UNIVERSITY OF BIRMINGHAM

e-science in the UK

Peter Watkins

Head of Particle Physics

University of Birmingham, UK

p.m.watkins@bham.ac.uk

Outline of Talk

- UK e-science funding
- e-science projects
- e-science infrastructure
- GridPP
- Recent developments

RCUK e-Science Funding

First Phase: 2001 –2004

- Application Projects
  - £74M
  - All areas of science and engineering
- Core Programme
  - £15M Research infrastructure
  - £20M Collaborative industrial projects
  - (+£30M Companies)

Second Phase: 2003 –2006

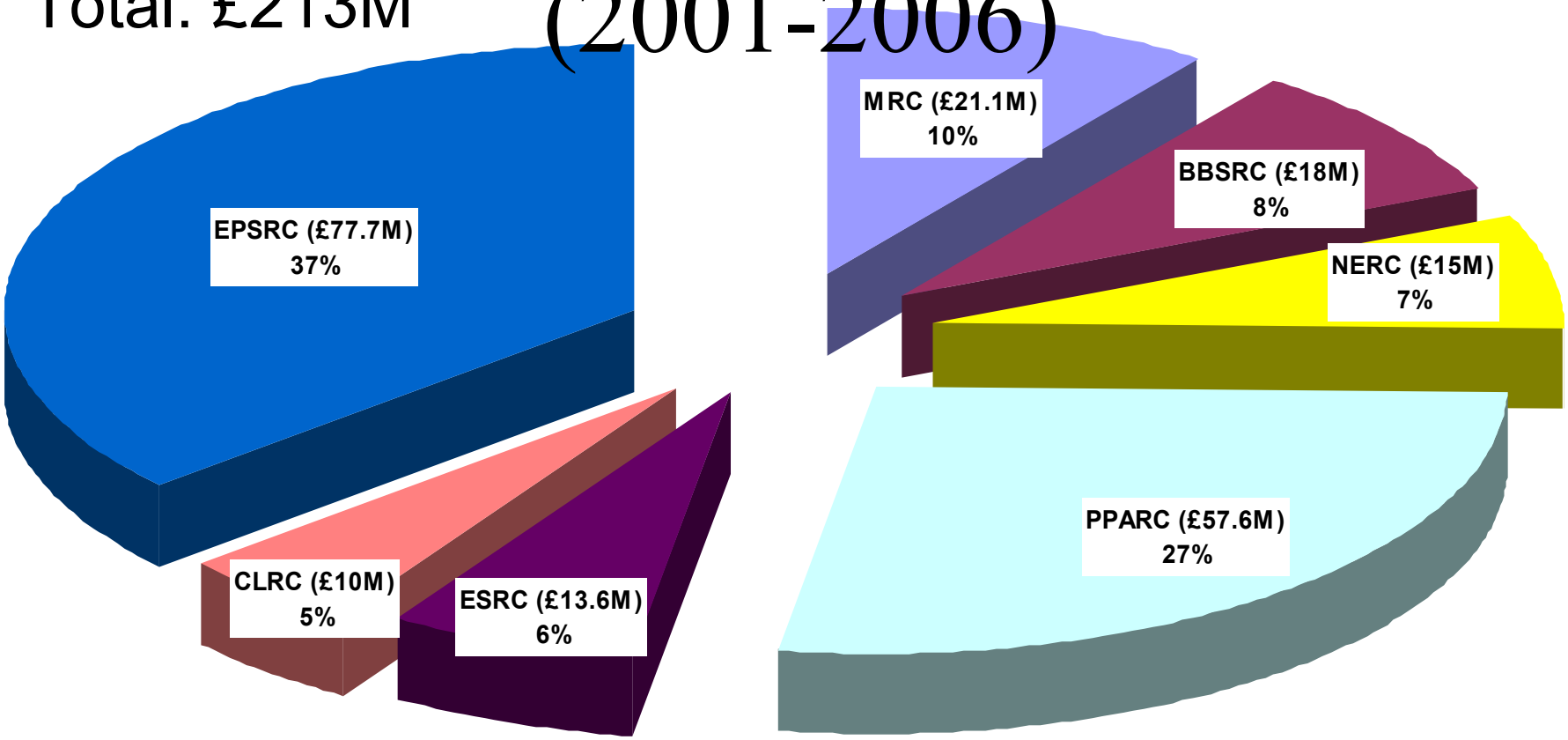
- Application Projects
  - £96M
  - All areas of science and engineering
- Core Programme
  - £16M Research Infrastructure
  - £10M DTI Technology Fund

The UK e-Science Experience: Phase 1

- All Research Council e-Science funds committed
  - e-Science pilots launched covering many areas of science, engineering and medicine
- UK e-Science Core Programme
  - DTI £20M for collaborative industrial R&D
- ➢ About 80 UK companies participating
- ➢ Over £30M industrial contributions
  - Engineering, Pharmaceutical, Petrochemical
  - IT companies, Commerce, Media

UK e-Science Budget (2001-2006)

Total: £213M



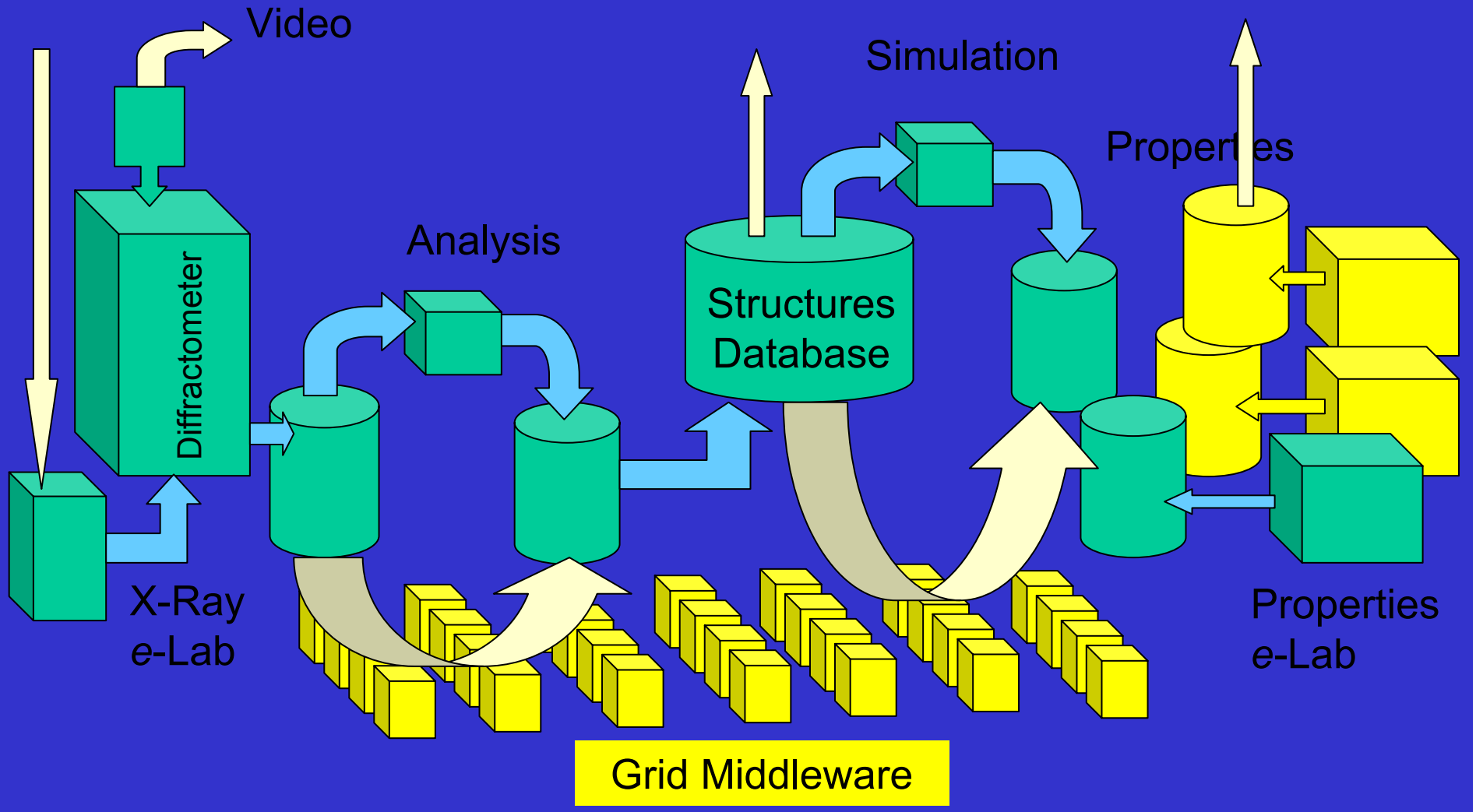
Source: Science Budget 2003/4 – 2005/6, DTI(OST)

EPSRC e-science projects (a few examples)

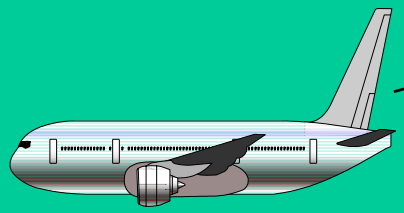
The e-Science Paradigm

- The Integrative Biology Project involves the University of Oxford (and others) in the UK and the University of Auckland in New Zealand
  - Models of electrical behaviour of heart cells developed by Denis Noble's team in Oxford
  - Mechanical models of beating heart developed by Peter Hunter's group in Auckland
- Researchers need to be able to easily build a secure 'Virtual Organisation' allowing access to each group's resources
  - Will enable researchers to do different science

