



the globus alliance
www.globus.org



UK e-Science Grid and EGEE

Malcolm Atkinson
&
Richard Sinnott
www.nesc.ac.uk

9th June 2004



| epcc |





the globus alliance

www.globus.org

The Primary Requirement ...



Enabling *People* to Work Together on Challenging Projects: Science, Engineering & Medicine



epcc



UNIVERSITY
of
GLASGOW

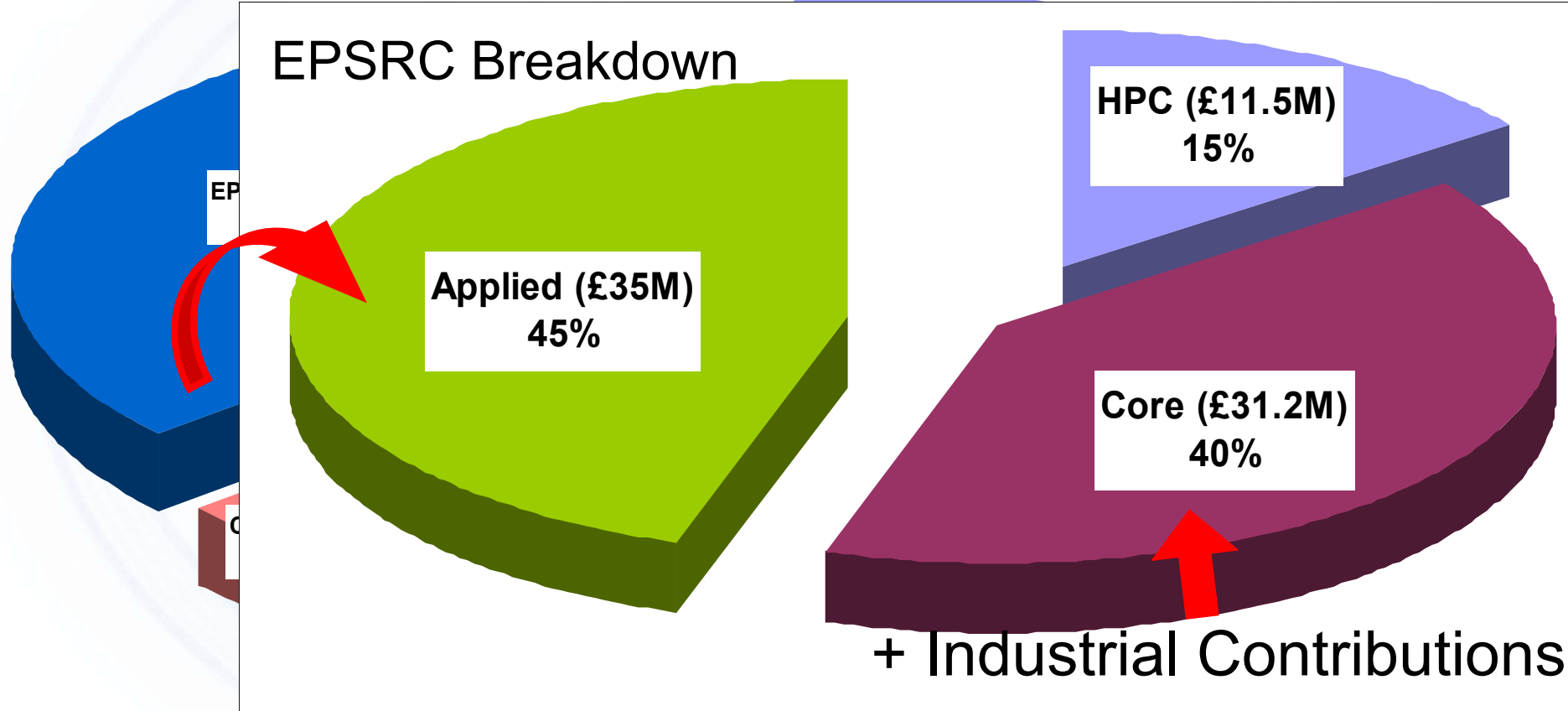


the globus alliance
www.globus.org

UK e-Science Budget (2001-2006)



