

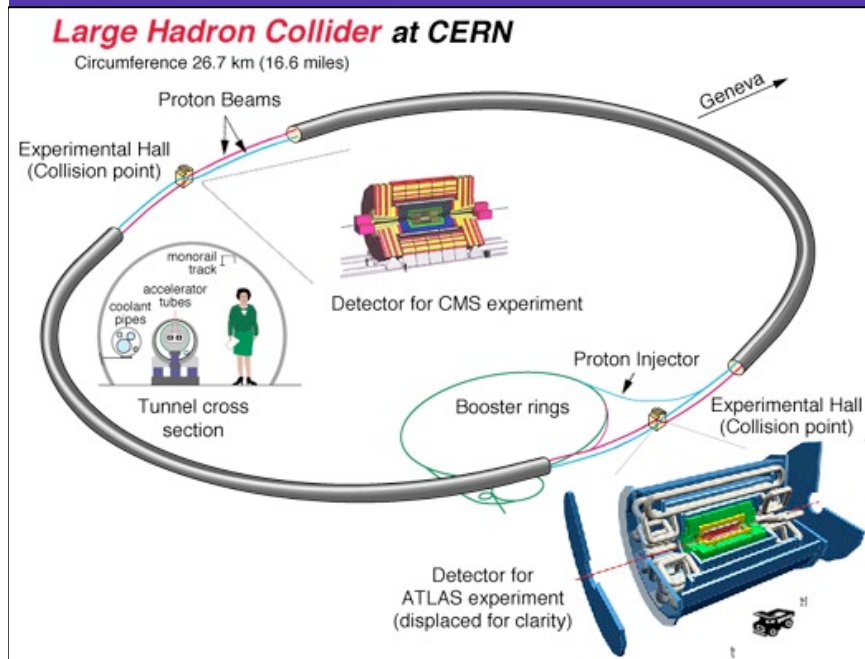
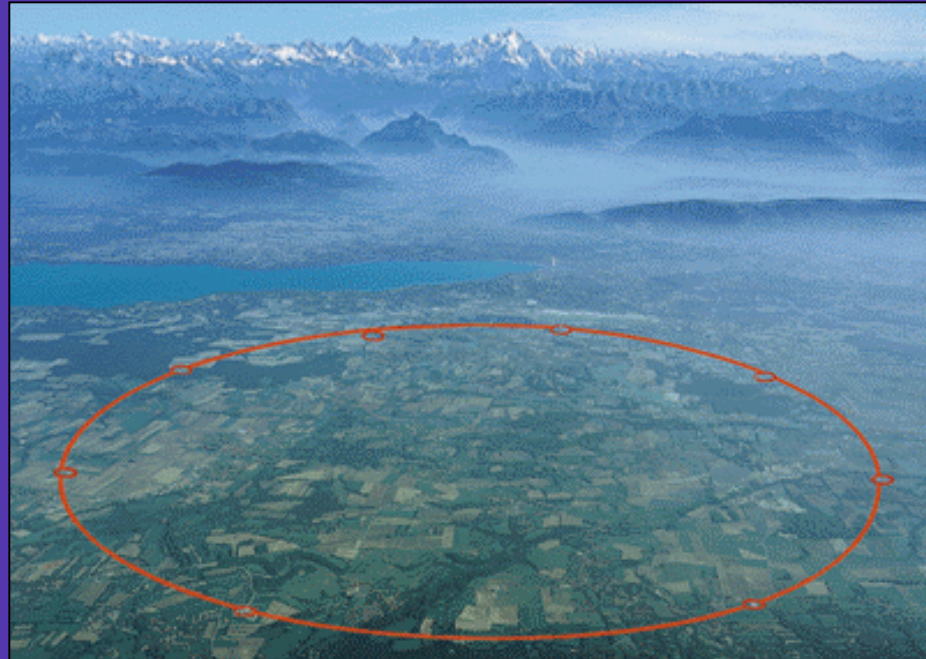
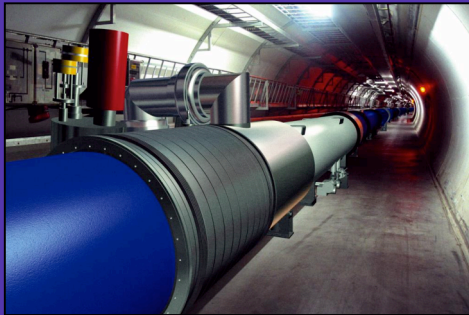
LHC Computing

Nick Brook

- The LHC & experiments
 - Requirements
 - Computing models
- Experiences so far
 - Interoperability
- LCG Baseline service group
 - Future requirements
- Summary

The CERN LHC

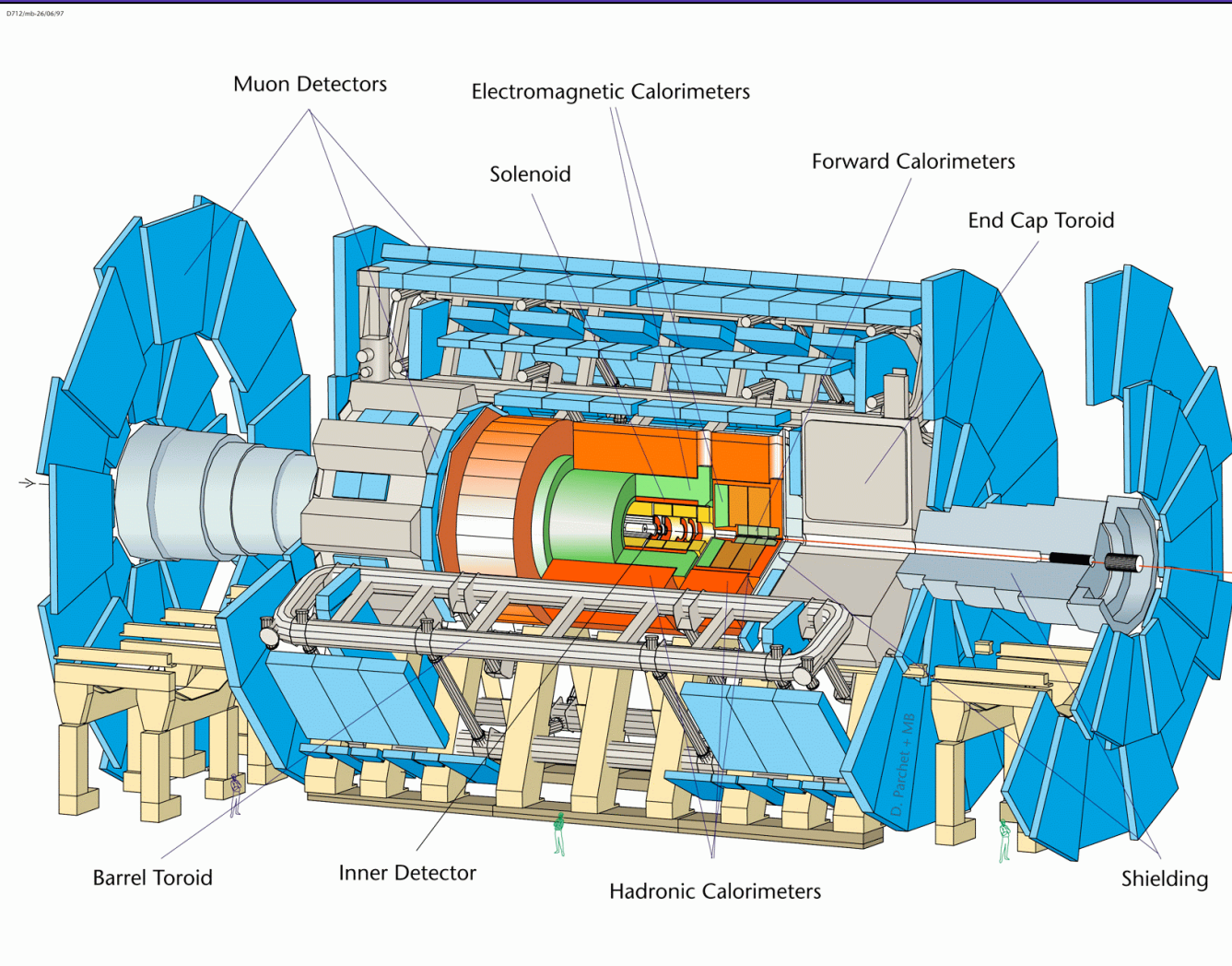
The world's most powerful particle accelerator
First (proton-proton) collisions due in 2007



4 Large Experiments



ATLAS Detector



7,000 tonnes
42m long
22m wide
22m high

(About the
height of a 5
storey
building)

2,000 Physicists
150 Institutes
34 Countries

LHC Physics Goals

What is mass ?

particles acquire their masses by interacting with another particle, the Higgs Boson

Is there supersymmetry ?

links the matter particles (the quarks and leptons) with the force particles (the gauge bosons) - "grand unified theory"

What is Dark Matter?

The discovery of supersymmetric particles could have important implications for cosmology

Where has all the antimatter gone?

very early moments after the Big Bang the universe should have contained equal amounts of matter and antimatter but the universe we see around us is made up almost entirely of matter

Why are there three "generations" of quarks and leptons?

