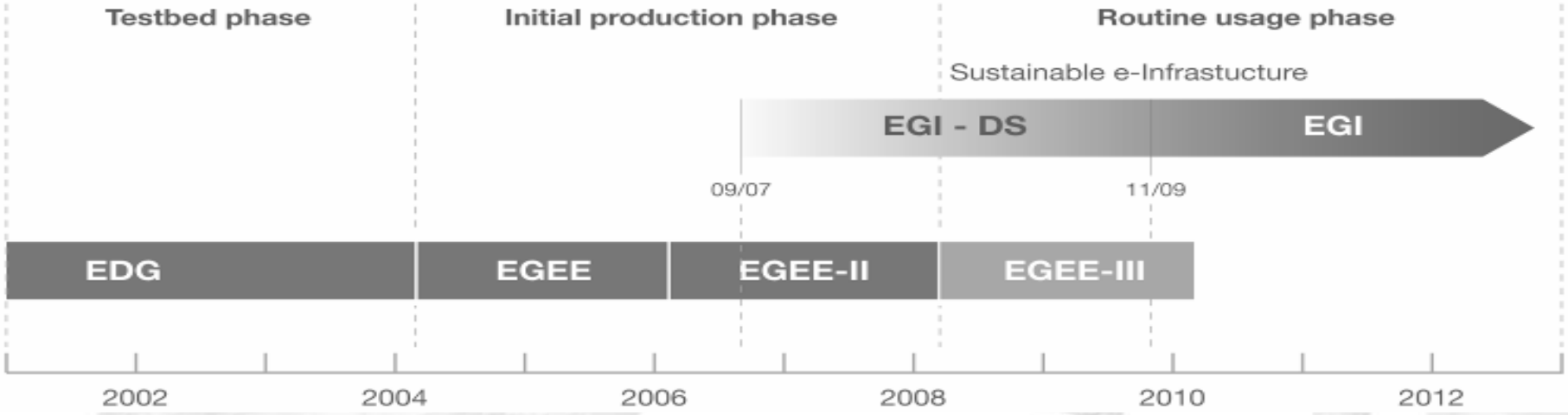
EGI: First Steps

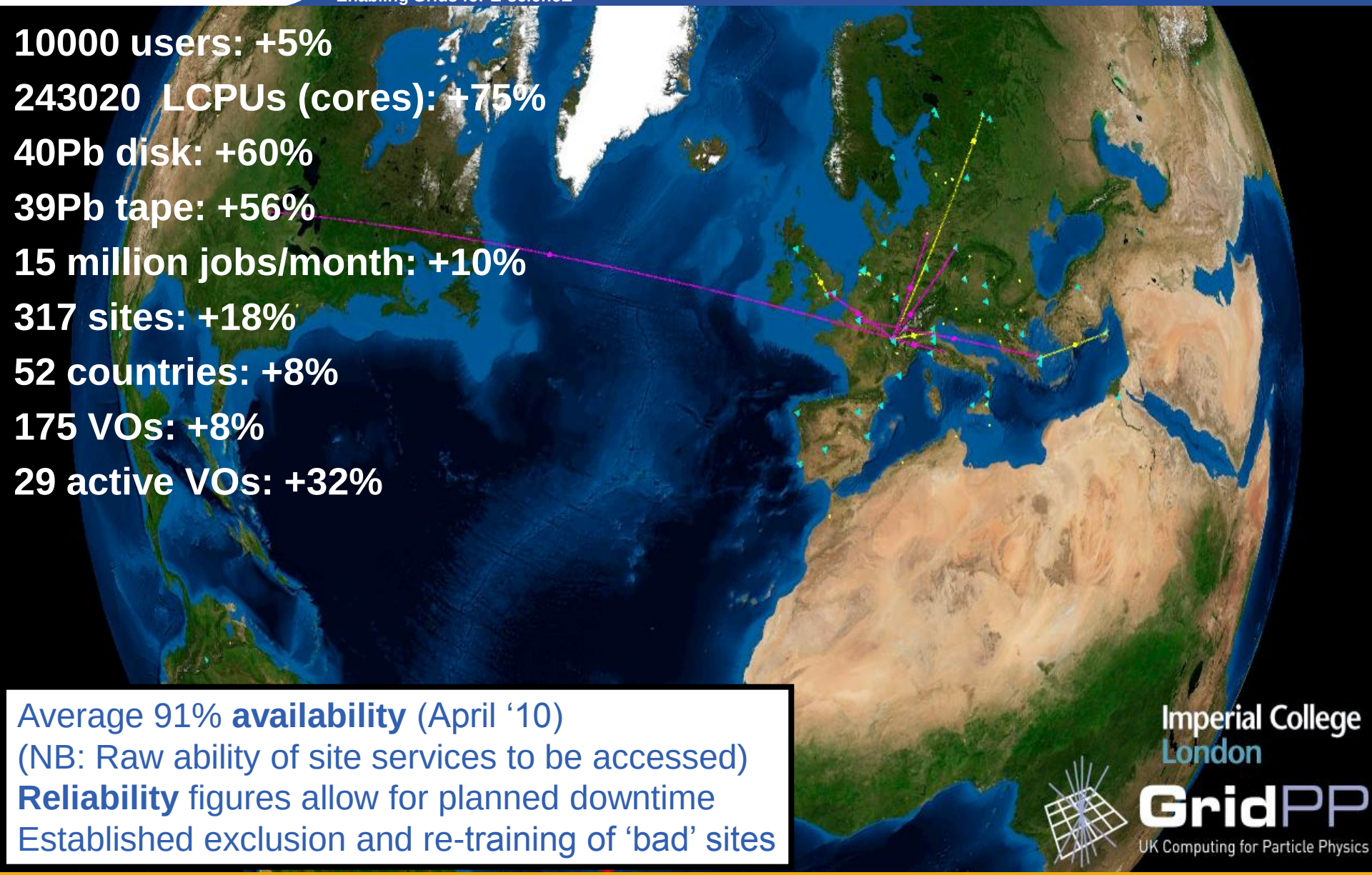
Steven Newhouse
Director EGI.eu
Project Director, EGI-InSPIRE

