

DC'04 News

Nick Brook

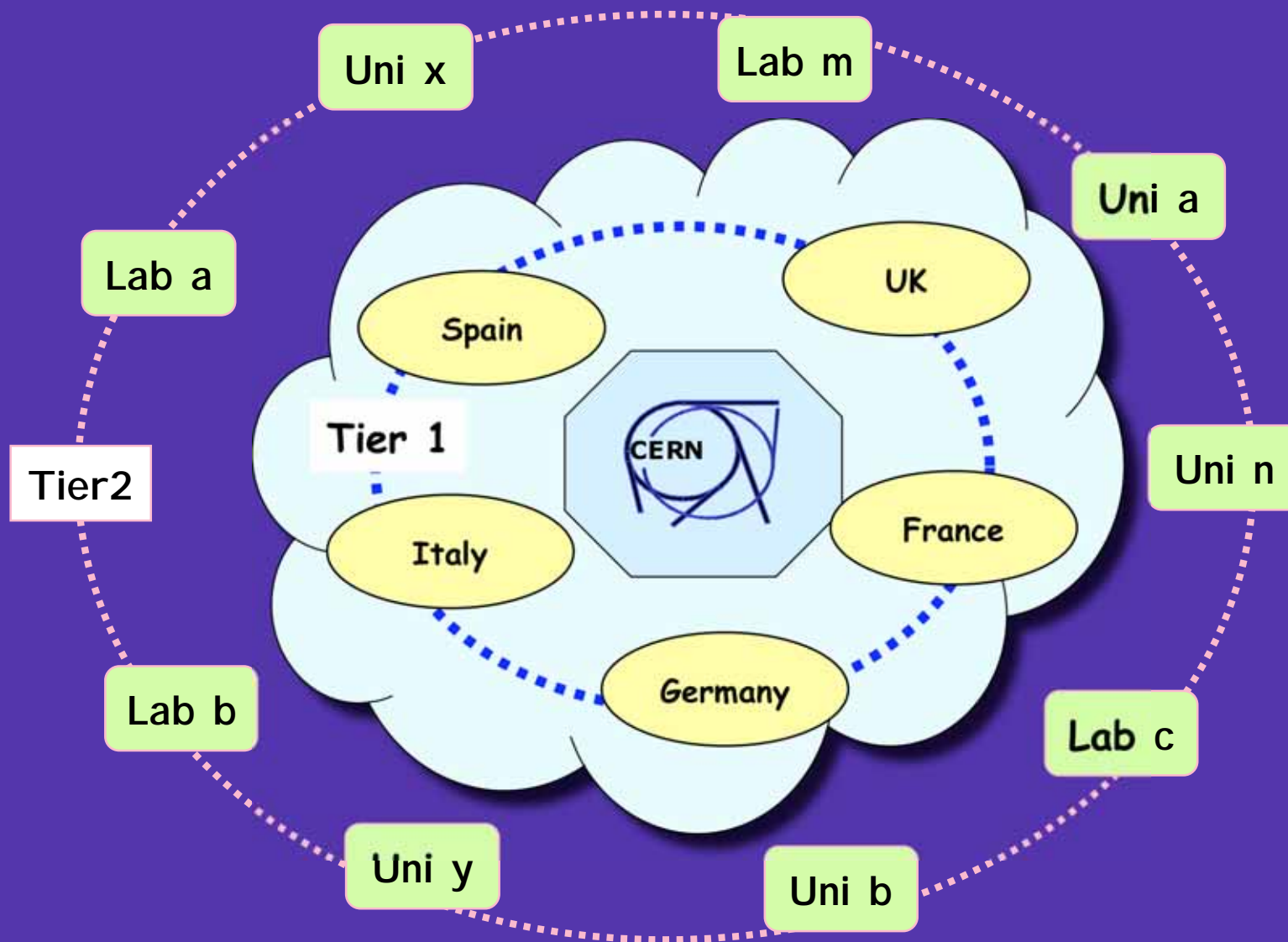


University of Bristol

Goals of DC'04

- Main goal: gather information to be used for writing the LHCb computing TDR/TP
 - Robustness test of the LHCb software and production system
 - Using software as realistic as possible in terms of performance
 - Test of the LHCb distributed computing model
 - Including distributed analyses in LCG
 - realistic test of analysis environment, need realistic analyses
 - Incorporation of the LCG application area software into the LHCb production environment
 - Use of LCG resources as a substantial fraction of the production capacity

