Geographical Information Services and Virtual Organisations

Mike Mineter
m.mineter@ed.ac.uk

5th July 2004
National Centre for e-Social Science, University of Manchester

Goal of talk

To explore the following:

- **Geographical information (GI) is created and used by collaborations**
  - By locating data relative to the earth multi-disciplinary science, decision-making and policy can be supported
- **Grids enable collaboration**
  - Supporting cooperation amongst diverse, geographically widespread organisations
  - Enable the integration of resources of people, data, computation
- **Therefore Grids set new horizons for the use of GI**

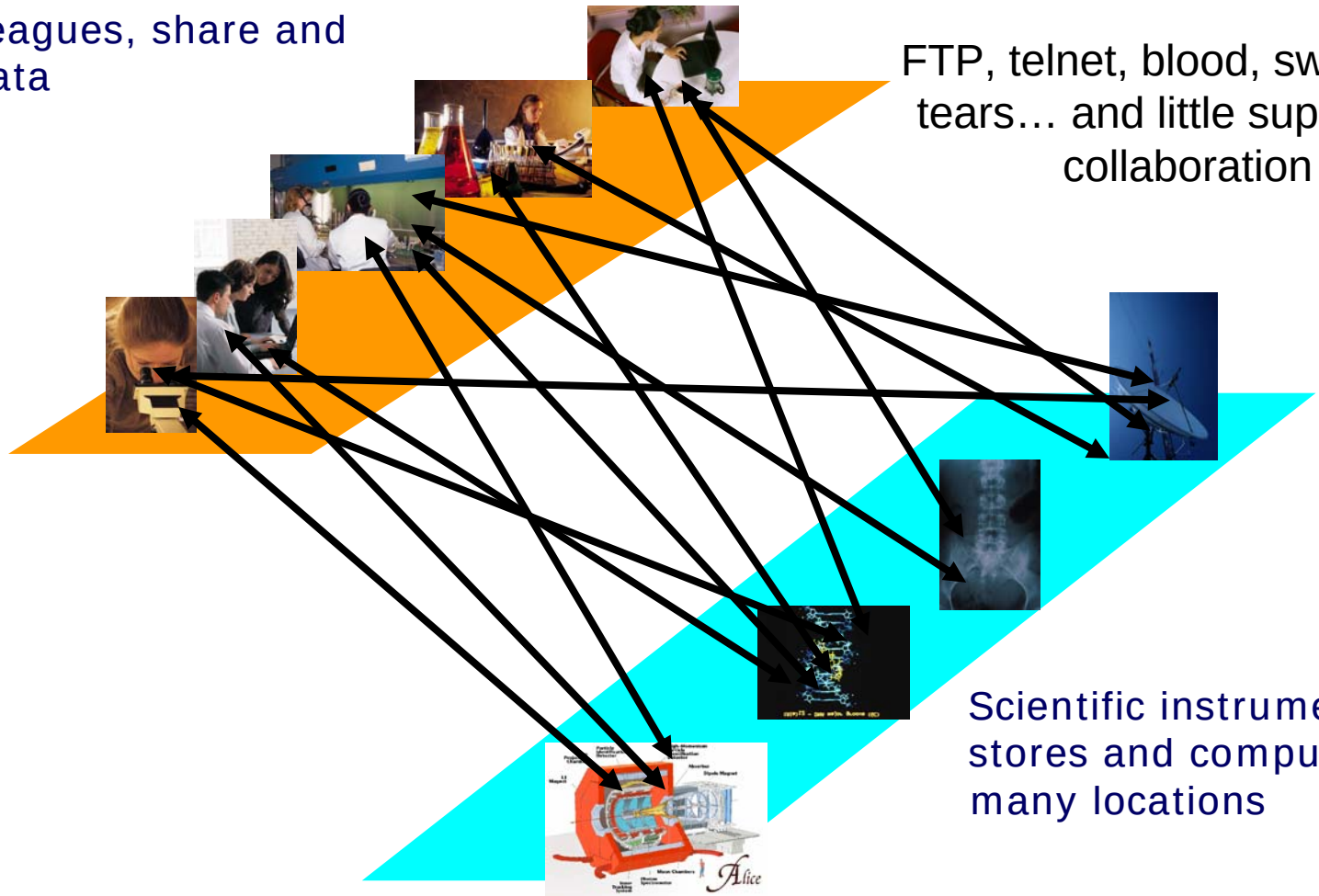
"Virtual organisations" and Service-orientation are key

