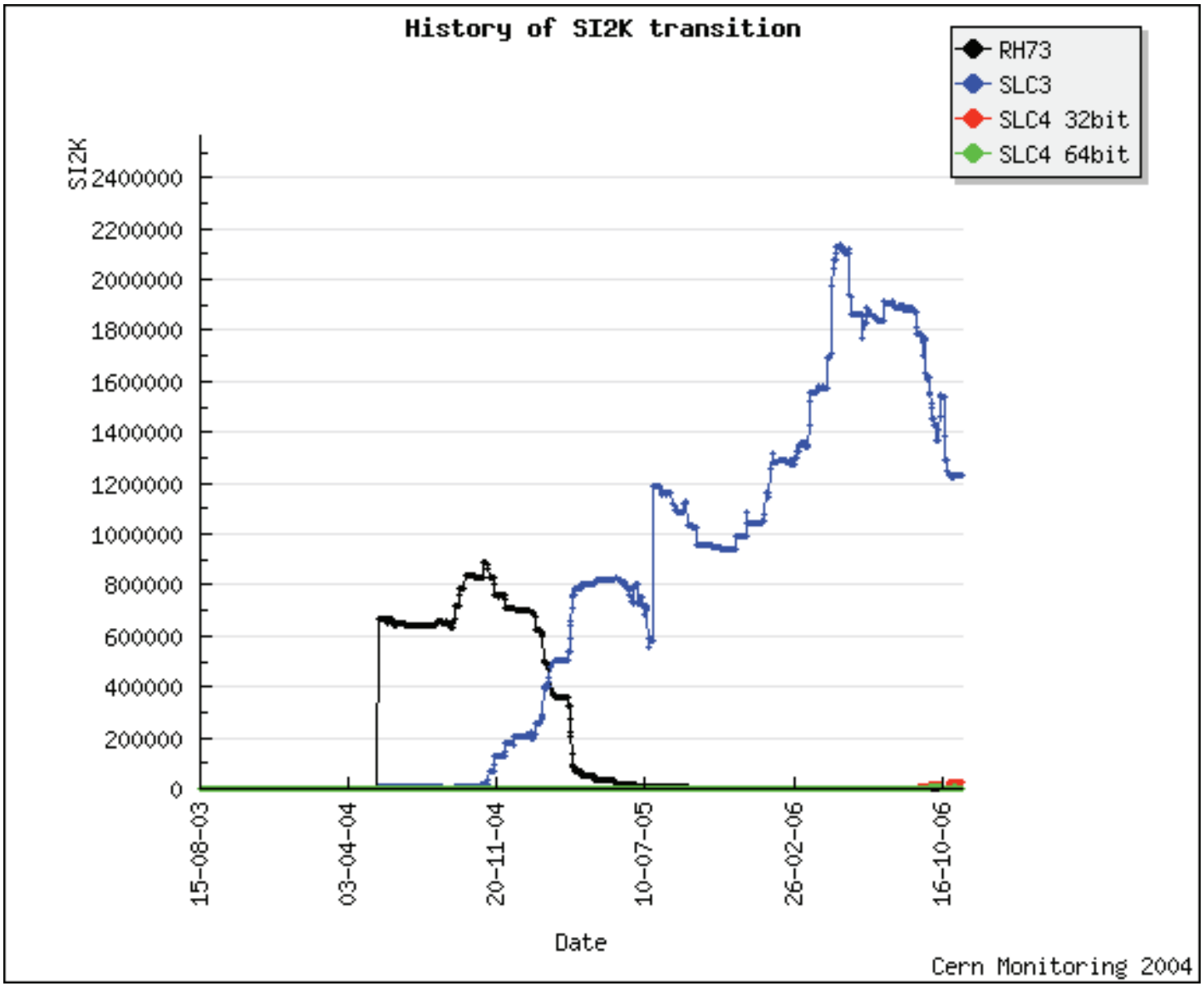
Configuration Hierarchy



Scalable s/w distribution…



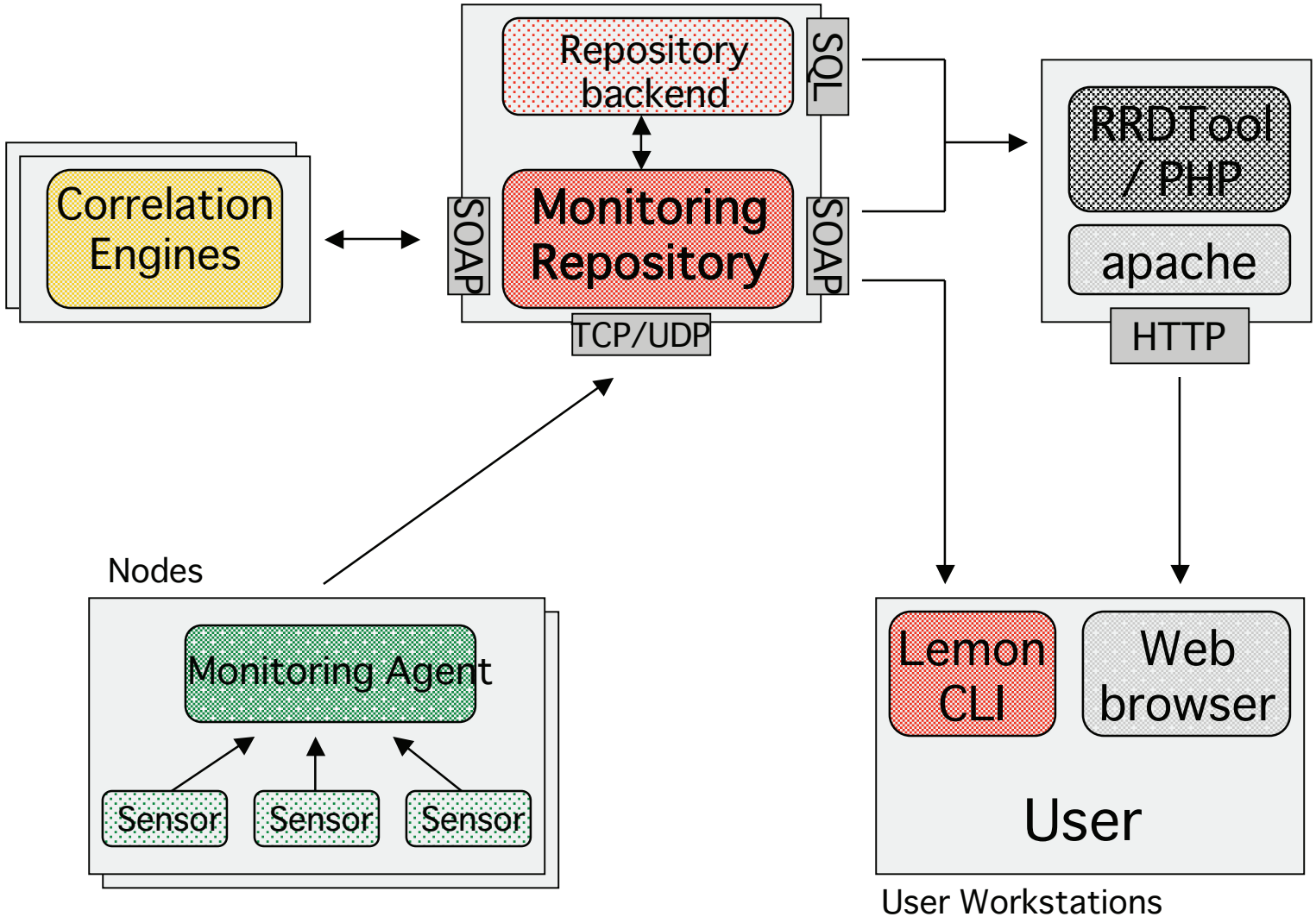
… in practice!


History of SI2K transition
RH73
SLC3
SLC4 32bit
SLC4 64bit
SI2K
2400000
2200000
2000000
1800000
1600000
1400000
1200000
1000000
800000
600000
400000
200000
0
15-08-03
03-04-04
20-11-04
10-07-05
26-02-06
16-10-06
Date
Cern Monitoring 2004

Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
    - Installation & Configuration
    - Monitoring
    - Workflow
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

Lemon


Repository backend
SQL
Correlation Engines
SOAP
Monitoring Repository
SOAP
TCP/UDP
RRDTool / PHP
apache
HTTP
Nodes
Monitoring Agent
Sensor
Sensor
Sensor
Lemon CLI
Web browser
User
User Workstations

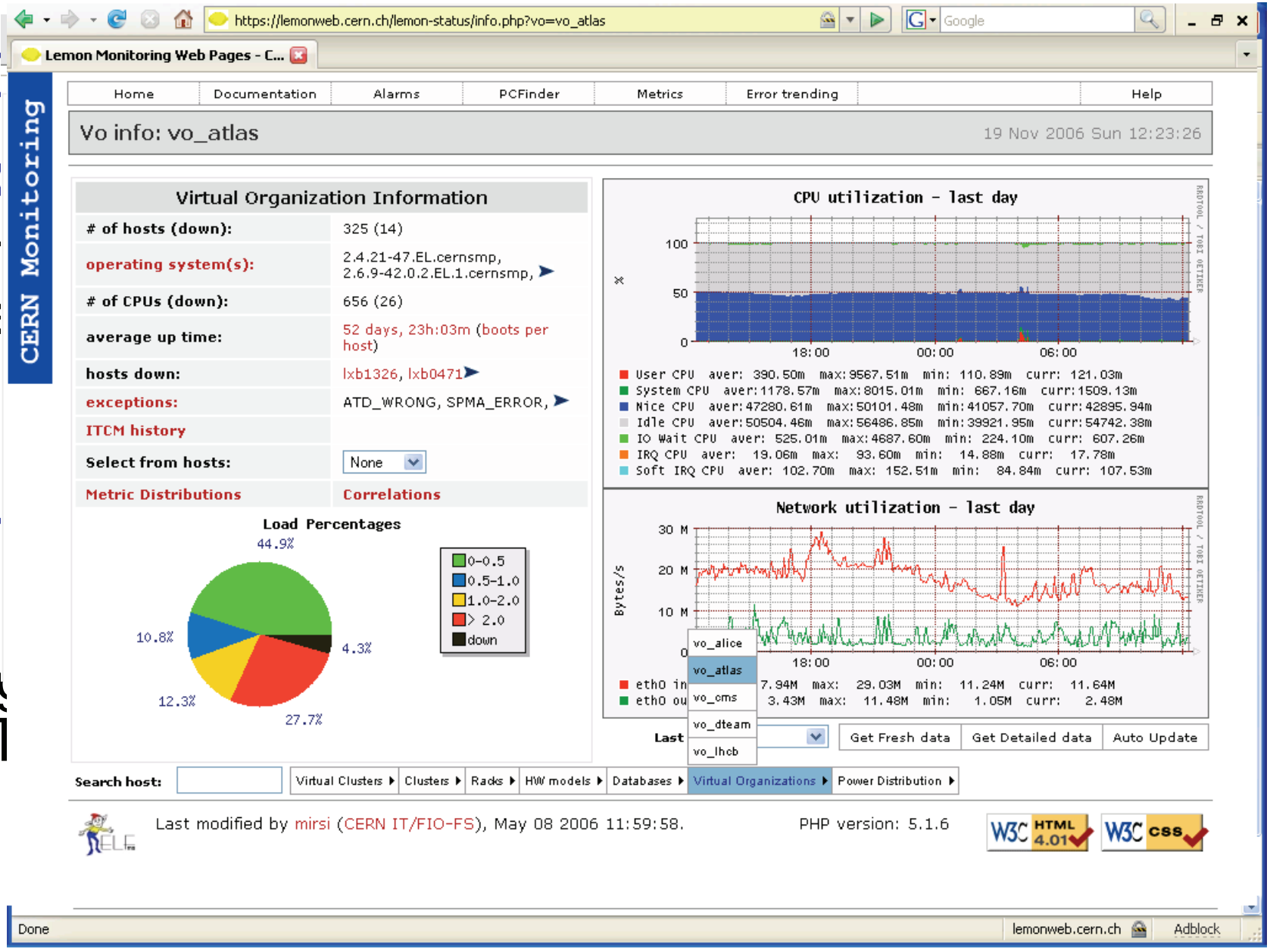
What is monitored

- Node based Lemon sensors cover all the usual system parameters and more
  - system load, file system usage, network traffic, daemon count, software version…
  - SMART monitoring for disks
  - Oracle monitoring
    - number of logons, cursors, logical and physical I/O, user commits, index usage, parse statistics, …
  - AFS client monitoring
  - …
- It is also possible to provide "non-node" sensors. At CERN these allow integration of
  - information from the building management system
    - Power demand, UPS status, temperature, …
      - and full feedback is possible (although not implemented): e.g. system shutdown on power failure
  - high level mass-storage and batch system details
    - Queue lengths, file lifetime on disk, …
  - hardware reliability data

Monitoring displays

Dynamic cluster definition

- As L
  - qua
  - spe
    - Th
      - we
      - in
- Lem
  - fly”
    - e.
      - AT
        -