Architecture of LHCb DC'04



Data Distributed

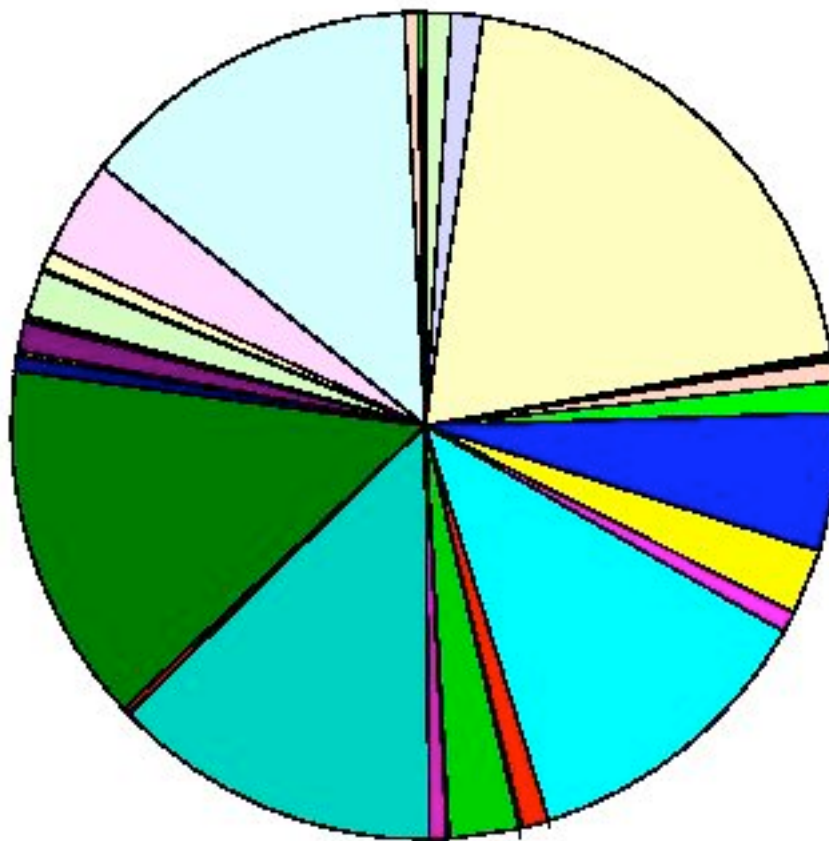
- All data DST stored at CERN
- Another copy stored at Tiers 1 centres:
 - PIC (Barcelona)
 - CNAF (Bologna) } CASTOR
 - RAL (Didcot)
 - FZK(Karlsruhe)
 - IN2P3/CNRS (Lyon)

Production Running

- DIRAC agent deployed on each LHCb dedicated site
- One single agent for LCG.
- Some sites offer both dedicated LHCb and LCG resources
 - Current production is dominated by non-LCG resources
 - Move to LCG once production on LCG more robust - on going process