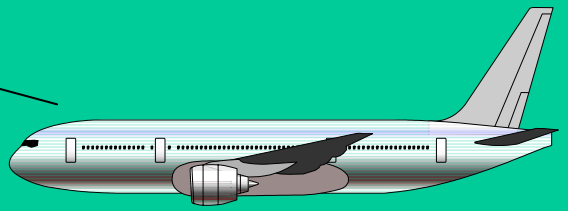
Comb-*e*-Chem Project



DAME Project



In flight data



Global Network
eg: SITA

Ground
Station

Airline

DS&S Engine Health Center

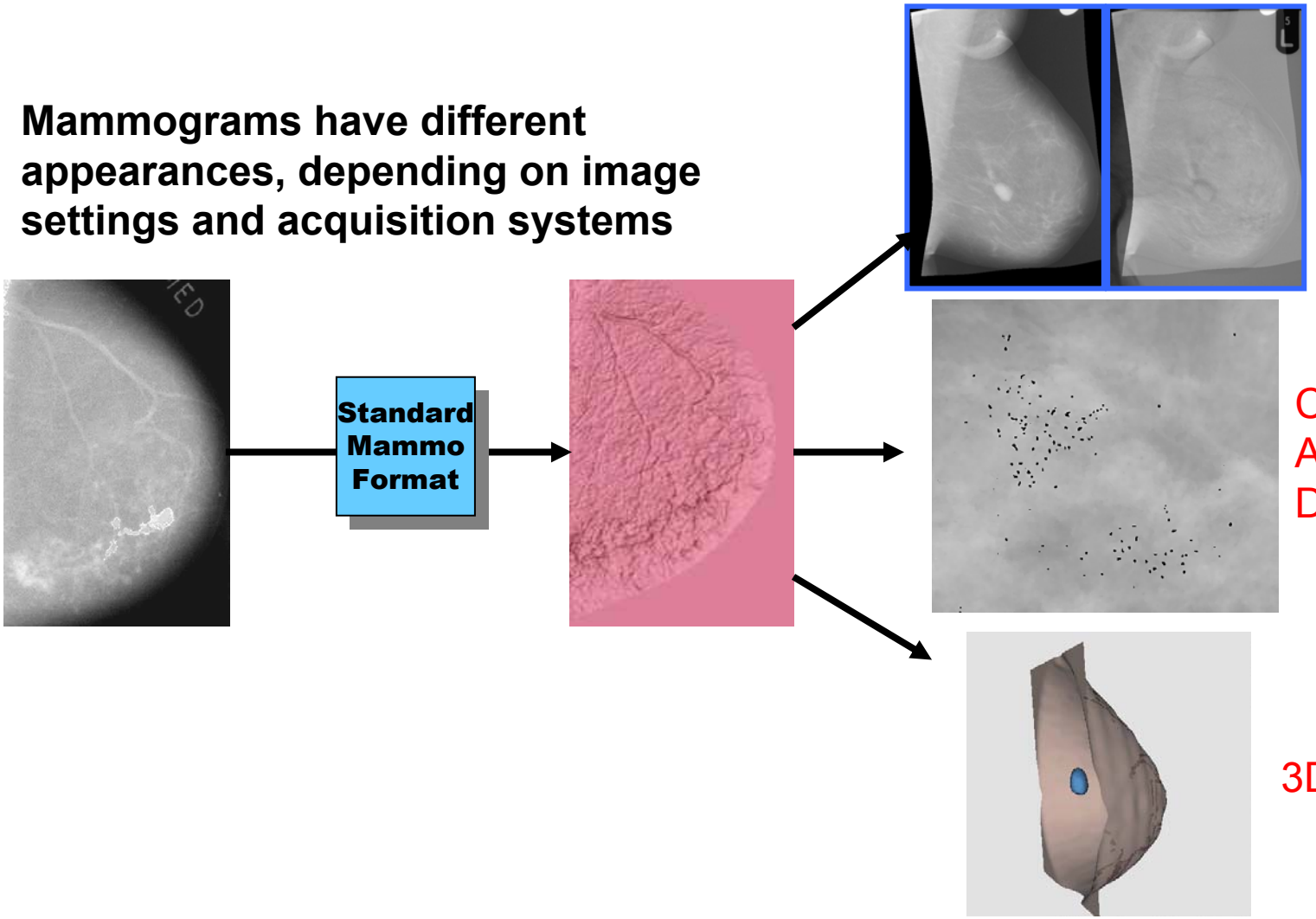
Data centre

Internet, e-mail, pager

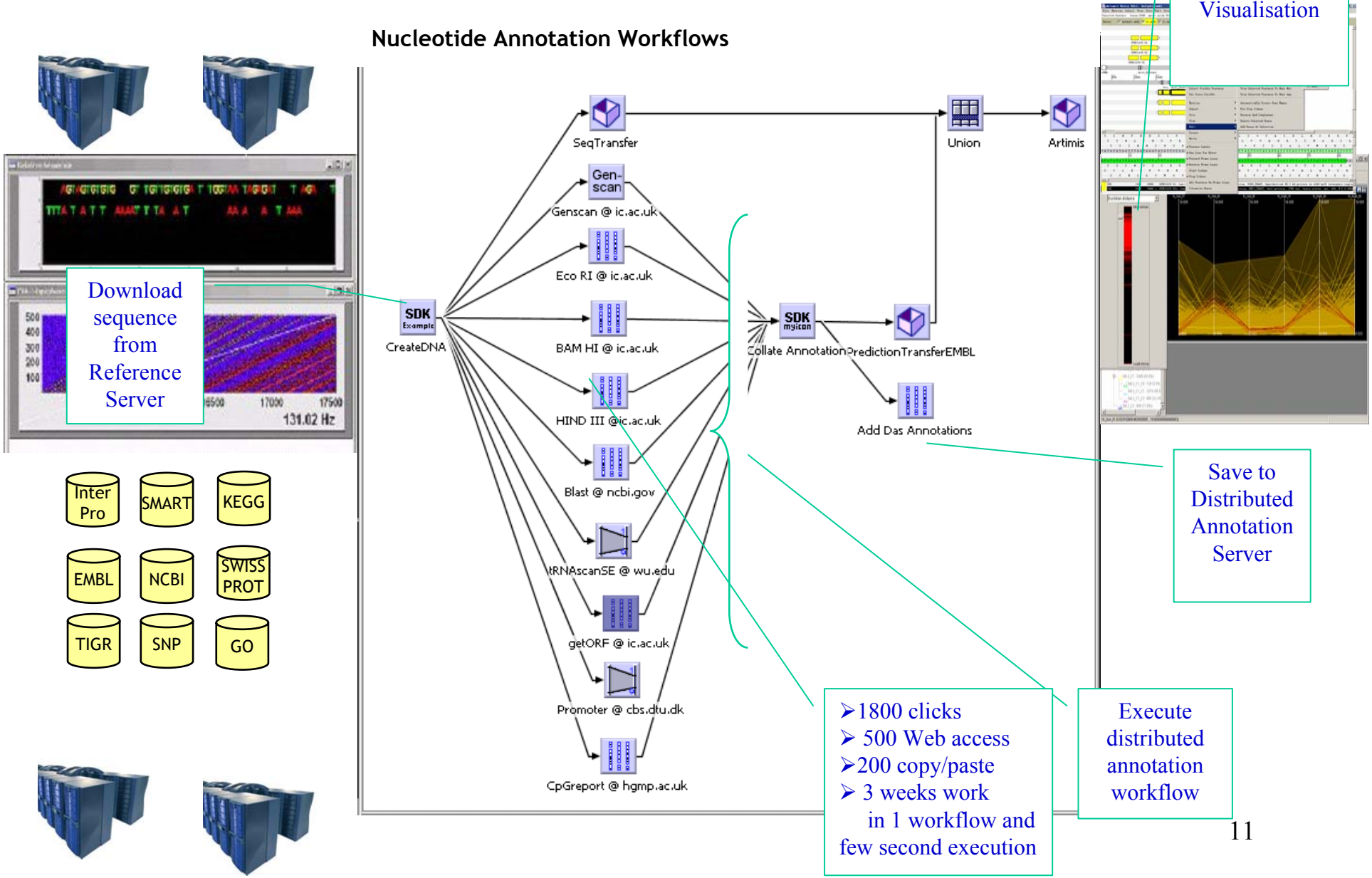
Maintenance Centre

eDiaMoND Project

Mammograms have different appearances, depending on image settings and acquisition systems



Discovery Net Project



UK Core e-Science programme -

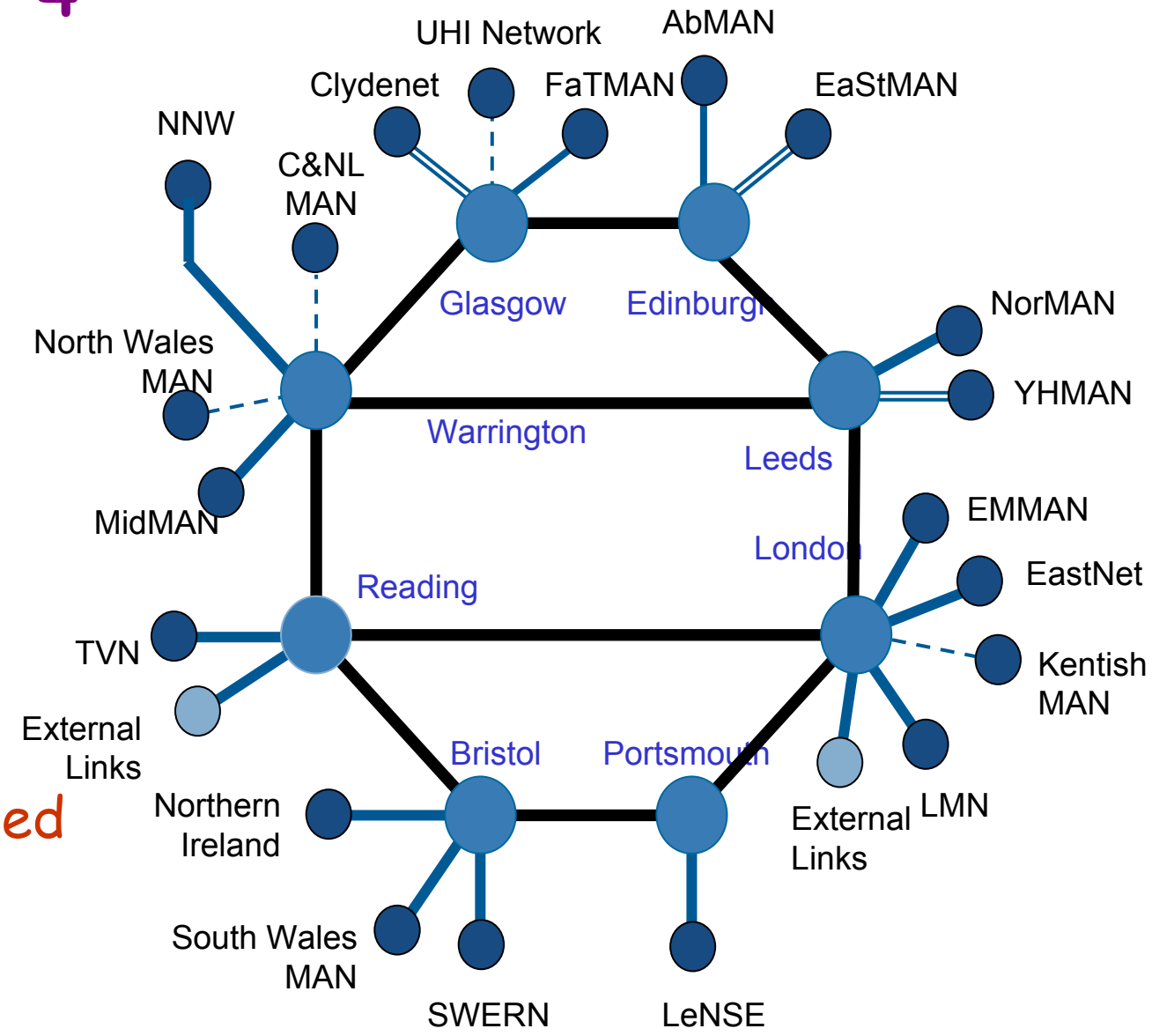
Access Grid

National e-science centre
NeSC – Edinburgh

- Help coordinate and lead UK e-Science
  - Community building & outreach
  - Training for UK and EGEE
- Undertake R&D projects
  - Research visitors and events
  - Engage industry (IBM, Sun, Microsoft, HP, Oracle, …)
  - Stimulate the uptake of e-Science technology

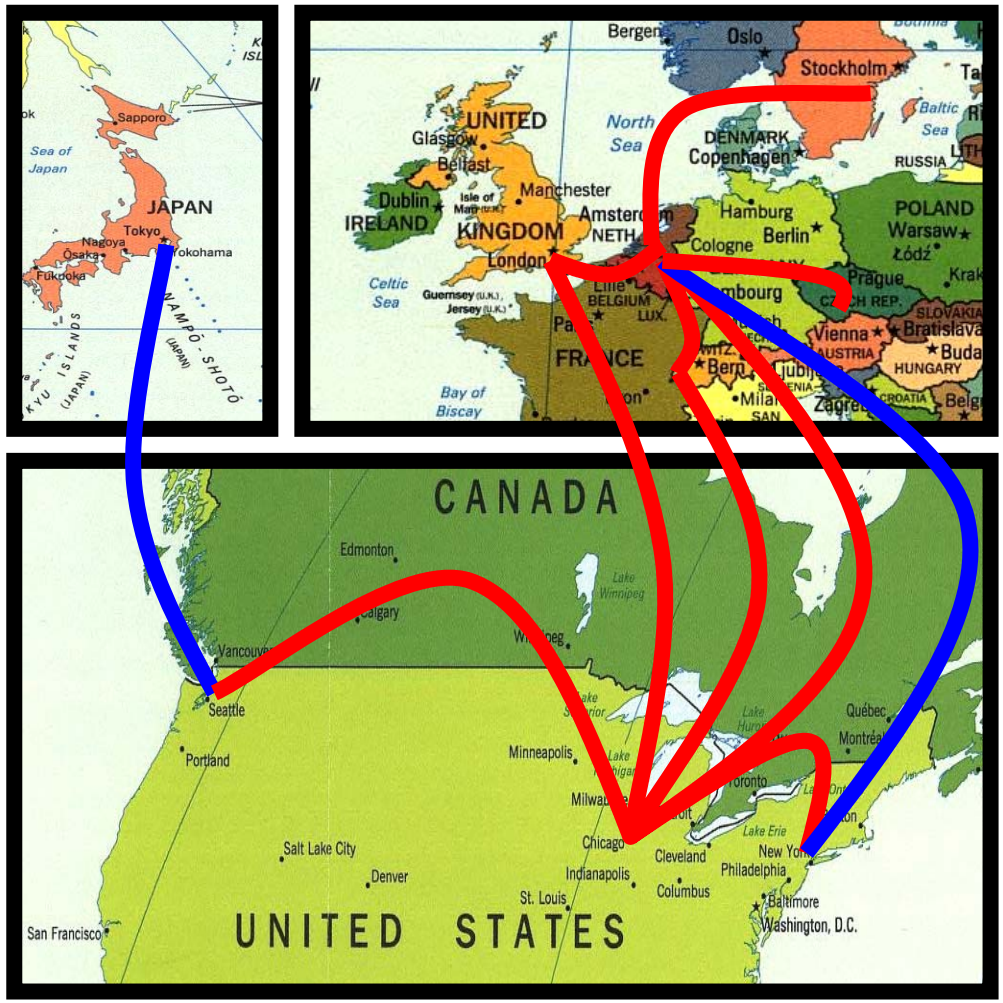
SuperJANET 4

- UK National Network
- Managed by UKERNA
- 10 Gbit/s core
- Regional Distribution
- IP Production Service
- Multicast enabled
- IPv6 rollout
- QoS Rollout
- MPLS Capable in core