Total: £213M



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



epcc





The e-Science Centres



e-Science Institute

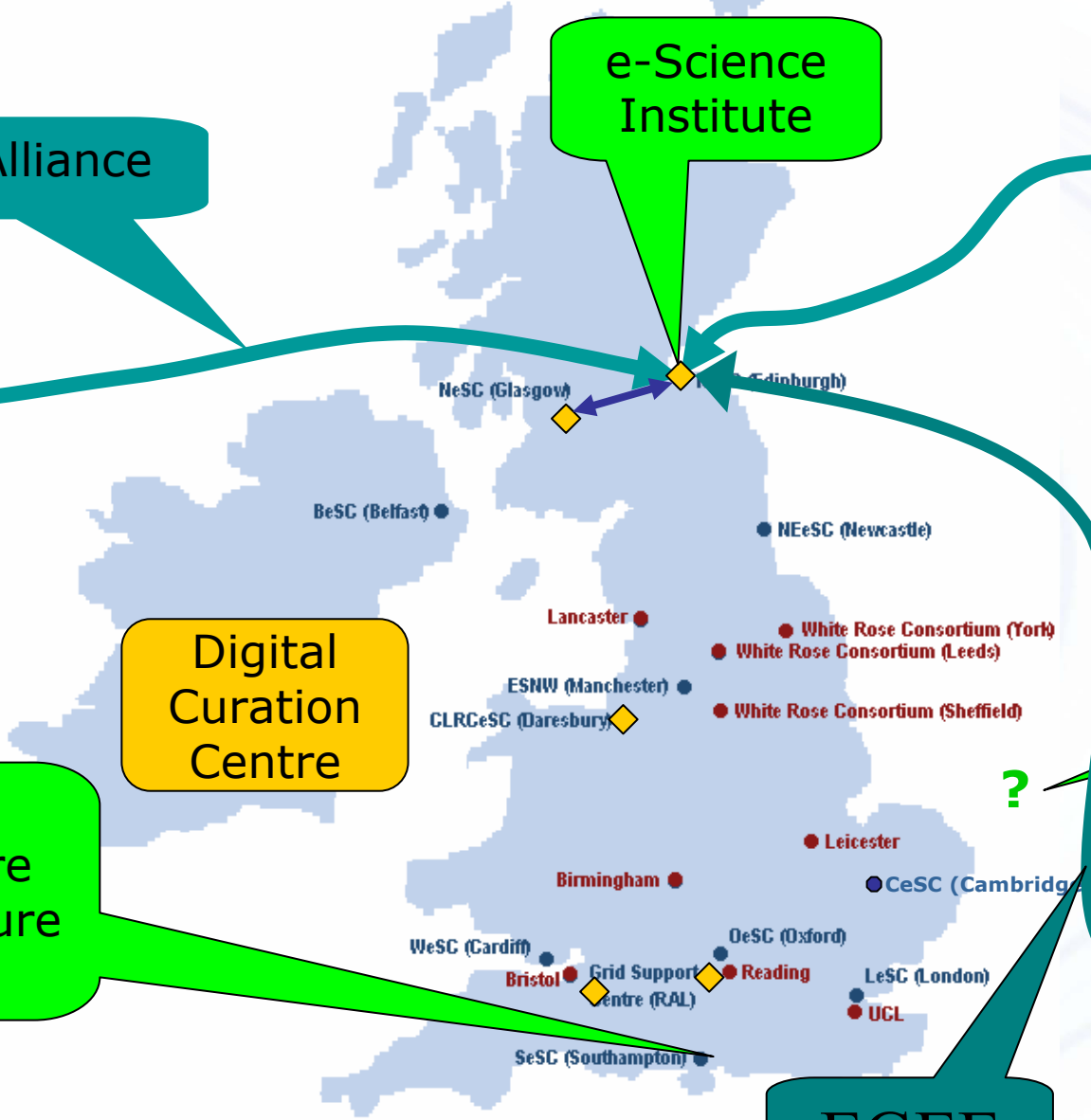
Globus Alliance

Grid Operations Centre

Digital Curation Centre

Open Middleware Infrastructure Institute

EGEE



epcc





the globus

The e-Science Grid: National Grid Service



1600 x CPU
AIX

512 x CPU
Irix

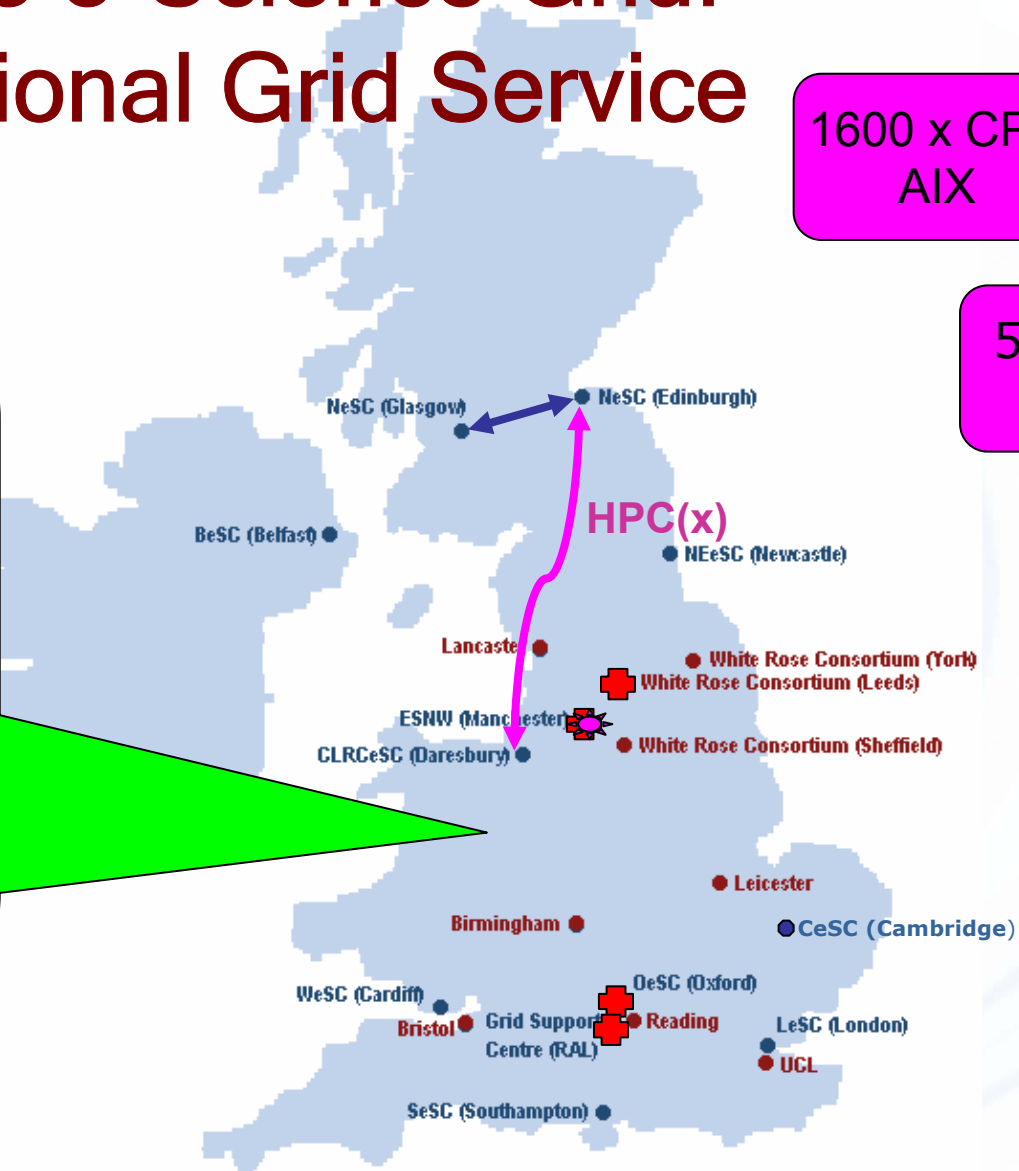
20 x CPU
18TB Disk
Linux

64 x CPU
4TB Disk
Linux

Engineering Task Force
(Contributions from e-Science Centres)

Grid Support Centre / Grid Operations Centre

OGSA Test Grid projects



epcc





the globus alliance

www.globus.org

UK e-Science Grid Status - Q3 2003



- Operational and heterogeneous Level-2 Grid based on Globus Toolkit 2
 - Demonstrated broad set of applications running across it
 - ▶ Monte Carlo simulations of ionic diffusion through radiation damaged crystal structures
 - ▶ Integrated Earth system modelling
 - ▶ BLAST on the Grid
 - ▶ Nimrod/G
 - ▶ DL_POLY and Portals
 - ▶ Grid Enabled Optimisation: Vibrations in Space Application to Satellite Truss Design
 - ▶ RealityGrid-Lite
 - ▶ Grid Integration Test Script Suite- from WP7



epcc





the globus alliance

www.globus.org

UK Level 2 Grid Project The Engineering Task Force They made it work



WP0 Project management

Rob Allan

WP1 Middleware development

Nick Hill

WP2 Grid information systems

Rob Allan

WP3 Authentication and the CA

Alistair Mills

WP4 User mgt and accounting

Steven Newhouse

WP5 Grid security

Jon Hillier

WP6 Grid platform deployment

Alistair Mills

WP7 Operational monitoring

David Baker

WP8 Applications

Simon Cox



epcc



UNIVERSITY
of
GLASGOW



the globus alliance
www.globus.org



UK e-Science Grids - the next steps

- **Prototype status ...**
 - Fragility?
 - Security ?
 - Reliability ?
 - Predictability ?
 - Dynamicity ?
 - Maintainability ?
- **Explore Research Grids based on OGSA**
- **Deliver production service – GT2 & LCG2**
 - NGS – initially a subset of sites
 - Grid Operations Centre – Director Neil Geddes
- **Many projects continue with GT2**