European Grid Infrastructure

- European Data Grid (EDG)
  - Explore concepts in a testbed
- Enabling Grid for E-sciencE (EGEE)
  - Moving from prototype to production
- European Grid Infrastructure (EGI)
  - Routine usage of a sustainable e-infrastructure



Production Infrastructure



10000 users: +5%
243020 LCPUs (cores): +75%
40Pb disk: +60%
39Pb tape: +56%
15 million jobs/month: +10%
317 sites: +18%
52 countries: +8%
175 VOs: +8%
29 active VOs: +32%

Average 91% **availability** (April '10)
(NB: Raw ability of site services to be accessed)
Reliability figures allow for planned downtime
Established exclusion and re-training of 'bad' sites

Imperial College London
GridPP
UK Computing for Particle Physics

What will EGI initially focus on?

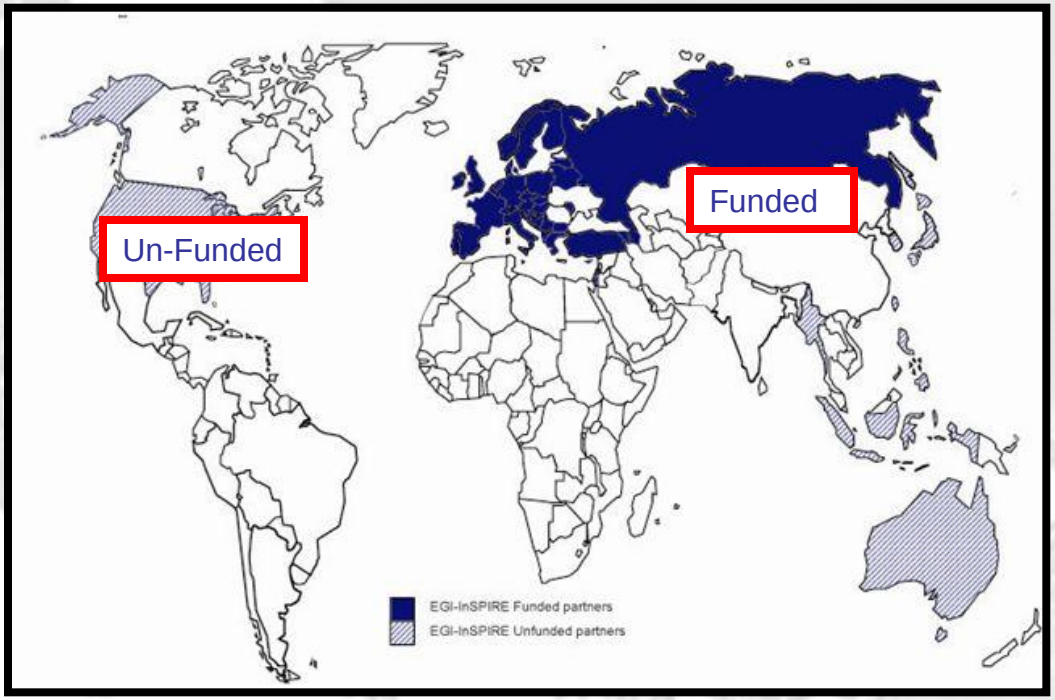
- Continue to provide a secure reliable generic infrastructure
  - Integrate resources based on gLite, UNICORE, ARC, Globus, ...
  - Leverage new technologies to provide more flexibility to users
- Support the user communities using the infrastructure
  - Assist and support the current EGEE communities
  - Engage with and support new structured communities
    - **e.g. ESFRI projects**
- Improve the efficiency of the infrastructure
  - The number of jobs, users & data continue to increase
  - Utilisation and effectiveness of the resources needs to match