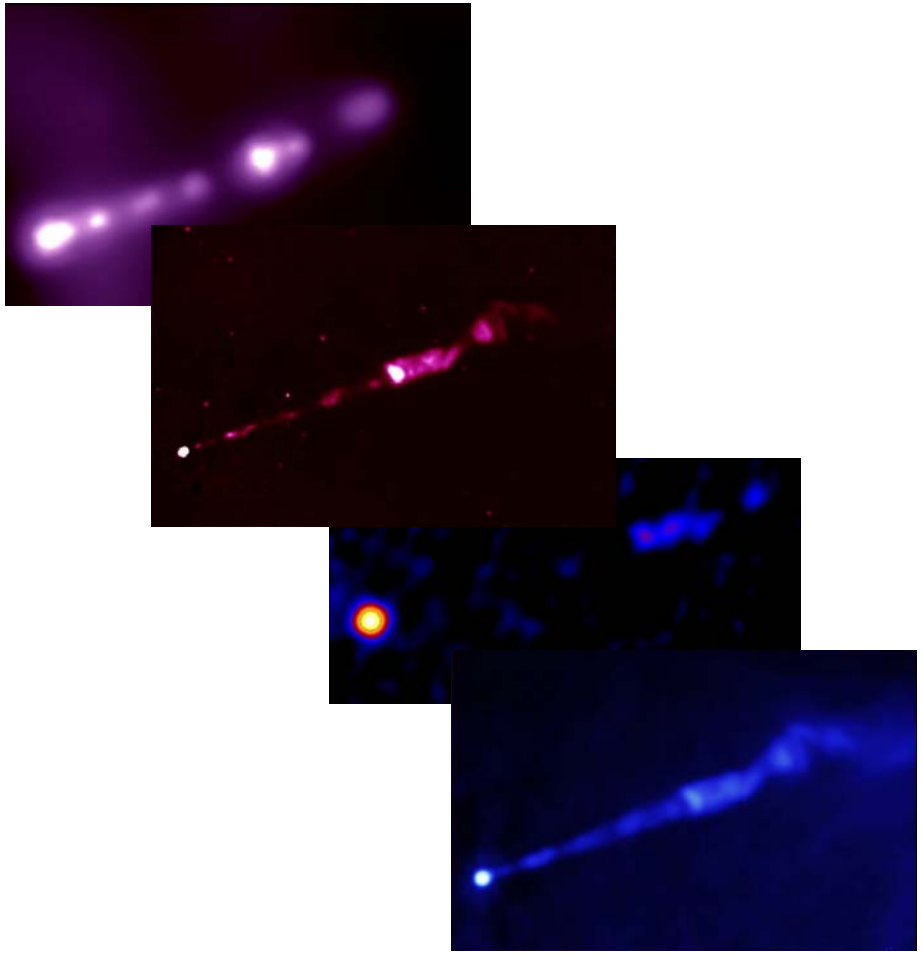


UKLIGHT is….

- UK Joining Global optical testbed
- R&D links to NL and USA
- 10 Gbit/s
- Available for leading edge network projects
- Designed to pilot new ways of networking
- Motivated on "lightpath" switching below IP layer



PPARC e-science
- two major projects



Powering the Virtual Universe

http://www.astrogrid.ac.uk

(Edinburgh, Belfast, Cambridge, Leicester, London, Manchester, RAL)

Multi-wavelength showing the jet in M87: from top to bottom – Chandra X-ray, HST optical, Gemini mid-IR, VLA radio. AstroGrid will provide advanced, Grid based, federation and data mining tools to facilitate better and faster scientific output.

Picture credits: "NASA / Chandra X-ray Observatory / Herman Marshall (MIT)", "NASA/HST/Eric Perlman (UMBC), "Gemini Observatory/OSCIR", "VLA/NSF/Eric Perlman (UMBC)/Fang Zhou, Biretta (STScI)/F Owen (NRA)"

The Virtual Observatory

- International Virtual Observatory Alliance

UK, Australia, EU, China, Canada, Italy, Germany, Japan, Korea, US, Russia, France, India

How to integrate many multi-TB collections of heterogeneous data distributed globally?

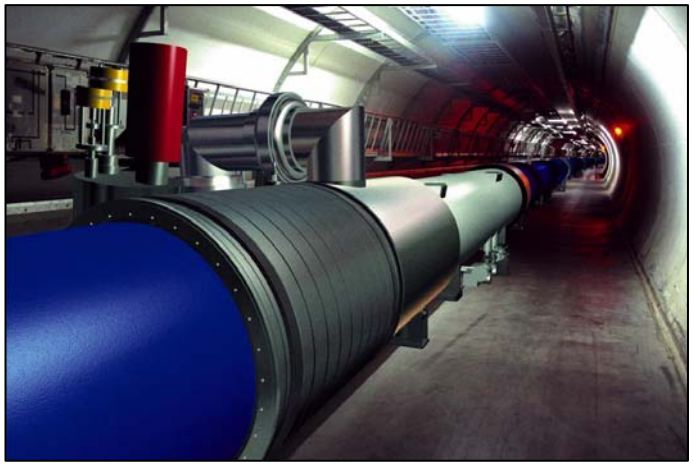
Sociological and technological challenges to be met

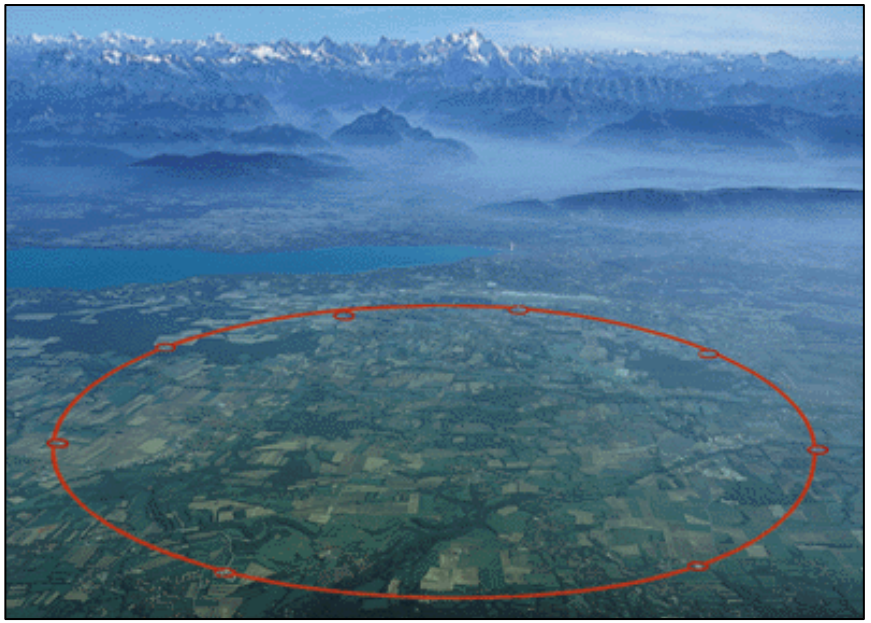
GridPP – Particle Physics Grid

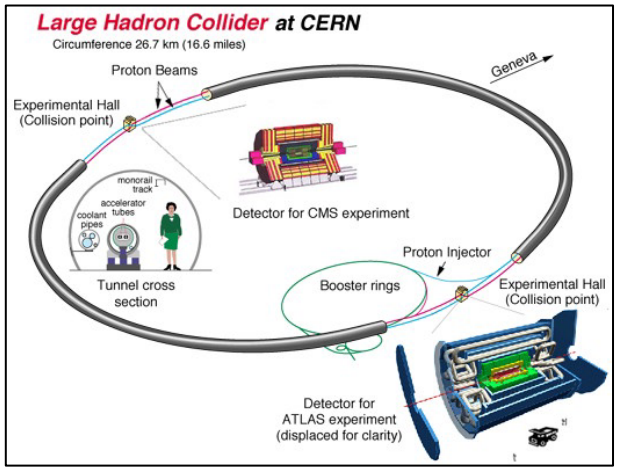
- 19 UK Universities, CCLRC (RAL & Daresbury) and CERN
- Funded by the Particle Physics and Astronomy Research Council (PPARC)
- GridPP1 - 2001-2004 £17m "From Web to Grid"
- GridPP2 - 2004-2007 £15m "From Prototype to Production"

The CERN LHC

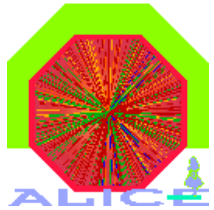
The world's most powerful particle accelerator - 2007







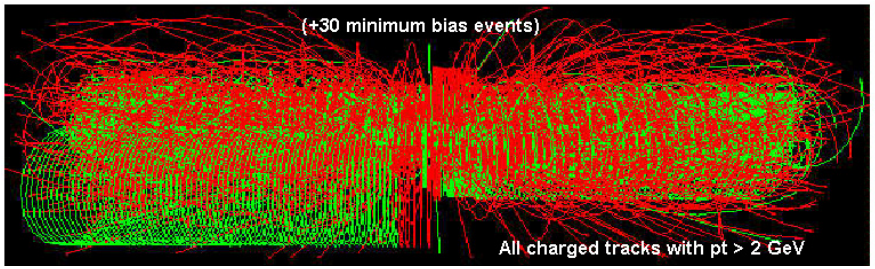
4 Large Experiments



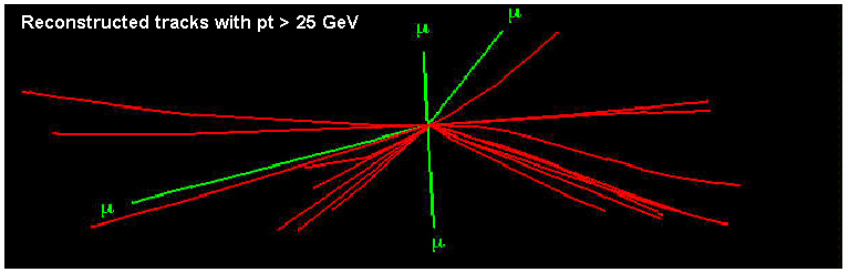
LHC Experiments

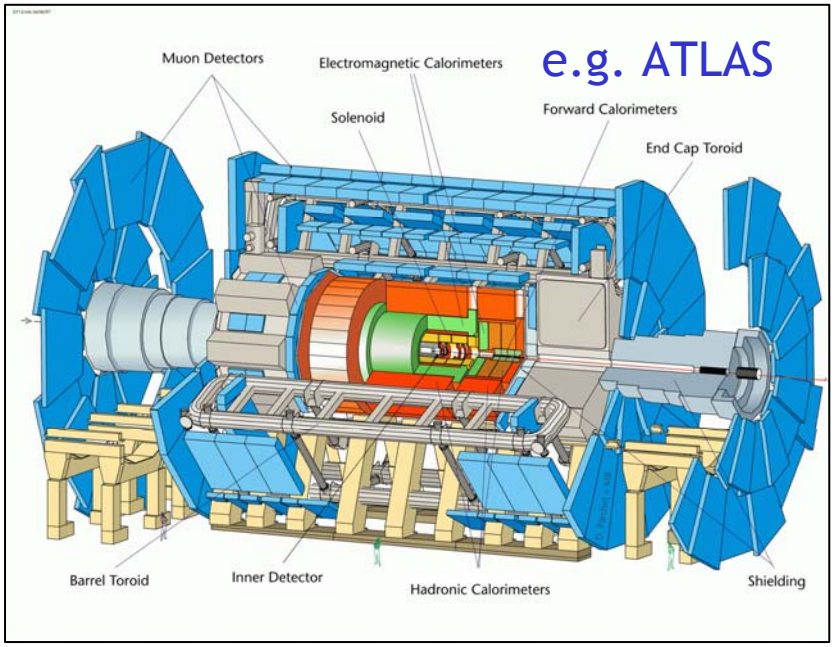
Searching for the Higgs Particle and exciting new Physics