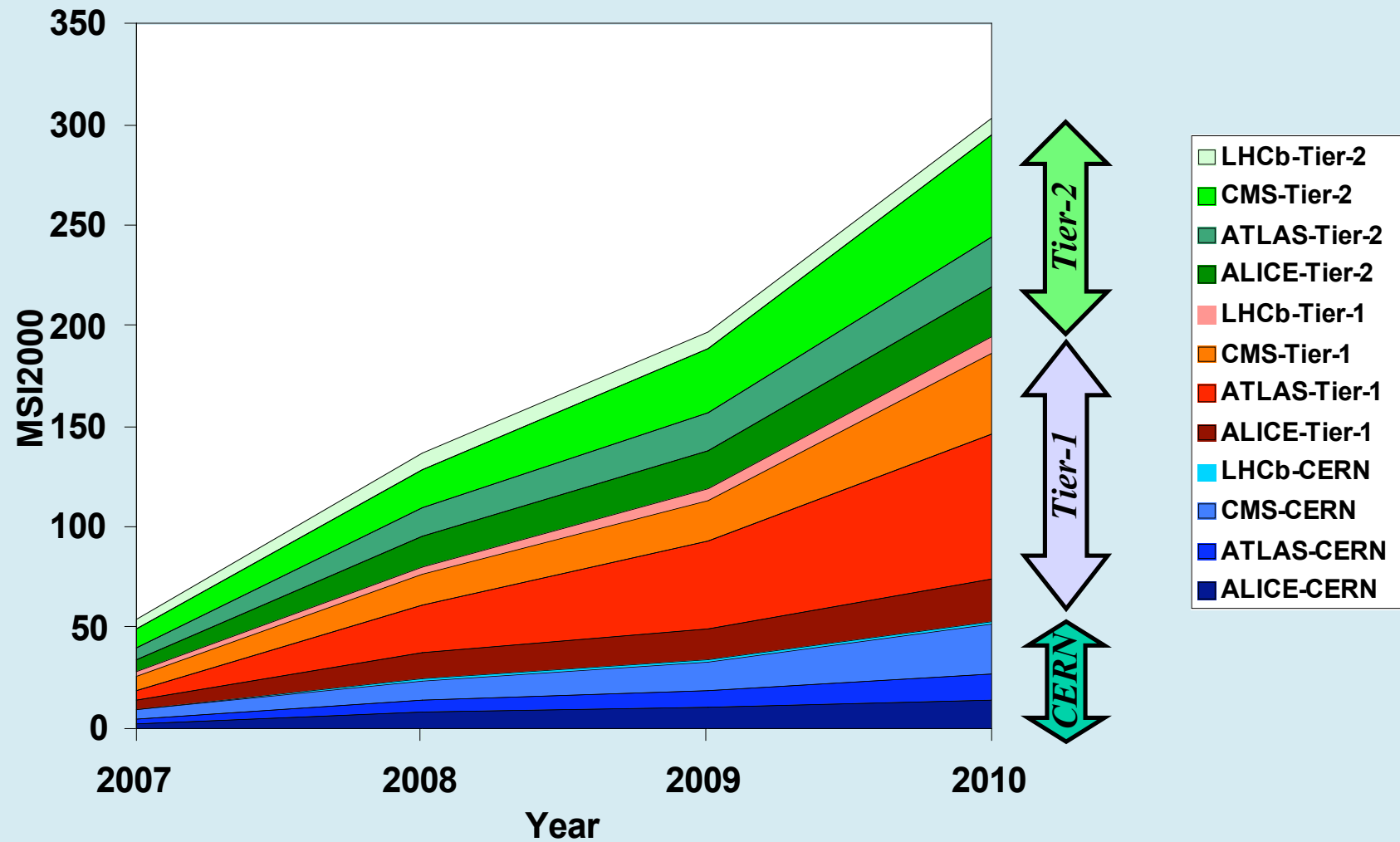
The answer to this question is probably linked to the answers to the other questions, and in particular to the ideas of supersymmetry and the resolution of the matter - antimatter problem.

Typical LHC experiment computing model

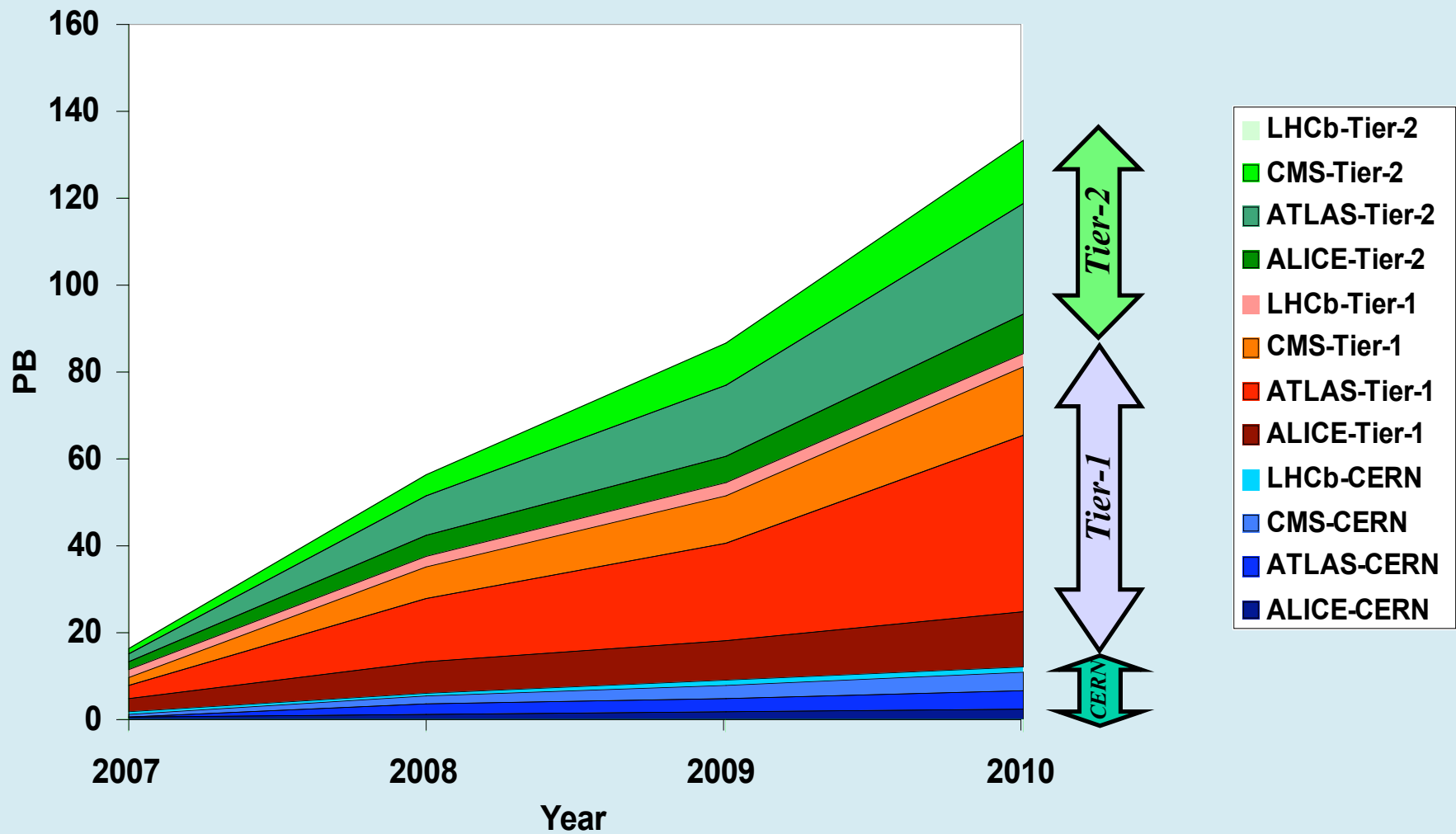
- CERN (Tier-0 centres)
 - First pass reconstruction, storage of one copy of RAW data from detectors, calibration data, 1st pass reconstructed data
- Large external computing centres+CERN (Tier-1 centres)
 - Reconstructions and Production-type analysis, storage of the second copy of RAW data and copy of all data to be kept, disk replicas of reconstructed data and analysis data
- Smaller external computing centres (Tier-2 centres)
 - Simulation and end-user analysis, disk replicas of analysis data