Use new technologies to make middleware selection and operation a domain specific decision

The EGI-InSPIRE Project

Integrated Sustainable Pan-European Infrastructure for Researchers in Europe

- A 4 year project with €25M EC contribution
  - Project cost €72M
  - Total Effort ~€330M
  - Effort: 9261



Project Partners (51)
- EGI.eu, 40 NGIs, 2 EIROs
- Asia Pacific (8 partners)

The EGI.eu Organisation

- Coordinating European DCI resources
  - Roadmap to integrate HTC, HPC, Data, Instruments, ...
  - Policy & services for a production infrastructure
- EGI.eu governed and owned by its stakeholders
  - EGI Council votes proportional to fees
    - **Fees currently set proportional to national income**
  - Builds on resources from within its stakeholders
- Located in the Amsterdam Science Park
  - Distributed staff (~45) with a core (~50%) in Amsterdam
    - **Human coordination in Amsterdam**
    - **Technical coordination with a few partners across Europe**

EGI means Innovation

- Deploy Technology Innovation
  - Distributed Computing continues to evolve
    - **Grids → Desktops → Virtualisation → Clouds →?**
- Enable Software Innovation
  - Provide reliable persistent technology platform
    - **Community tools built on the deployed technology**
- Support Research Innovation
  - Infrastructure for data intensive science
    - **Support for Virtual Research Communities (e.g. ESFRI)**

EEF - European E-Infrastructure Forum

Forum for the discussion of principles and practices to create synergies for distributed Infrastructures

(http://www.einfrastructure-forum.eu/**)**

- **Goal:** seamless interoperation of leading e-Infrastructures serving the European Research Area

- **Requirements Questionnaire**
  - **Total of 28 ESFRI (+ eNMR and NeuGrid) projects consulted**
  - **Report published:**
    - http://www.einfrastructure-forum.eu/documents/EEF-report
- **First step in what EEF believes will be an iterative process**
- **Address any common requirements by harmonising existing services so users get a consistent access to all e-infrastructure resources**

First pass analysis of ESFRI requirements (all sectors) identified the following common areas

- Single sign-on
  - Consistent access to resources
- Virtual organisations
  - Collaboration mechanisms
- Persistent storage
  - Long term preservation of data and its access
- Data Management services
  - Movement and access
- Standards
  - Web services
- Workflows
  - Access to HPC/grid/network resources (compute & data) across Europe
- Training & consultancy
  - Generic access to the infrastructure and domain specific
- Global scope
  - Integration with activity beyond Europe

EEF - European E-Infrastructure Forum

- Single sign-on – ensure same identify can be used across network/HPC/grid by harmonising policies for Authentication, Authorization and eventually Accounting and Auditing
  - Needs technical developments and a review of policies by all parties

- Virtual organisations – harmonise support across HPC & grids
  - Review policies and potentially some middleware developments needed

- Security incident handling (e.g. cooperating security incident response group)
  - Integrated security groups

- Workflows – support of access to HPC/grid/network resources (compute & data) by a variety of workflow engines

- Global scope – beyond Europe: leverage existing EEF connections and contacts with sister e-infrastructures around the world

EEF - European E-Infrastructure Forum

- Persistent storage
  - Work with middleware & data mgmt providers so PIDs are supported
- User support (e.g. interoperating problem handling procedures)
- Training & consultancy (contribute to customised training events/material organised by ESFRI projects)
  - EGI relies on NGIs for material and trainers
- Standards – offer web service interfaces for all relevant e-infrastructure services
  - Will require some middleware developments
- Cloud / volunteer computing integration
  - Rely on 3rd party projects (StratusLab, EDGI etc.)

Summary

- EGI & EGI.eu:
  - Provide a sustainable production e-infrastructure
  - EGI.eu is now a legal entity based in Amsterdam
  - Supported transition for 4 years through EGI-InSPIRE
- Bridging e-Infrastructures
  - Need to providing an integrated environment
    - **Both for end-users & operations**
- Contact: director@egi.eu

EGI Technical Forum

14-17th September 2010 in Amsterdam