Outline

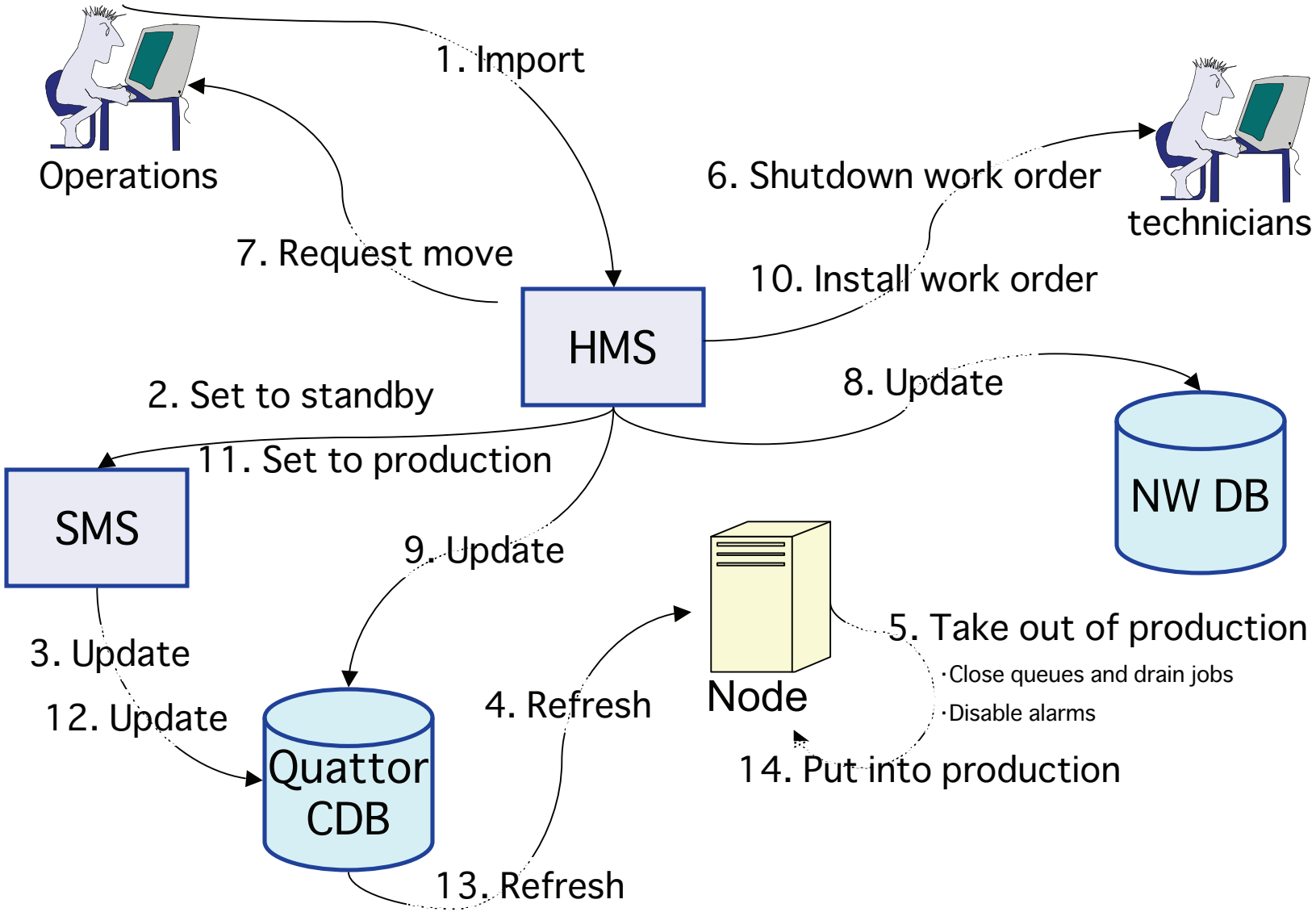
- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
    - Installation & Configuration
    - Monitoring
    - Workflow
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

LHC Era Automated Fabric

LEAF is a collection of workflows for *high level* node hardware and state management, on top of Quattor and LEMON:

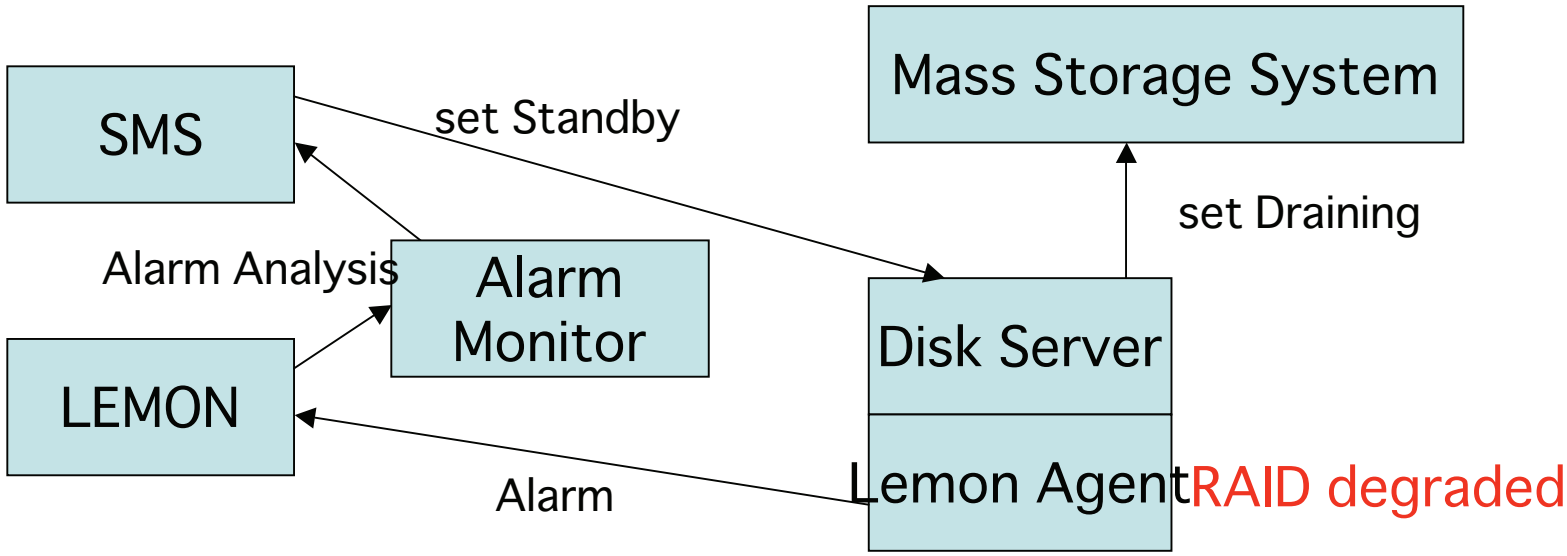
- HMS (Hardware Management System):
  - Track systems through all *physical* steps in lifecycle eg. installation, moves, vendor calls, retirement
  - Automatically requests installs, retires etc. to technicians
  - GUI to locate equipment physically
  - HMS implementation is CERN specific, but concepts and design should be generic
- SMS (State Management System):
  - Automated handling (and tracking of) high-level configuration steps
    - Reconfigure and reboot all LXPLUS nodes for new kernel and/or physical move
    - Drain and reconfig nodes for diagnosis / repair operations
  - Issues all necessary (re)configuration commands via Quattor
  - extensible framework _ plug-ins for site-specific operations possible

LEAF workflow example



Integration in Action

- Simple
  - Operator alarms masked according to system state
- Complex
  - Disk and RAID failures detected on disk storage nodes lead automatically to a reconfiguration of the mass storage system:



Draining: no new connections allowed; existing data transfers continue.

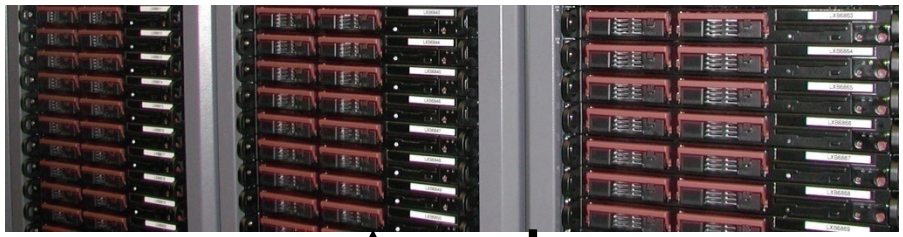
Box Mgmt Successes & Challenges

- At CERN, the ELFms toolkit has allowed us to cope with a significant increase in box count with reduced staffing levels.
- We have confidence the software will scale further
  - although changes needed (e.g. to cope with virtualisation).
- Large scale farm operation, though, remains a challenge!
  - ramp-up, purchasing, h/w failures, …
    - (even if we are not at the Google scale)

Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

Dataflows and rates



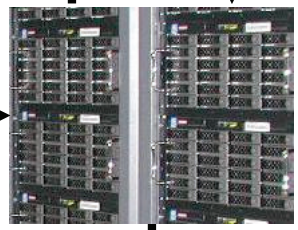
Scheduled work only!

700MB/s

420MB/s

700MB/s

(1600MB/s)



1120MB/s

(2000MB/s)

Averages! Need to be able to support 2x for recovery!

1430MB/s

Remember this figure

Volumes & Rates

- 15PB/year. Peak rate to tape >2GB/s
  - 3 full SL8500 robots/year
- Requirement in first 5 years to reread all past data between runs
  - 60PB in 4 months: 6GB/s
- Can run drives at sustained 80MB/s
  - 75 drives flat out merely for controlled access
- Data Volume has interesting impact on choice of technology
  - Media use is advantageous: high-end technology (3592, T10K) favoured over LTO.

Access Patterns

- Multiple use cases…
  - Sustained transfer to remote site
    - WAN visibility; I/O intensive
  - Rapid transfer of data set to CPU node
    - LAN access; I/O intensive
  - Long running analysis access to data on server
    - LAN access, low I/O, long duration
- …all have different footprint on disk servers

Mass Storage Systems @ CERN

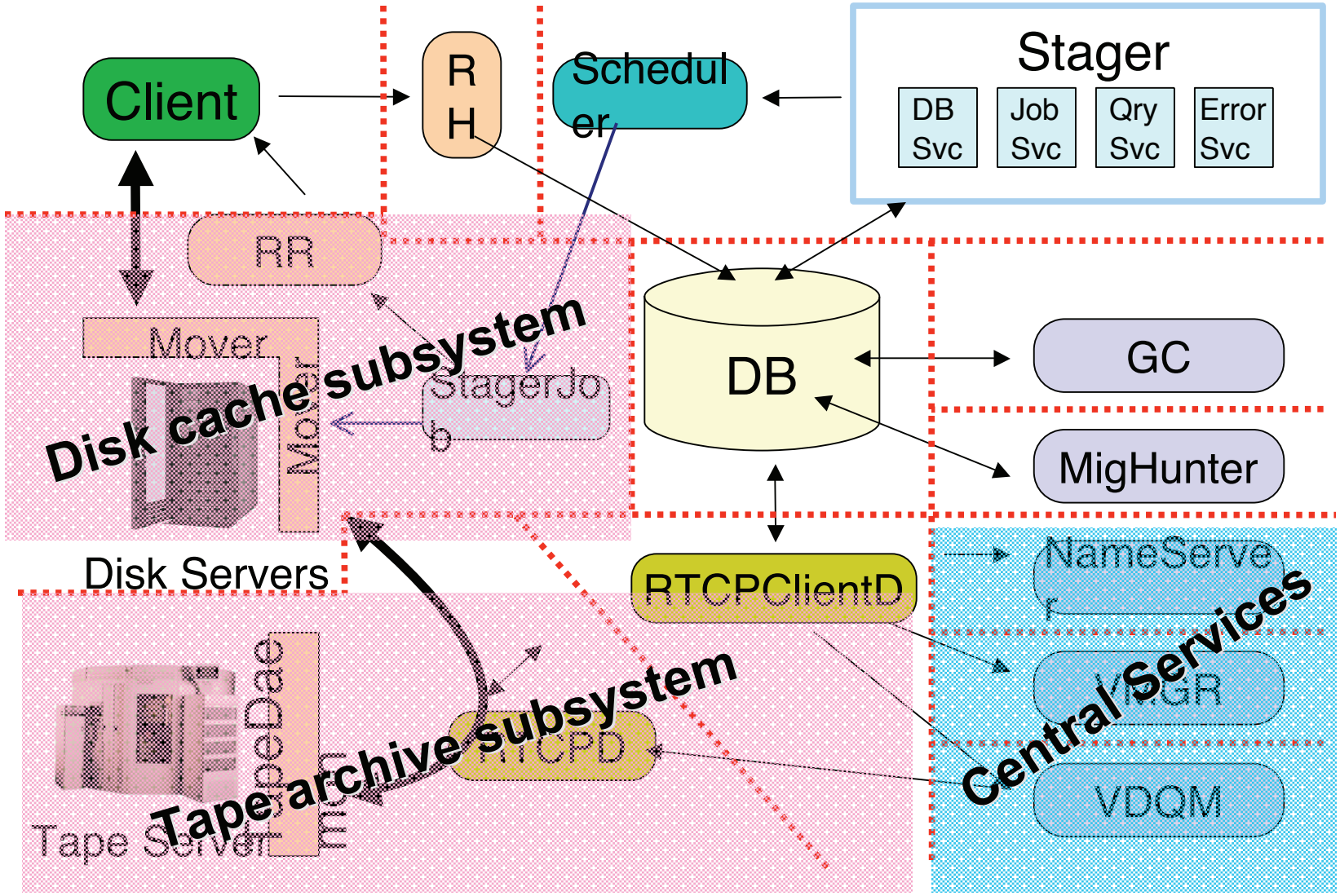
- Commercial Mass Storage Systems have been evaluated at CERN, but with little success. Key systems evaluated were
  - Lachmann/Legent OSM
    - Still in use at DESY, but interest at CERN much reduced due to lack of long-term support (DESY provide their own support)
  - IBM's HPSS
    - In use at SLAC, BNL (US labs) and IN2P3 (French Computer Centre)
    - Experience at CERN showed random access to files (a major use case) was poor; addressing this required additional software and disk buffers
      - At the time, HPSS also required a DCE infrastructure and had limited O/S & hardware support.
  - IEEE "vision" of companies providing pluggable components of an overall system didn't work out in practice; we ended up with single vendors providing all the components…