Outline

- **Grids**
  - What are Grids?
  - Grids and Virtual Organisations (VO)
    - The VO in particle physics and in EGEE
- **GIS**
  - From programs to service-orientation
  - Standards: the Open GIS Consortium
- **Towards Geospatial services on Grids**
- **What are the key issues?**

Before Grids

Researchers in many locations need to interact with colleagues, share and access data



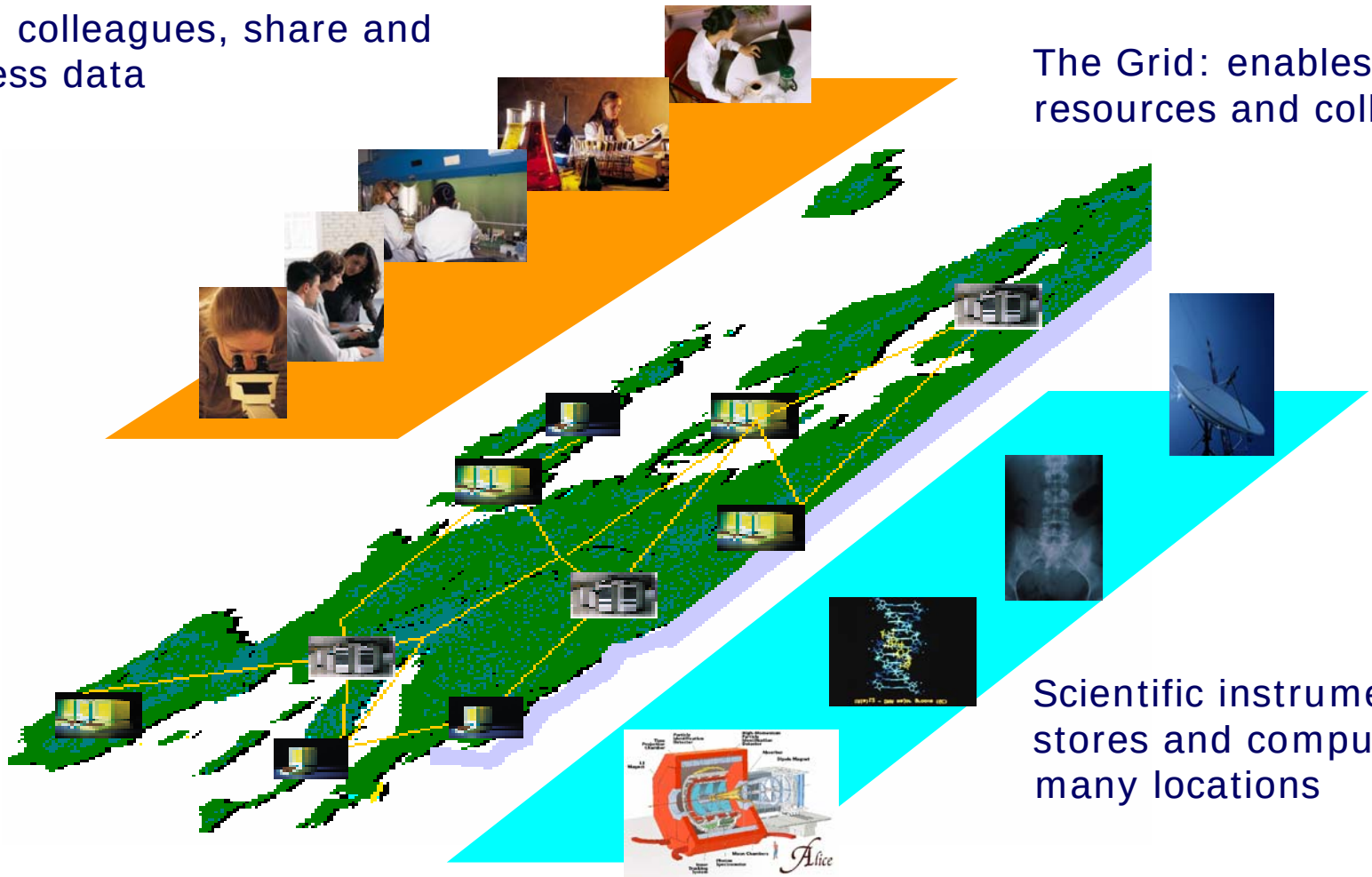
FTP, telnet, blood, sweat and tears… and little support for collaboration