The future *is* based on WSRF – Converge by mid 2005



epcc





the globus alliance

www.globus.org

Projects achieve results



- Across all disciplines
- More than 100 projects
 - Applications 70% core M/W 30%
 - Approximately \$50 million from Industry collaboration
 - ▶ >70 companies
- Using UK e-Science Grid / NGS or Local Grids
- Or using their own “grids”
- The Grid idea is thriving
- It is still hard work to get started
- It is still very hard to push the technical limits
 - E.g. Teragyroid



Success depended on a strong rallying cry



epcc



UNIVERSITY
of
GLASGOW



Enabling Grids for
E-science in Europe

www.eu-egee.org

*Research Grid Experience Panel, GGF11
9 June 2004*

EGEE & LCG2 Experience

**Malcolm Atkinson
&
Ian Bird**

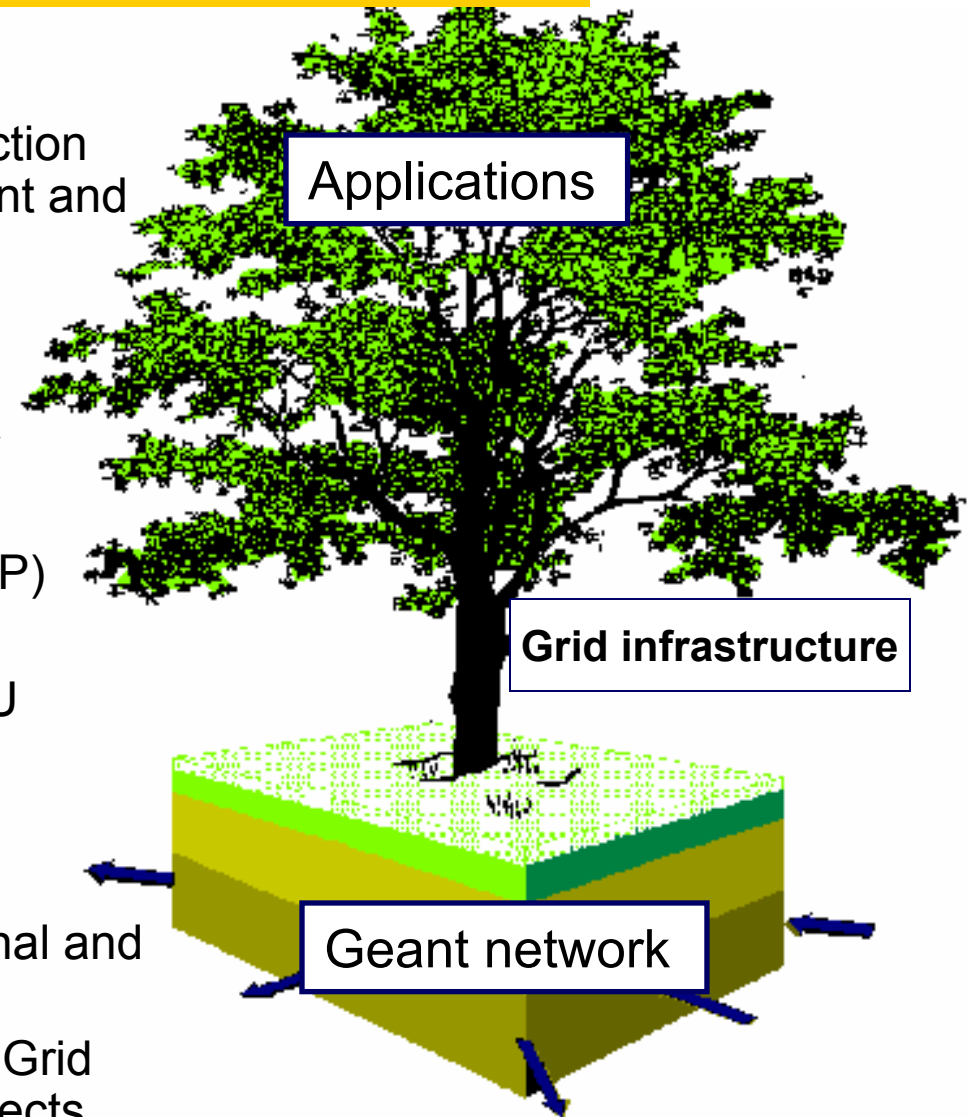


EGEE is a project funded by the European Union under contract IST-2003-508833