Tier-1 & Tier-2 centres are defined by the
Level of Service provision

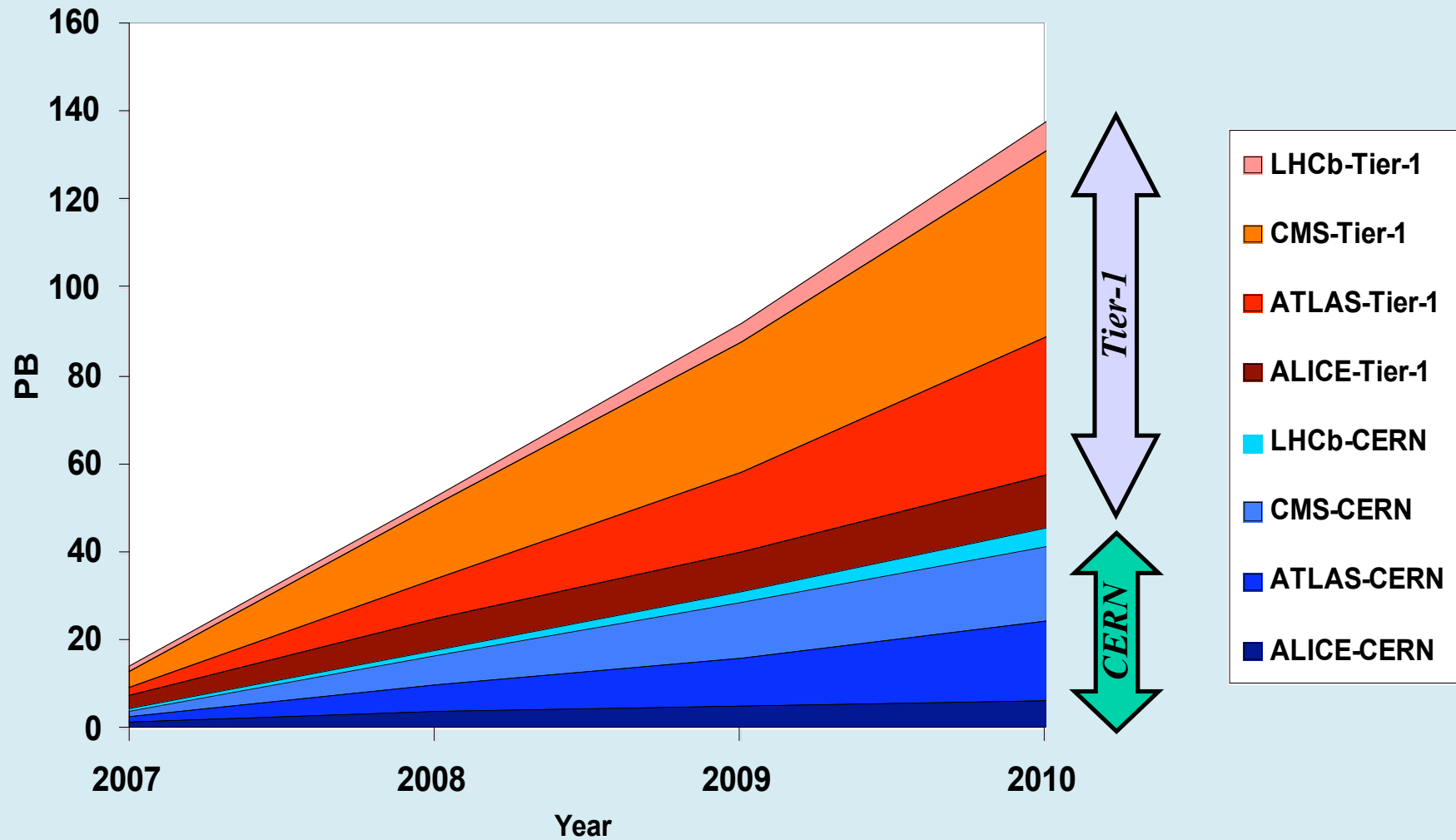
CPU Requirements



Disk Requirements

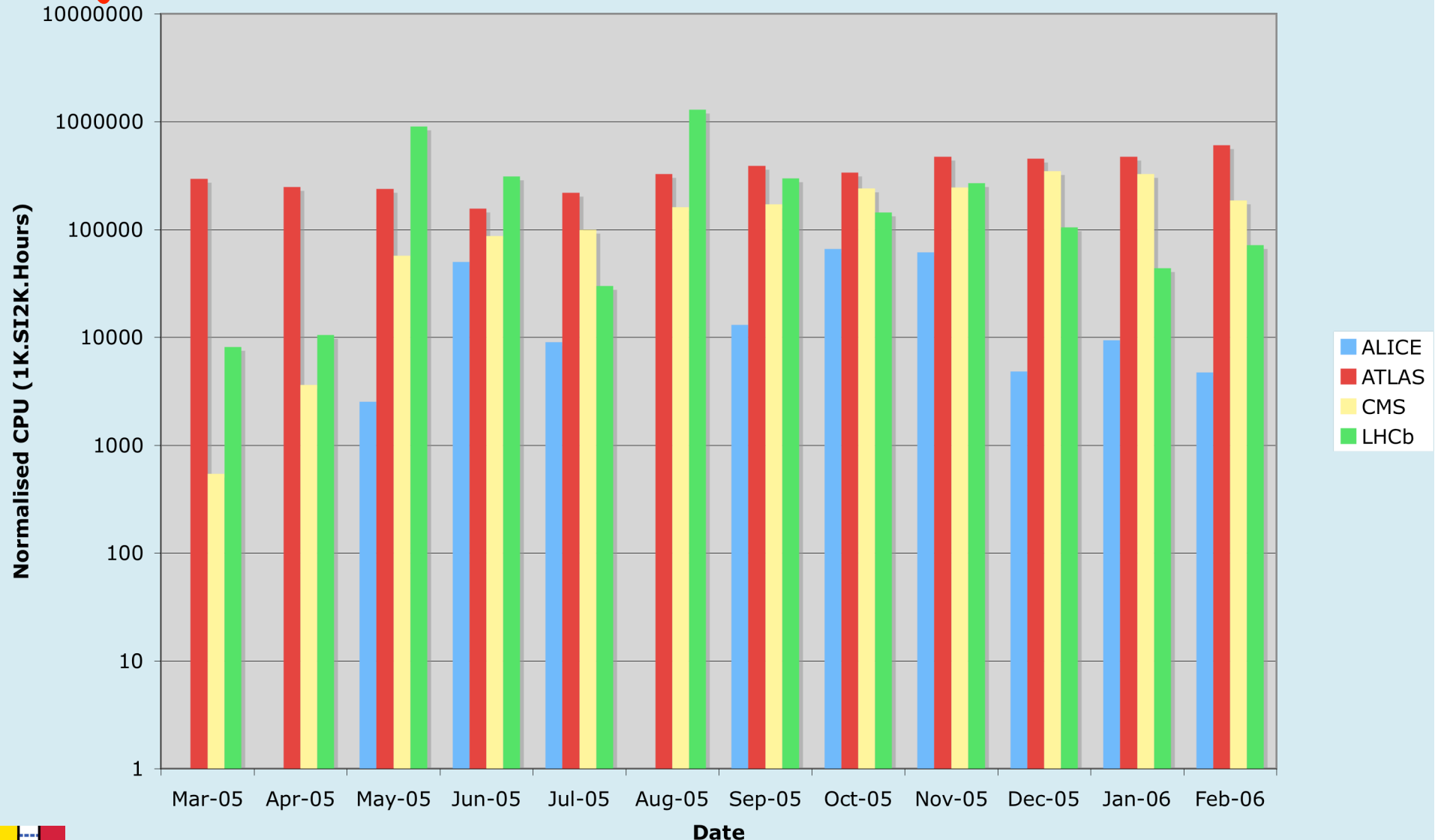


Tape Requirements



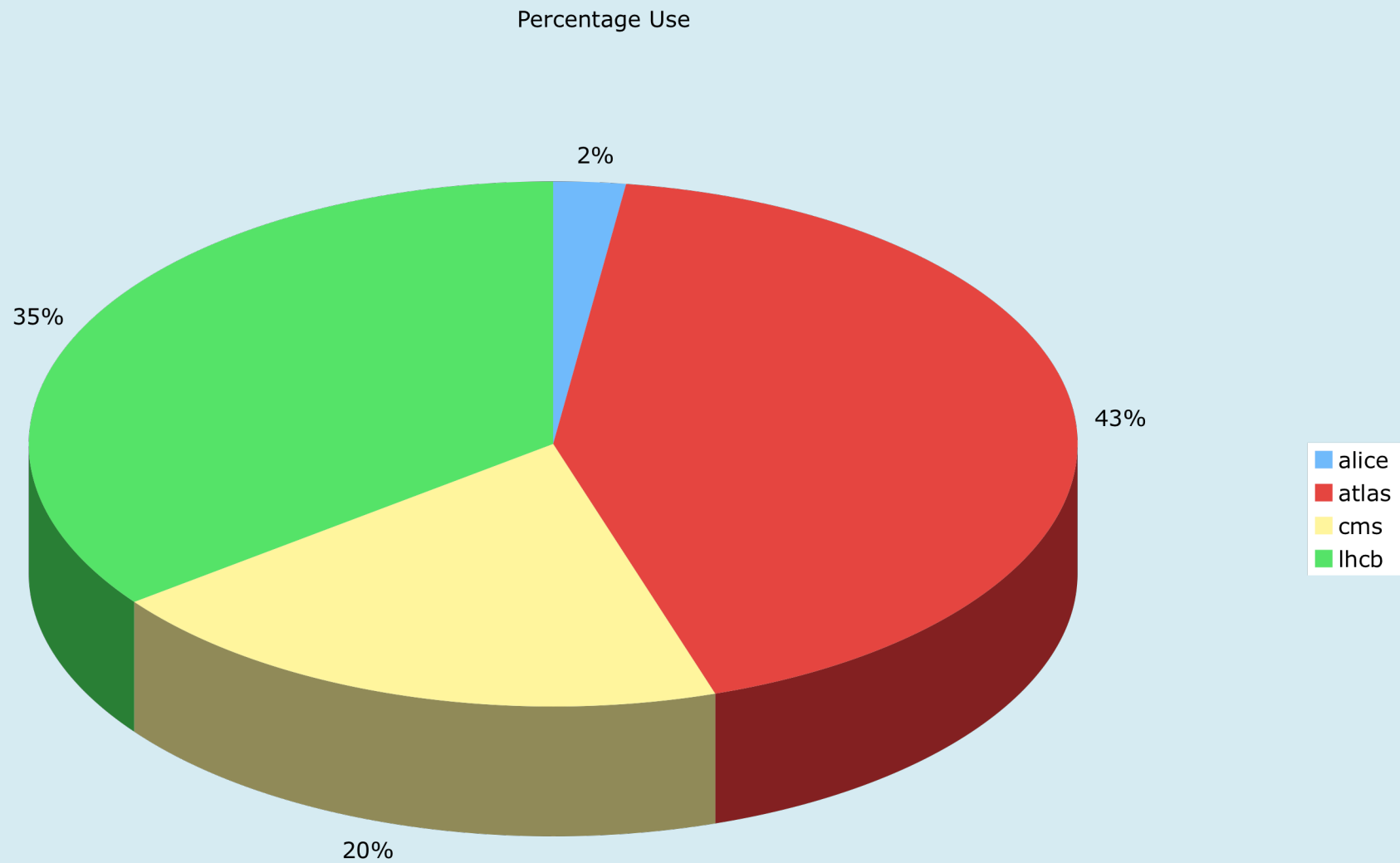
LCG/EGEE Usage by LHC Experiments

Major use of Grid so far has been for Monte Carlo simulation

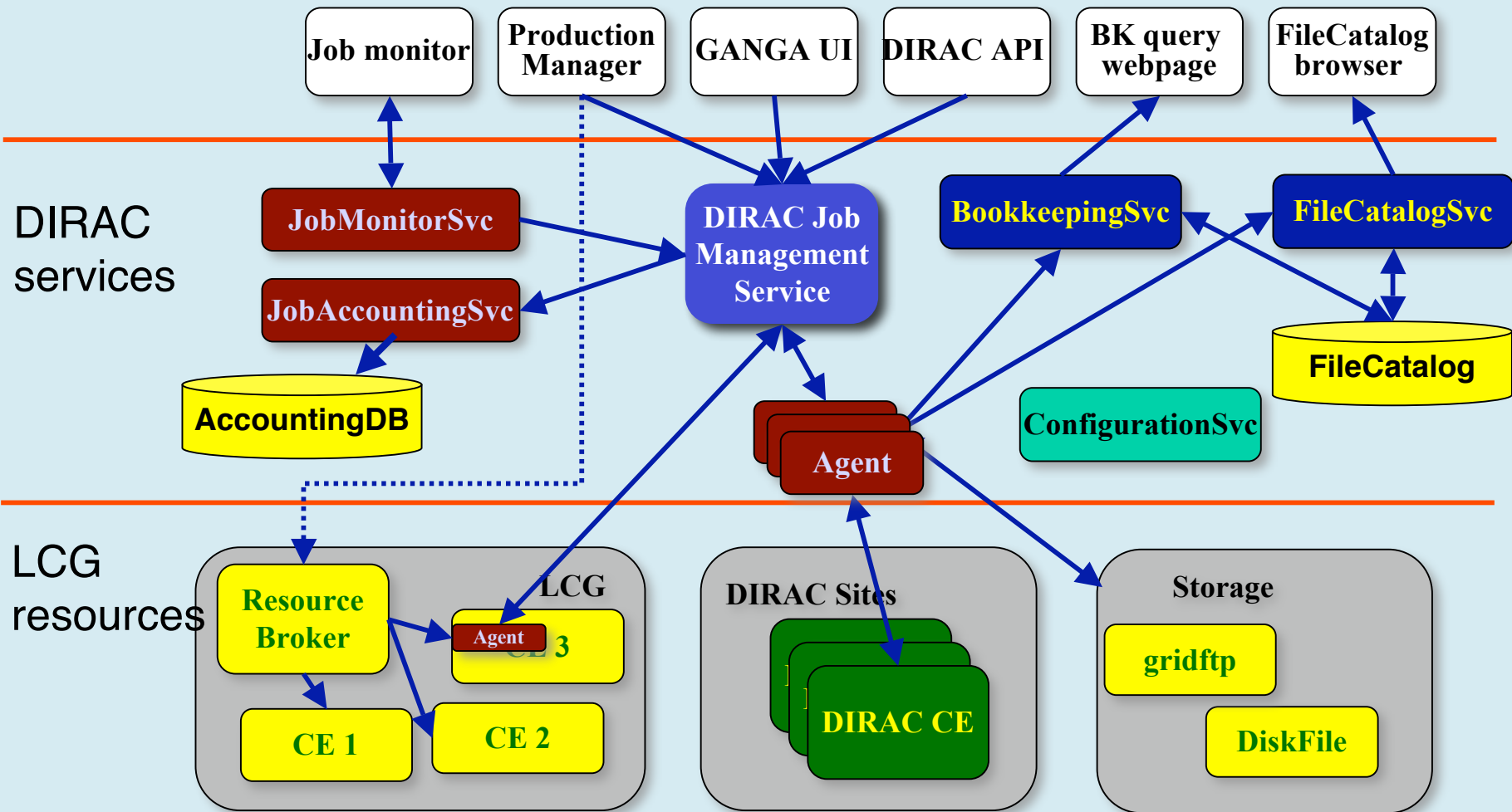


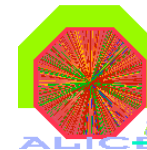
1st EGEE User forum - CERN, 1st March'06