Production Snapshot

Total Running Jobs: 2980
DIRAC: 63% LCG: 37%

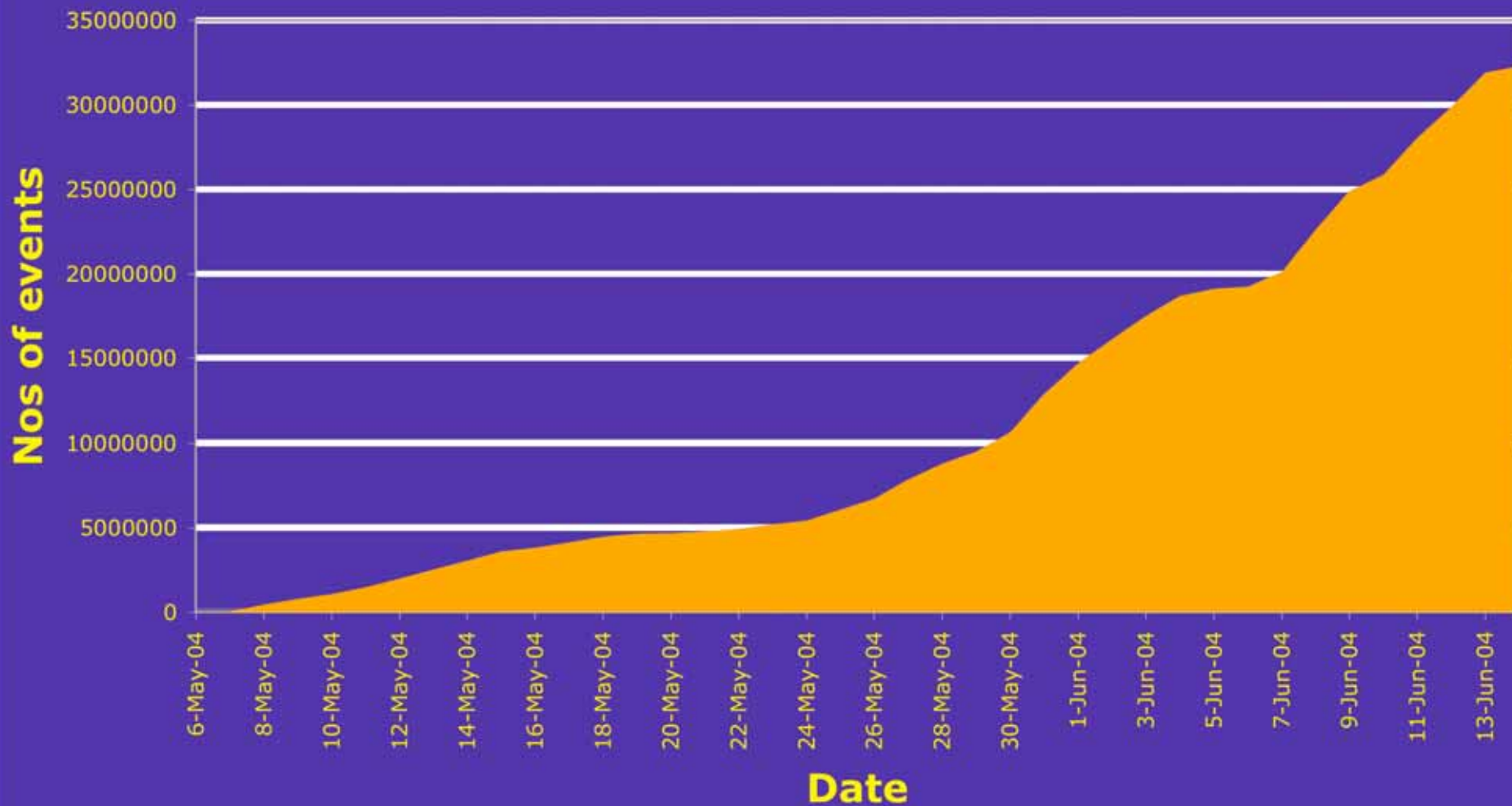


DIRAC.Barcelona.es	1.006%
DIRAC.Bologna.it	1.241%
DIRAC.CERN.ch	19.93%
DIRAC.CracowAgu.pl	0.167%
DIRAC.IF-UFRJ.br	0.134%
DIRAC.IHEP-Protvino.ru	0.805%
DIRAC.IHEP2-Protvino.ru	1.208%
DIRAC.ITEP-Moscow.ru	5.402%
DIRAC.Imperial.uk	2.583%
DIRAC.JINR-Dubna.ru	0.805%
DIRAC.Karlsruhe.de	11.87%
DIRAC.LHCBOONLINE.ch	1.040%
DIRAC.Liverpool.uk	0.134%
DIRAC.Lyon.fr	2.718%
DIRAC.Manno.ch	0.067%
DIRAC.Oxford.uk	0.033%
DIRAC.Santiago.es	0.671%
DIRAC.ScotGrid.uk	12.88%
DIRAC.Zurich.ch	0.268%
LOG.CERN.ch	14.12%
LOG.CNAF.it	0.604%
LOG.Cambridge.uk	0.201%
LOG.Imperial.uk	1.107%
LOG.Krakow.pl	0.134%
LOG.Legnaro.it	0.134%
LOG.Milano.it	1.812%
LOG.NCU.tw	0.100%
LOG.NIKHEF.nl	0.738%
LOG.PIC.es	3.825%
LOG.RAL.uk	13.42%
LOG.Torino.it	0.536%
LOG.Triumpf.ca	0.201%
LOG.USC.es	0.067%