Scientific instruments, data stores and computers in many locations

The Grid Vision

Researchers in many locations need to interact with colleagues, share and access data

The Grid: enables sharing of resources and collaboration



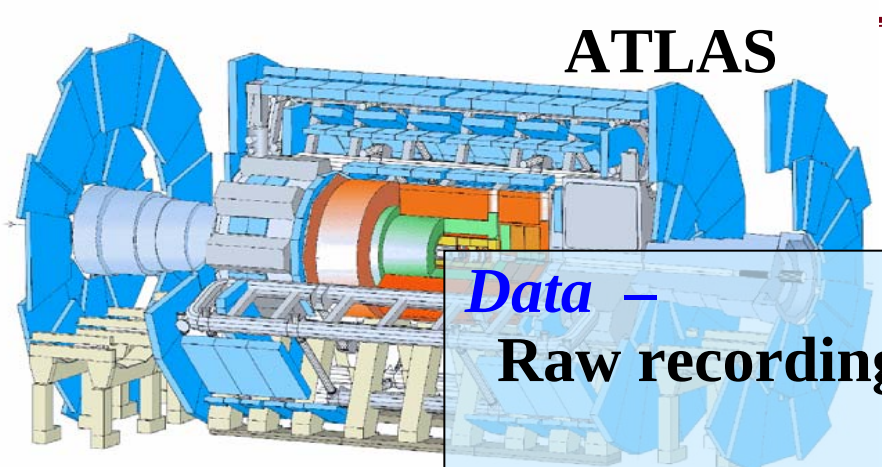
Scientific instruments, data stores and computers in many locations

Slide derived from EDG / LCG tutorials

Grids and Virtual Organisations

- **A Grid provides an infrastructure that coordinates flexible, secure, sharing of resources (e.g. computers, data, storage)**
  - for one or more virtual organisations (VO)
  - assuming the absence of central control and of trust relationships.
- **A VO is a collection of collaborating, geographically widespread individuals and institutions**

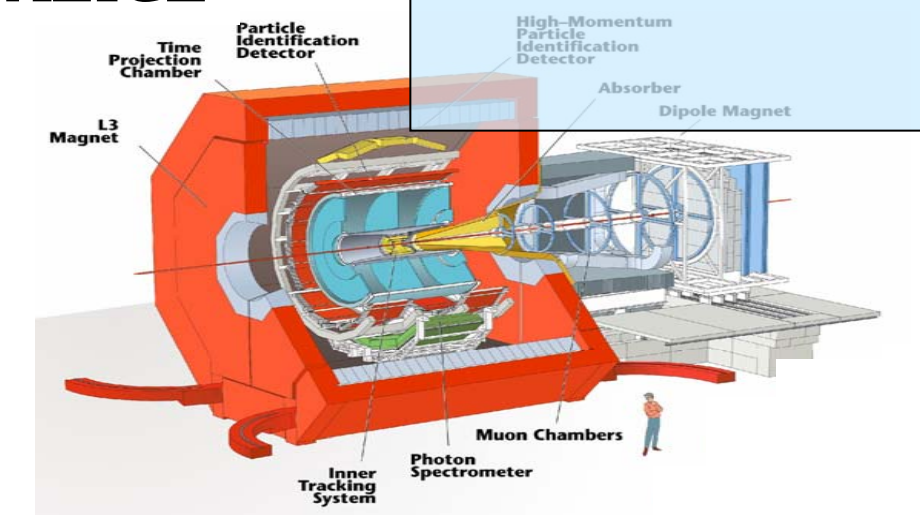
Particle physics: *the* "traditional" Grid VO

