Overview of the Application Migration report DNA 4.3.1

Frank Harris (CERN/Oxford)

Top Level of TOC (9 chapters)

1) Introduction
2) Executive Summary

3) Migration report for HEP (data challenges and ARDA)
4) Migration report for Biomedicine
5) Migration report for Generic applications
6) Suggestions for meeting needs of new user communities on EGEE
7) Conclusions

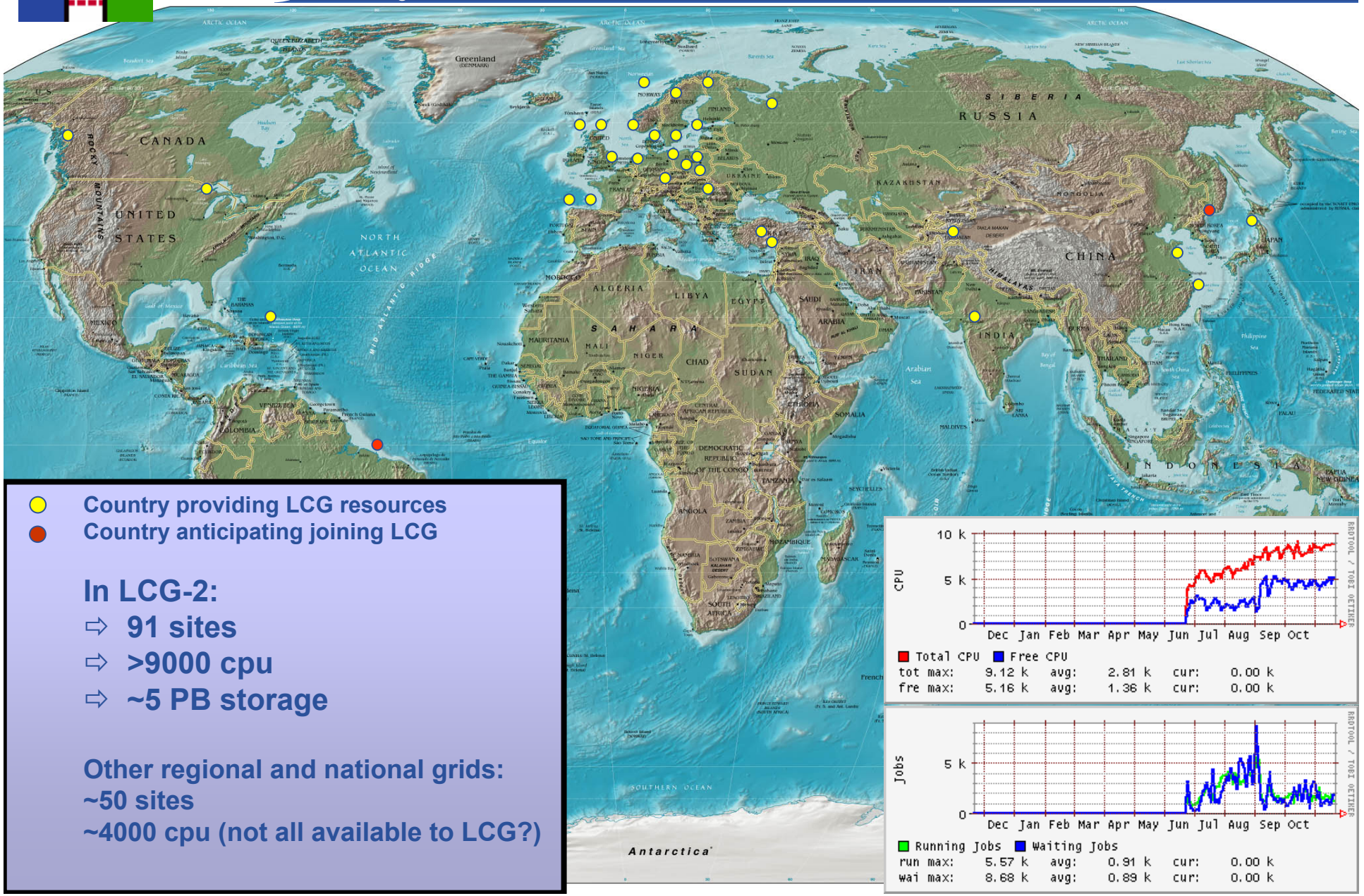
8) References (over 40)
9) List of figures and tables