EGEE manifesto: Enabling Grids for E-science in Europe

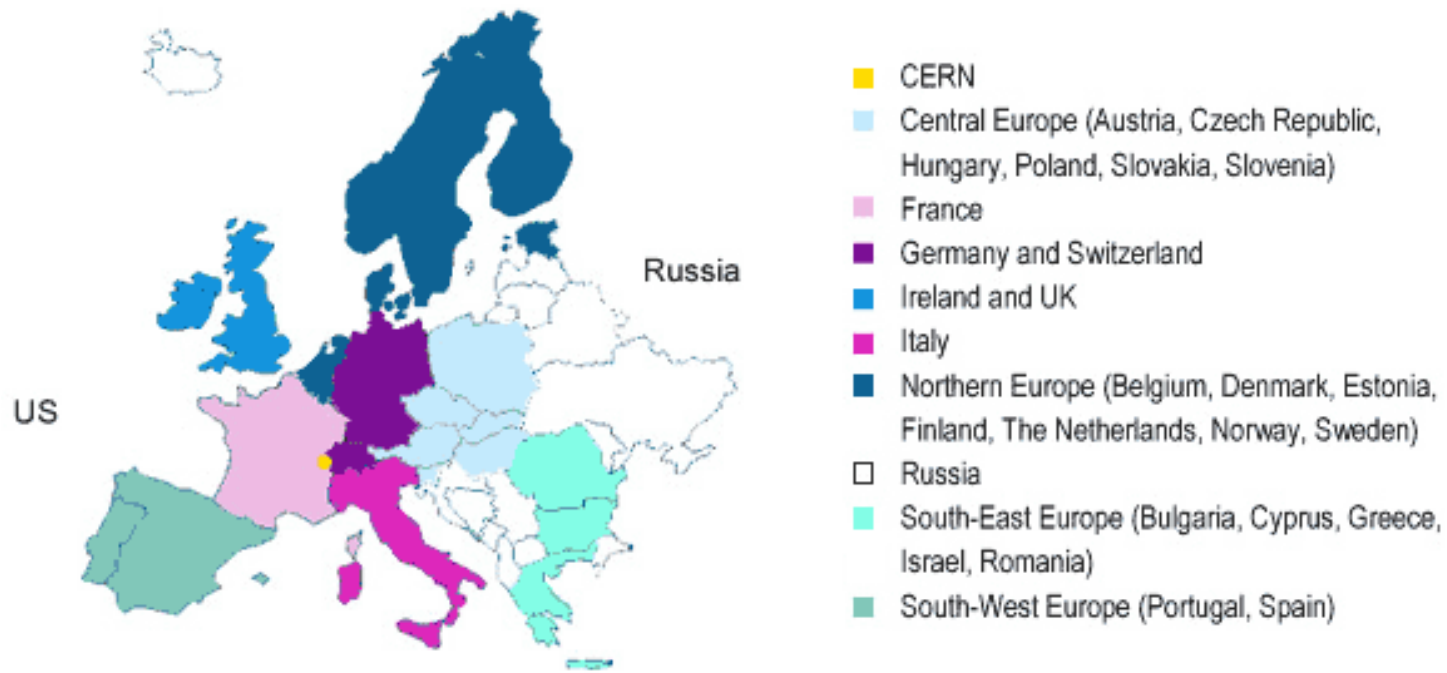


- **Goal**
 - Create a wide European Grid production quality infrastructure on top of present and future EU RN infrastructure
- **Build On:**
 - EUEU and EU member states major investments in Grid Technology
 - International connections (US and AP)
 - Several pioneering prototype results
 - Large Grid development teams in EU require major EU funding effort
- **Approach**
 - Leverage current and planned national and regional Grid programmes
 - Work closely with relevant industrial Grid developers, NRENs and US-AP projects



EGEE: Partners

- Leverage national resources in a more effective way for broader European benefit
- 70 leading institutions in 27 countries, federated in regional Grids