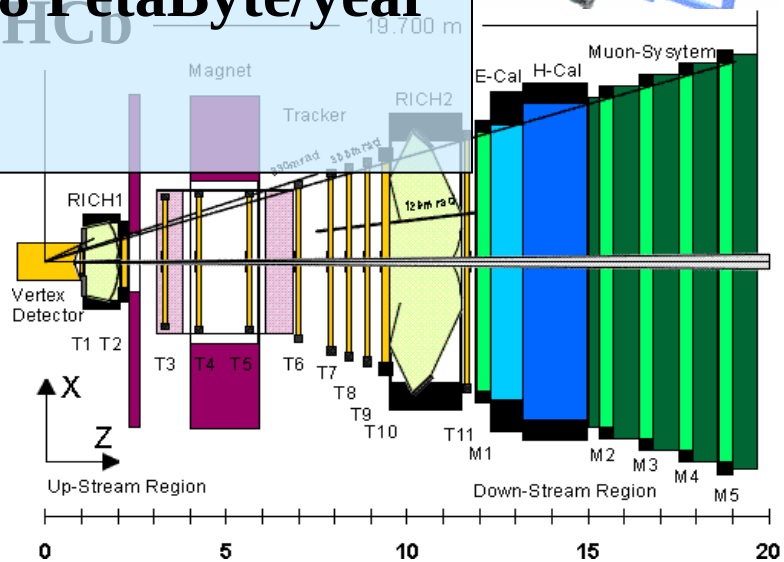


Data –

Raw recording rate 0.1 – 1 GByte/s

Accumulating at 5-8 PetaByte/year



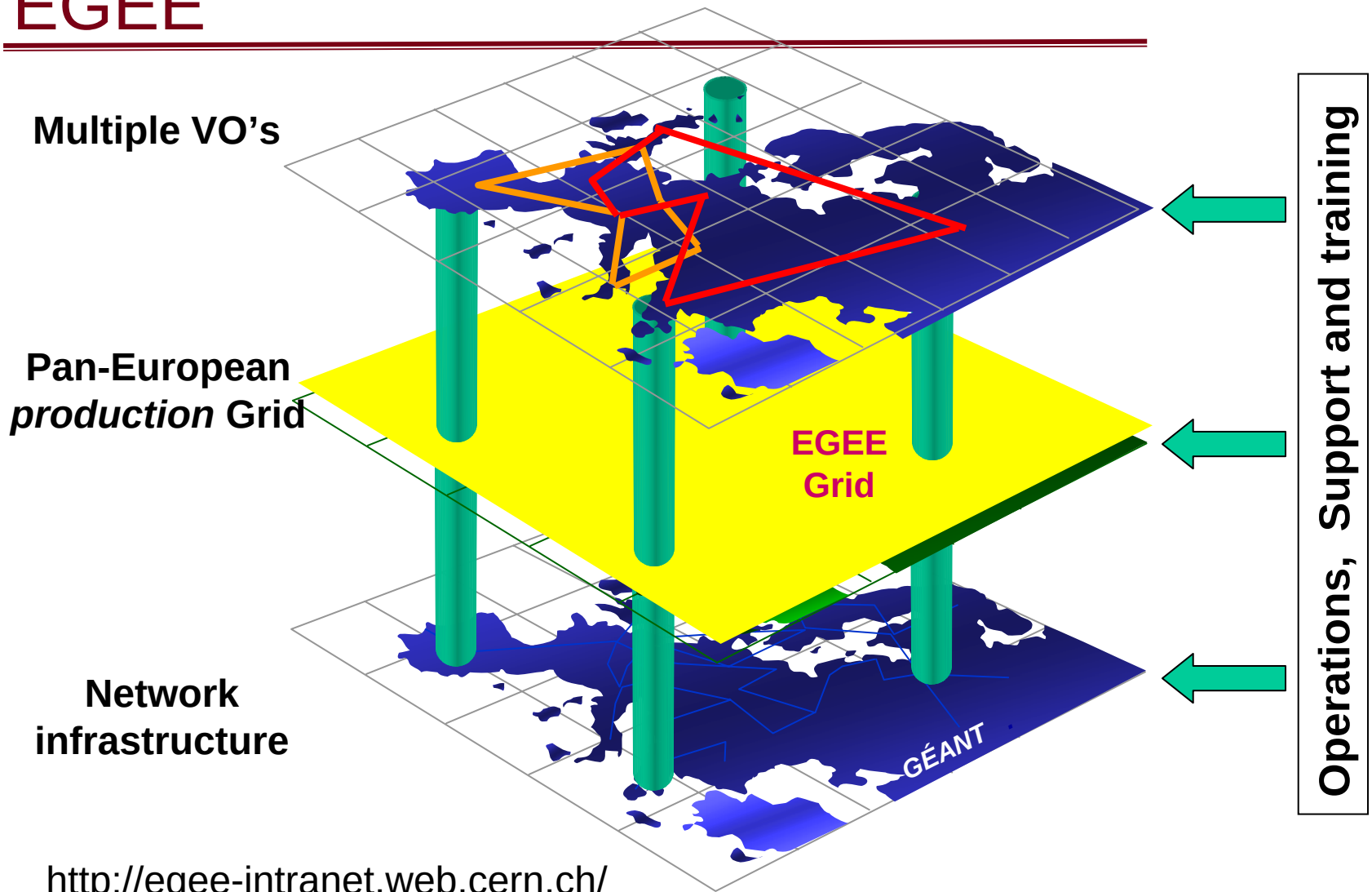


Slide derived from EGEE / LCG presentations

The *Large Hadron Collider* Computing Grid

- **LHC Computing Grid:**
  - 22 Countries
  - 58 Sites
  - 3800 cpu
- **5 VO's: one for each experiment + 1 for developers**
- **Relatively simple VO structure:**
  - Particle physicists can only engage with these experiments by using the LHC Computing Grid (LCG)
  - VO's will be needed for years
- **Massive data volumes, need for high throughput computing**
- **http://lcg.web.cern.ch/LCG/**

Enabling Grids for eScience in Europe
EGEE



http://egee-intranet.web.cern.ch/

slide modified from EOGEO conference talk by Bernhard Fabianek, European Commission, Information Society Directorate-General

EGEE: adding a VO

EGEE has a formal procedure for adding new user communities. Currently, practicalities include:

- **Negotiation with one of the Regional Operations Centres**
- **Seek balance between the resources contributed by a VO and those that they consume.**
- **Resource allocation will be made at the VO level.**
- **Many resources need to be available to multiple VOs : shared use of resources is fundamental to a Grid**

Virtual Organisations: issues for GI

- **GIS users in general are in more complex VO's, more flexibly formed, for various lifetimes, than is the case for many traditional users of Grids – e.g. collaboration might be for specific purposes, not lasting many years**
- **To what extent:**
  - do "traditional" grid VOs contrast with VOs that use GI ?
  - do the differences matter ?
- **Temptation is to address technical issues, and forget the context of a VO**

Outline

- **Grids**
  - What are Grids?
  - Grids and Virtual Organisations (VO)
    - The VO in particle physics and in EGEE
- **GIS**
  - From programs to service-orientation
  - Standards: the Open GIS Consortium
- **Towards Geospatial services on Grids**
- **What are the key issues?**

Geographical Information Systems

- **Progressing from *programs* that store, process, display GI towards**
  - services based on open standards (OGC, ISO…)
- **Current standards**
  - focus primarily on data representation and data interoperability
  - Have already had tremendous impact!!
- **Ongoing work:**
  - Stronger data models (e.g. to allow integration with 3-D and 4-D (time) data)
  - Support for chained geoprocessing Web Services