Starting from this event



Looking for this 'signature'

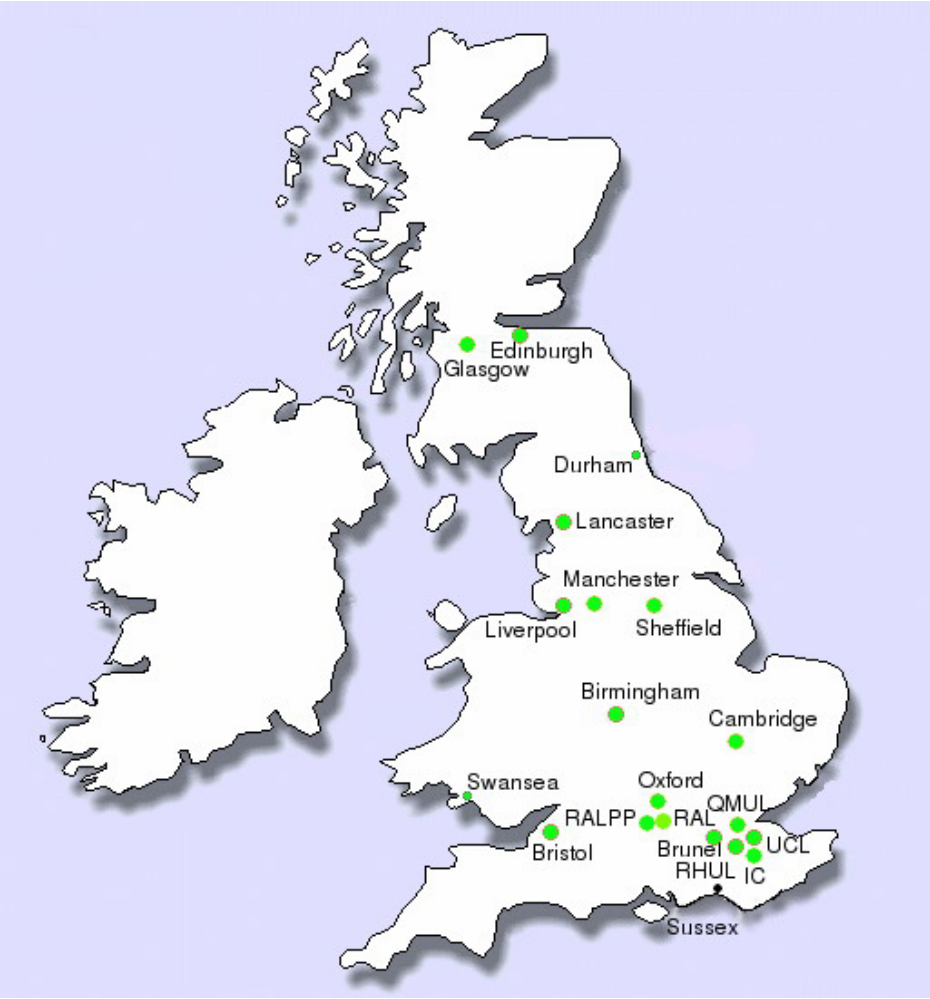




- $> 10^8$ electronic channels
- 8×10^8 proton-proton collisions/sec
- 2×10^{-4} Higgs per sec
- 10 Petabytes of data a year
- (10 Million GBytes = 14 Million CDs)

ATLAS CAVERN (end 2004)

UK GridPP

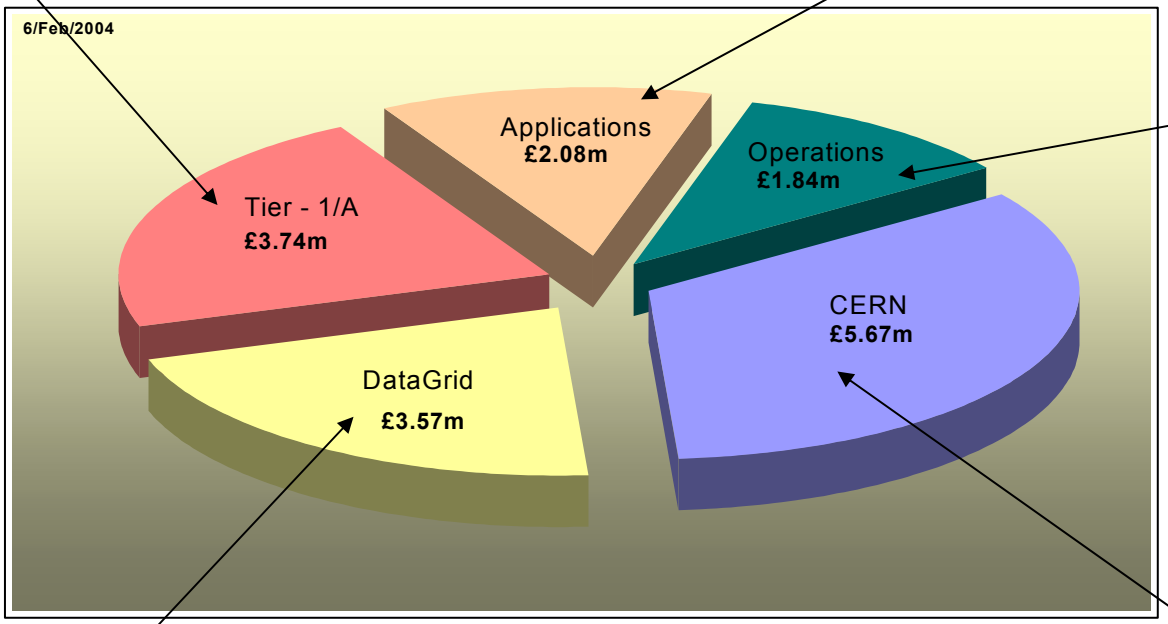

Edinburgh
Glasgow
Durham
Lancaster
Manchester
Liverpool
Sheffield
Birmingham
Cambridge
Swansea
Oxford
QMUL
RALPP
RAL
Bristol
Brunel
UCL
RHUL
IC
Sussex

GridPP1 Areas

GridPP
UK Computing for Particle Physics

UK Tier-1/A Regional Centre
Hardware and Manpower

Grid Application Development
LHC and US Experiments +
Lattice QCD



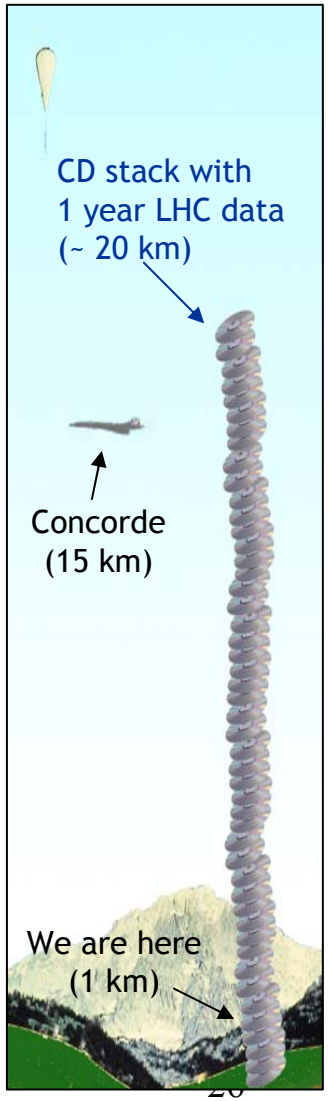
Management
Travel etc

European DataGrid (EDG)
Middleware Development

LHC Computing Grid Project (LCG)
Applications, Fabrics, Technology
and Deployment

Challenges

- Scaling to full size ~10,000 → 100,000 CPUs
- Stability, Robustness etc
- Security
- Sharing resources (in RAE environment!)
- International Collaboration
- Continued funding beyond start of LHC!

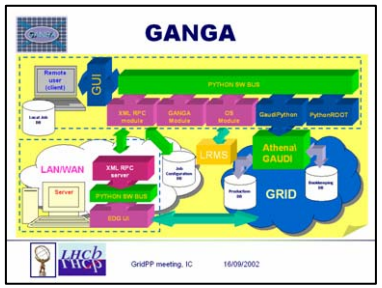


Application Development

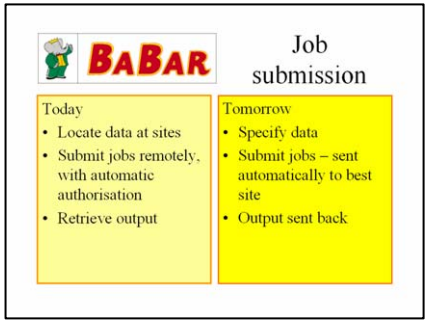
GridPP
UK Computing for Particle Physics