Structure of presentation

- For each main application area we summarise
  - Achievements
  - Major Issues arising and a forward look
- Pilot Applications (established from EDG)
  - HEP data challenges and ARDA prototypes
  - Biomedicine
- New applications selected in July 2004
  - Earth science (well rehearsed from EDG)
  - Computational Chemistry(coming from EU COST project)
  - Astro-particle physics (strong background in Crossgrid)
- CONCLUSIONS

Computing Resources in LCG – Nov 2004

Enabling Grids for E-sciencE



- Country providing LCG resources
- Country anticipating joining LCG

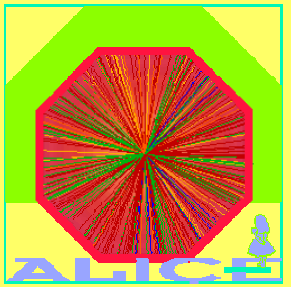
In LCG-2:

- ⇨ 91 sites
- ⇨ >9000 cpu
- ⇨ ~5 PB storage

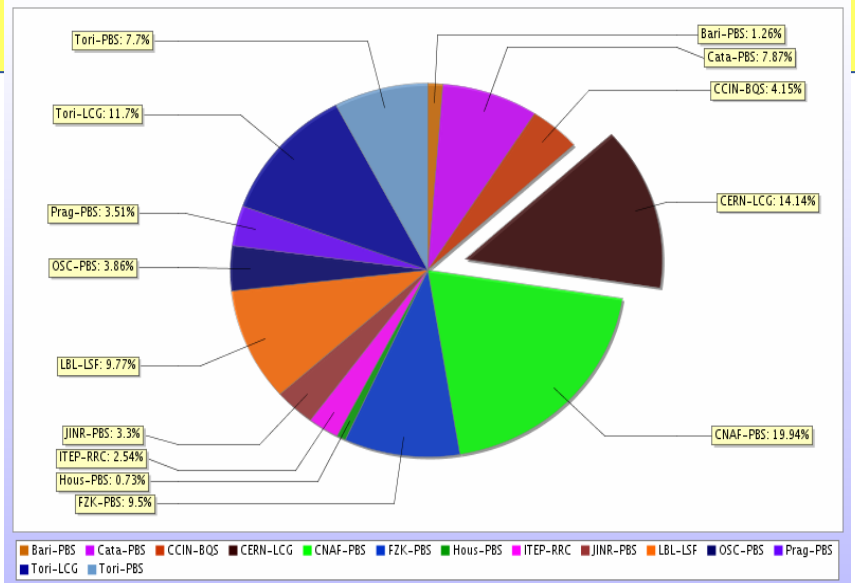
Other regional and national grids:
~50 sites
~4000 cpu (not all available to LCG?)

Data Challenges – ALICE

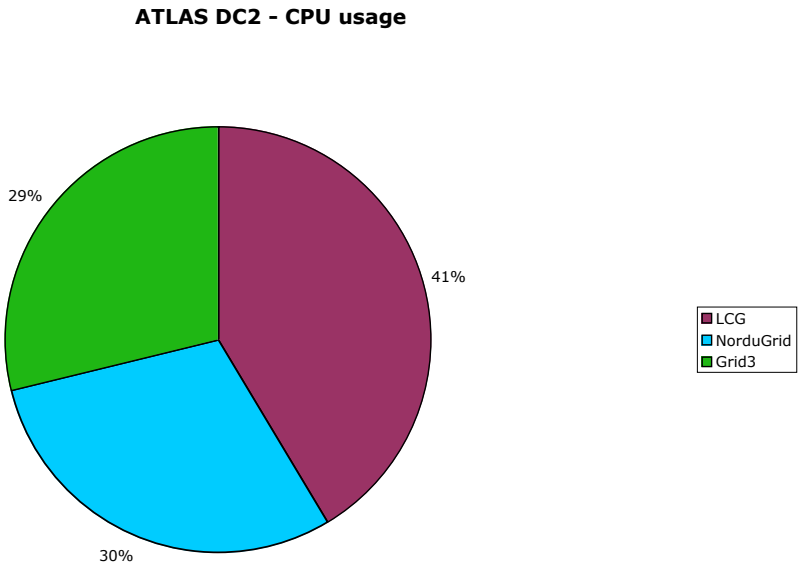
- Phase I
  - 120k Pb+Pb events produced in 56k jobs
  - 1.3 million files (26TByte) in Castor@CERN
  - Total CPU: 285 MSI-2k hours (2.8 GHz PC working 35 years)
  - ~25% produced on LCG-2