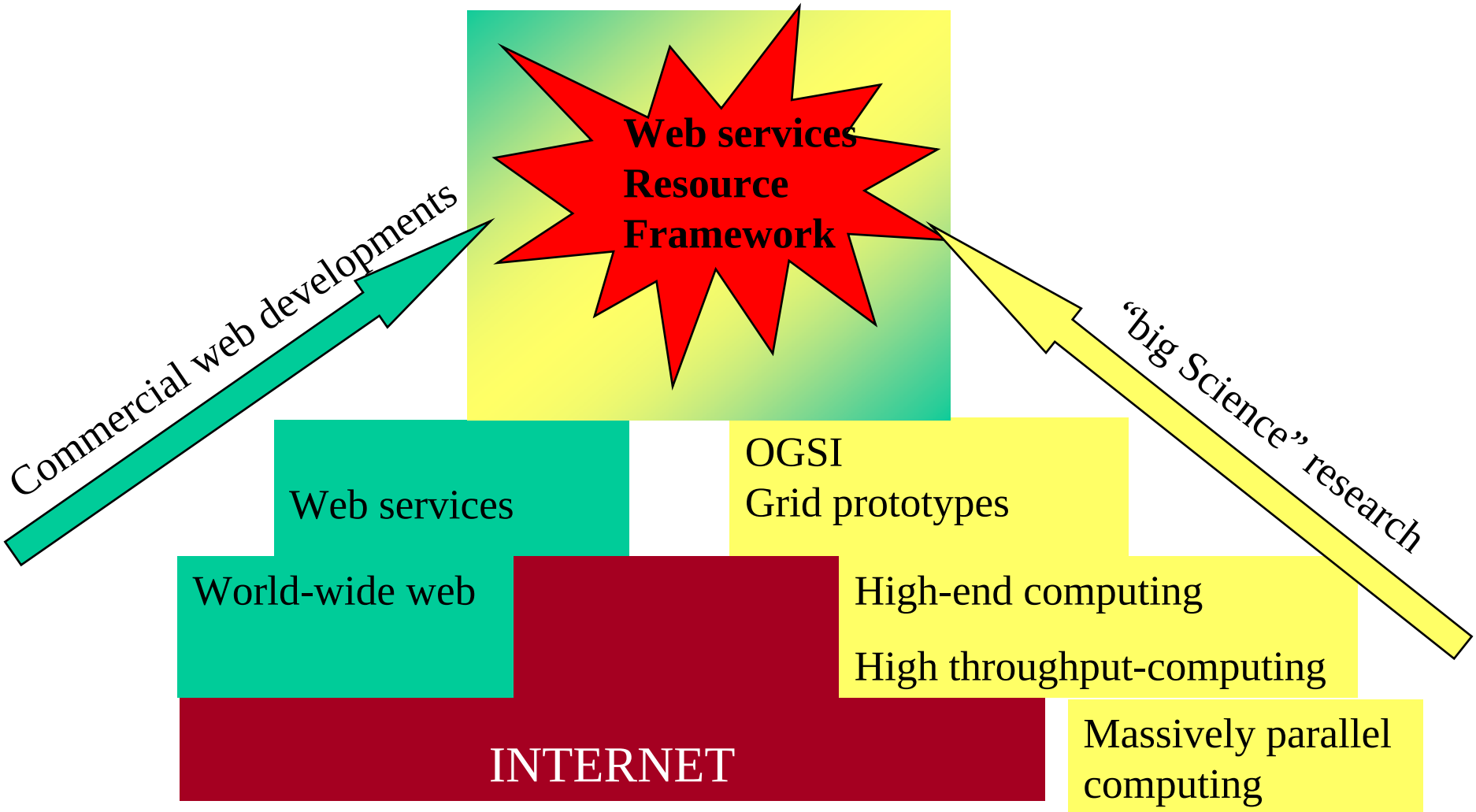
Effect of move to WS and open standards

- **Service orientation with web services is growing**
- **Standards exist (even if development continues)**
- **Stimulates development of new software including open source toolkits…**
- **…including Grid-based GI services**