LHCb

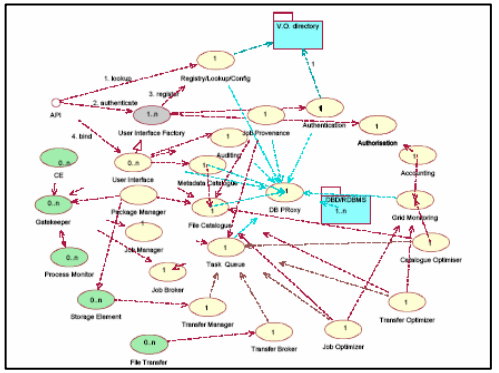
GANGA



BaBar



AliEn → ARDA



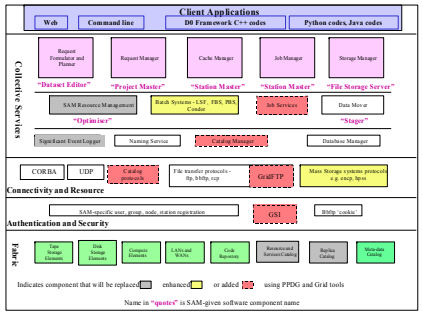
ALICE

UKQCD collaboration

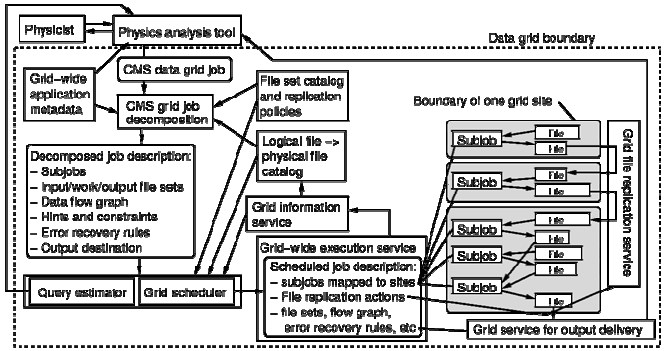
D0

CDF

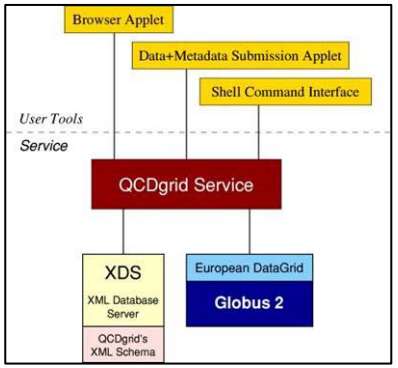
SAMGrid



CMS



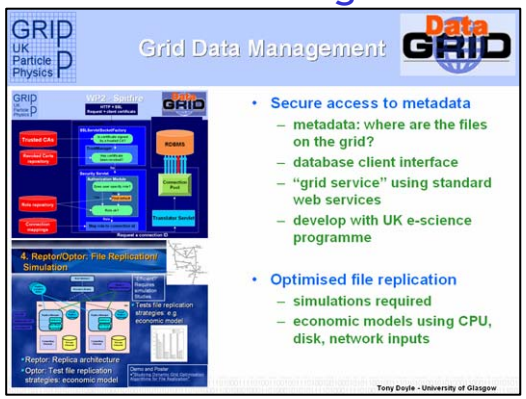
Lattice QCD



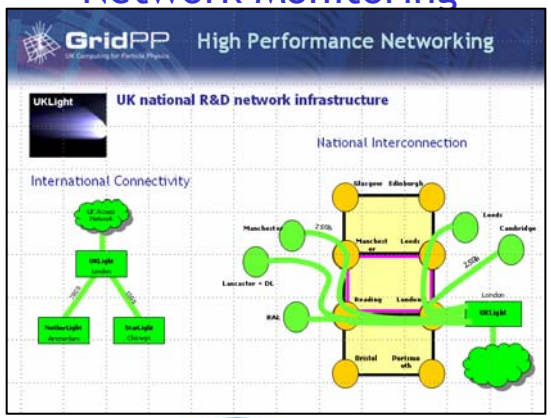
Middleware Development

GridPP
UK Computing for Particle Physics

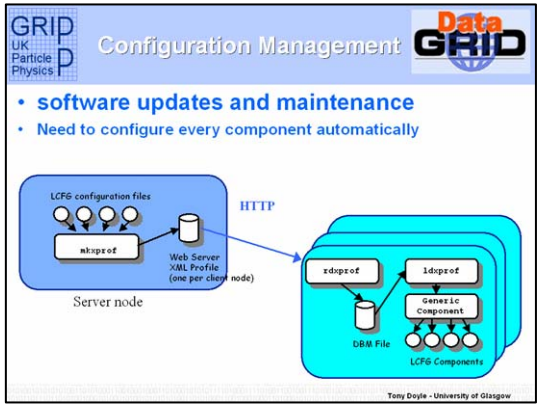
Grid Data Management



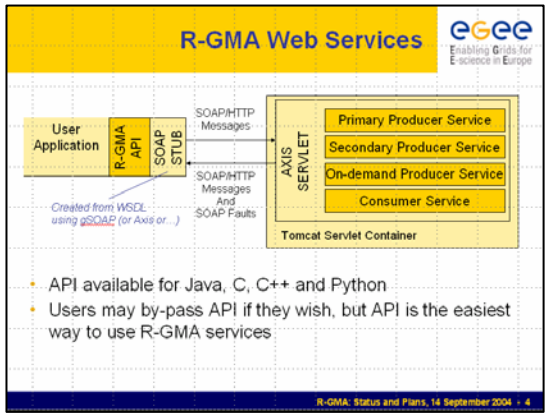
Network Monitoring



Configuration Management



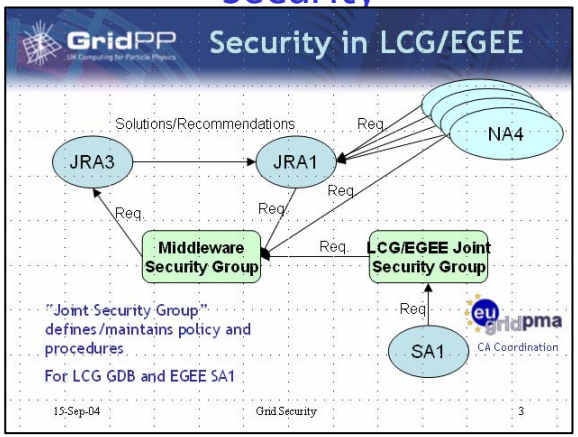
Information Services



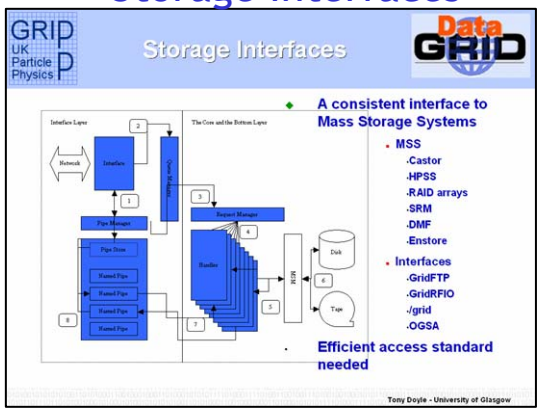
the globus project

GGF

Security

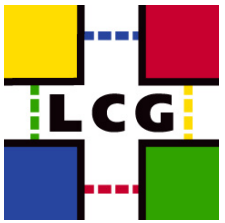


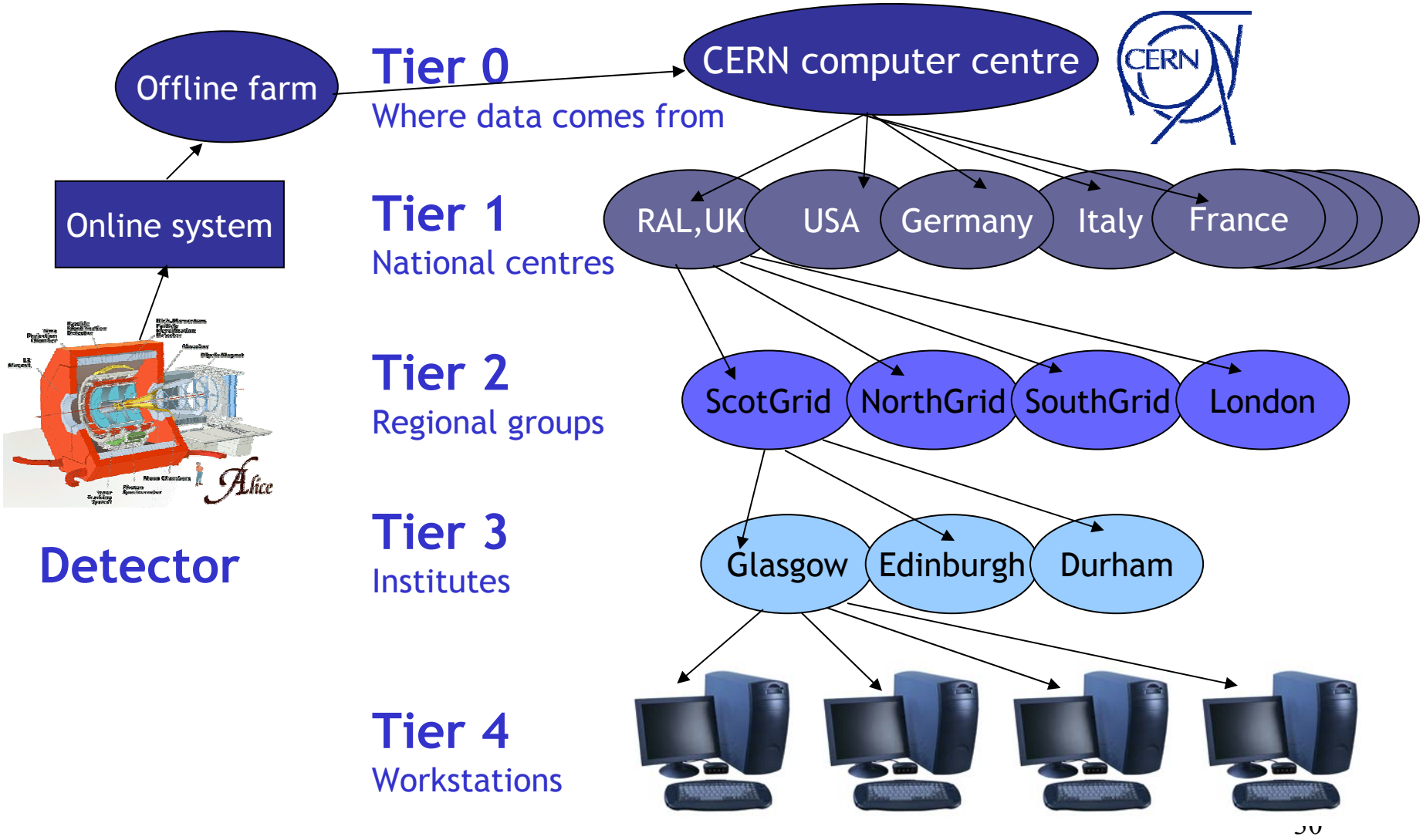
Storage Interfaces



International Collaboration

- EU DataGrid (EDG) 2001-2004
  - Middleware Development Project
- LHC Computing Grid (LCG)
  - Grid Deployment Project for LHC
- EU Enabling Grids for e-Science in Europe (EGEE) 2004-2006
  - Grid Deployment Project for all disciplines
- US and other Grid projects
  - Interoperability




Offline farm
Online system
Detector
Tier 0
Where data comes from
CERN computer centre
CERN
Tier 1
National centres
RAL,UK
USA
Germany
Italy
France
Tier 2
Regional groups
ScotGrid
NorthGrid
SouthGrid
London
Tier 3
Institutes
Glasgow
Edinburgh
Durham
Tier 4
Workstations

UK Tier-2 Centres

GridPP
UK Computing for Particle Physics

ScotGrid

Durham, Edinburgh, Glasgow

NorthGrid

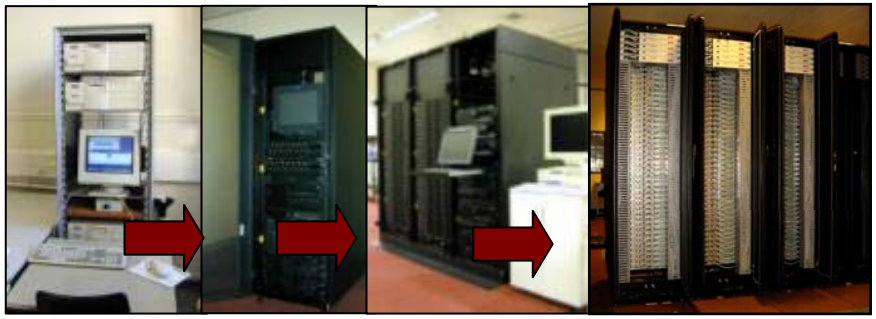
Daresbury, Lancaster, Liverpool, Manchester, Sheffield