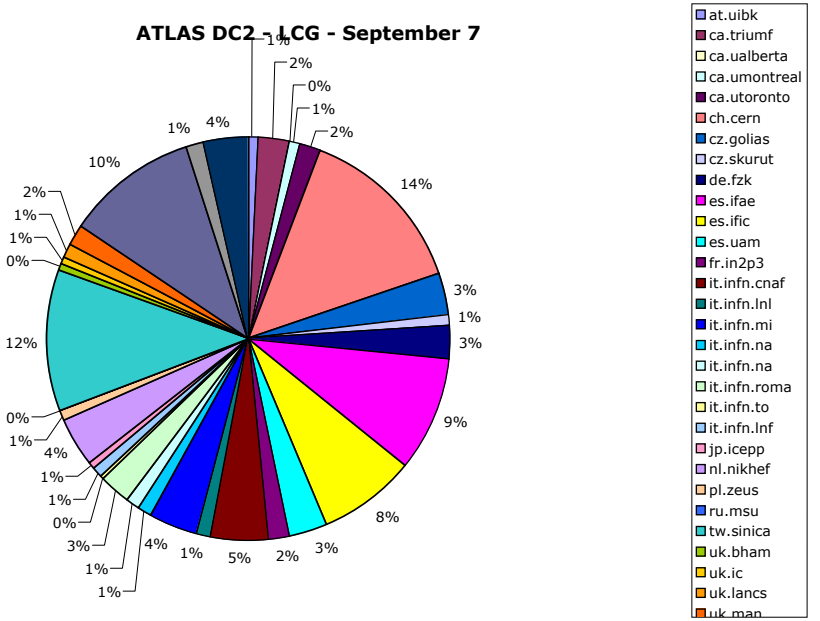


- **Phase II**
  - 1 million jobs, 10 TB produced, 200TB transferred ,500 MSI2k hours CPU
  - ~15% on LCG-2



Data Challenges – ATLAS

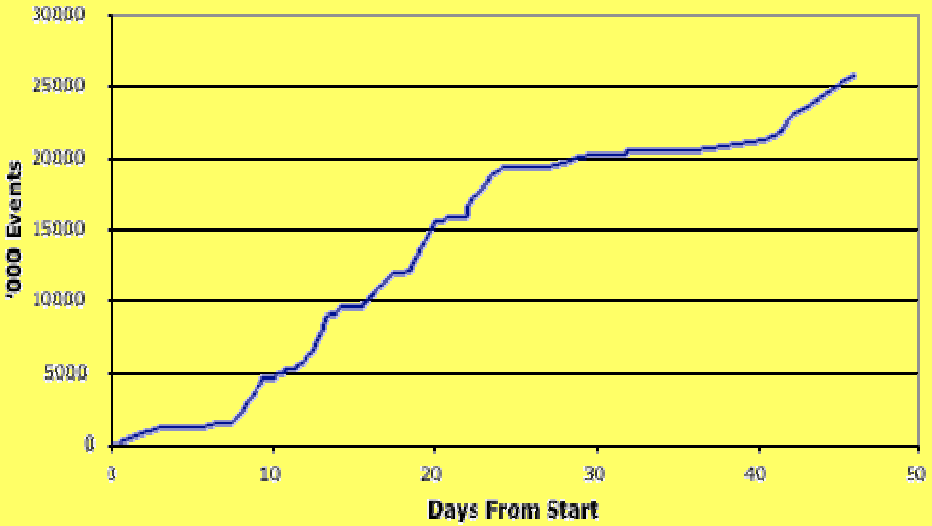




- Phase I
  - **7.7 Million events fully simulated (Geant 4) in 95.000 jobs**
  - **22 TByte**
  - **Total CPU: 972 M SI-2k hours**
  - **>40% produced on LCG-2 (used LCG-2, GRID2003, NorduGrid)**