EGEE Activities

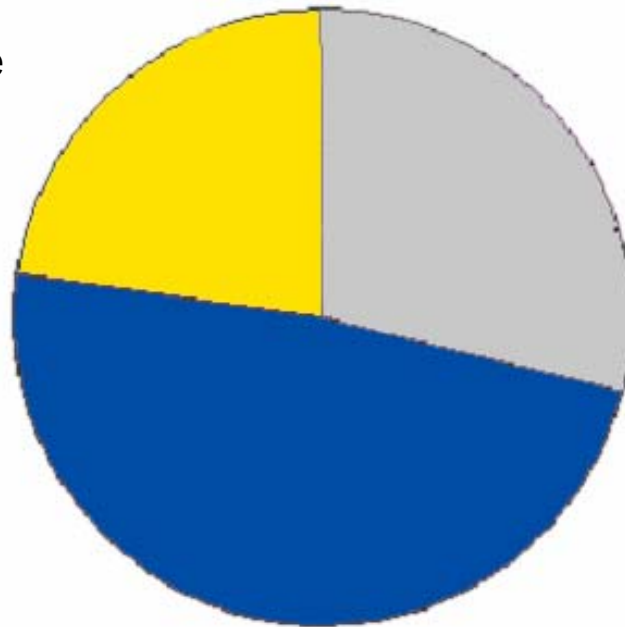
24% Joint Research

JRA1: Middleware Engineering and Integration

JRA2: Quality Assurance

JRA3: Security

JRA4: Network Services Development



48% Services

SA1: Grid Operations, Support and Management

SA2: Network Resource Provision

28% Networking

NA1: Management

NA2: Dissemination and Outreach

NA3: User Training and Education

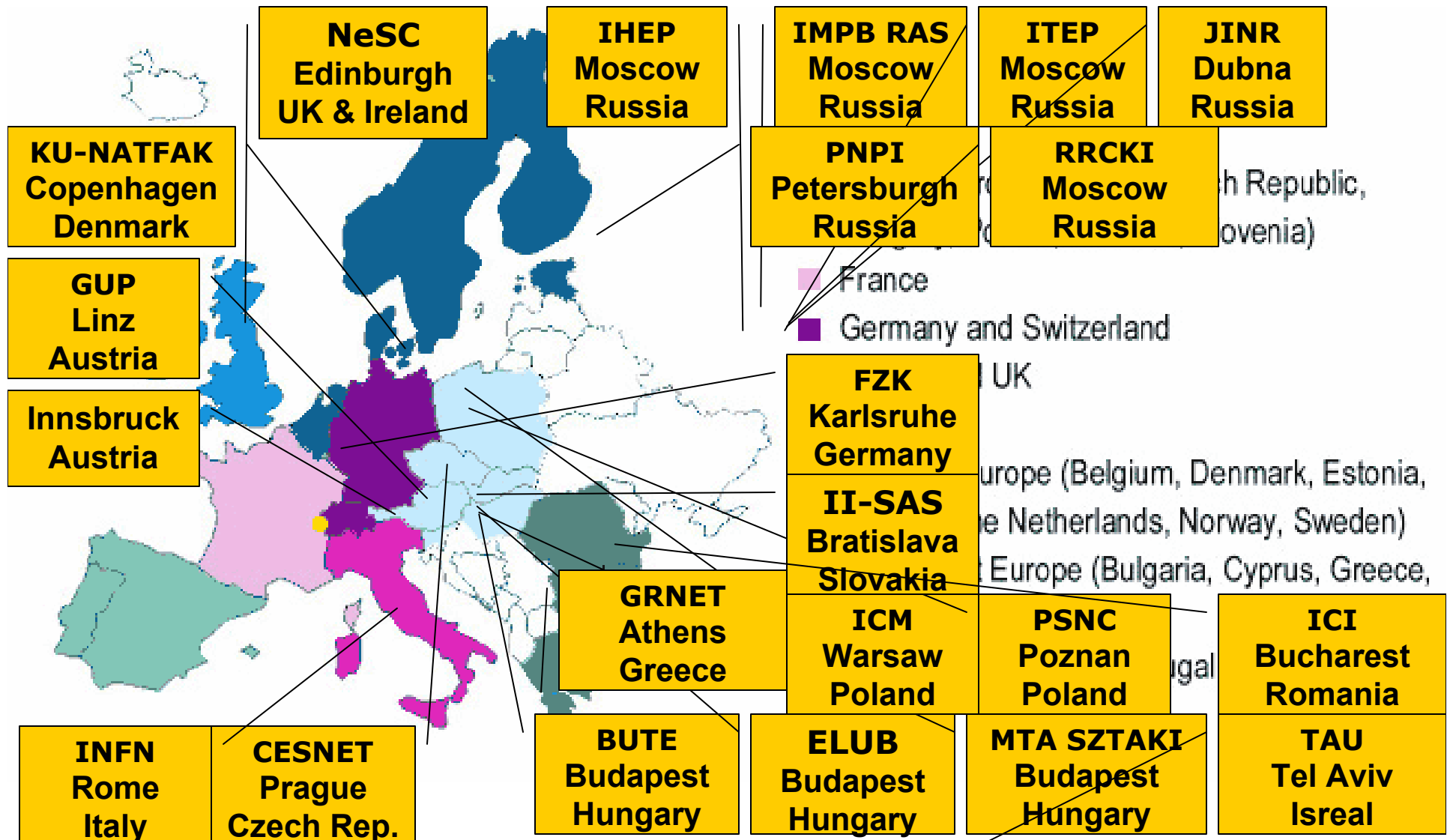
NA4: Application Identification and Support