LCG/EGEE Usage by LHC Experiments



Example Use of EGEE Resources





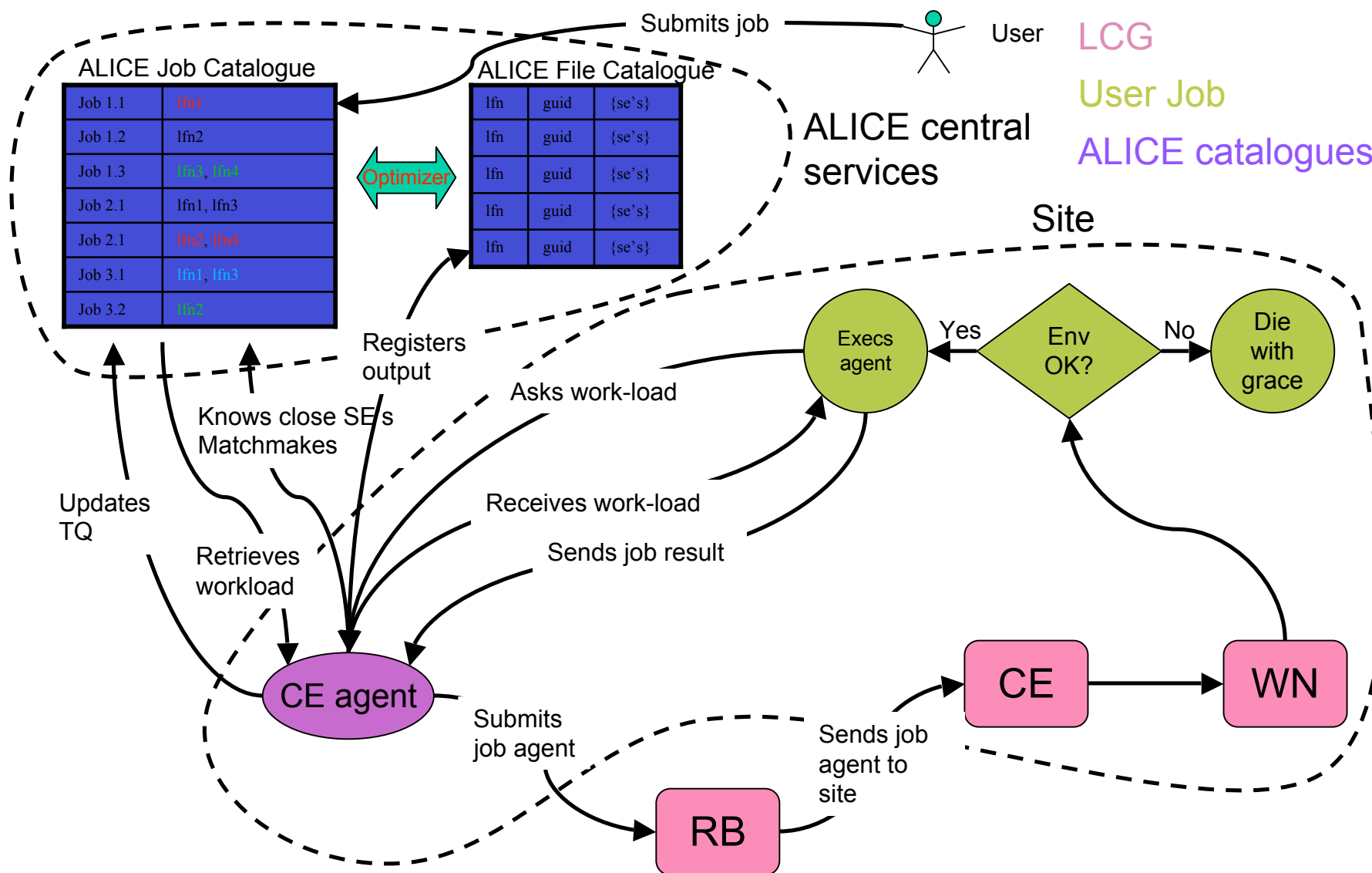
Job submission

VO-Box

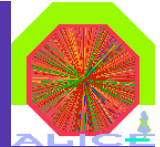
LCG

User Job

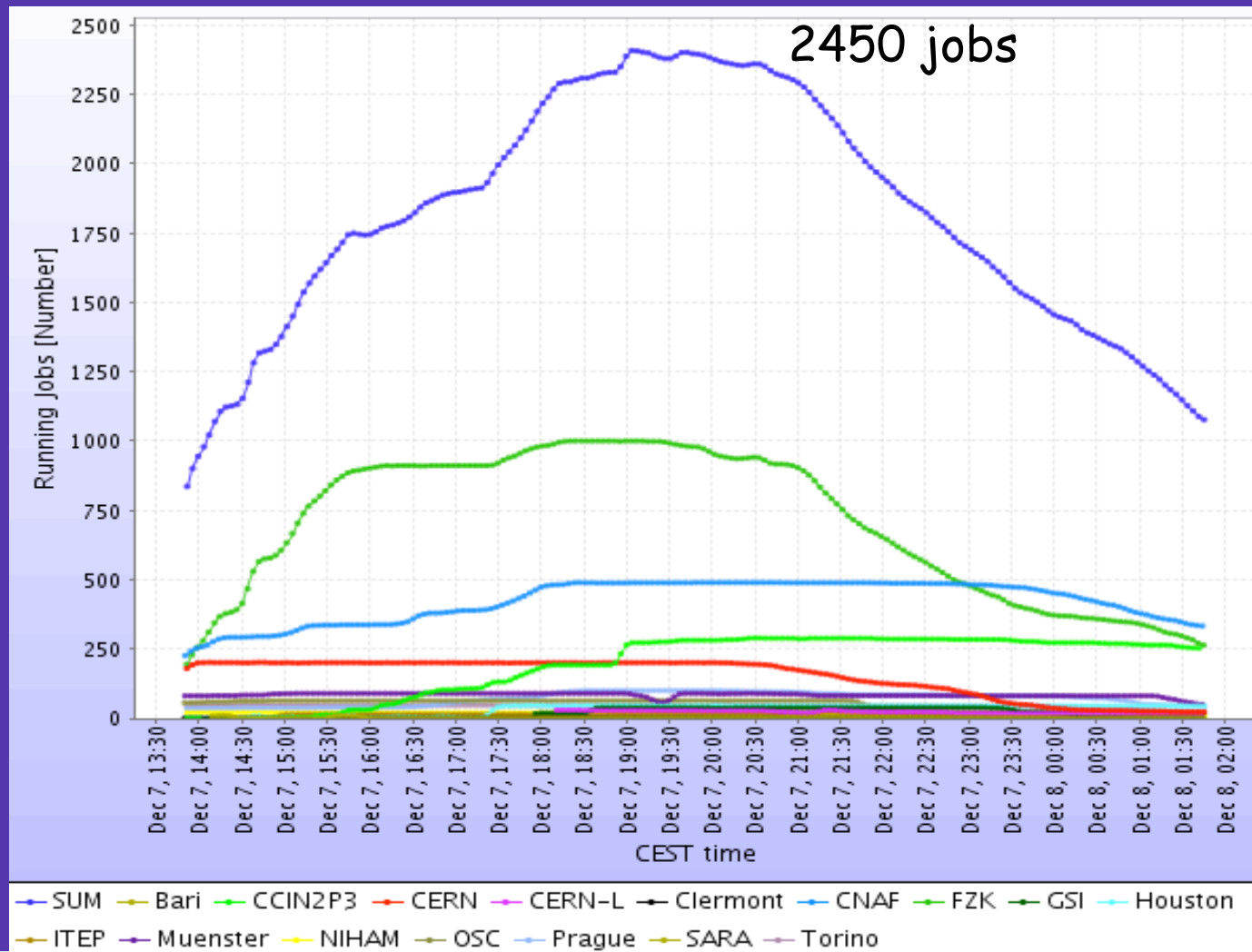
ALICE catalogues



Status of production



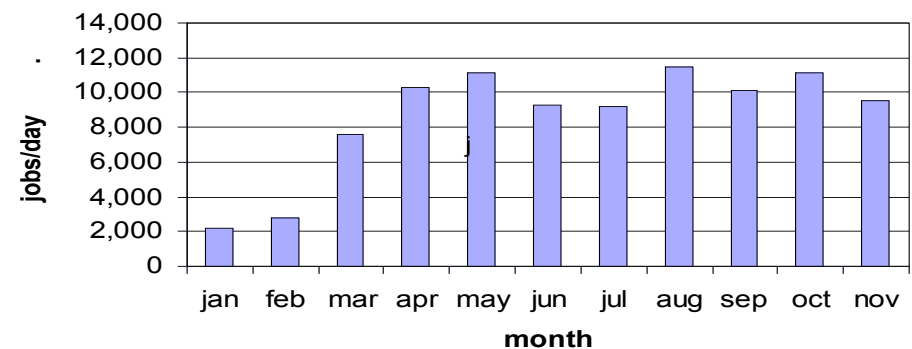
Production job duration: 8 ½ hours on 1KSi2K CPU, output archive size: 1 GB (consists of 20 files)



Production Grid

- Basic middleware
- A set of *baseline services* agreed and initial versions in production
- All major LCG sites active
- Grid job failure rate 5-10% for most experiments,
down from ~30% in 2004
- Sustained 10K jobs per day
- > 10K simultaneous jobs during prolonged periods

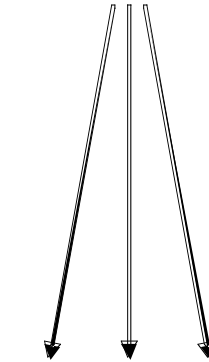
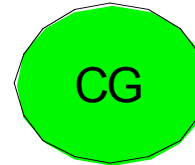
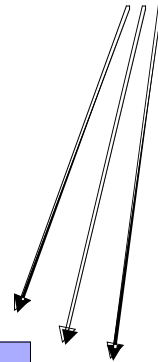
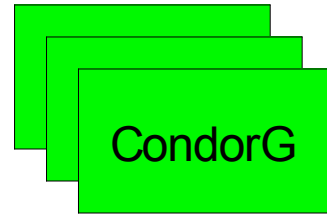
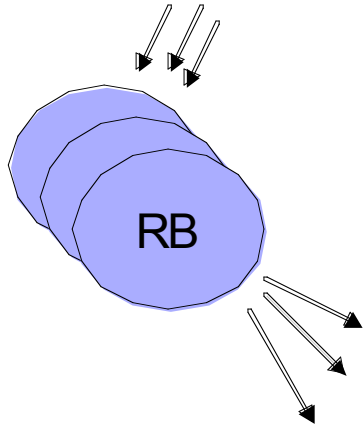
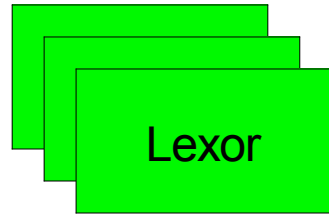
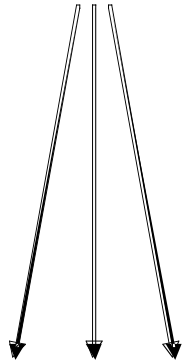