Data Challenges – CMS

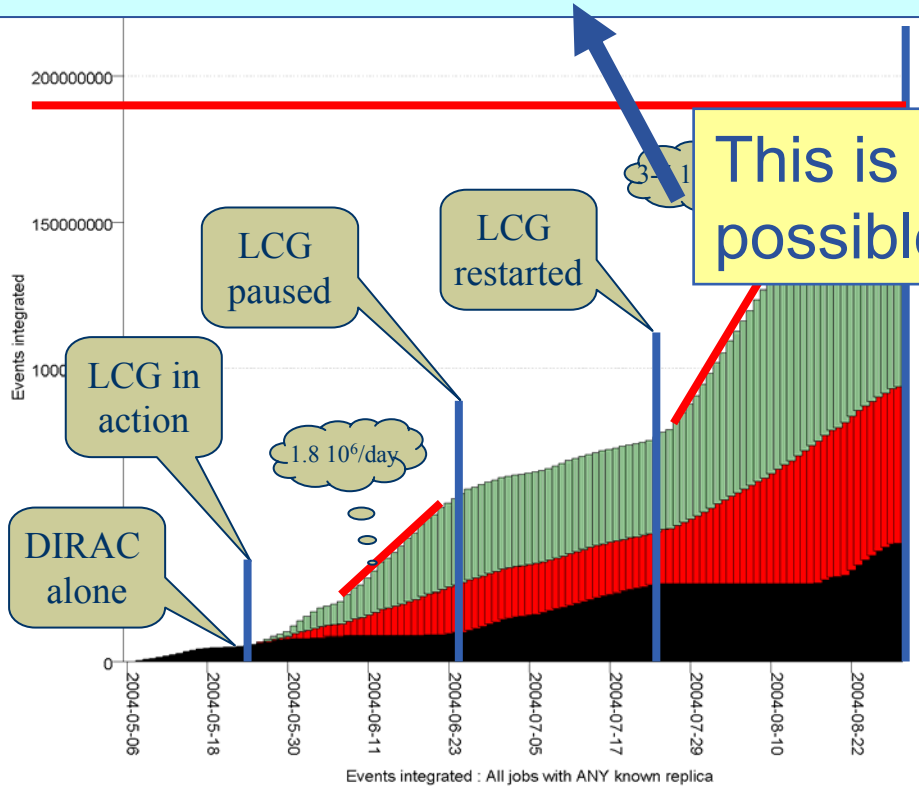
- ~30 M events reconstructed at Tier-0
- 25Hz reached for flow to analysis in Tier-1
  - (only once for a full day)
  - RLS, Castor, control systems, T1 storage, …



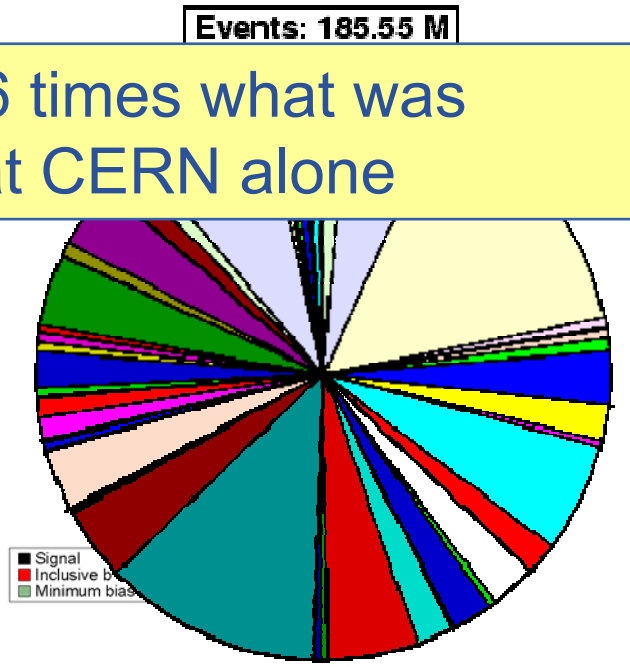
- Not a CPU challenge, but a full chain demonstration
- Pre-challenge production in 2003/04
  - **70 M Monte Carlo events (30M with Geant-4) produced**
  - **Classic and grid (CMS/LCG-0, LCG-1, Grid3) productions**

Data Challenges – LHCb

- Phase I
  - 186 M events 61 TByte
  - Total CPU: 424 CPU years (43 LCG-2 and 20 DIRAC sites)
  - Up to 5600 concurrent running jobs in LCG-2



This is 5-6 times what was possible at CERN alone



D0 MC efficiency on LCG2 since Xmas

CE	Success	Failed
bohr0001.tier2.hep.man.ac.uk	237	3
cclcgceli01.in2p3.fr	-	14
grid-ce.physik.uni-wuppertal.de	-	-
gridkap01.fzk.de	2564	19
golias25.farm.particle.cz	198	15
heplnx131.pp.rl.ac.uk	246	4
lcgce02.gridpp.rl.ac.uk	293	10
mu6.matrix.sara.nl	397	7
tbn18.nikhef.nl	154	2
Total	4089	74