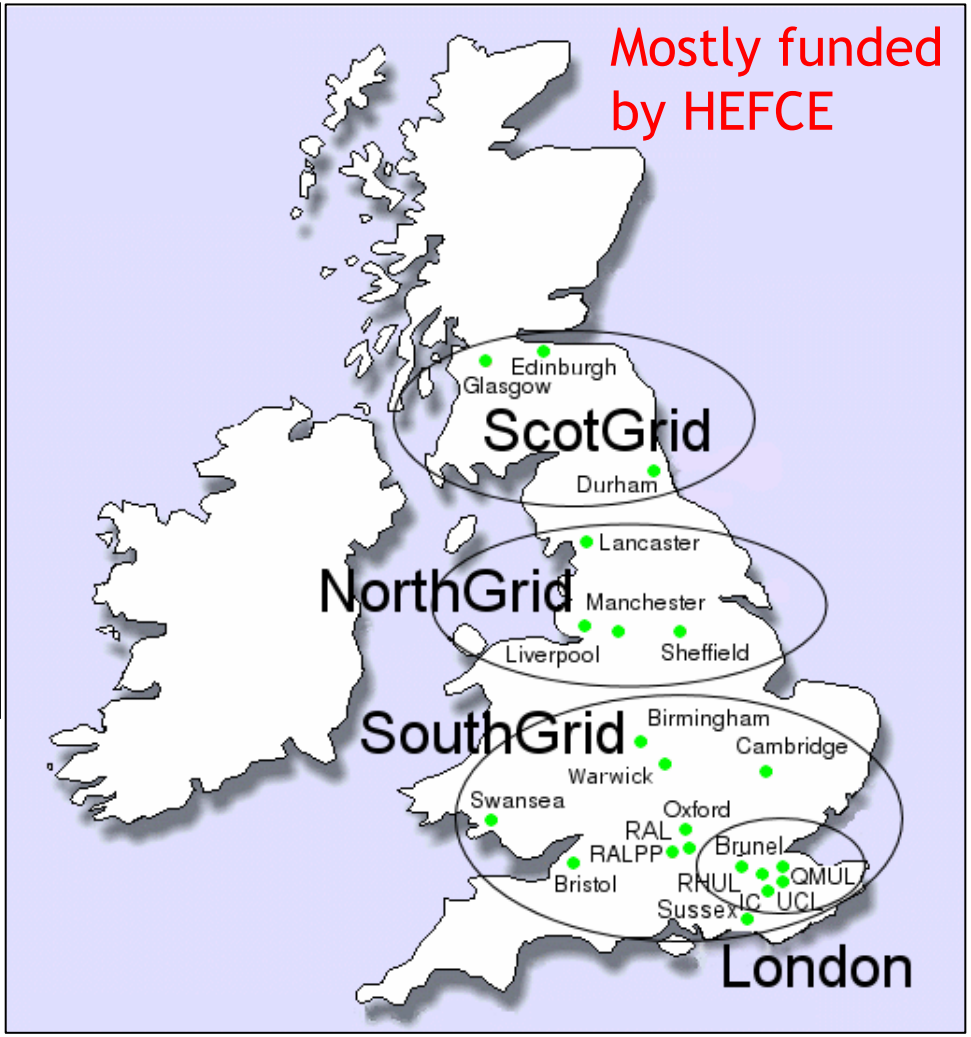
SouthGrid

Birmingham, Bristol, Cambridge, Oxford, RAL PPD, Warwick

LondonGrid

Brunel, Imperial, QMUL, RHUL, UCL

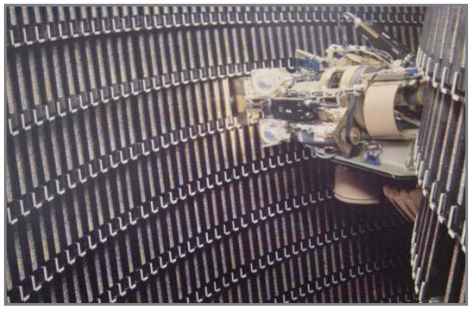




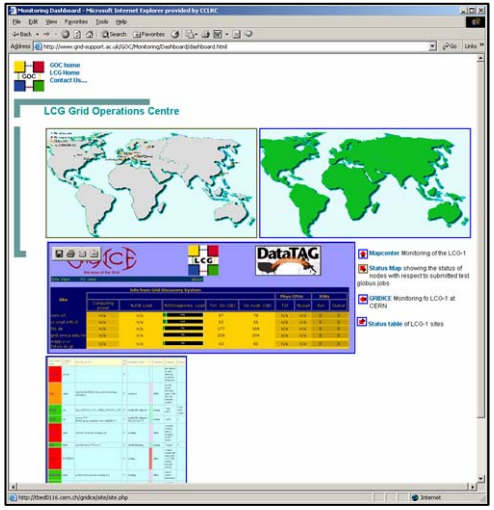
UK Tier-1/A Centre Rutherford Appleton Lab

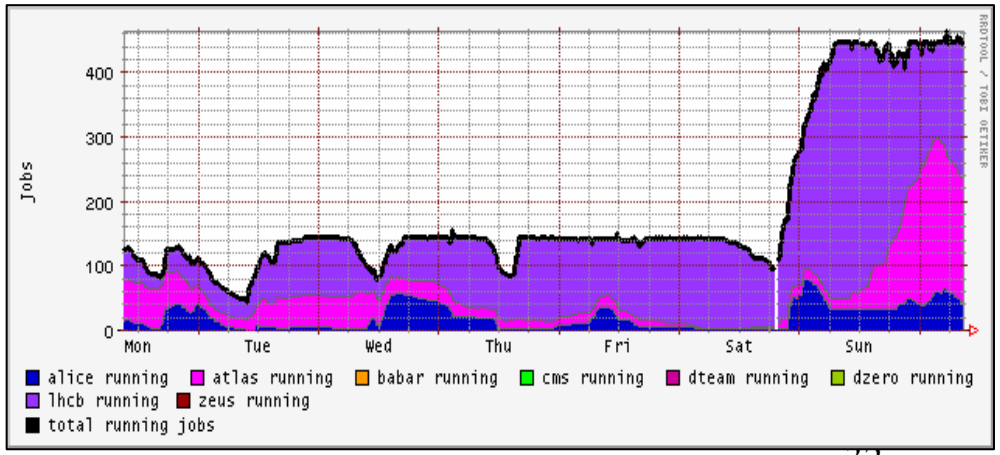
- High quality data services
- National and International Role
- UK focus for International Grid development

- 700 Dual CPU
- 80 TB Disk
- 60 TB Tape (Capacity 1PB)