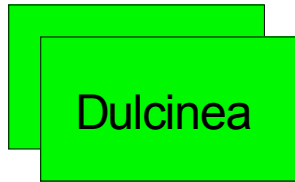
**Average number of jobs/day
EGEE Grid - 2005**





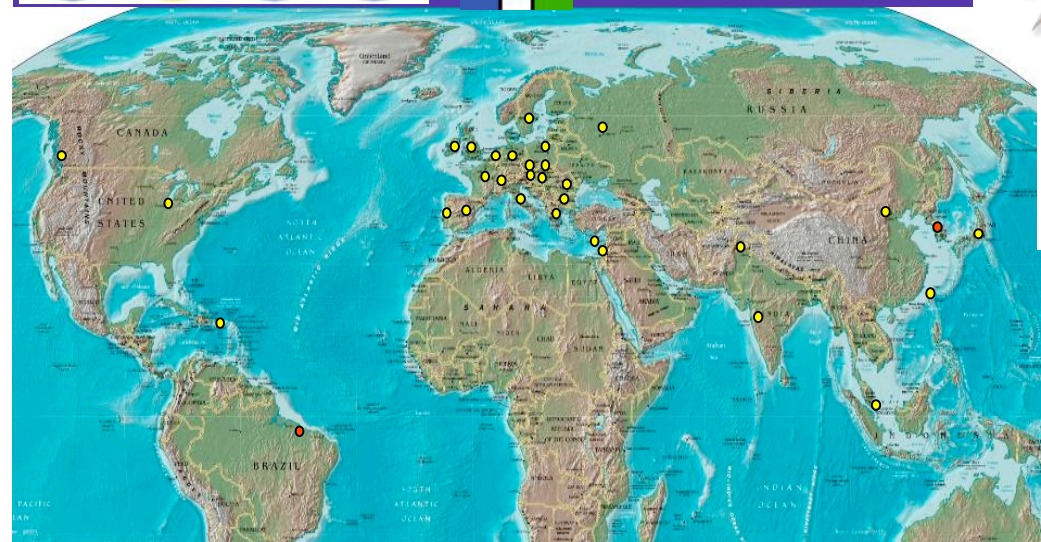
ATLAS Prodsys

ProdDB

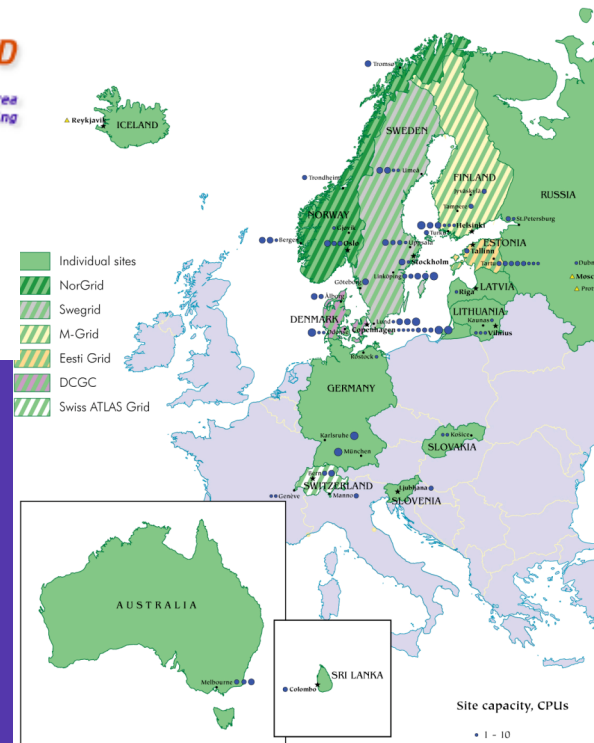


Worldwide LHC Computing Grid
Distributed Production Environment for Physics data Processing





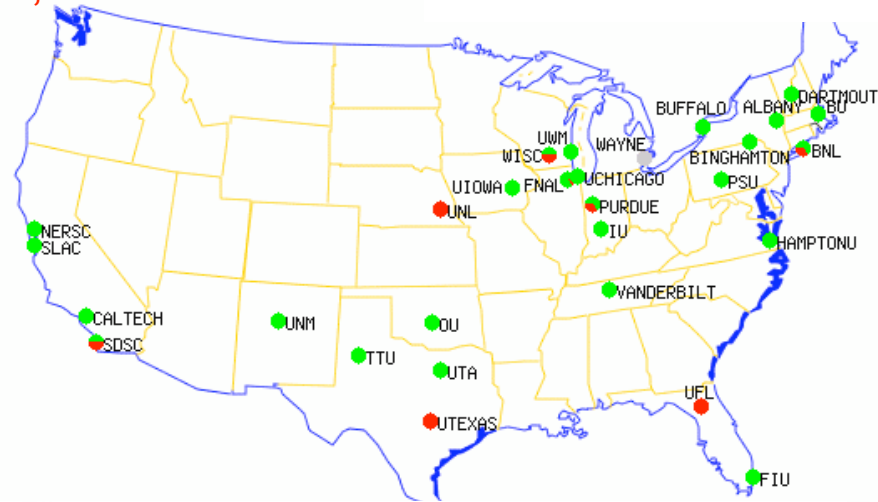
50 sites
13 countries
> 5000 CPU's



EGEE/LCG-2 grid:
174 sites, 40 countries
>17,000 processors,
~5 PB storage

country	sites	country	sites	country	sites
Austria	2	India	1	Russia	10
Belgium	2	Israel	2	Singapore	1
Bulgaria	4	Italy	25	Slovakia	3
Canada	6	Japan			
China	1	Korea			
Croatia	1	Netherlands			
Cyprus	1	Macedonia			
Czech Republic	2	Pakistan			
France	8	Poland			
Germany	8	Portugal			
Greece	6	Puerto Rico			
Hungary	1	Romania			

46 CEs, 15459 CPUs
6 SEs



Open Science Grid

Interoperability is
a major issue

Interoperability

EGEE - OSG:

- Job submission demonstrated in both directions
- Done in a sustainable manner
- EGEE WN tools installed as a grid job on OSG nodes