- … and so CERN developments became more and more capable leading to Castor: CERN Advanced Storage System

Key Castor Features

- Database Centric
  - Stateless agents; can restart easily on error
  - No direct connection from users to critical services
- Scheduled Access to I/O
  - No overloading of disk servers
    - Per-server limit set according to type of transfer
      - servers can support many random access style accesses, but only a few sustained data transfers
  - I/O requests can be scheduled according to priority
    - Fair shares access to I/O just as for CPU
    - Prioritise requests from privileged users

Castor Architecture



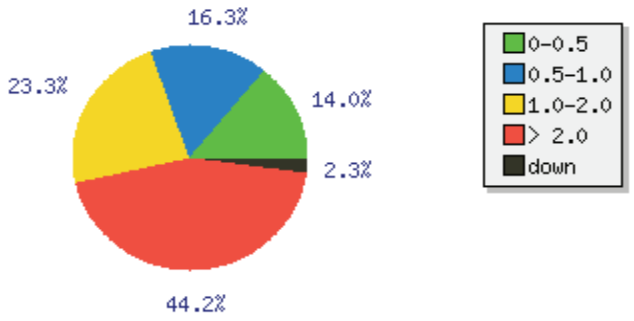
Detailed view

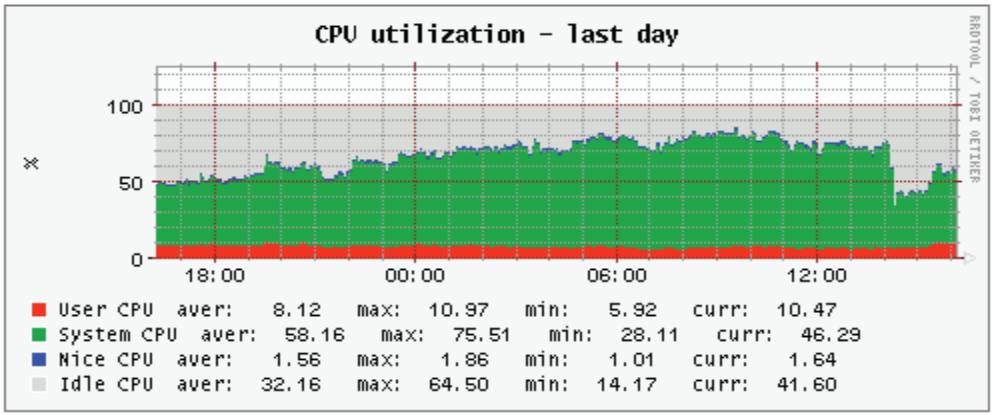
Castor Performance

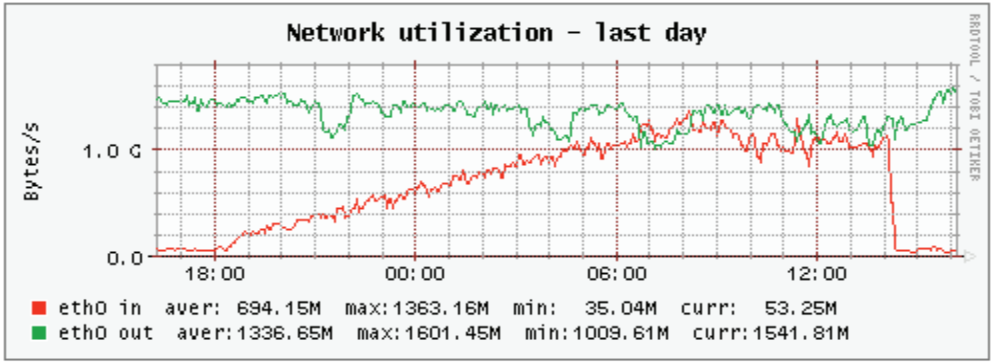
Cluster info: c2sc4 subcluster wan — 19 Apr 2006 Wed 16:14:27

Cluster Information	
# of hosts (down):	43 (1)
operating system(s):	2.6.9-34.EL.cernsmp
# of CPUs (down):	51 (2)
average up time:	20 days, 18h:50m (boots per host)
hosts down:	lxfsra3004
exceptions:	RPC_STATD_WRONG, MIRROR_BROKEN,
ITCM history	View template
Select from hosts:	None
Metric Distributions	Correlations

Load Percentages



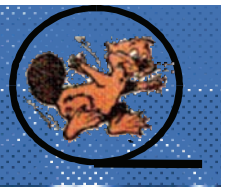




Last day | Get Fresh data | Get Detailed data | Auto Update

Sustained transfer from disk of 1.2GB/s as data import ramps

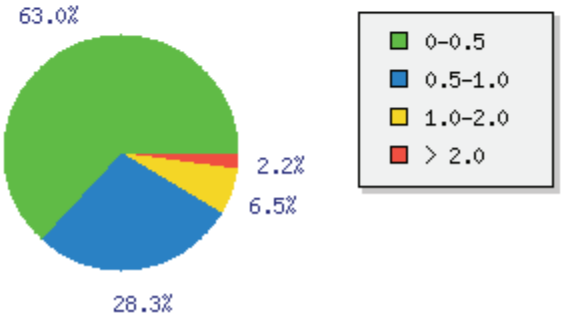
Castor Performance

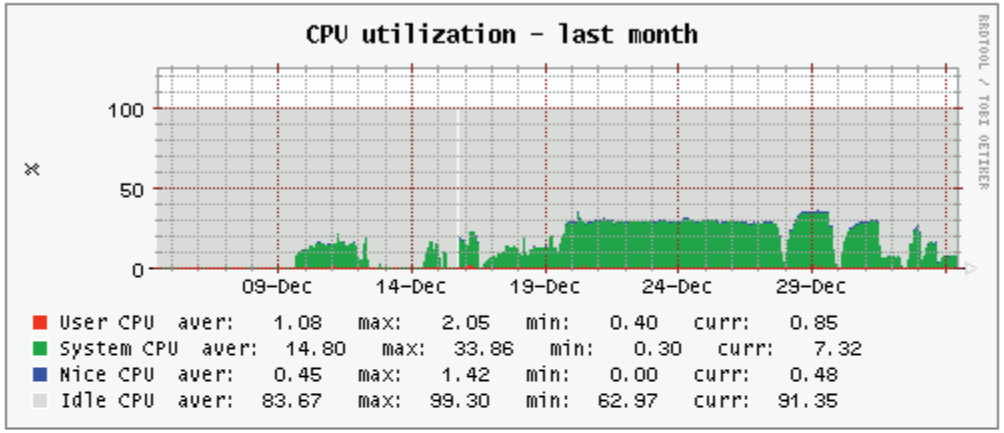


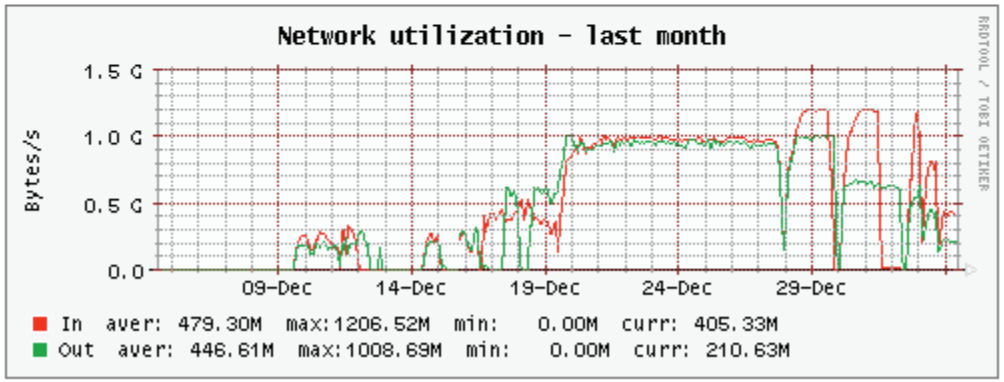
Cluster info: ITDC — 03 Jan 2006 Tue 10:19:12

Cluster Information	
# of hosts (down):	46 (0)
operating system(s):	2.4.21-37.EL.cernsmp
# of CPUs (down):	62 (0)
average up time:	47 days, 14h:14m (boots per host)
hosts down:	none
exceptions:	FILESYSTEM_ERROR
ITCM history	View template
Select from hosts:	None
Metric Distributions	Correlations

Load Percentages







Last month | Get Fresh data | Get Detailed data | Auto Update

Sustained transfer of incoming data to tape at 1GB/s

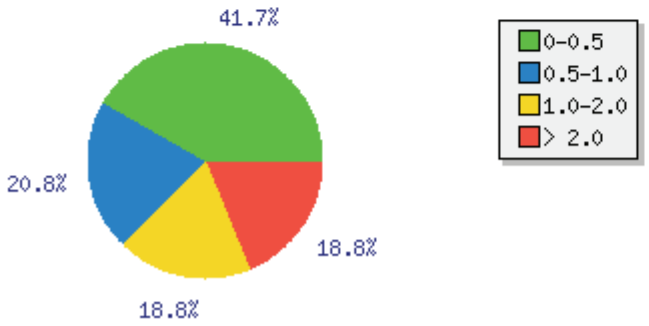
Note the dates! Failed hardware was left down.

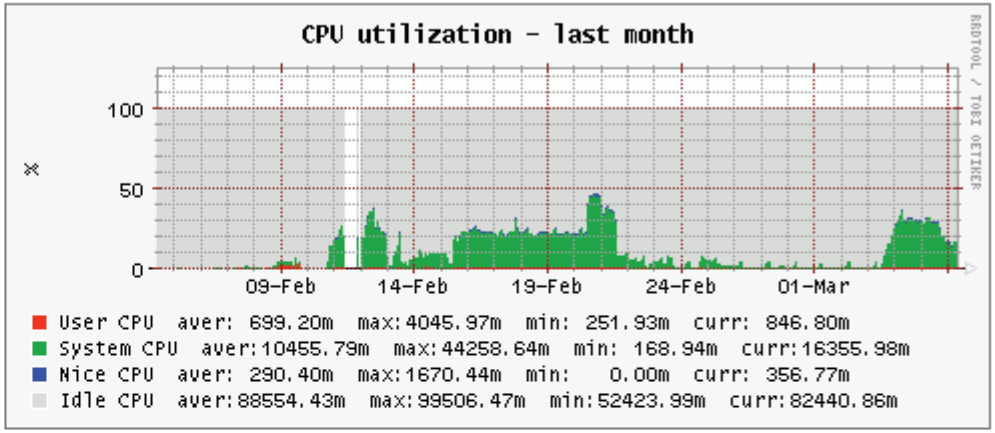
Castor Performance

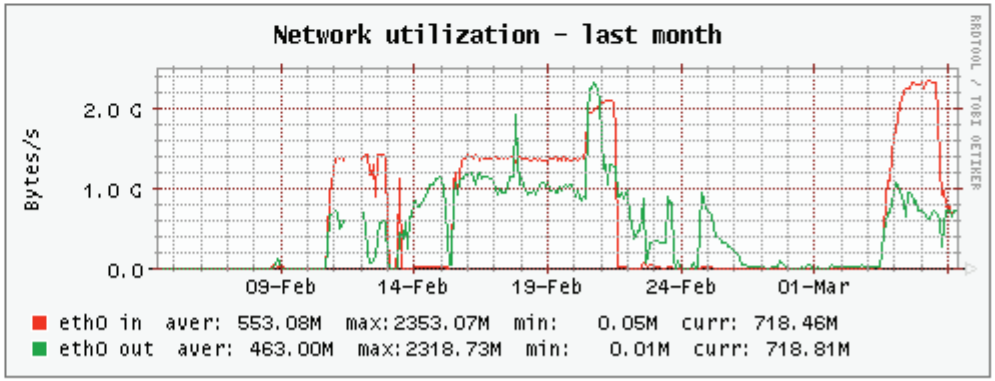
Cluster info: castor2 subcluster ITDC — 06 Mar 2006 Mon 08:15:45