Efficiency 98 %

Issues arising from DCs and a forward look

- **Global efficiency taking all 'losses' into account has been ~60-80% up to September (N.B. similar experience in other international grids)**
  - Data management needs improvement (especially metadata handling-CMS had serious problems)
  - Site certification and monitoring was a major problem for LHC experiments***(note D0 now obtain 98%+ efficiency with ~ 10 controlled sites)***
  - Monitoring jobs and trouble-shooting is very hard
- **We move now to experiments wanting to expand user community to do individual data analyses (must have good efficiency)**
  - Experiments looking to gLite
- **Experiments have been very happy with their relations with LCG support (? Scalability with more users?)**

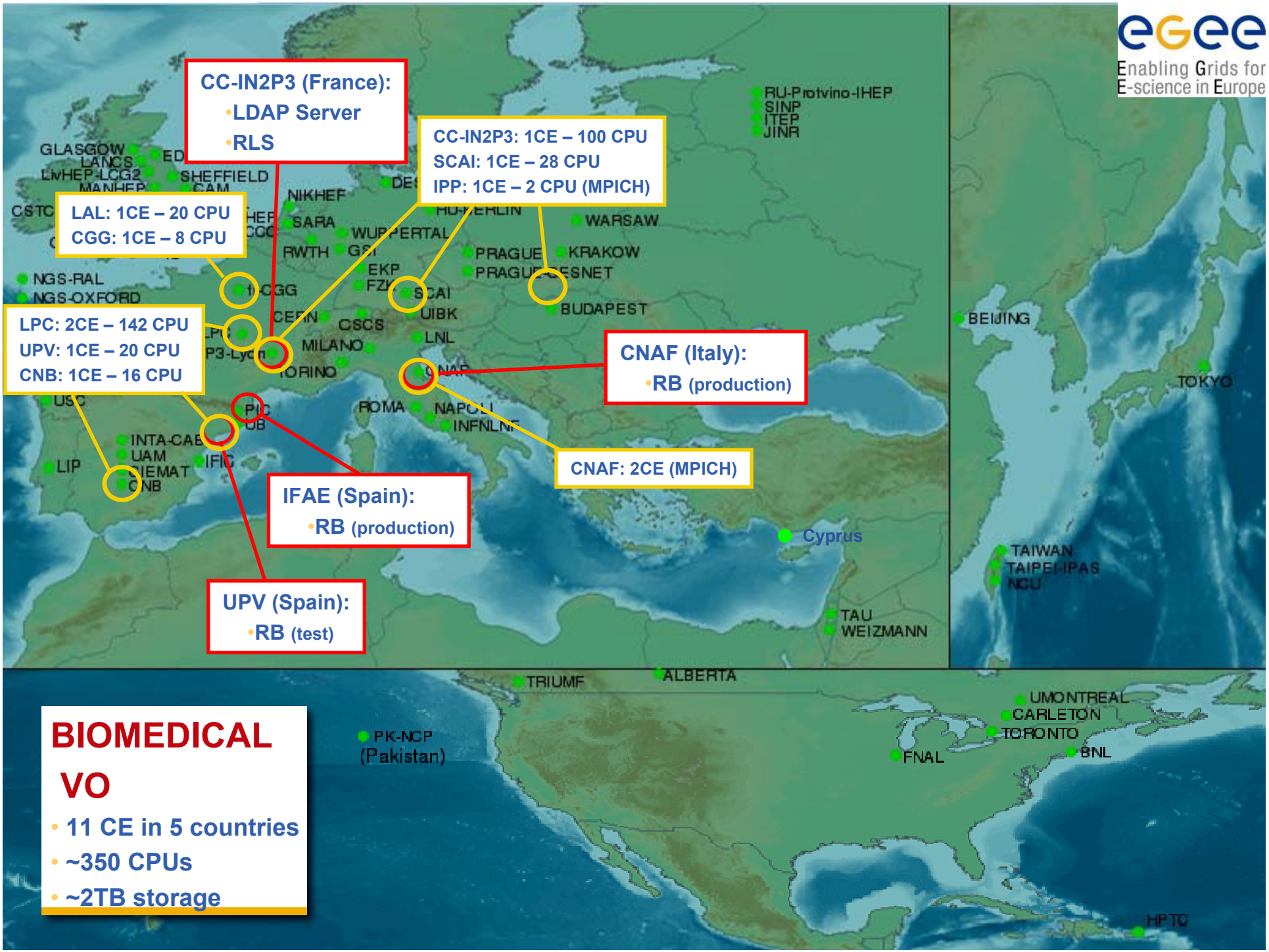
ARDA Prototypes overview

LHC Experiment	**Main focus**	**Basic prototype component**	**Experiment analysis application framework**	**Middleware prototype**
LHCb	GUI to Grid	**GANGA**	DaVinci	**gLite**
ALICE	Interactive analysis	**PROOF** **ROOT**	AliROOT	**gLite**
ATLAS	High level service	**DIAL**	Athena	**gLite**
CMS	Use of maximum native gLite functionality	**Aligned with the APROM activity**	ORCA	**gLite**

gLite components tested by ARDA and issues arising

- **The access procedure and the VOMS service behind**
- **3 shell-like mechanisms (one from original Alien shell, one built on top of gLiteIO service, and one provided by ARDA Common Application Layer)**
- **The file catalogue systems together with gLiteIO**
  - One system derived from Alien file catalogue
  - Another a new system implementing gLite Fireman interface
- **Workload management System**
  - Taskqueue system derived from Alien
  - New WMS derived from EDG
- **Job Splitting (ARDA does merging)**
- **Package Manager**