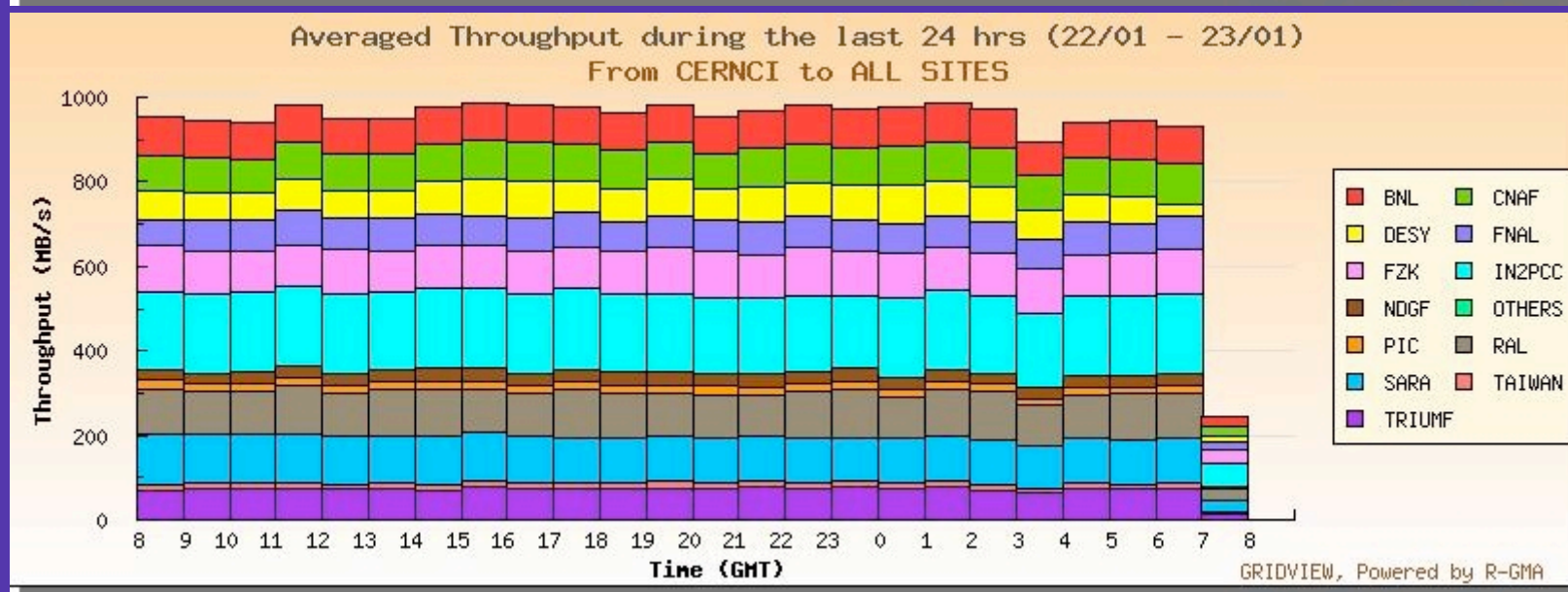
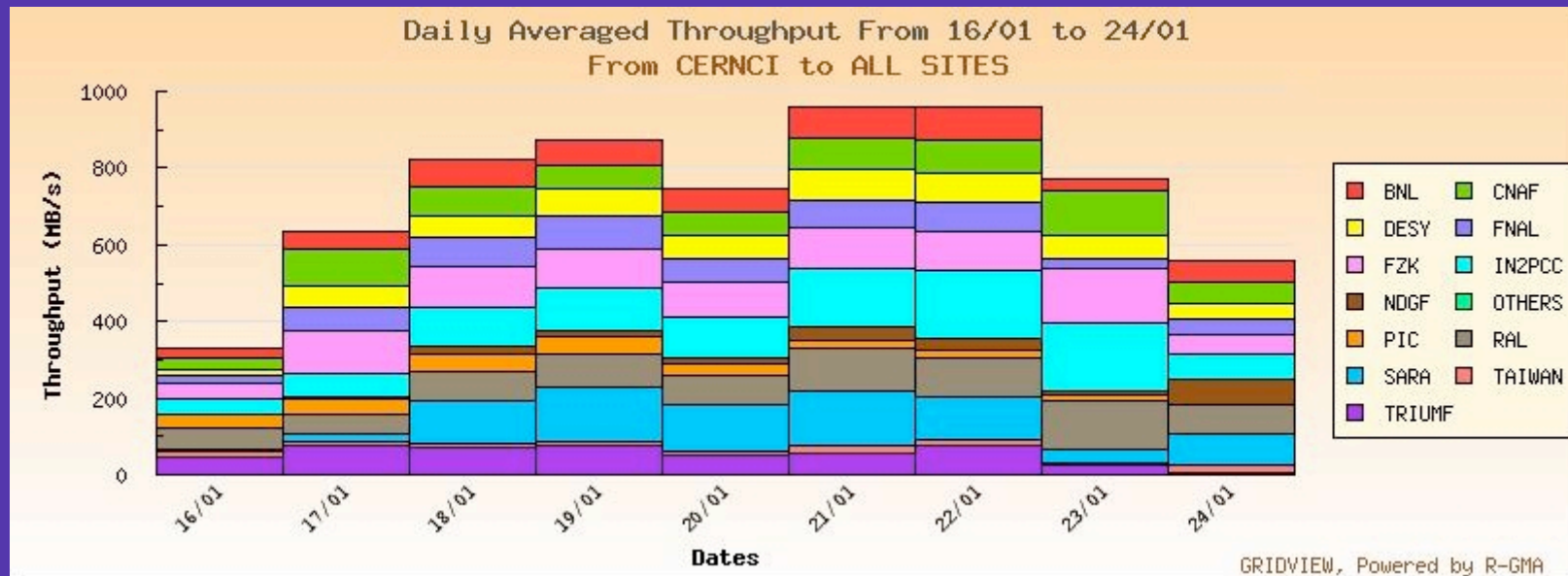
EGEE - ARC:

- Longer term want to agree standard interfaces to grid services
- Short term:
 - EGEE→ARC: Try to use Condor component that talks to ARC CE
 - ARC→EGEE: discussions with EGEE WMS developers to understand where to interface
- Default solution: NDGF acts as a gateway

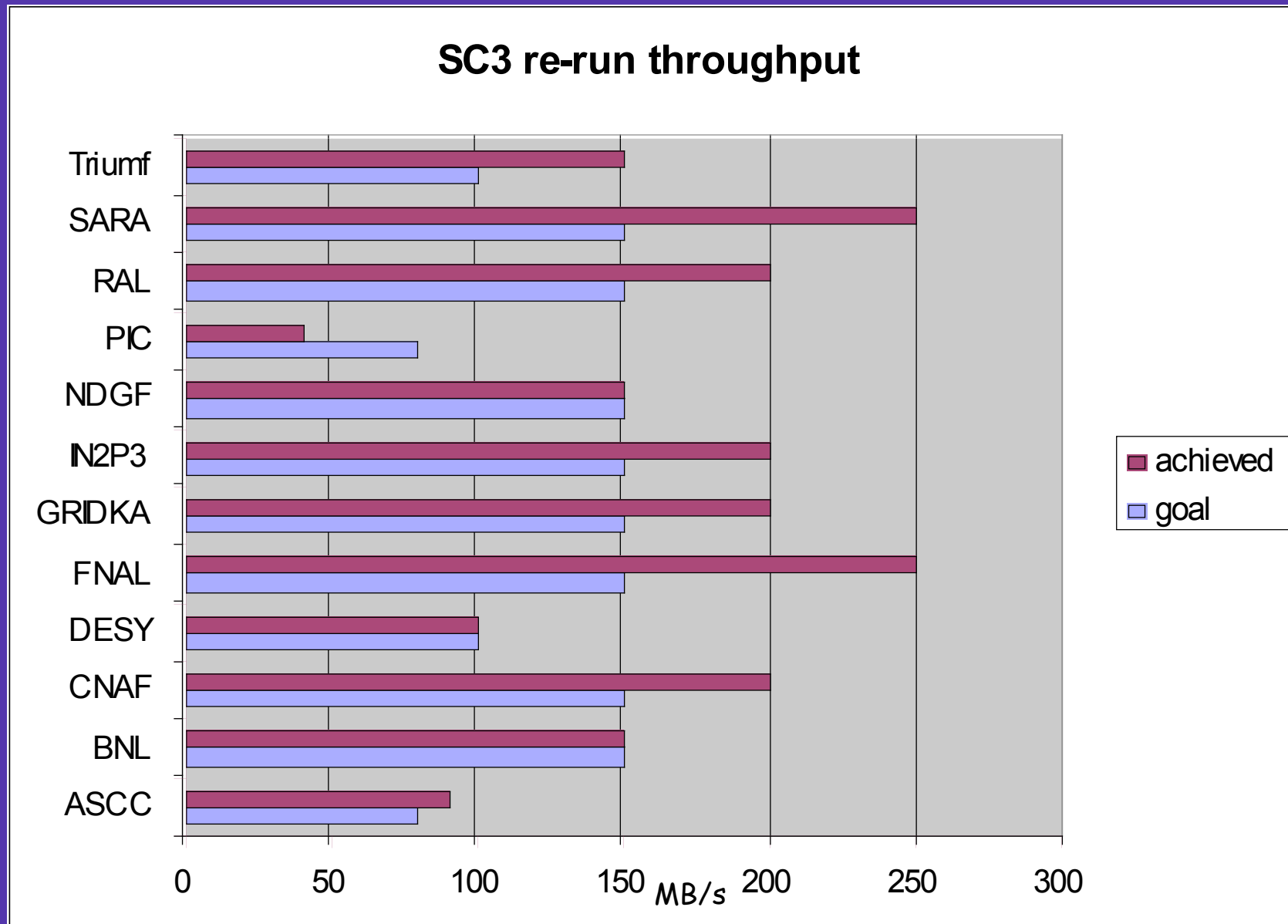
In both cases:

- Catalogues are experiment choices - generally local catalogues use local grid implementations