Jun 14 2004, 12:20

Production Snapshot

Data Produced



DIRAC sites

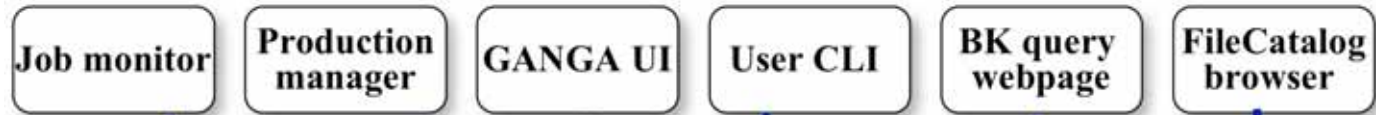
- 14 sites in production

- 4 Russia
- 4 UK
- 2 Spain
- 2 Switzerland
- 1 Italy
- 1 France
- 1 Poland
- 1 Germany
- 1 Brazil
- 2 CERN

- Barcelona
- Bologna
- Cern (x2 incl online farm)
- Cracow
- If-ufrj
- I hep-protvino(x2)
- Imperial
- I tep-moscow
- Jlnr-dubna
- Karlsruhe
- Liverpool
- Lyon
- Manno
- Oxford
- Santiago
- ScotGrid
- Zurich

DIRAC Status

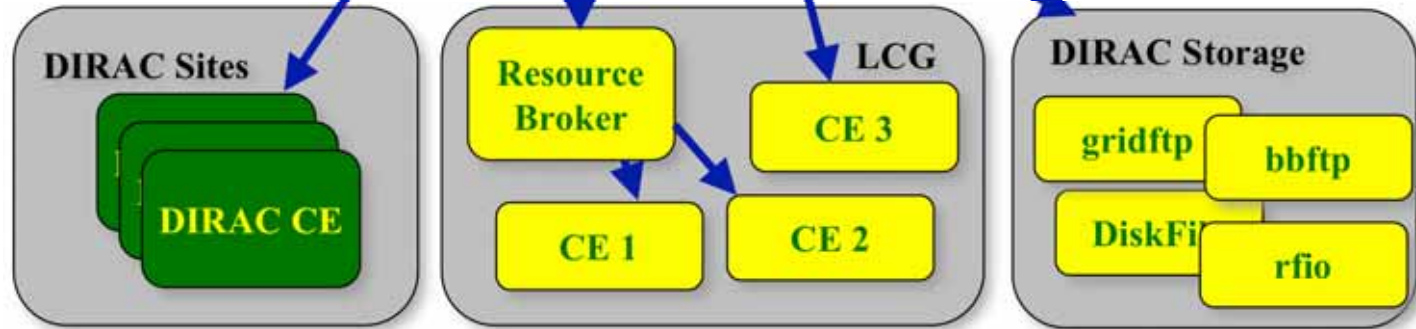
User interfaces



DIRAC services



DIRAC resources



DIRAC & File Catalogs

- Using both LHCb Bookkeeping and AliEn File catalog:
 - Getting experience
 - Redundancy
 - Behaving reliably
- Plans to use LCG RLS
 - Tools being developed to cross populate/register

LCG Production experience

- invaluable central LCG support
- No major problems with LCG
 - Very few jobs failing due to LCG problem
- **File Transfers !** - problems transfer with BBFTP, SFTP, GridFTP (not just a LCG problem)
 - This has led to many failed jobs
- Debugging problems is very time consuming and difficult
 - Lack of returned info & need to involve local LCG ops.