- **DEPLOYMENT ISSUE**
  - Obtaining sufficient resources for realistic testing on testbed
  - 'Early' move of gLite software to experiment sites for 'realistic' testing (e.g. in Atlas analysis data challenge)


CC-IN2P3 (France):
•LDAP Server
•RLS
CC-IN2P3: 1CE – 100 CPU
SCAI: 1CE – 28 CPU
IPP: 1CE – 2 CPU (MPICH)
LAL: 1CE – 20 CPU
CGG: 1CE – 8 CPU
LPC: 2CE – 142 CPU
UPV: 1CE – 20 CPU
CNB: 1CE – 16 CPU
CNAF (Italy):
•RB (production)
CNAF: 2CE (MPICH)
IFAE (Spain):
•RB (production)
UPV (Spain):
•RB (test)
BIOMEDICAL VO
• 11 CE in 5 countries
• ~350 CPUs
• ~2TB storage
RU-Protvino-IHEP
SINP
ITEP
JINR
GLASGOW
SHEFFIELD
NIKHEF
SARA
WUPPERTAL
RWTH
PRAGUE
KRAKOW
WARSAW
BUDAPEST
NGS-RAL
NGS-OXFORD
CSCS
MILANO
LNL
ROMA
PIC
INTA-CAB
UAM
IFIC
LIP
CIEMAT
CNB
Cyprus
TAU
WEIZMANN
BEIJING
TOKYO
TAIWAN
TAIPEI-IPAS
NCU
TRIUMF
ALBERTA
UMONTREAL
CARLETON
TORONTO
BNL
FNAL
PK-NCP
(Pakistan)
HPTC

Biomedical production jobs on EGEE

RC	machine_time used (s)	#jobs
LPC-IN2P3	81 760 675	17 399
CC-IN2P3	750 580	2 259
LAL-IN2P3	2 785 370	1 186
UPV	1 864	48
CNB	18 248 246	1 939
CGG	701 921	567
SCAI*	0	0
IPP*	0	0
CNAF*	0	0
Total	**104 248 656**	**23 398**
(hours)	**~ 29.000**	
(days)	**> 1.200**	

* **These sites have just configured their cluster to accept the Biomed VO**

- **31 registered users**
- **12 Labs/ institutes in**
- **3 countries**

Biomedical applications

- 3 High Performance applications ported on LCG2
  - SiMRI3D: medical image simulation
  - xmipp_MLRefine: molecular structure analysis
  - GATE: radiotherapy planning
- 3 High Throughput applications ported on LCG2
  - CDSS: clinical decision support system
  - GPS@: bioinformatics portal
  - gPTM3D: radiology images analysis
- New applications to join in the near future
  - Especially in the field of drug discovery

Biomed Issues (solved and unsolved)