NA5: Policy and International Cooperation

Emphasis in EGEE is on operating a production grid and supporting the end-users

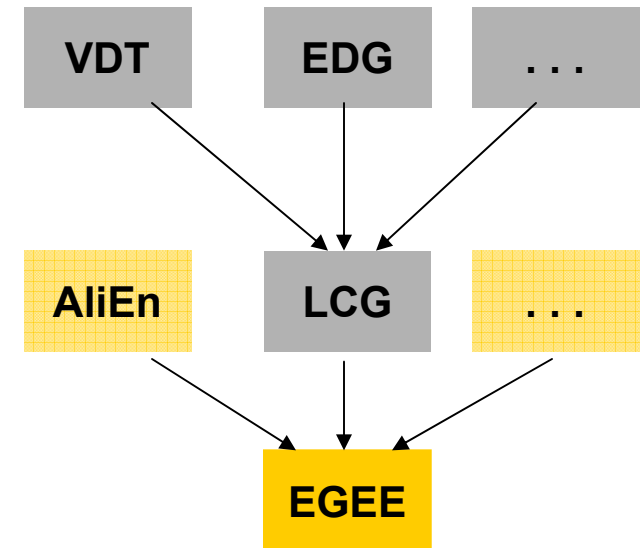
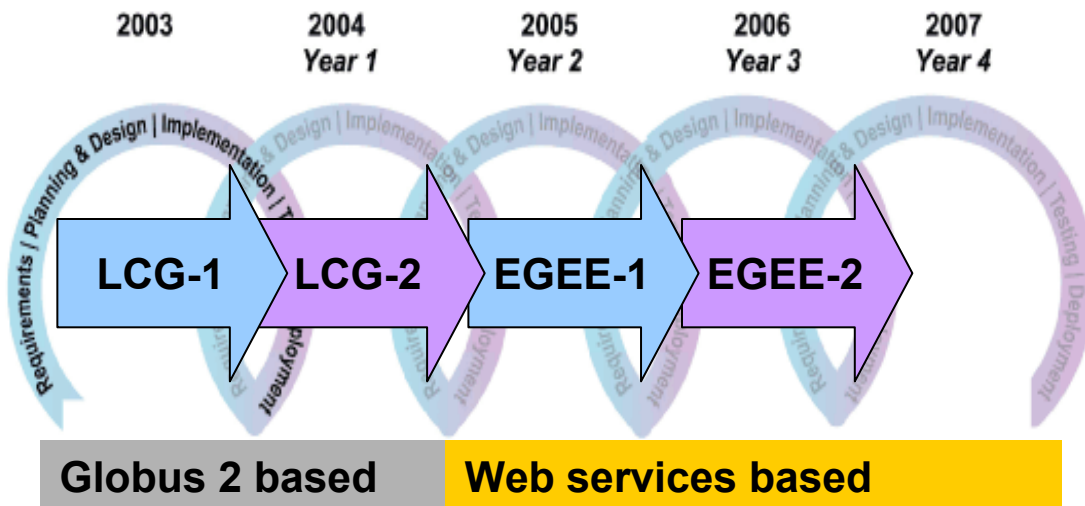
31 million Euros over 2 years