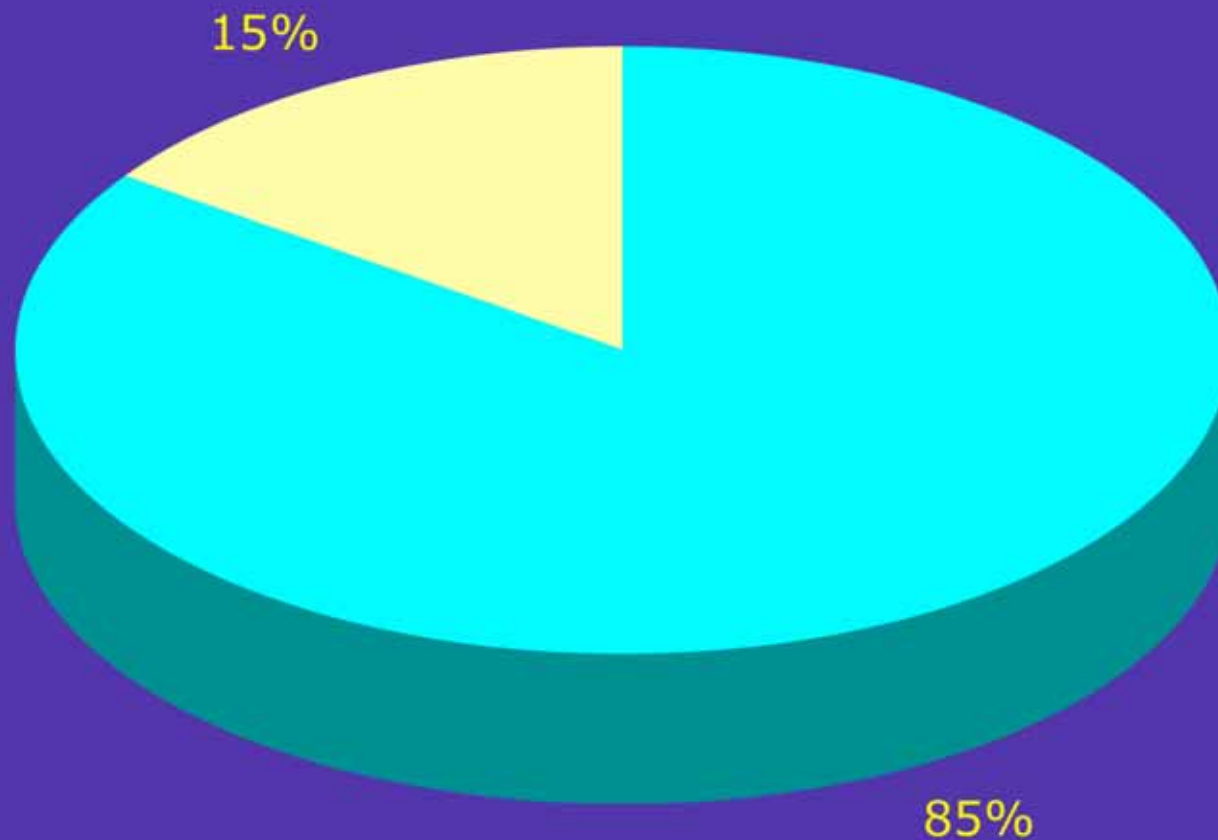
LCG Production experience

- **Normalised Queue Time Limit**
 - an ad-hoc solution for this problem being implemented
 - CERN only site that advertises non-WallClocktime - perhaps correct approach but exception causes headache
- **Tools to handle large number of production jobs**
 - Job submission
 - Status via API's - C++ & Python
- **Issues with proxies for long jobs**
 - Jobs are submitted with 3 day long proxy and "some" are aborted with "Proxy expired" reason after few hours.

LCG Production experience

Event Production

■ DIRAC ■ LCG



LCG Production experience

- 14 LCG sites

- 4 Italy
- 3 UK
- 2 Spain
- 1 Switzerland
- 1 Taiwan
- 1 USA
- 1 Germany
- 1 Poland
- 1 Canada
- CERN

- Cambridge
- CERN
- CNAF
- Imperial
- Krakow
- Legnaro
- Milano

- NCU
- Nikhef
- PIC
- RAL
- Torino
- Triumf
- USC

Also have used: Karlsruhe, FNAL

In process of adding: Toronto &
KFKI (Hungary)

Good working relationship with site managers
where we have had problems

Next steps

- Reprocessing jobs:
 - E.g. stripping
 - Need scheduling dependant on the input data
 - interrogates File Catalog then chooses destination site(s)
 - Input data specified as logical file names
 - Physical file names are resolved locally automatically using POOL/ROOT mechanisms

Next steps - preparation for analysis

- DaVinci application is packaged and usable as any other Gaudi application:
- DaVinci job for DIRAC or LCG is prepared in the same way:
 - GANGA
 - Command line submission:
 - After user job options and dll's are prepared to be shipped with a job

Summary

- DC'04 started with brand new DIRAC
 - components allow flexible use for different scenarios
- MC production, with data replication to Tier1 centres, running relative smoothly
 - No major show stoppers but a number of niggling problems are manpower intensive
- LCG-2 beginning to be fully incorporated in production
- More complex scenarios need to be prepared for: reprocessing and analysis.