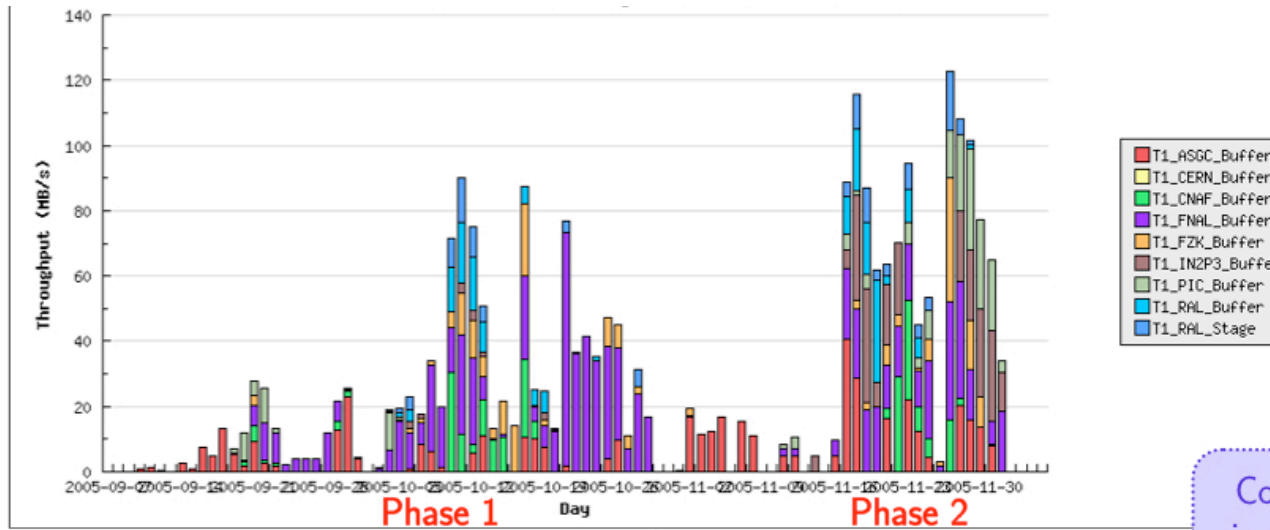
Recent Service Challenges - throughput phase



Recent Service Challenges - throughput phase



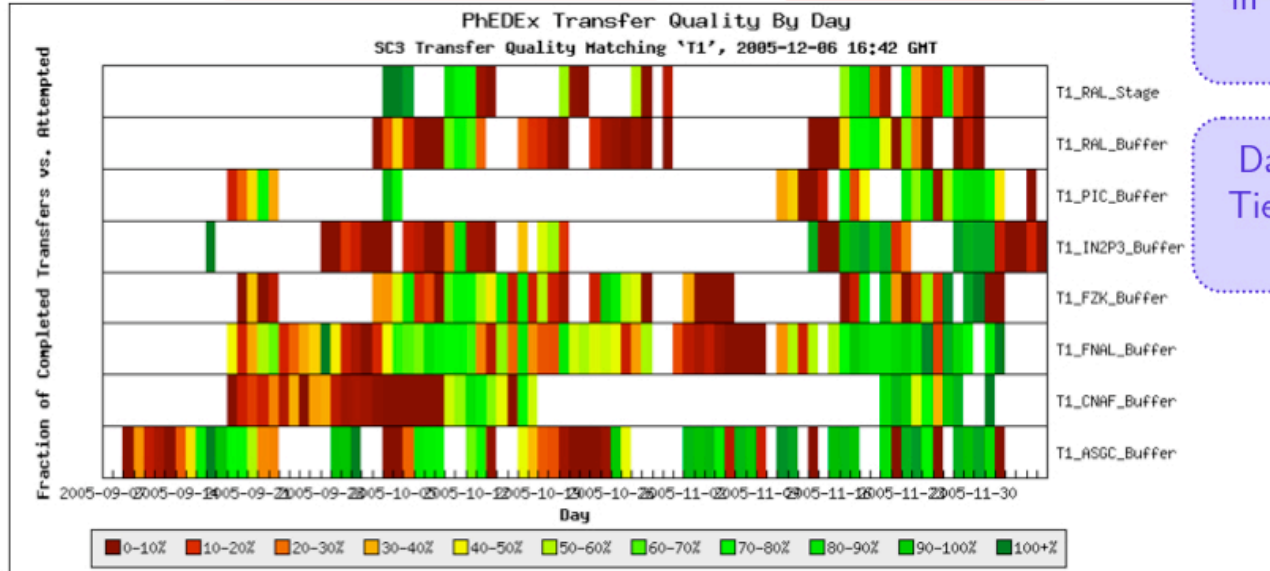
Recent Service Challenges - experiment experiences



Tier-1 WAN Transfers

Considerable improvement in transfer rate and quality for service phase 2.

Daily aggregate rate from Tier 0 to Tier 1s peaked at 120 MB/s.



SC3 summary - expt perspective

Extremely useful for shaking down sites, experiment systems & WLCG

- Many new components used for the 1st time in anger
- Need for additional functionality in services
 - F(ile) T(ransfer) S(ervice), L(CG) Fi(le) C(atalog), S(torage) R(esource) M(anager), ...

Reliability seems to be the major issue

- MSS at CERN - still ironing out problems, but big improvements
- Coordination issues
- Problems with sites and networks
 - MSS, security, network, services...

FTS:

- For well-defined site/channels performs well after tuning
- Timeout problems dealing with accessing data from MSS

SRM:

- Limitations/ambiguity in functionality for v1.1





Ganga

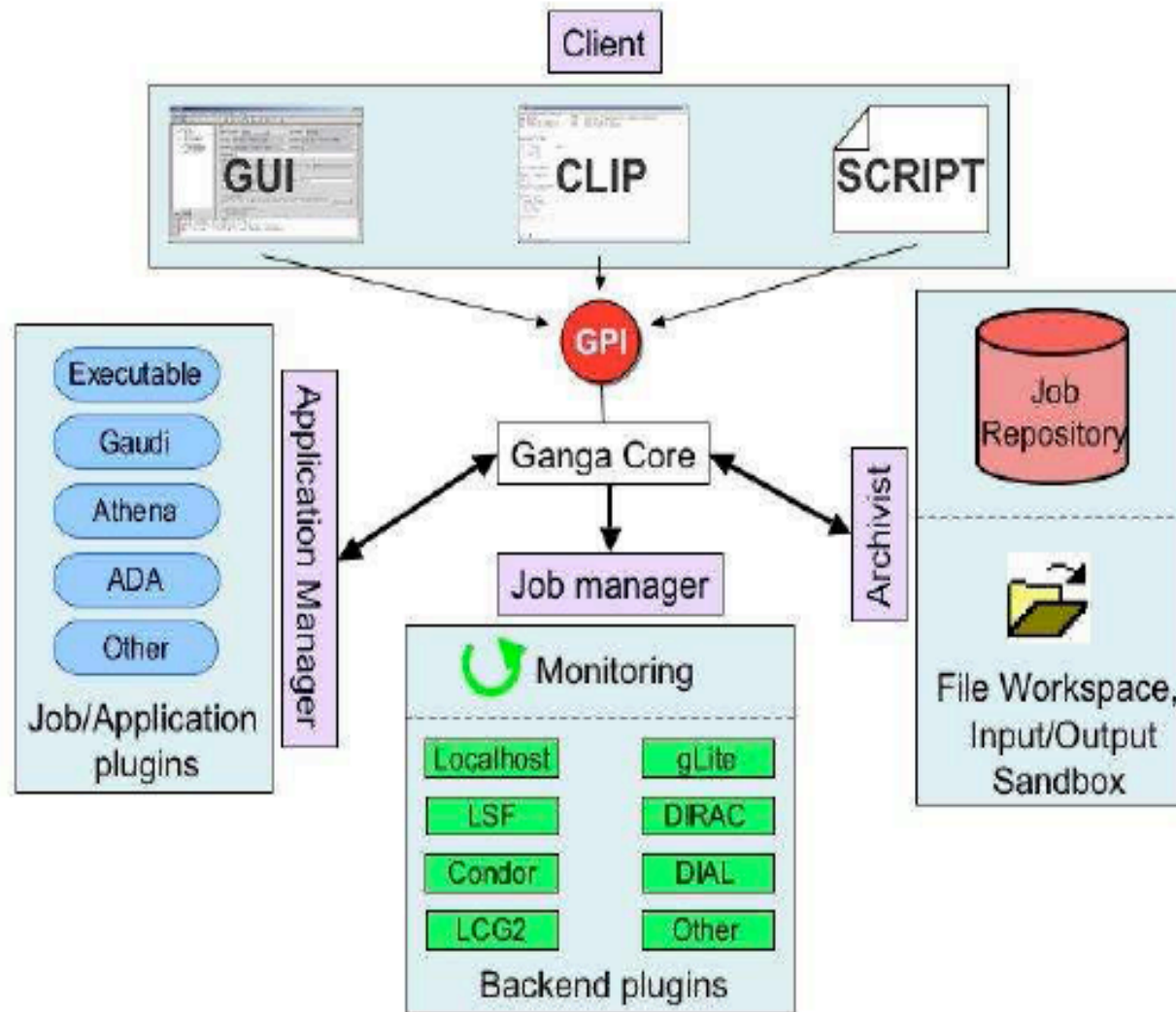


LHCb
LHCb

- Designed for data analysis on the Grid
 - LHCb will do all its analysis on T1's
 - T2's mostly for simulation
- System should not be general - we know all main use cases
 - Use prior knowledge
 - Identified use pattern
- Aid user in
 - Bookkeeping aspects
 - Keeping track of many individual jobs
- Developed in cooperation between LHCb and ATLAS with EGEE support