Cluster Information	
# of hosts (down):	48 (0)
operating system(s):	2.4.21-37.EL.cernsmp, 2.4.21-37.0.1.EL.cernsmp
# of CPUs (down):	88 (0)
average up time:	38 days, 7h:43m (boots per host)
hosts down:	none
ITCM history	View template
Select from hosts:	None
Metric Distributions	Correlations

Load Percentages



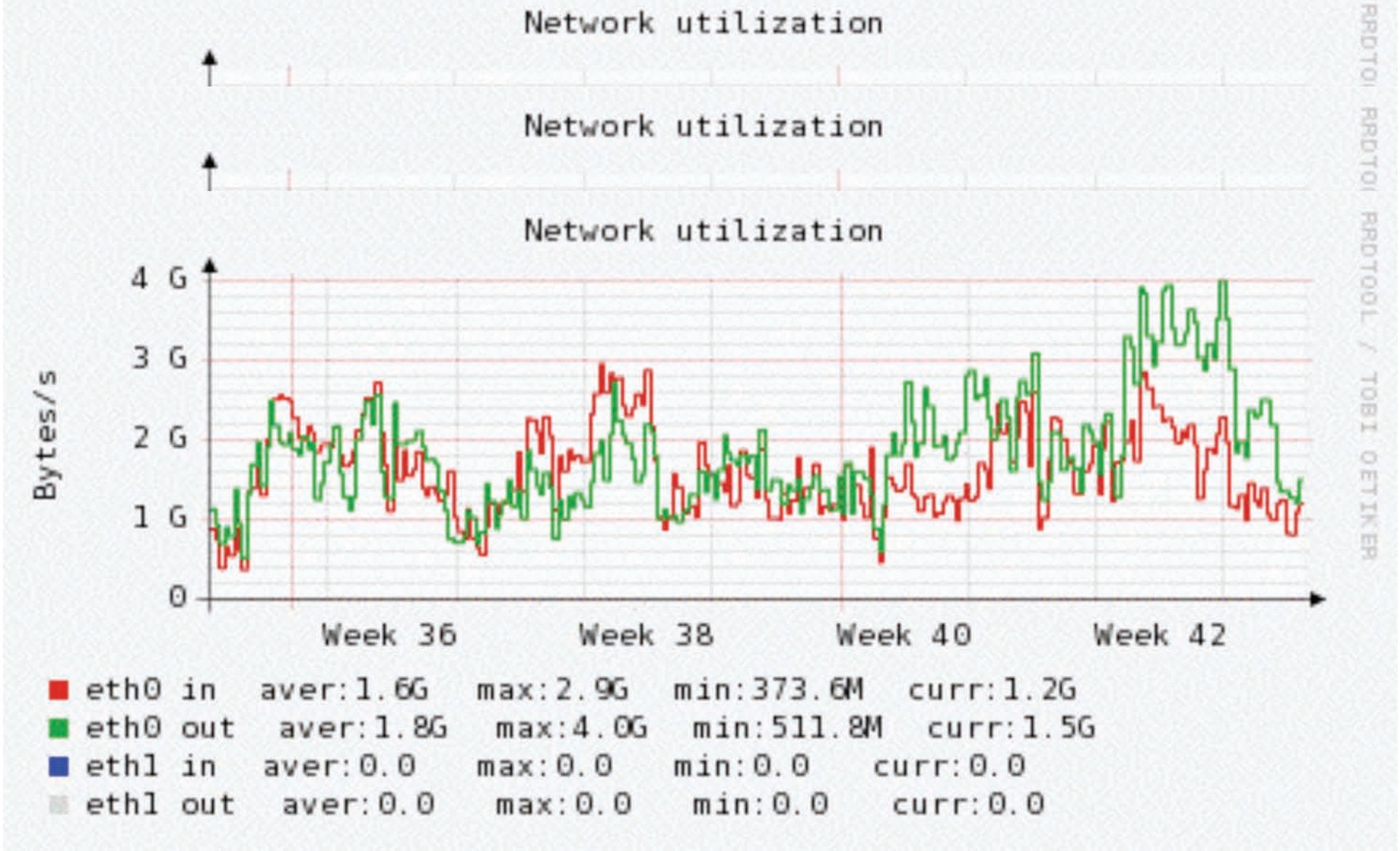




Last month | Get Fresh data | Get Detailed data | Auto Update

Peak transfer of incoming data to tape at over 2GB/s

Castor Performance



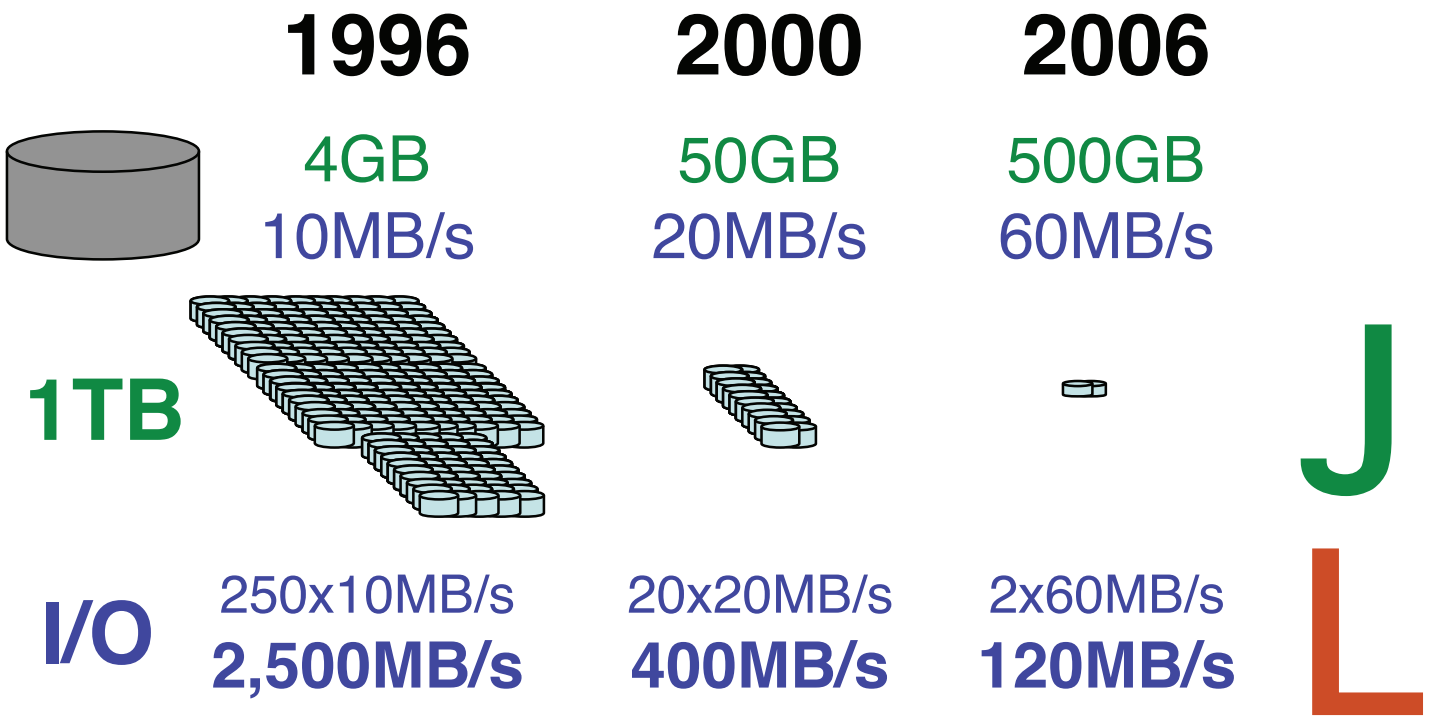
More Data Challenges

- Long data lifetime
- Disk capacity vs I/O rates
- File sizes
- Multiple Mass Storage Systems
- Organised Data Export

Long lifetime

- LEP, CERN's last accelerator, started in 1989 and shutdown 10 years later.
  - First data recorded to IBM 3480s; at least 4 different technologies used over the period.
  - All data ever taken, right back to 1989, was reprocessed and reanalysed in 2001/2.
- LHC starts in 2007 and will run until at least 2020.
  - What technologies will be in use in 2022 for the final LHC reprocessing and reanalysis?
- Data repacking required every 2-3 years.
  - Time consuming
  - Data integrity must be maintained

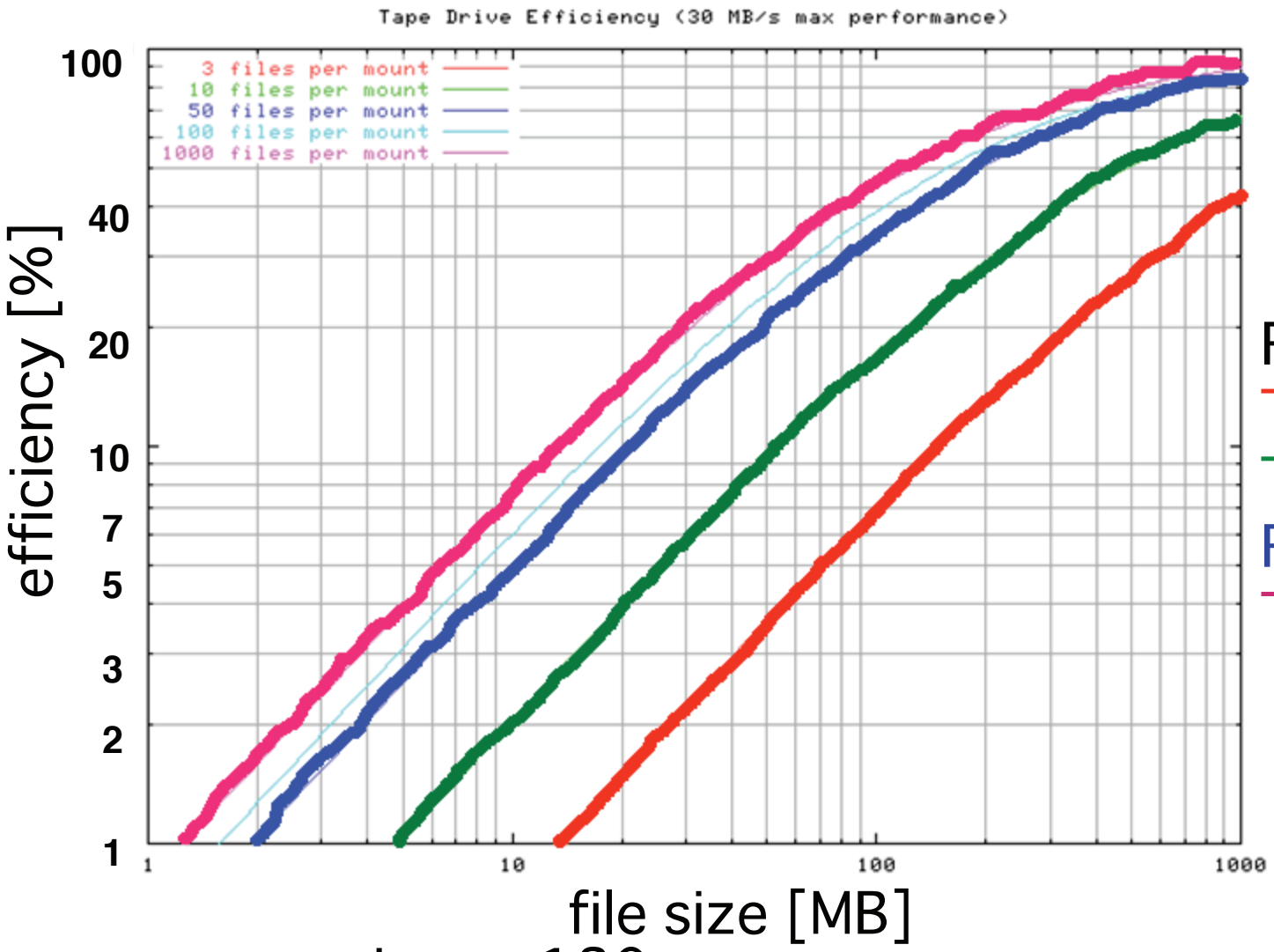
Disk capacity & I/O rates



CERN now purchases two different storage server models: capacity oriented and throughput oriented.

- fragmentation increases management complexity
- (purchase overhead also increased…)

(File) Size Matters!