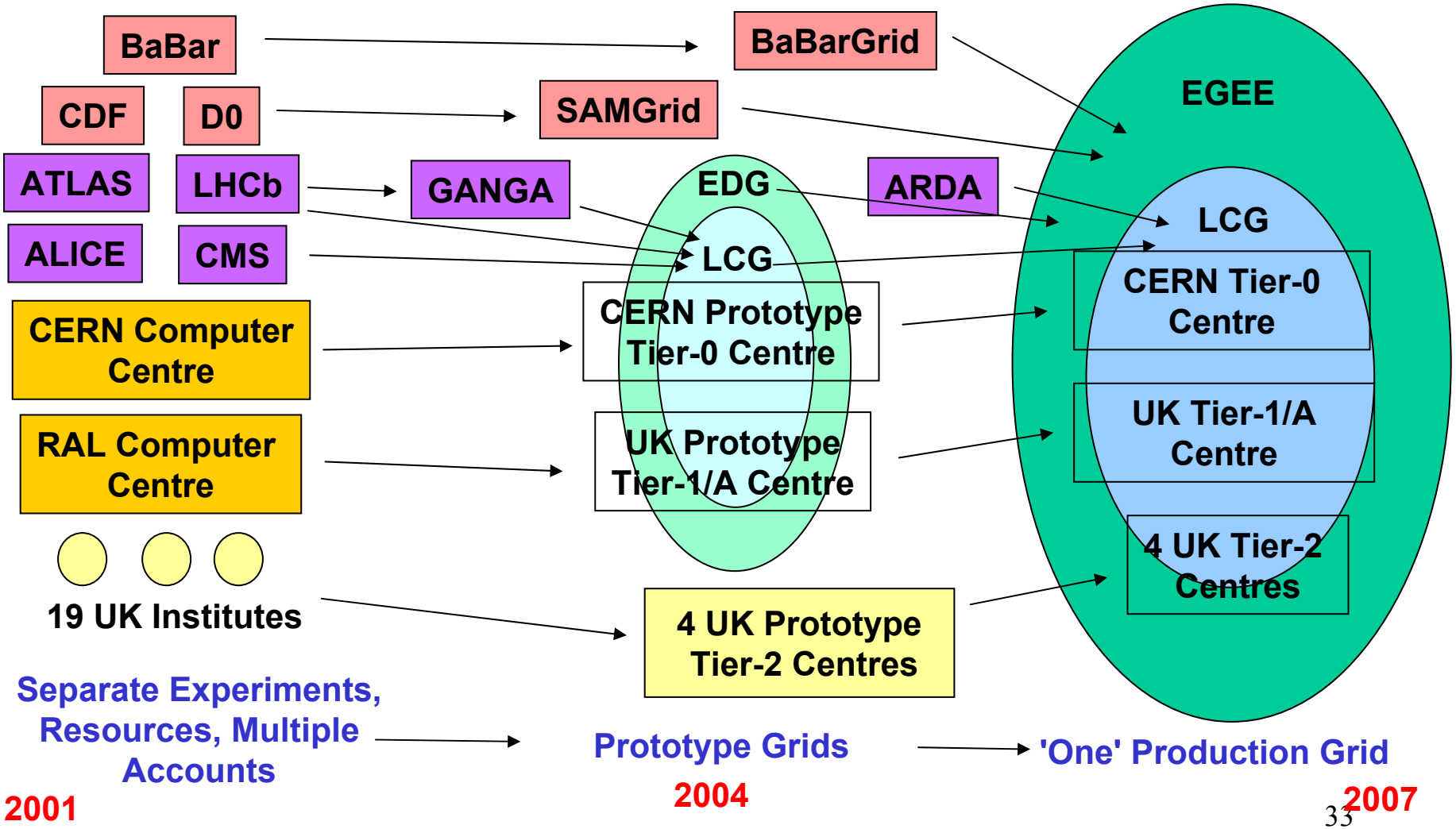


Grid Operations Centre





Prototype to Production Grid

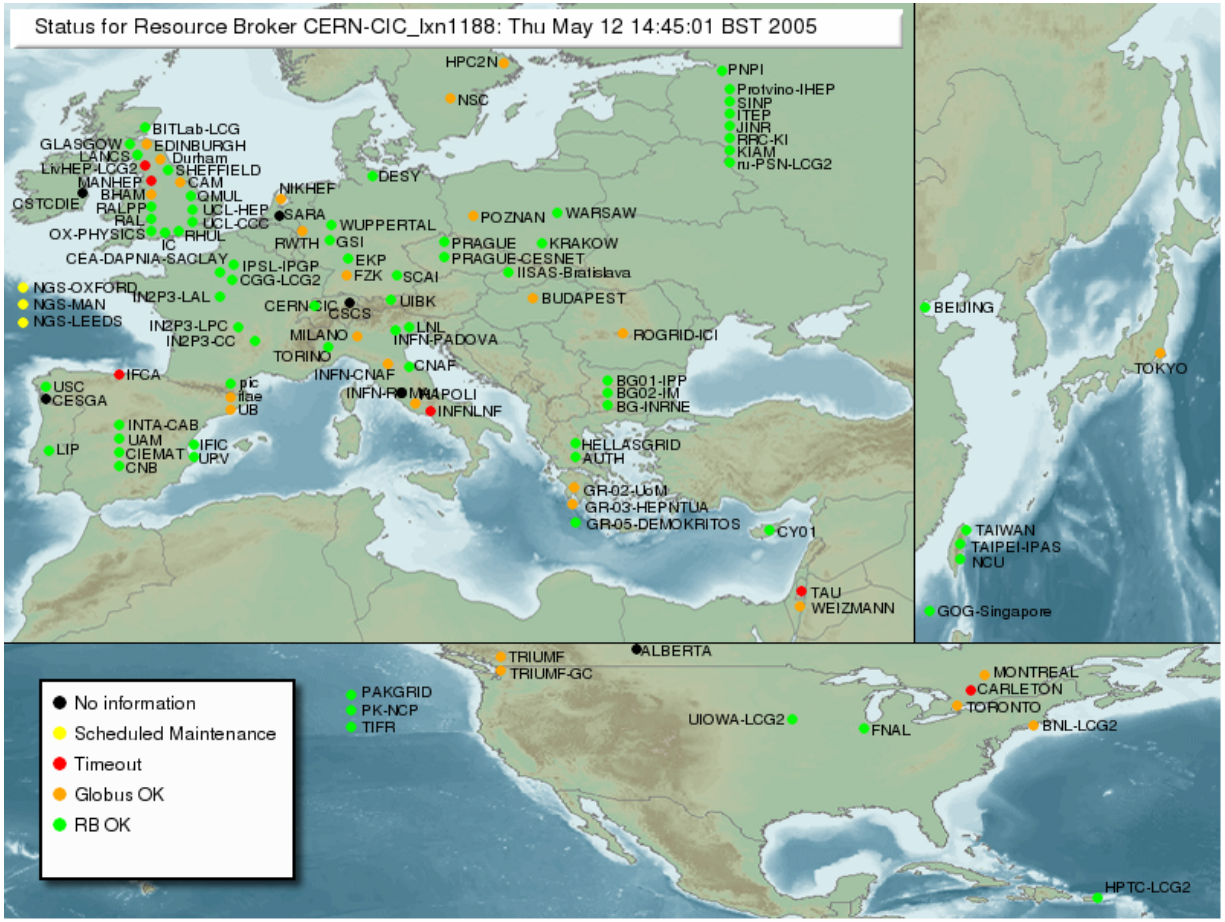


GridPP is active part of LCG

20 Sites

2,740 CPUs

67 Tbytes storage



GridPP
UK Computing for Particle Physics

LCG

20 Sites

2,740 CPUs

67 Tbytes storage

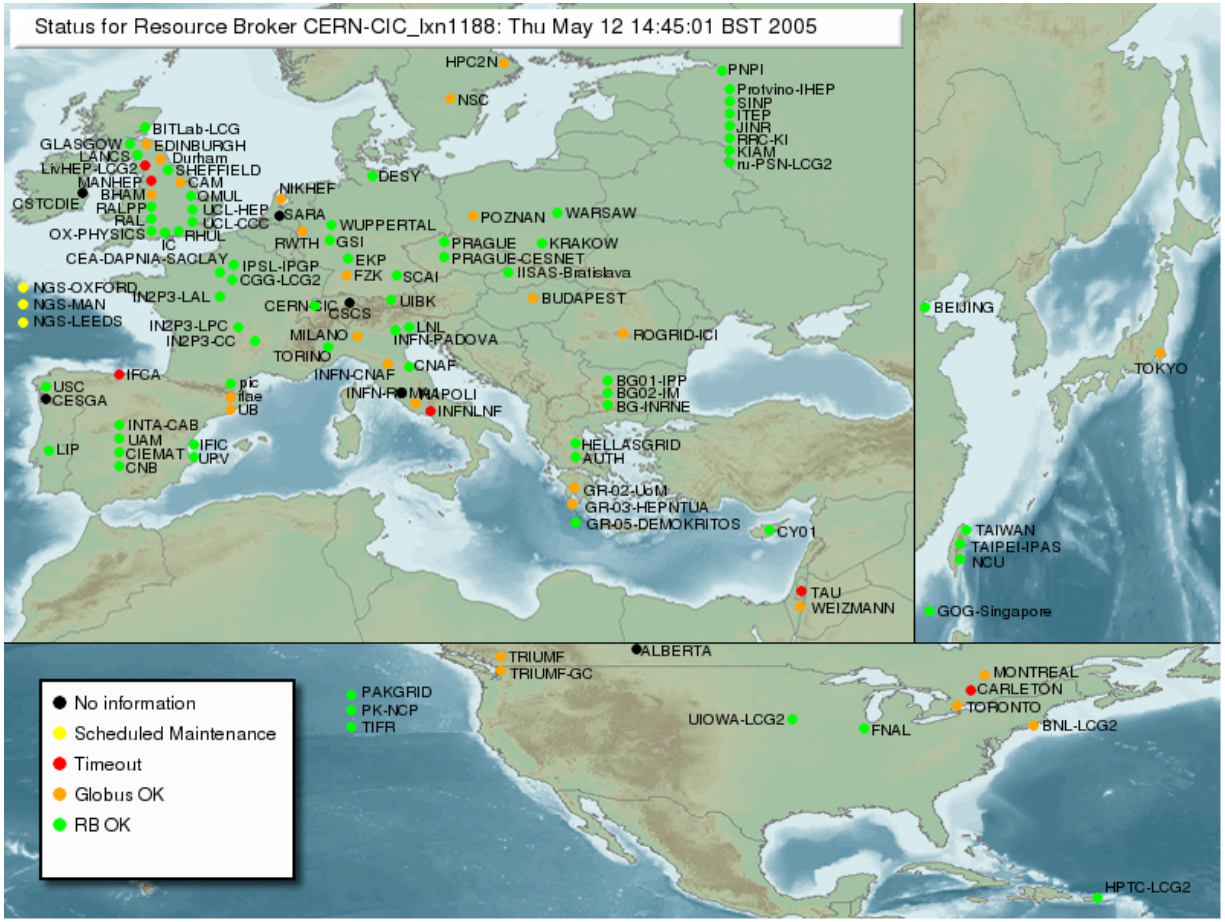
- GridPP is part of LCG

138 Sites

34 Countries

13,784 CPUs

4402 Tbytes storage



http://www.gridpp.ac.uk

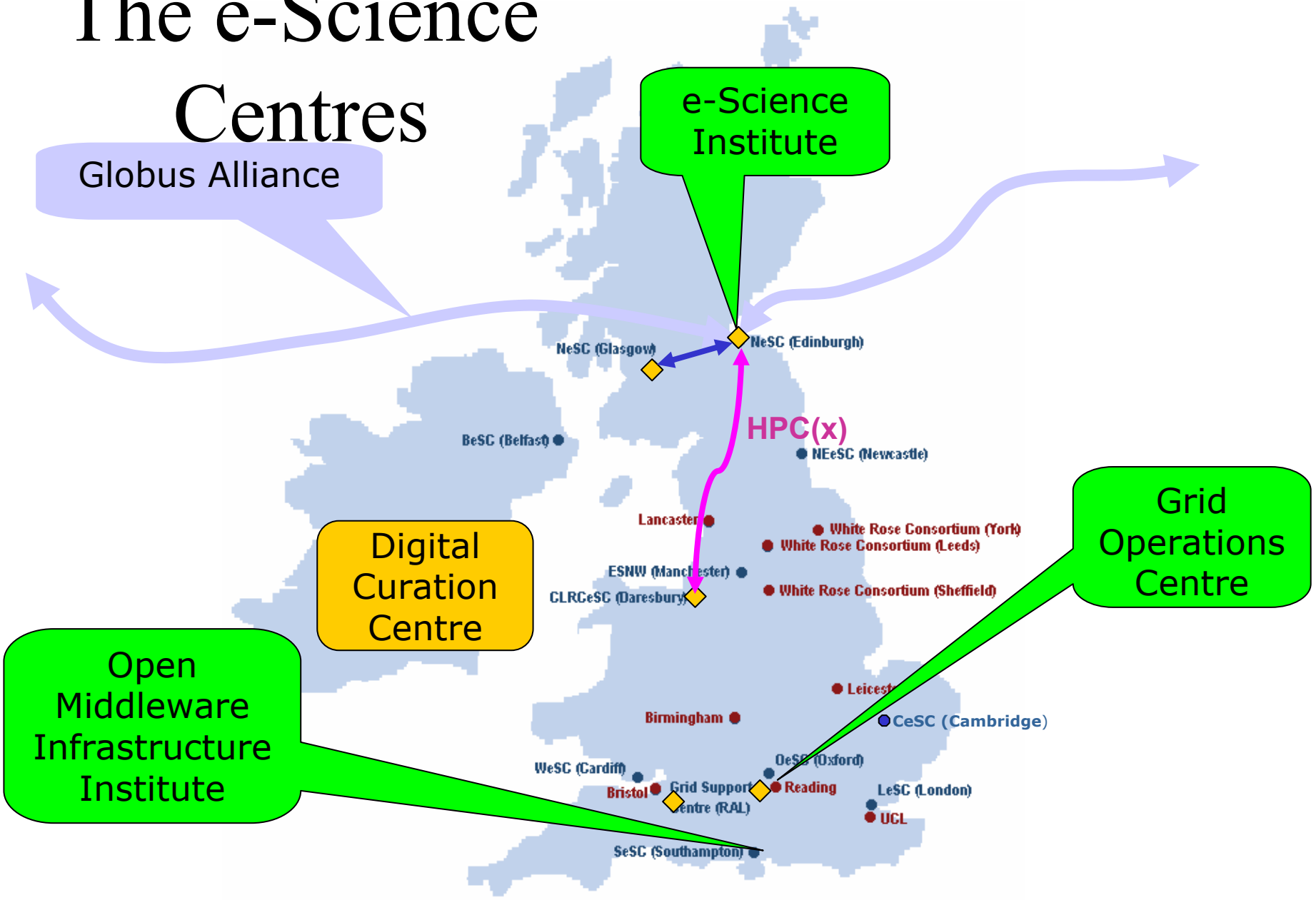
GridPP
UK Computing for Particle Physics

UK Core e-Science: Phase 2

Three major new activities:

1. National Grid Service and Grid Operation Support Centre
2. Open Middleware Infrastructure Institute for testing, software engineering and repository for UK middleware
3. Digital Curation Centre for R&D into long-term data preservation issues

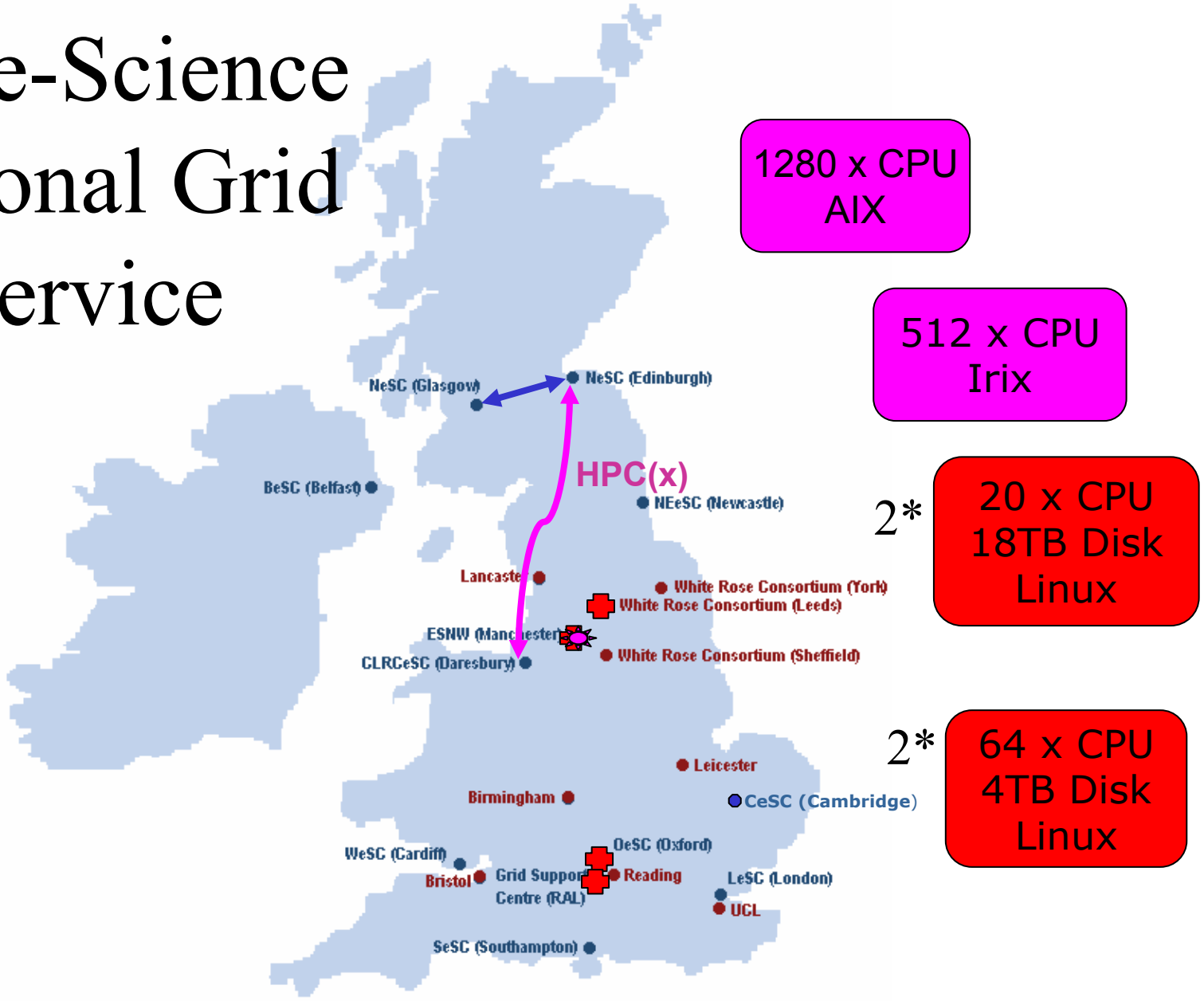
The e-Science Centres



UK National Grid Service(NGS)