- **MPI enable nodes accessible to the biomed VO (largely SOLVED)**
  - Italy,France,Germany,Bulgaria
- **Stable RLS service (SOLVED as of Nov 24)**
- **Fine grain security(UNSOLVED…need VOMS + new middleware)**
  - File access control. Encryption. Anonymization.
  - Will soon become critical to enlarge the community of medical users.
  - **Start by getting VOMS server for Biomedicine - urgent request**
- **Getting accounting information is very painful (ongoing development..) – we need to define better the efficient retrieval of information**
- **Batch oriented LCG2 infrastructure well adapted for High Perf applications**
  - **... but not so good for dealing with High Throughput ones (CDSS, GPS@, gPTM3D)**

forward look with Biomed Task Force

- **Objectives**
  - provide application oriented support to Biomed community **(analogy to very successful LCG/EIS group)**
- **Members and support**
  - 5 biomed engineers
  - SA1 experts…….
  - (Access to GGUS)
- **Will continue production work on LCG service and testing of gLite**
  - Testing security related software is very important

Earth Science Achievements

- **The ESR (Earth Sciences Research) VO** at SARA was created in July 2004 and is functional now using LCG resources. 17 registered users from 6 countries.
- **The EGEODE (Expanding GEOsciences on DEmand) VO** was created at IN2P3 (Lyon) in mid-October for CGG and Geocluster partners. Developing now prior to migration into EGEE production service
- **An important EGEODE application has successfully been deployed on GILDA** and demonstrated at the Second EGEE Conference in The Hague using the GENIUS portal
- **Production of ozone profiles from the satellite experiment GOME** and their validation by using LIDAR data has been run on EGEE production service

MAJOR ES ISSUES

- **Metadata catalogue**
  - To search in a large number of files
    - For GOME algorithm store ~10,000,000 profiles of 12kb in a year
  - To share and exchange data
  - Restricted access to the metadata
- **Strict data policy to access any data set (this is really important)**
- **Improve efficiency of production processing for GOME on LCG (currently 40%)**
- **Access control to licensed software (e.g. Geocluster)**

Next steps for ES

- **Other applications are being ported successfully on EGEE or GILDA**
  - **in Solid Earth Physics (Earth Quake mechanisms in Asia)**
  - **In Hydrology for management of coastal waters in the Mediterranean region**
- Metadata developments (discussions going on now.).
  - test OGSA–DAI - successful
  - Web interface in gLite to be tested
- Development of the VOs (sites and services)
- Licensed s/w on sites (secure server or multi keys for access control)
- Secure access to data (wait for new software..)
- VOMS installation (after review)

Achievements in Comp Chemistry

- A cluster of 13 nodes + CE + SE + VOMS server has been deployed in GILDA for dedicated use by CompChem.
- The Grid based Molecular Simulator(GEMS) has been ported onto the GILDA test cluster and interfaced to GENIUS
- The CompChem VO has been activated
- Work in hand now to move to production service