The Training Team Consortium



EGEE Implementation

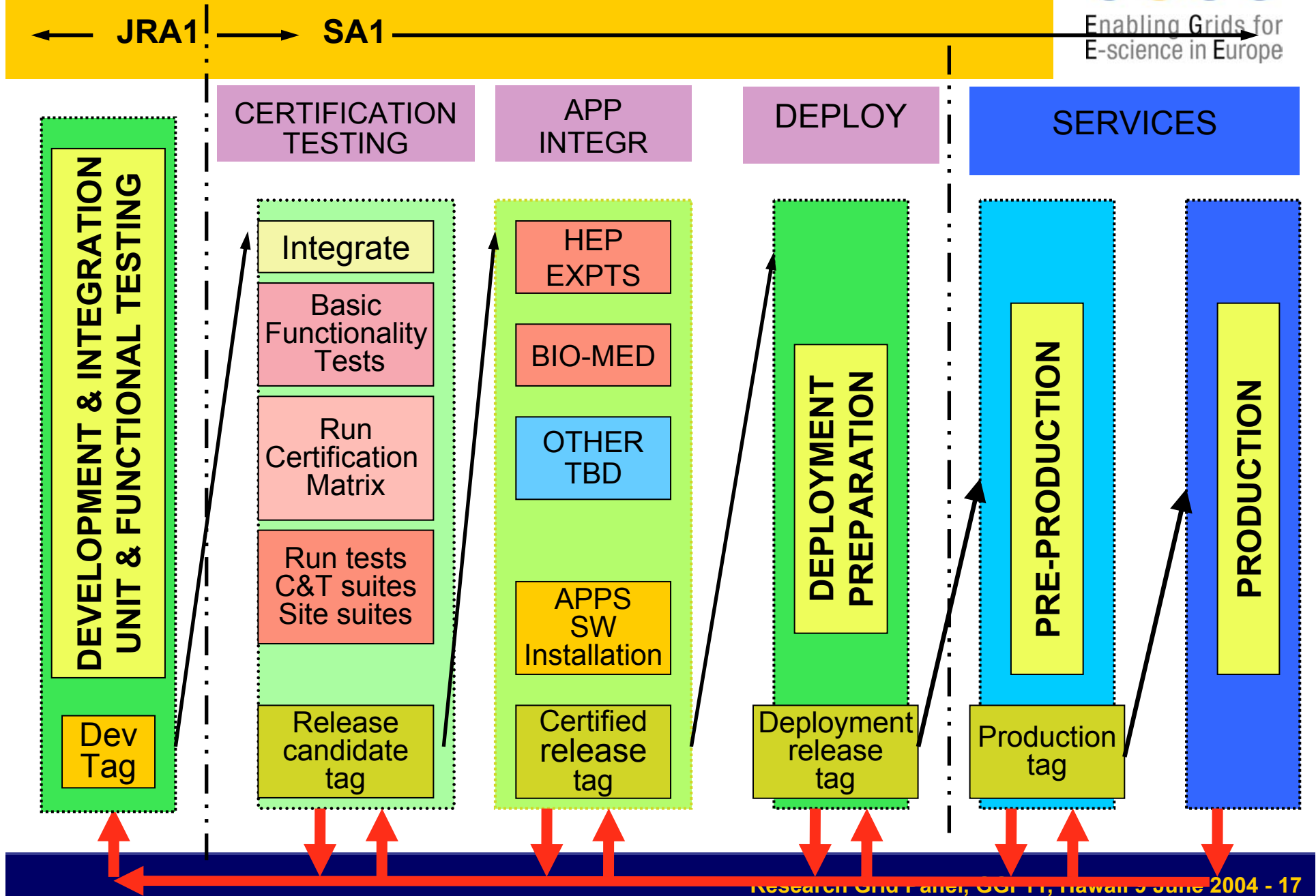
- From day 1 (1st April 2004)
 - Production grid service based on the **LCG infrastructure** running **LCG-2** grid middleware (SA)
 - LCG-2 will be maintained until the new generation has proven itself (fallback solution)
- In parallel develop a “**next generation**” grid facility (*JRA*)
 - Produce a new set of grid services according to evolving standards (Web Services)
 - Run a development service providing early access for evaluation purposes



Some observations - middleware

- LCG took close to 1 year to make existing middleware into something close to production quality
- Found existing middleware:
 - Was not well tested
 - Did not handle exceptions
 - Assumed that networks, other services would always work
 - This is a distributed system!
 - Did not address reliability
 - Did not address scalability
 - Did not address application required functionality
 - Use-cases often do not describe exactly how a service will be used → underlying architecture sometimes not appropriate
 - Could not easily be integrated into existing computing infrastructures
 - Often assumed full control of dedicated test-beds – this is not the real situation

Certification, Testing and Release Cycle



LCG Certification

- Significant investment in certification and testing process and team
 - Skilled people capable of system-level debugging, tightly coupled to VDT, Globus, and EDG teams
 - Needs significant hardware resources
 - This was essential in achieving a robust service
- Making production quality software is
 - Expensive
 - Time consuming
 - Not glamorous!
 - ... and takes very skilled people with a lot of experience
- Message for developers
 - One good idea at a time – test, test, test
 - Exception handling is not the exception – it is the normal mode of operation in a distributed system!
 - Reference implementations are (by definition) NOT production quality
 - Expect components to be replaced by the deployers!

Experiences in deployment

- LCG covers many sites (~60) now – both large and small
 - Large sites – existing infrastructures – need to add-on grid interfaces etc.
 - Small sites want a completely packaged, push-button, out-of-the-box installation (including batch system, etc)
 - Satisfying both simultaneously is hard – requires very flexible packaging, installation, and configuration tools and procedures
 - A lot of effort had to be invested in this area
- Richness of batch systems does not match simple gatekeeper model
 - Many queues, heterogenous clusters, hierarchical “fair-share” scheduling and polices versus simple assumptions at gatekeeper
- Current model of Information system with unique schema does not scale
 - Does not allow for differences between sites and full richness of available reesources
- Data management systems still have a long way to go

Summary

- This was a list of problems – but in the end we are quite successful
 - System is stable and reliable
 - System is used in production
 - System is reasonably easy to install now – 60 sites
 - Now have a basis on which to incrementally build essential functionality
- This infrastructure forms the basis of the initial EGEE production service

Sites in LCG-2/EGEE-0 : June 4 2004