Files per mount
Three
Ten
Fifty
Thousand

tape mount time ~120s
file overhead ~5s (400MB)

Multiple Mass Storage

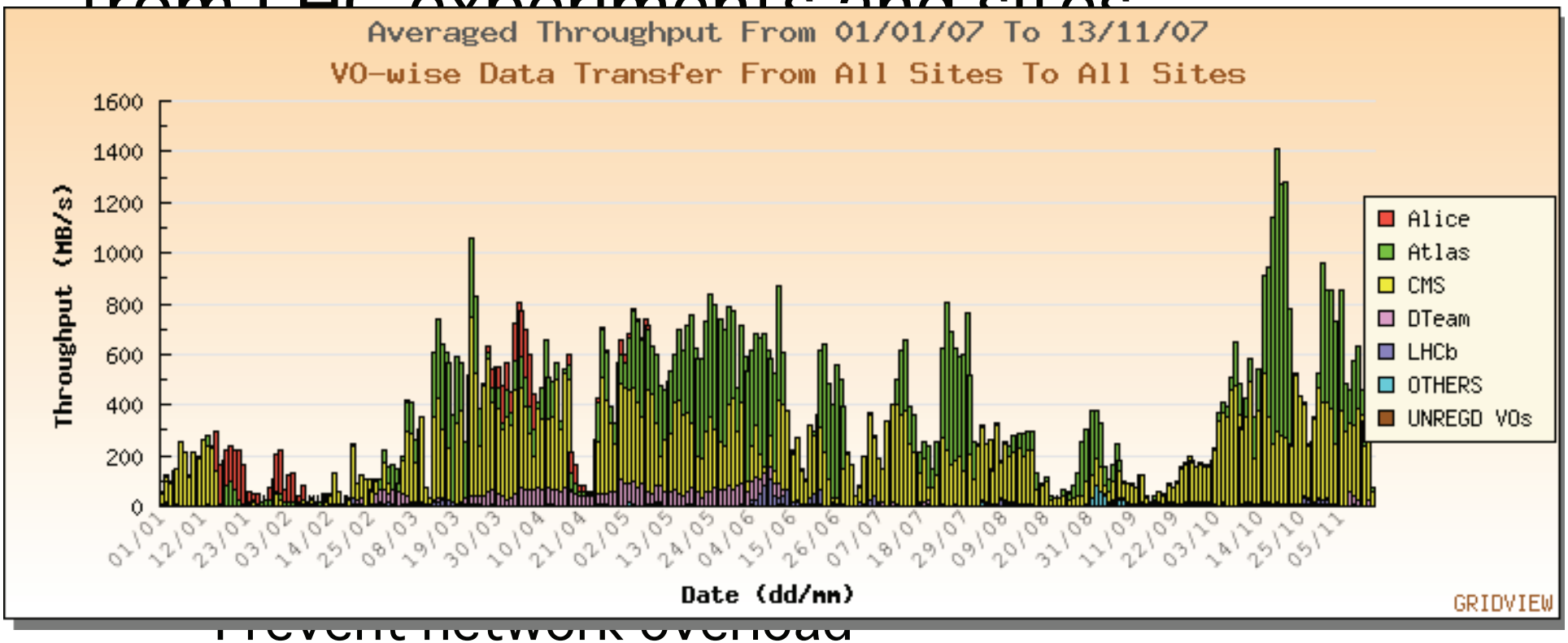
- CASTOR is not the sole MSS for LHC
  - Fermilab's dCache is used at many sites; DPM, a disk-only storage manager is also common.
- Users, of course, don't want to know…
  - … and experiment code needs to run at many sites
  - SRM, the Storage Resource Manager, provides a common interface layer to the various mass storage systems
    - See http://sdm.lbl.gov/srm-wg/
  - These multiple and independent implementations of the interface **all talk to each other.**
  - Key element to have successful…

Organised Data Export

- LHC experiments need to ship data between sites
  - Raw data export, Analysis Data updates, Monte Carlo data import, …
- This is complicated at our scale
  - with petabytes of data transferred, a 0.1% failure rate can't be easily rescued or followed up manually
  - Sites policies (e.g. fraction of resources allocated to a given VO) must be respected
  - …

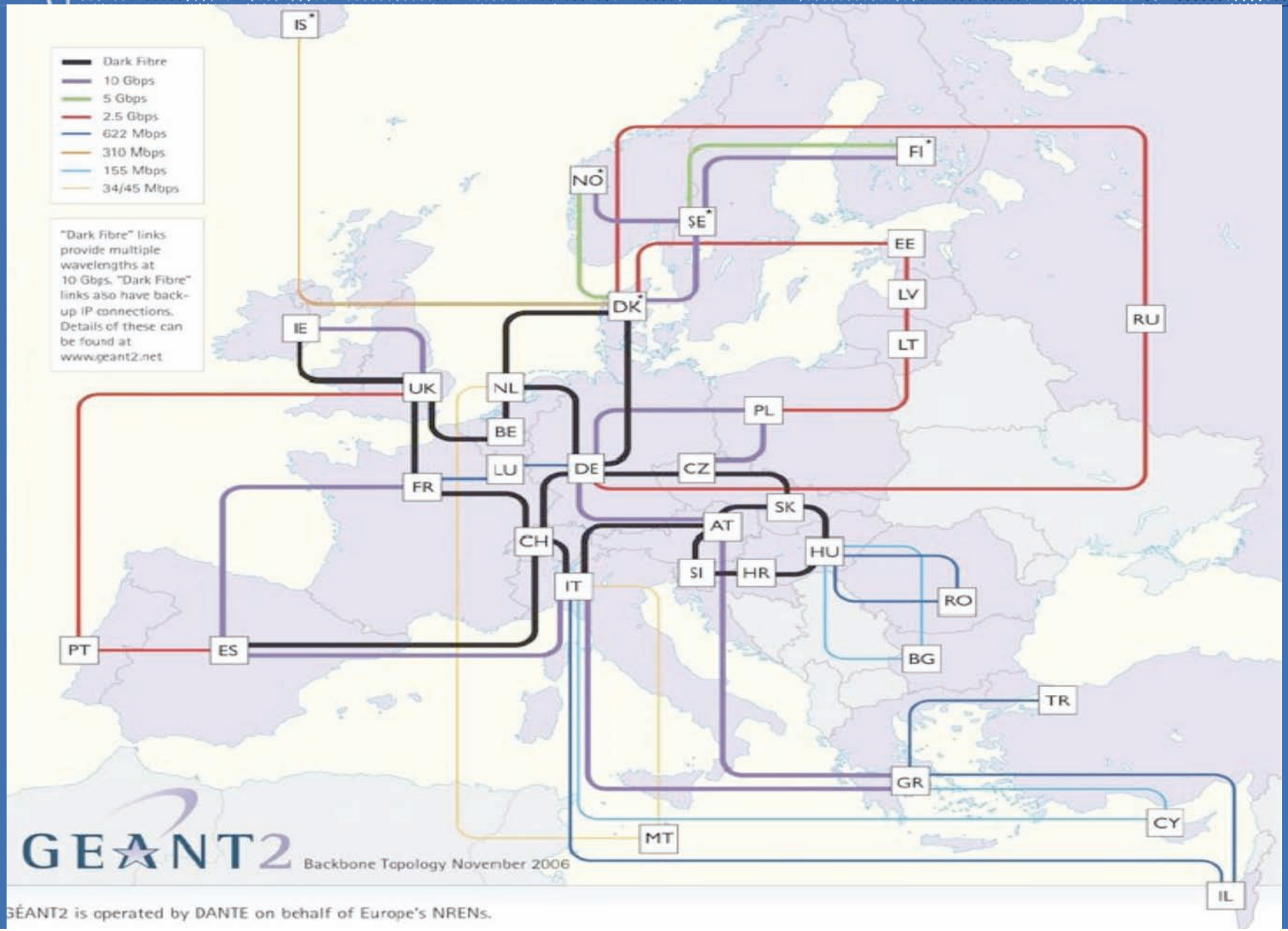
The File Transfer Service

- Developed as part of the EGEE data management activity to meet requirements from LHC experiments and sites



- Prevent network overload
  - *(But no management on the network level)*
- Prevent storage overload

… and we needed a network



Data Successes & Challenges

- Successes:
  - We have an advanced Mass Storage System at CERN able to meet the demanding requirements for Data Acquisition and export.
  - Large scale data transfers between sites are becoming routine.
- But yet to demonstrate
  - exports for multiple experiments simultaneously
  - operations for large scale user analysis
    - most work so far has been controlled "production"

Outline

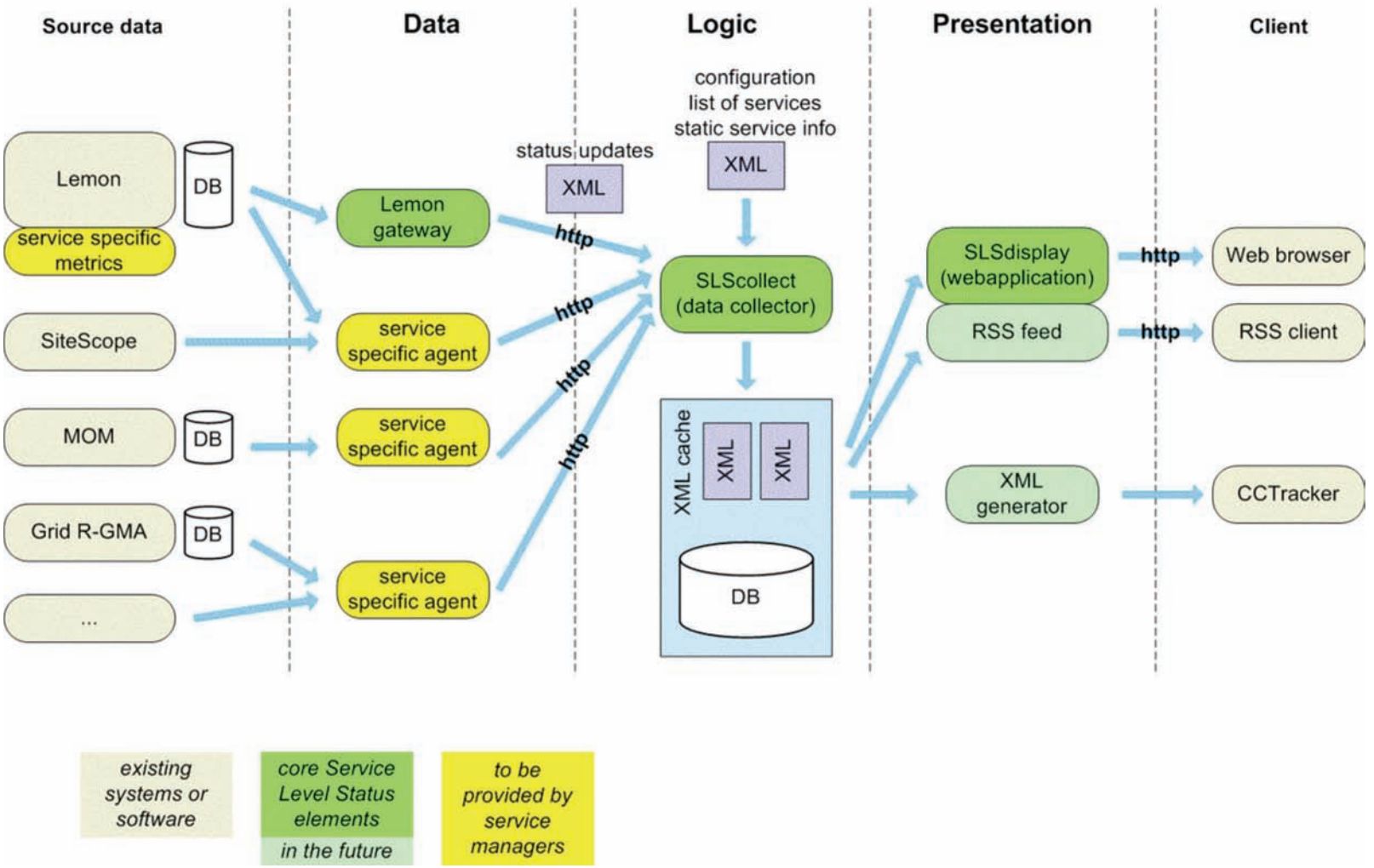
- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