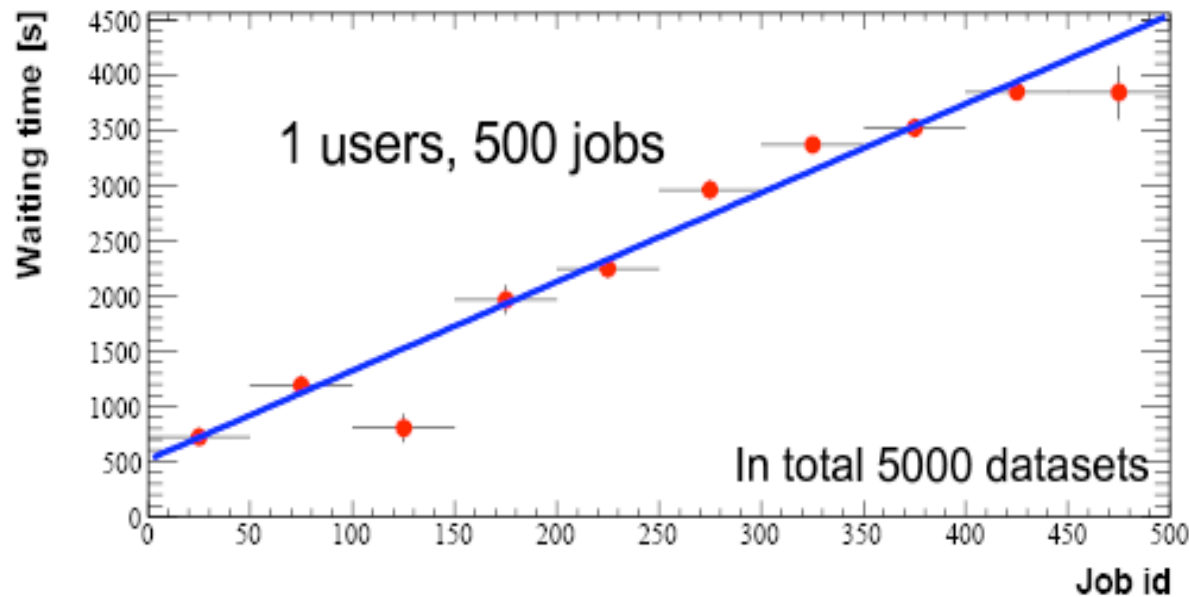
Ganga Architecture



Implementation:
Pure python
~20k lines of code

Performance: Throughput

An analysis of 5M events entered into the system at the same time.



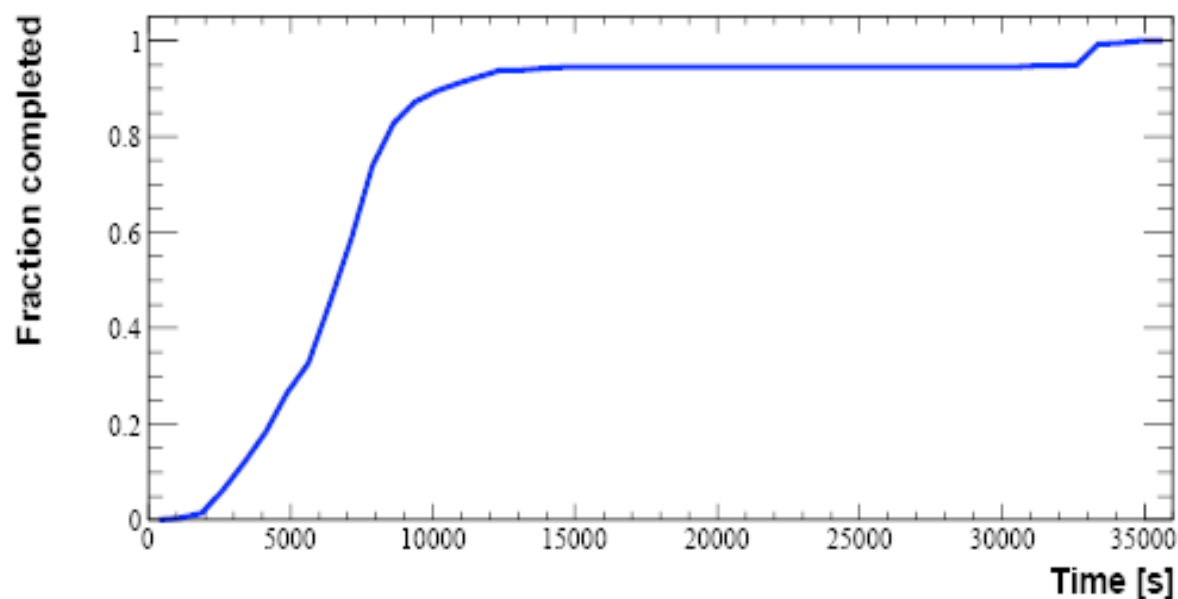
$$\text{wait [s]} = 517 + 8.06 \times \text{job id}$$

517 s (8 min) before first job starts is dominated by latency in jobs for LCG and software installation.

Queue time afterwards is dominated by the time it takes for WMS system to submit agents to LCG.

Performance: Throughput

Look at the time it takes before the results are back



90% of results are back within less than 3 hours

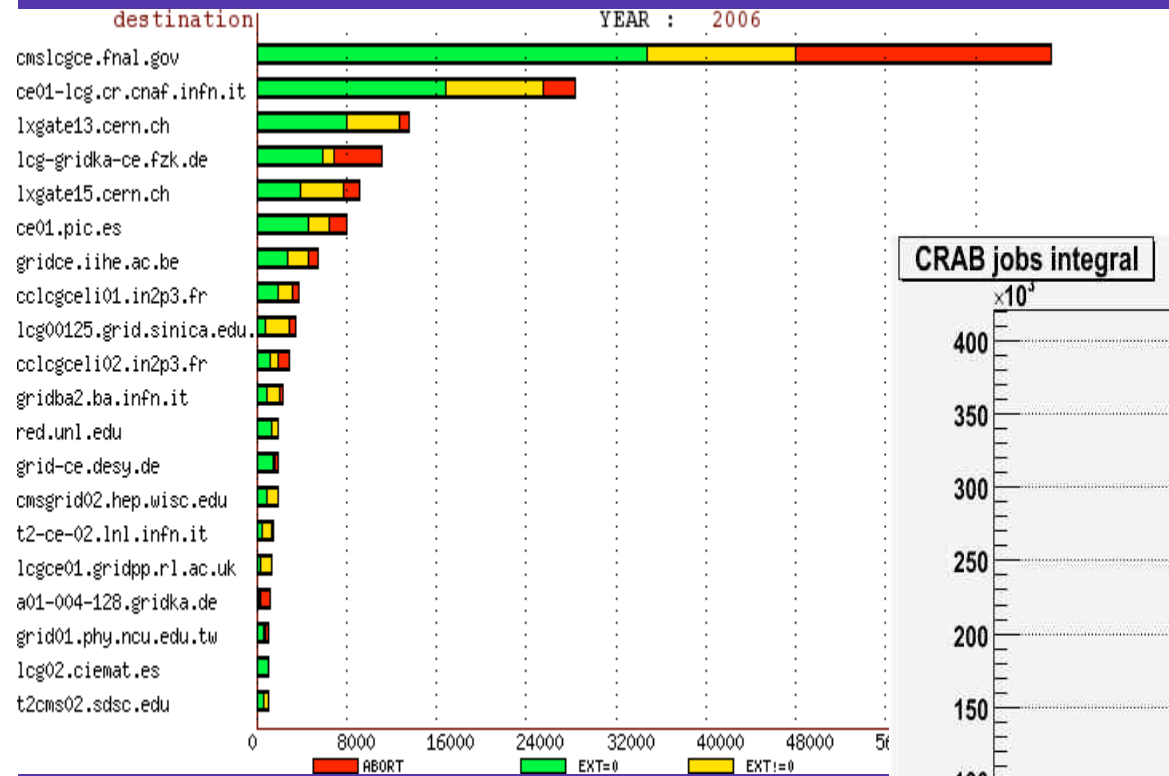
95% in 4 hours

100% in 10 hours

Last 5% caused by Tier 1 site with data access problems



CMS Analysis on the Grid



CRAB jobs so far



Most accessed sites since July 05

Many 10's of thousands of jobs run to produce results for CMS technical design report



LCG baseline service working group -goals

- Experiments and regional centres agreement on baseline services
 - Support the computing models for the initial period of LHC
 - Thus must be in operation by September 2006.
- The services concerned are those that
 - supplement the basic services
 - (e.g. provision of operating system services, local cluster scheduling, compilers, ..)
 - and which are not already covered by other LCG groups
 - such as the Tier-0/1 Networking Group or the 3D Project.
- Not a middleware group - focus on what the experiments need & how to provide it
- Expose experiment plans and ideas



Baseline services

- Nothing really surprising here - but a lot was clarified in terms of requirements, implementations, deployment, security, etc

- Storage management services
 - Based on SRM as the interface
- Basic transfer services
 - gridFTP, srmCopy
- Reliable file transfer service
- Grid catalogue services
- Catalogue and data management tools
- Database services
 - Required at Tier1,2
- Compute Resource Services
- Workload management

- VO management services
 - Clear need for VOMS: roles, groups, subgroups
- POSIX-like I/O service
 - local files, and include links to catalogues
- Grid monitoring tools and services
 - Focussed on job monitoring
- VO agent framework
- Applications software installation service
- Reliable messaging service
- Information system

Service Challenge 4 - the Pilot LHC Service

- Start service June'06
- Must be able to support a demonstration of the complete chain
- Data Acquisition → Tier-0 → Tier-1
data recording, calibration, reconstruction
- simulation, batch and end-user analysis
Tier-1 ↔ Tier-2 data exchange
- Service metrics → agreed (MoU) service levels
- Extension of the service to most Tier-2 sites
- Production LCG service by October'06

Summary

- Steady progress in understanding and fixing problems
 - Not everything is there at the moment
 - Every Service Challenge has increasingly ambitious objectives and an increasing number of problems to be fixed!
- Experiments, Sites, m/w developers all work in the same direction and with the same objectives / priorities
- (Some) experiment specific solutions are unavoidable and should be supported in a form acceptable to all parties
 - First collision due in 2007 & experiments need a stable system beforehand to ensure their systems work
- Proper prioritisation of experiment requirements is now mandatory in order to satisfy them
 - Ongoing activity with EGEE TCG
- Progress needs to continue to allow different flavoured Grid to interoperate