Issues related to ComChem VO

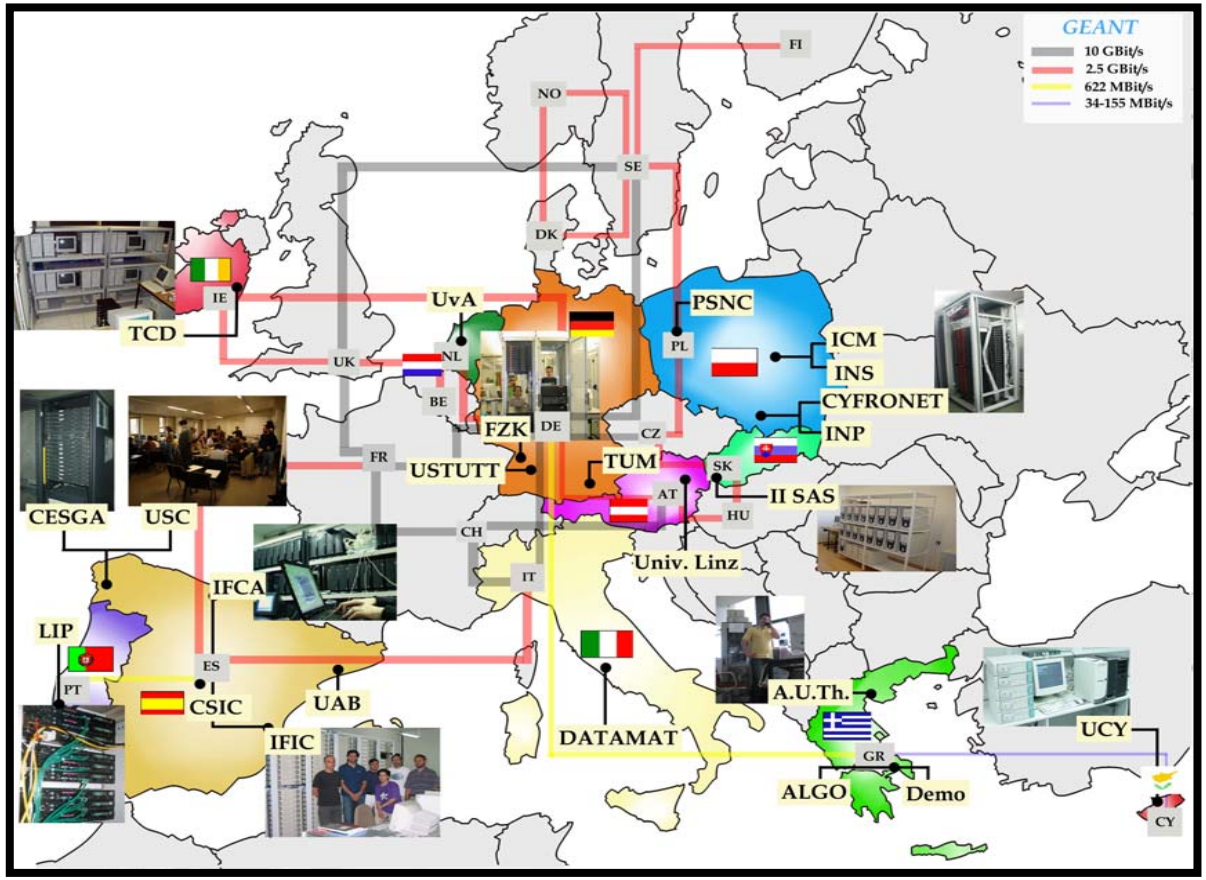
- **The implementation of the prototype has emphasised the following key requirements**
  - **Interactive Access** to the Grid through Sockets
    - We use the supported set of ports: 20000-25000, **left open by the firewall** in all nodes of the Grid
    - We need a **small latency** to schedule the job on the Grid
  - **Parallel** and distributed nature of most CompChem programs
    - **MPI prerequisite** in all nodes
    - We need the **highest possible number of Working Nodes**
  - **Commercial Licences** of some software
    - Some software requires a Commercial Licence even if used for Research purposes.
    - A strategy for acquiring the licences for the users of the Grid should be defined.

Next steps for Comp Chem

- Starting the deployment in EGEE using the sites currently in GILDA
- A consortium agreement will be signed with the participating institutions.
- Will contribute to the Grid with new resources and expertise.
- A new COST D23 Action called GridChem will be started next year following the ending initiative MetaChem.

MAGIC @ CrossGrid and EGEE

- Have been using LCG-2 middleware with CrossGrid extensions
- 16 sites in Europe



The first prototype for a MC production system for the MAGIC telescope exists
→A Europeanwide system is being developed together with CNAF, PIC and GridKA in EGEE

Issues arising from original work by MAGIC

- Some small technical matters which were resolved with help from GILDA team
- Found documentation difficult to find
- Definitely need simple ‘EGEE for dummies’!
- Procedure in EGEE for getting more resources, if only for a data challenge, (at Den Haag regions pledged support....)

The plans for the future and next steps

- **A Magic Virtual Organisation already exists in EGEE**
  - VO server is hosted by SARA/NIKHEF
  - Successful first running in GILDA
- **Developments underway**
  - CNAF will support the Magic VO with a Resource Broker
  - PIC will support the Magic VO with storage and the RLS CNAF, PIC and GridKA will provide CPU
  - GILDA can be used for the first test as well
- **Others sites can join the MC data challenge Feb 2005**
  - How to organise this?

CONCLUSIONS

- **The 2 pilot applications (HEP and Biomedicine)** have made excellent progress in the migration of applications to the production service and the large scale of this service - excellent relations with LCG
- **The 3 new application areas** made significant progress suing GILDA testbed, and are ready to go to production service (Earth science already did)
- **We look now to maintain the improvements on the production service** in key areas of data management, site certification and monitoring,job monitoring and trouble-shooting, and accounting
- **The move to use of gLite software will need to be controlled very carefully** since service must maintain high level of performance (i.e. not go backwards!)