A Complex Overall Service

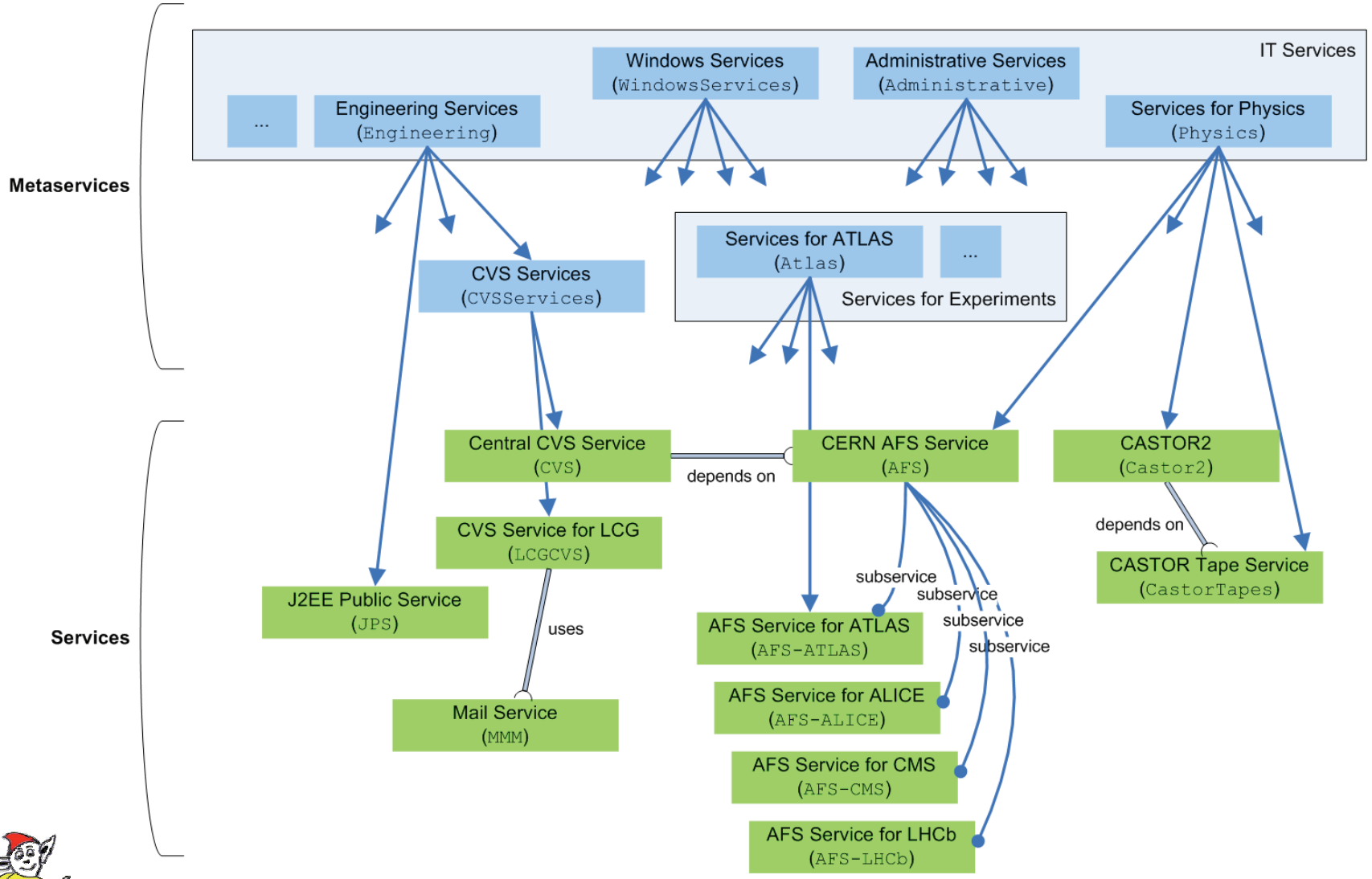
- Site managers understand systems (we hope!).
- But do they understand the service?
  - and do the users?
  - and what about cross site issues?
    - Are things working?
    - If not, just where is the problem?
      - how many different software components, systems and network service providers are involved in a data transfer site X to site Y?

User Status Views @ CERN

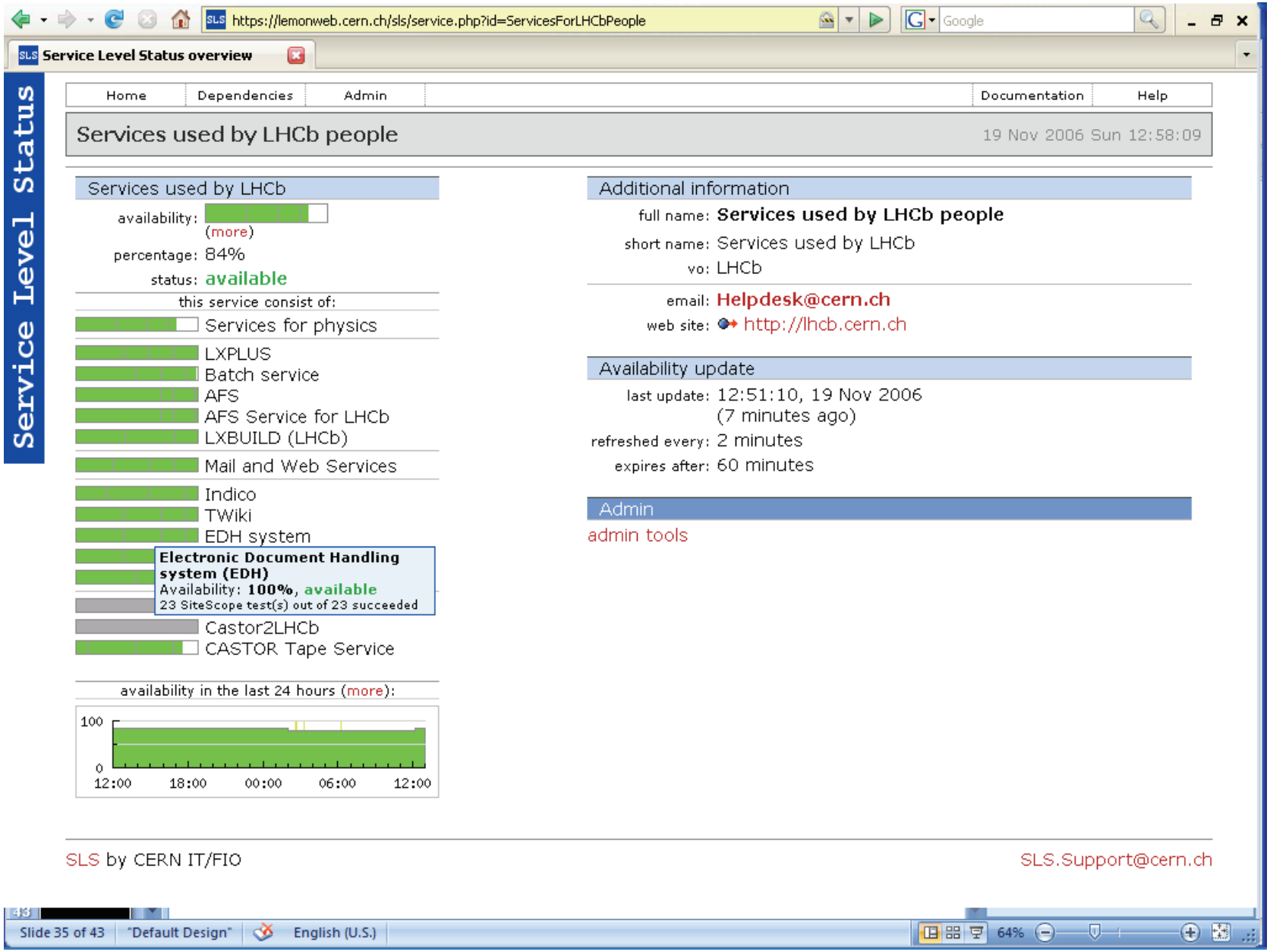
SLS Architecture



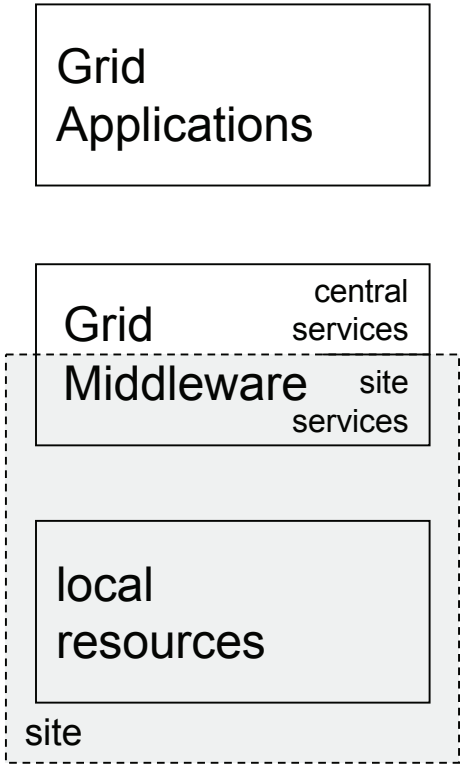
SLS Service Hierarchy



SLS Service Hierarchy

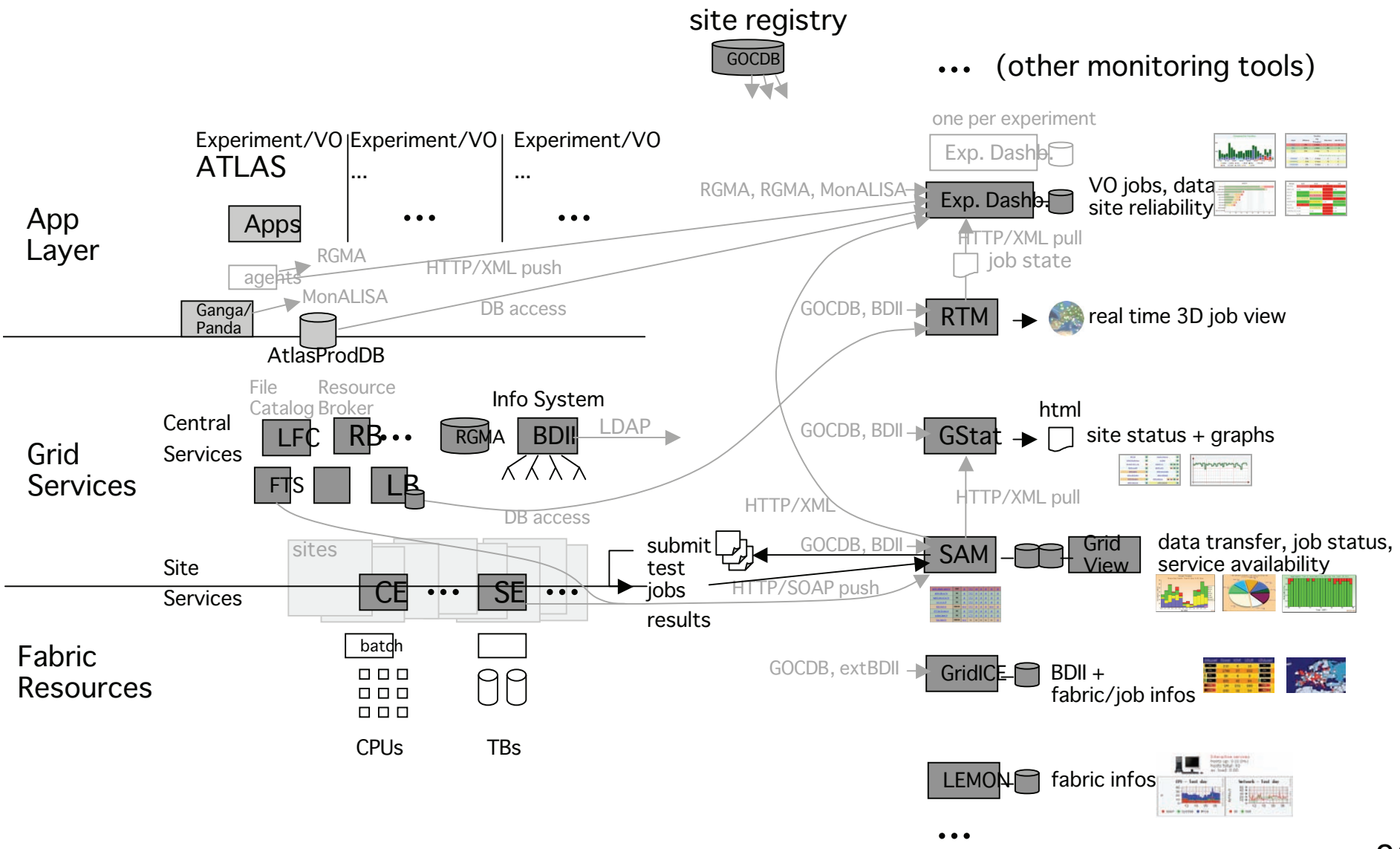


WLCG Grid Monitoring



Domain	Monitoring Tools in use
Application monitoring	Experiment Dashboards ...
Grid Services monitoring	GStat SAM/GridView GridICE GridPP Real Time Monitor ...
Local monitoring	Lemon/SLS Nagios Ganglia ...