- From April 2004, offers free access to
  two 128 processor compute nodes and
  two data nodes
- Initial software is based on GT2 via VDT and LCG releases plus SRB and OGSA-DAI
- Plan to move to Web Services based Grid middleware this summer

The e-Science National Grid Service



The UK Open Middleware Infrastructure Institute (OMII)

- Repository for UK-developed Open Source ‘e-Science/Cyber-infrastructure’ Middleware
- Documentation, specification,QA and standards
- Fund work to bring ‘research project’ software up to ‘production strength’
- Fund Middleware projects for identified ‘gaps’
- Work with US NSF, EU Projects and others
- Supported by major IT companies
- ➢ Southampton selected as the OMII site

Digital Curation Centre (DCC)

- In next 5 years e-Science projects will produce more scientific data than has been collected in the whole of human history
- In 20 years can guarantee that the operating and spreadsheet program and the hardware used to store data will not exist
  - Research curation technologies and best practice
  - Need to liaise closely with individual research communities, data archives and libraries
- Edinburgh with Glasgow, CLRC and UK office for library networking (Bath) selected for this work

The UK e-Science Programme

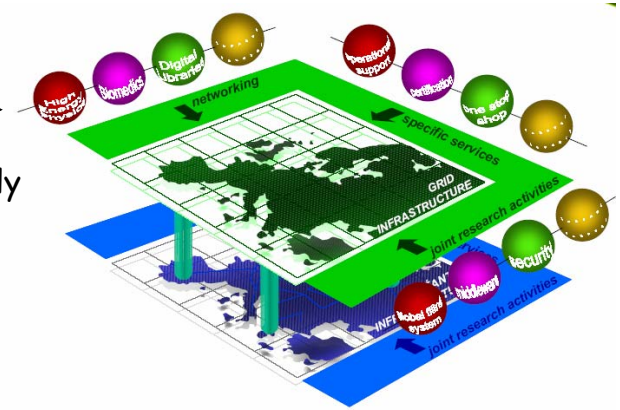
- An exciting portfolio of Research Council e-Science projects
  - Beginning to see e-Science infrastructure deliver in several areas
  - Astronomy, Chemistry, Bioinformatics, Engineering, Environment, Healthcare, Particle Physics ….
- The UK has very strong industrial component
  - UK companies contributing over £30M
  - Engineering, Pharmaceutical, Petrochemical, IT companies, Commerce, Media, …
- Needs continued support from Research Councils with identifiable e-Science funding after 2006

UK e-Infrastructure

Users get common access, tools, information,
Nationally supported services, through NGS

GridPP UK Computing for Particle Physics
ScotGRID
ULGrid
Regional and Campus grids

GOSC
NGS
Integrated internationally

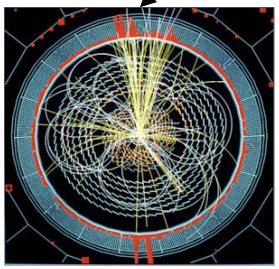


EMBL-EBI European Bioinfo

diamond

JISC
VRE, VLE, IE

UKERNA


LHC


ISIS TS2


HPCx + HECtoR

VISTA

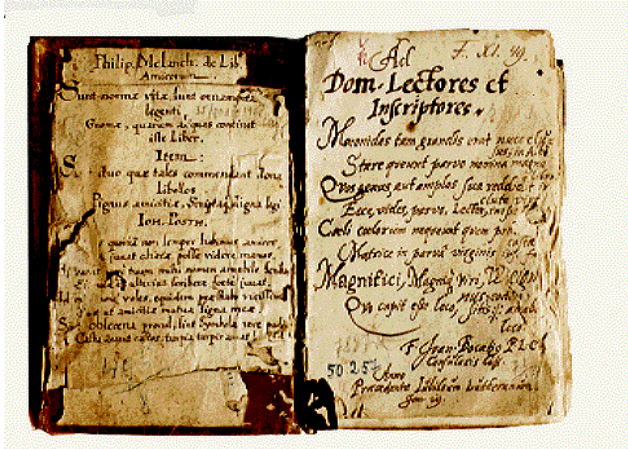
Summary

- Use common tools and support where possible for GRIDPP and NGS
- Strong industrial component in UK e-science
- Research Council support with identifiable e-Science funding after 2006 is very important
- Integrate e-science infrastructure and posts into the computing services at Universities

Acknowledgements

Thanks to previous speakers on this topic for their slides and to our hosts for their invitation to this interesting meeting

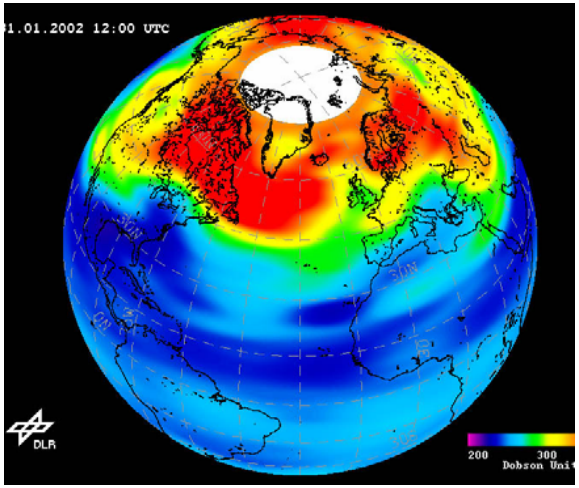
Digitization of almost anything

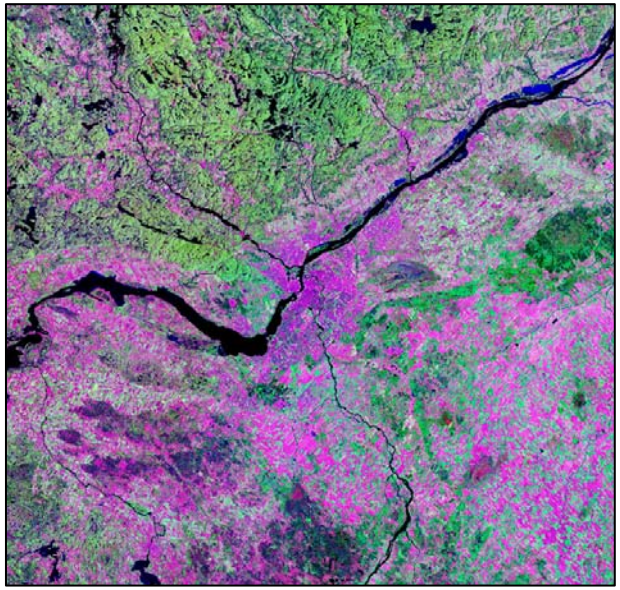


To create Digital Libraries and Museums

1 TB/day

Ozone map



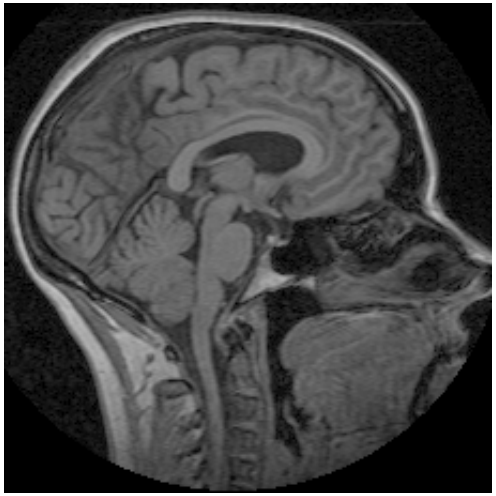


Ottawa

Trafalgar Square

PLANETARY VISIONS SATELLITE IMAGEMAP

Dynamic Brain Atlas



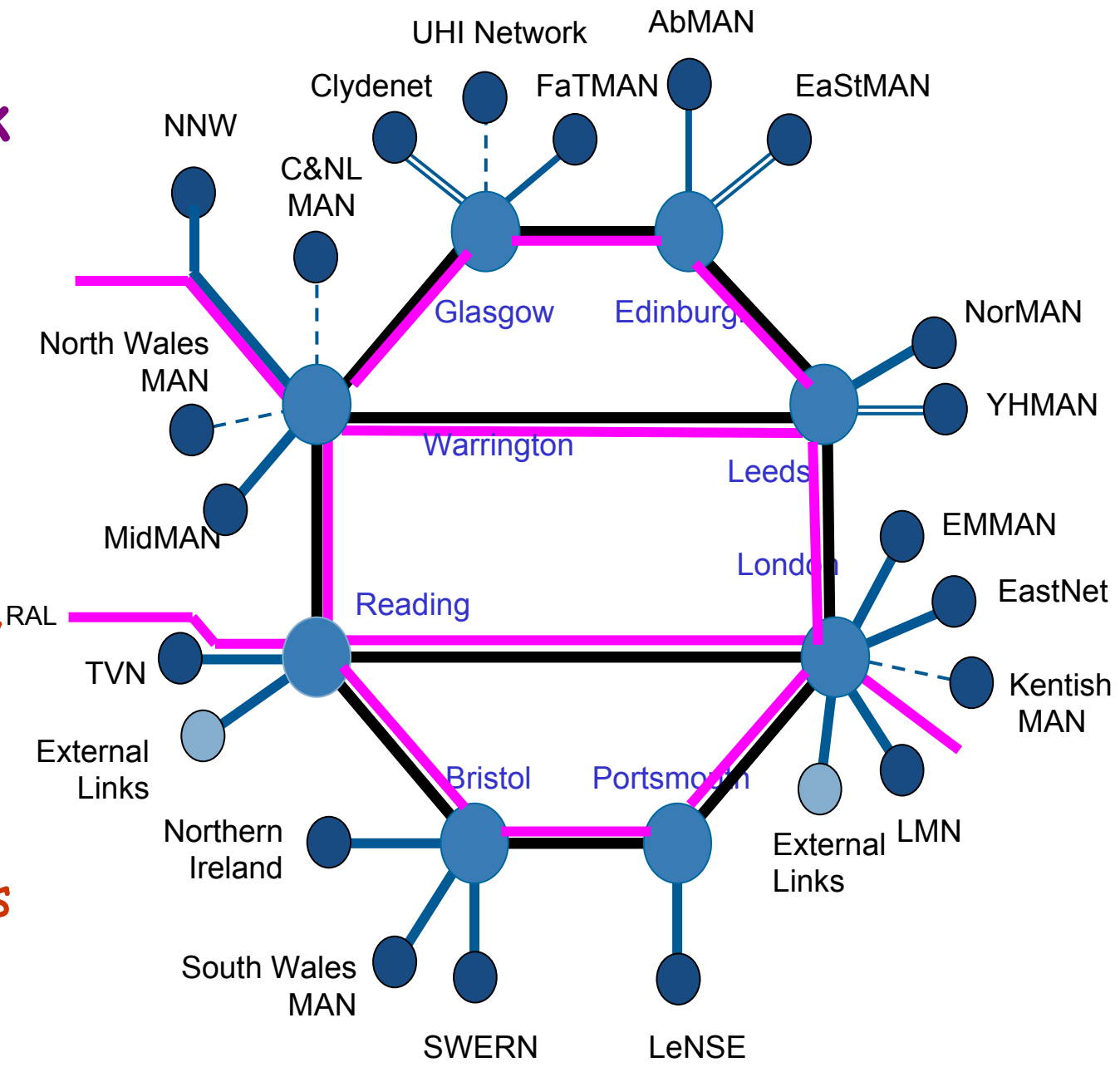
Scanning

Breast Screening

Remote Consultancy

SuperJANET R&D Network

- Parallel R&D Network
- Ethernet infrastructure
- 1 – 10 Gbit/s
- Can connect any UK institute in principle
- Currently connects 3 sites



UKLight – showing connections to selected International peer facilities