Grid Monitoring Landscape



New visualizations for the

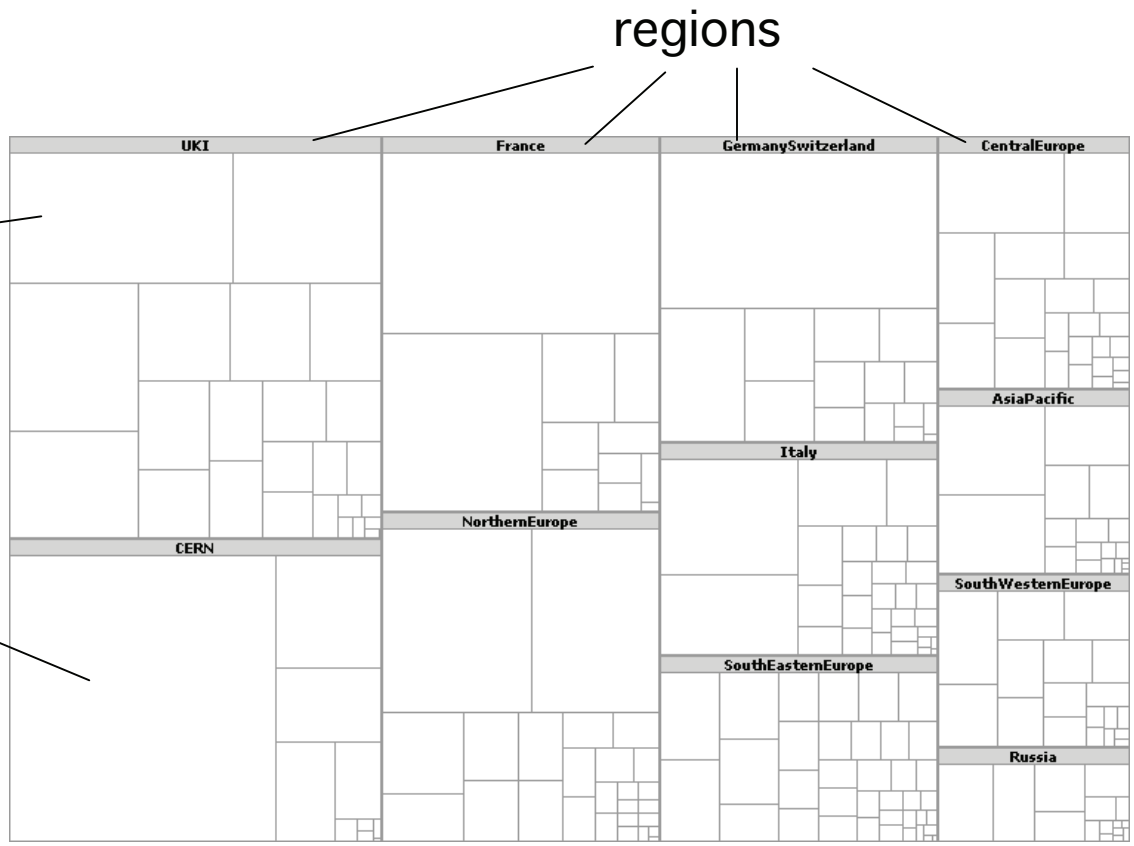
- Grid monitoring data is complex!
  - And there are many sites…
- Current tools visualize data by sorted tables, bar charts, etc.
- Difficult to present an easy to understand **top-level view** which provides
  - quick, action oriented oversight and insight
  - help understand job failures and availability patterns

Can new visualizations help?

GridMap Visualization

- Idea
  - visualize the Grid by using *Treemaps*
    (Grid + Treemap = *GridMap*)

- Example *GridMap*

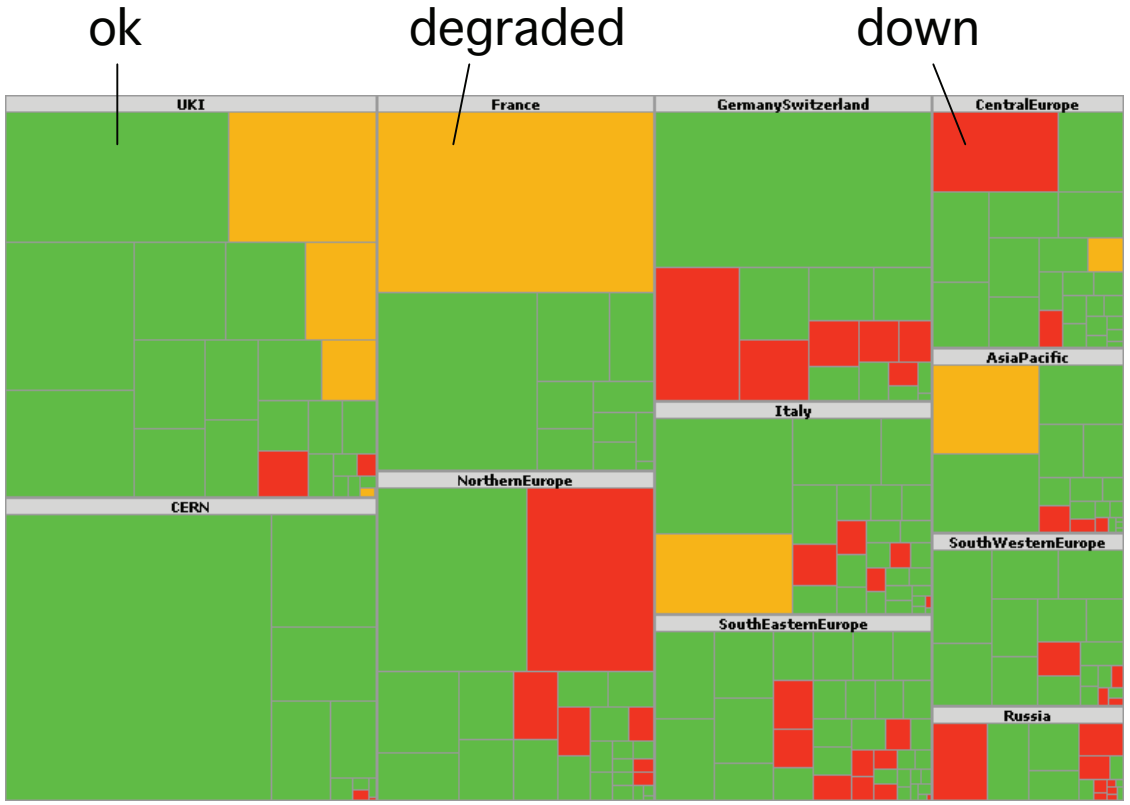


Size of rectangle is e.g.
- size of site (#CPUs)
- #running jobs
- ...

GridMap Visualization

- Idea
  - visualize the Grid by using *Treemaps*
    (Grid + Treemap = *GridMap*)

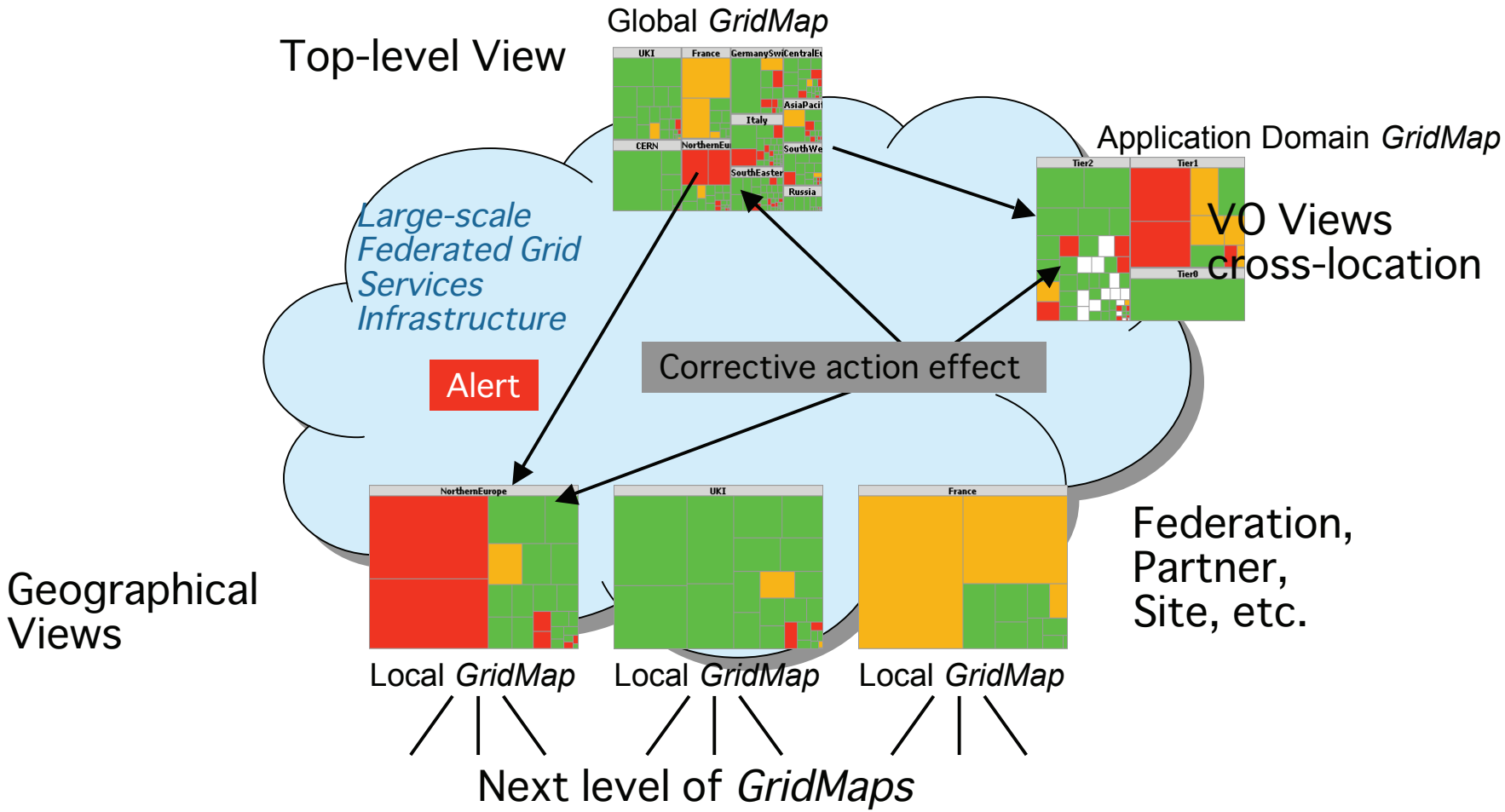
- Example *GridMap*



Colour of rectangle is e.g.
- SAM status of site / service
- Availability of site / service
- ...

Multiple Views

- *GridMaps* can be used for *top-level*, *geographical* and *VO* views



Trends

- Trends can be understood by looking at a sequence of *GridMaps*

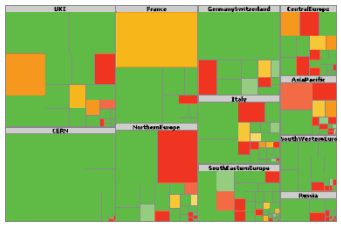
Site Availability over time:

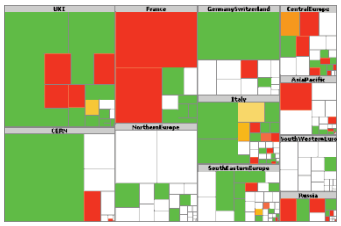
More Views

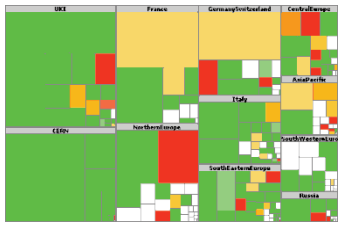
- Correlations of metrics can be discovered by switching between different views

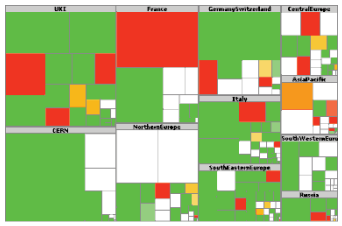
Site Availability from different VO perspectives:

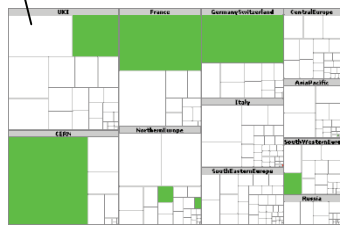
sites without colour do not support the VO

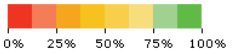

OPS


Alice

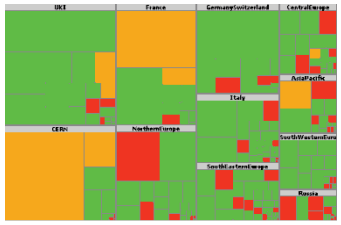

Atlas

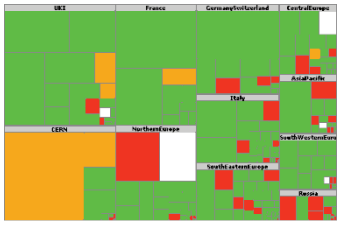

CMS

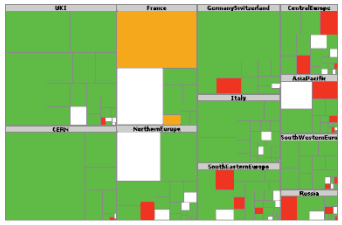

LHCb

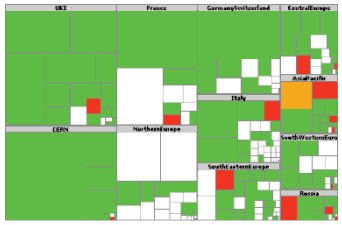


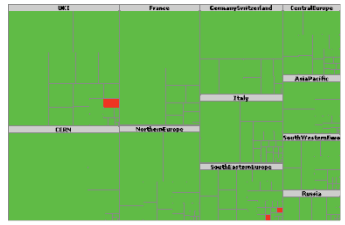
Status of different Site Services:

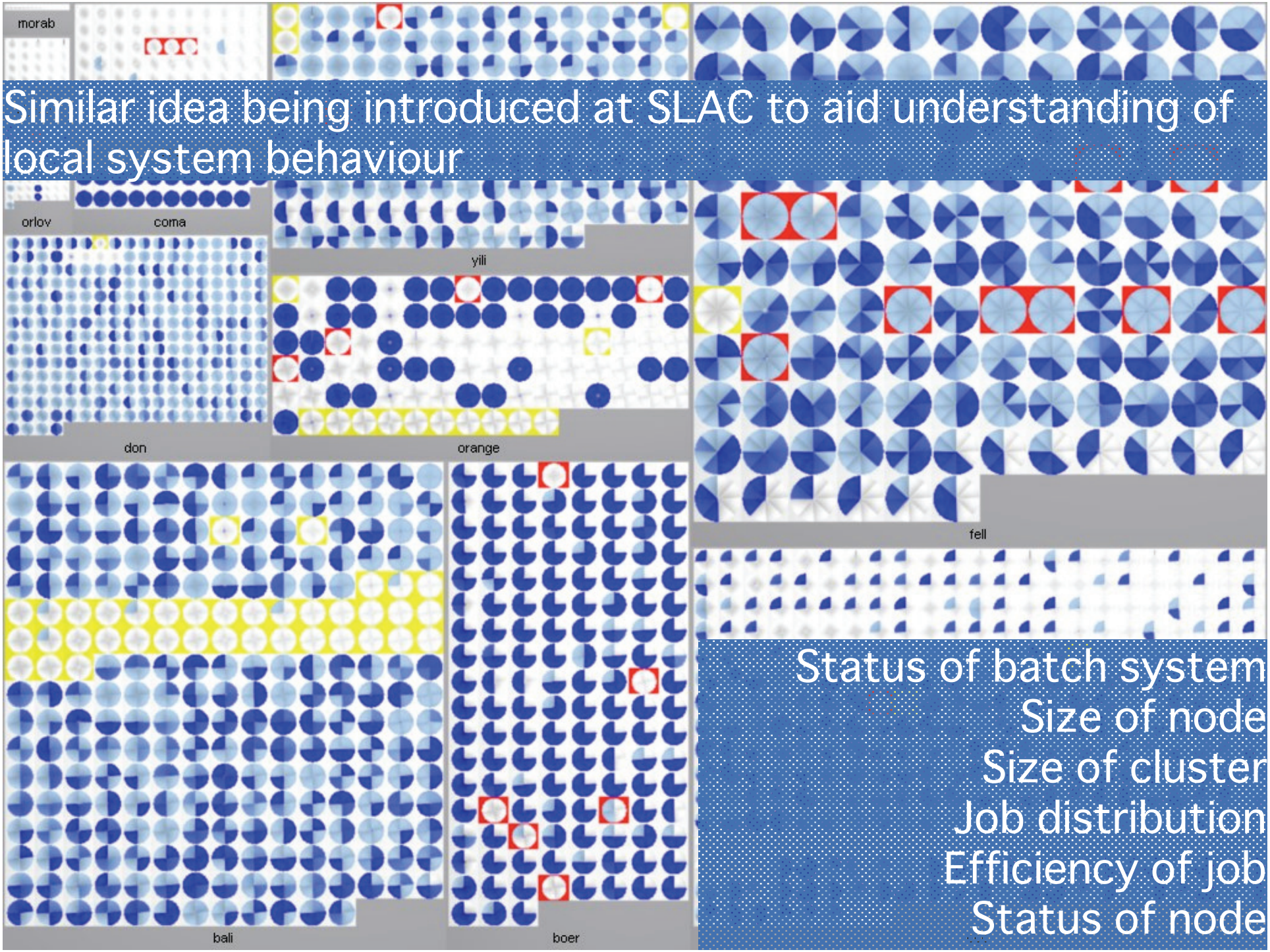

Overall Site


CE

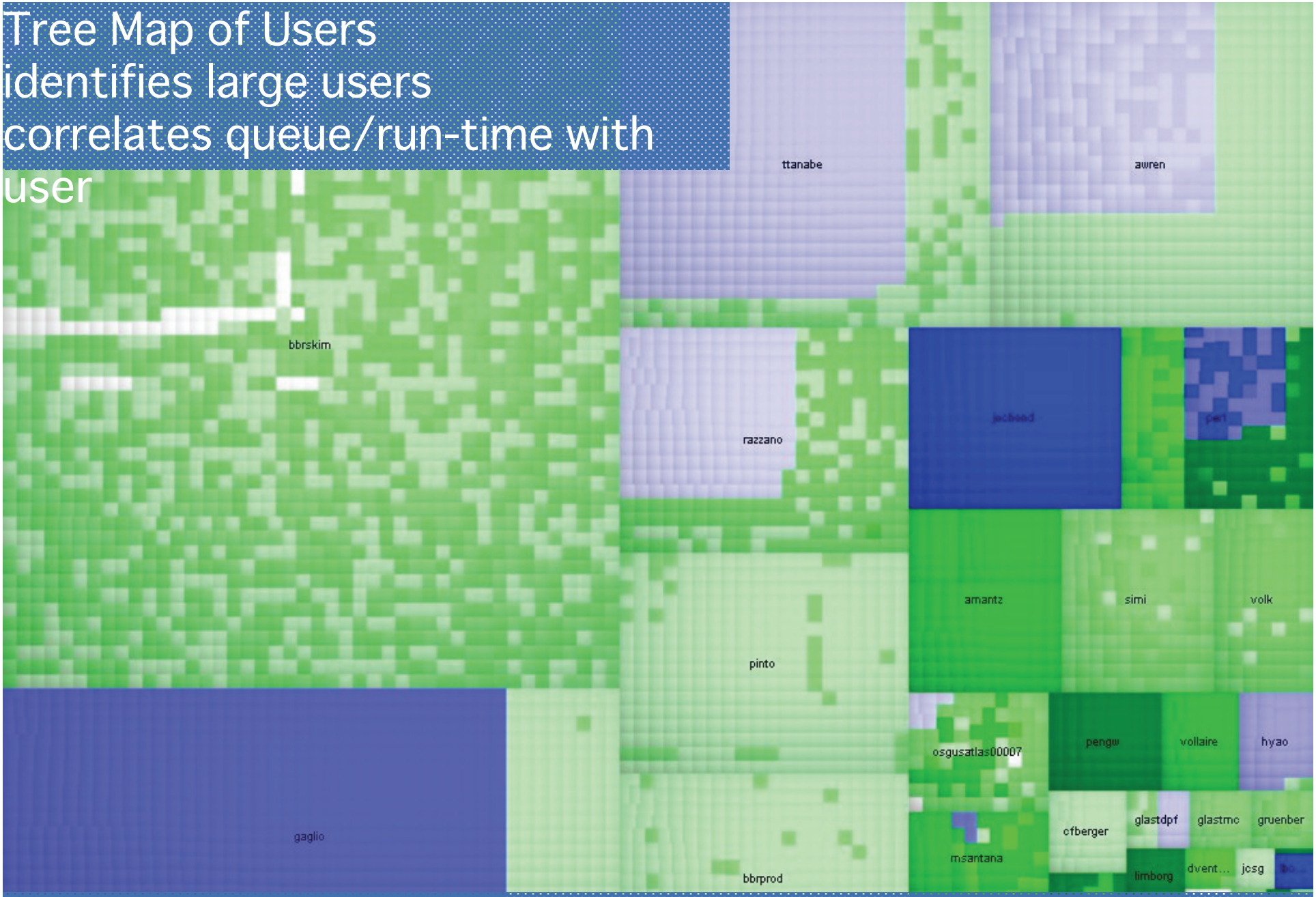

SE


SRM


site BDII

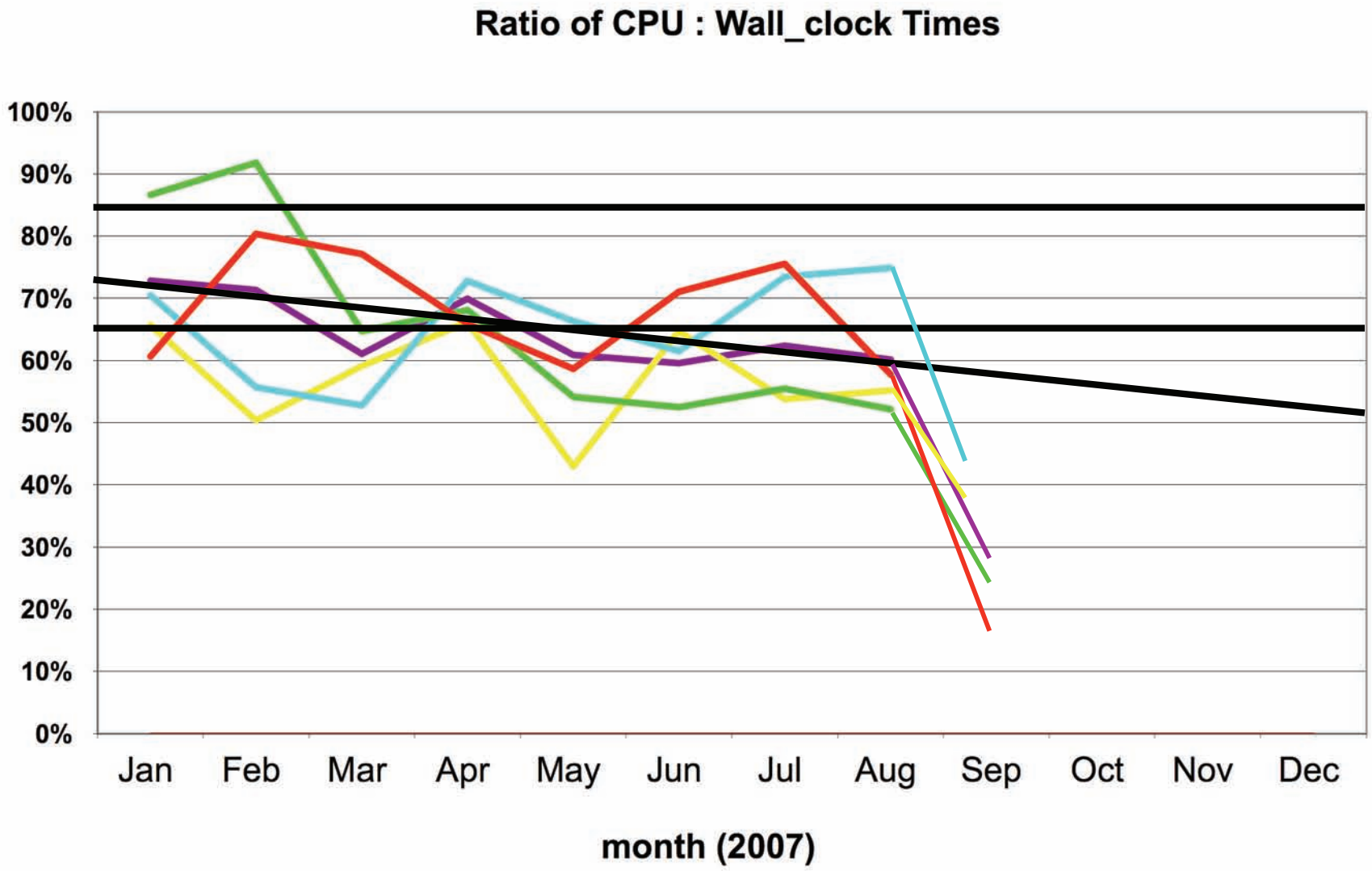

morab
Similar idea being introduced at SLAC to aid understanding of local system behaviour
orlov
coma
yili
don
orange
fell
bali
boer
Status of batch system
Size of node
Size of cluster
Job distribution
Efficiency of job
Status of node

Tree Map of Users identifies large users correlates queue/run-time with user



Uses Magnaview, a commercial data exploration tool (www.magnaview.nl)
See also

So, how to solve this?



Outline

- Introduction to CERN and Experiments
- LHC Computing
- Challenges
  - Capacity Provision
  - Box Management
  - Data Management and Distribution
  - What's Going On?
- Summary/Conclusion

Summary

- Immense Challenges & Complexity
  - Data rates, developing software, lack of standards, worldwide collaboration, …
- Considerable Progress in last ~5-6 years
  - WLCG service exists
  - Petabytes of data transferred
  - …
- But real data is nearly here…
  - Will the system cope with chaotic analysis?
  - Will we understand the system enough to identify problems—and fix underlying causes?
  - Major "Dress Rehearsals" in Feb & May 2008
    - last chance to shake system down before operation
- Answer(s) at LISA '08?

Thank You!