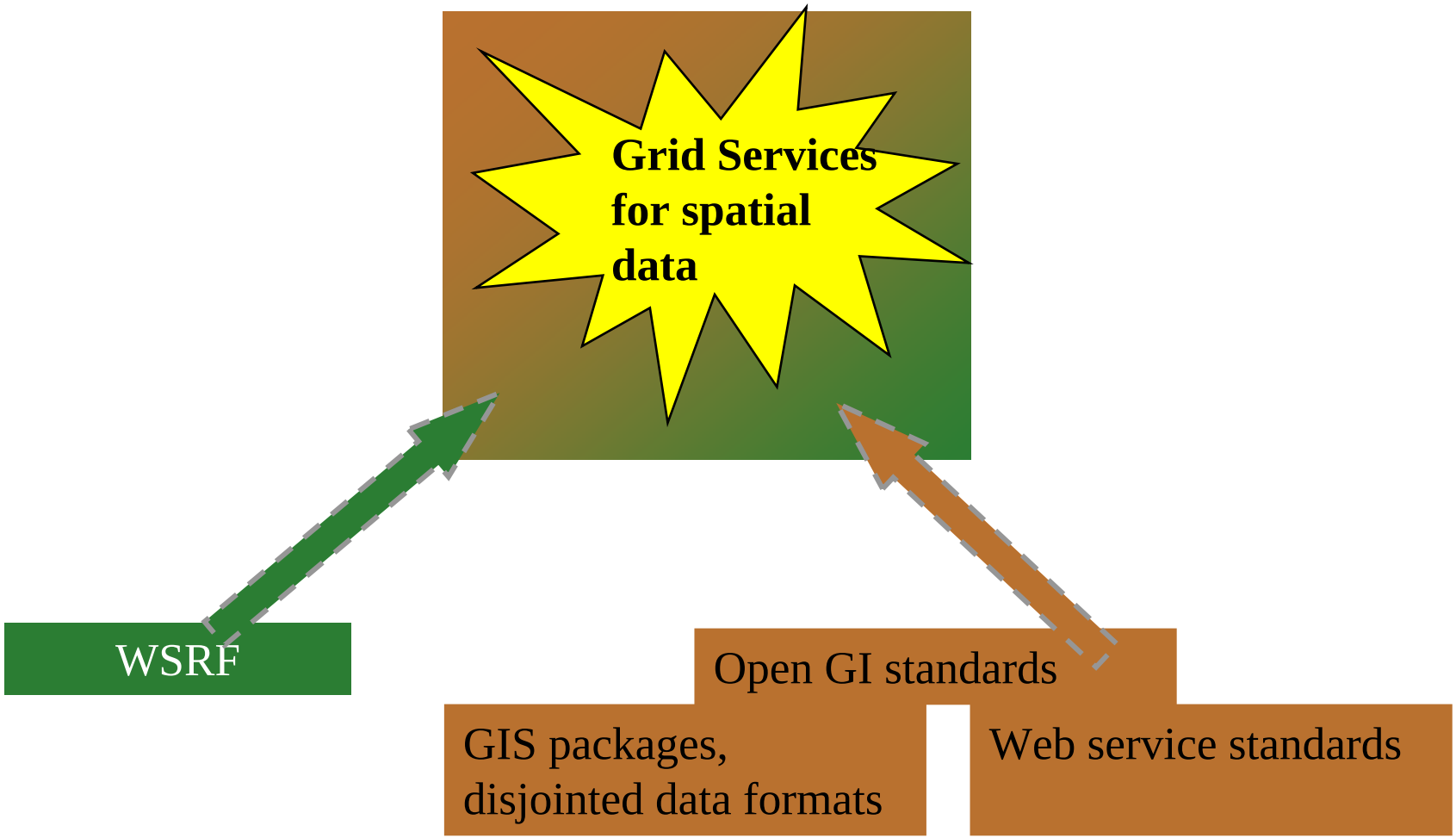
Outline

- **Grids**
  - What are Grids?
  - Grids and Virtual Organisations (VO)
    - The VO in particle physics and in EGEE
- **GIS**
  - From programs to service-orientation
  - Standards: the Open GIS Consortium
- **Towards Geospatial services on Grids**
- **What are the key issues?**

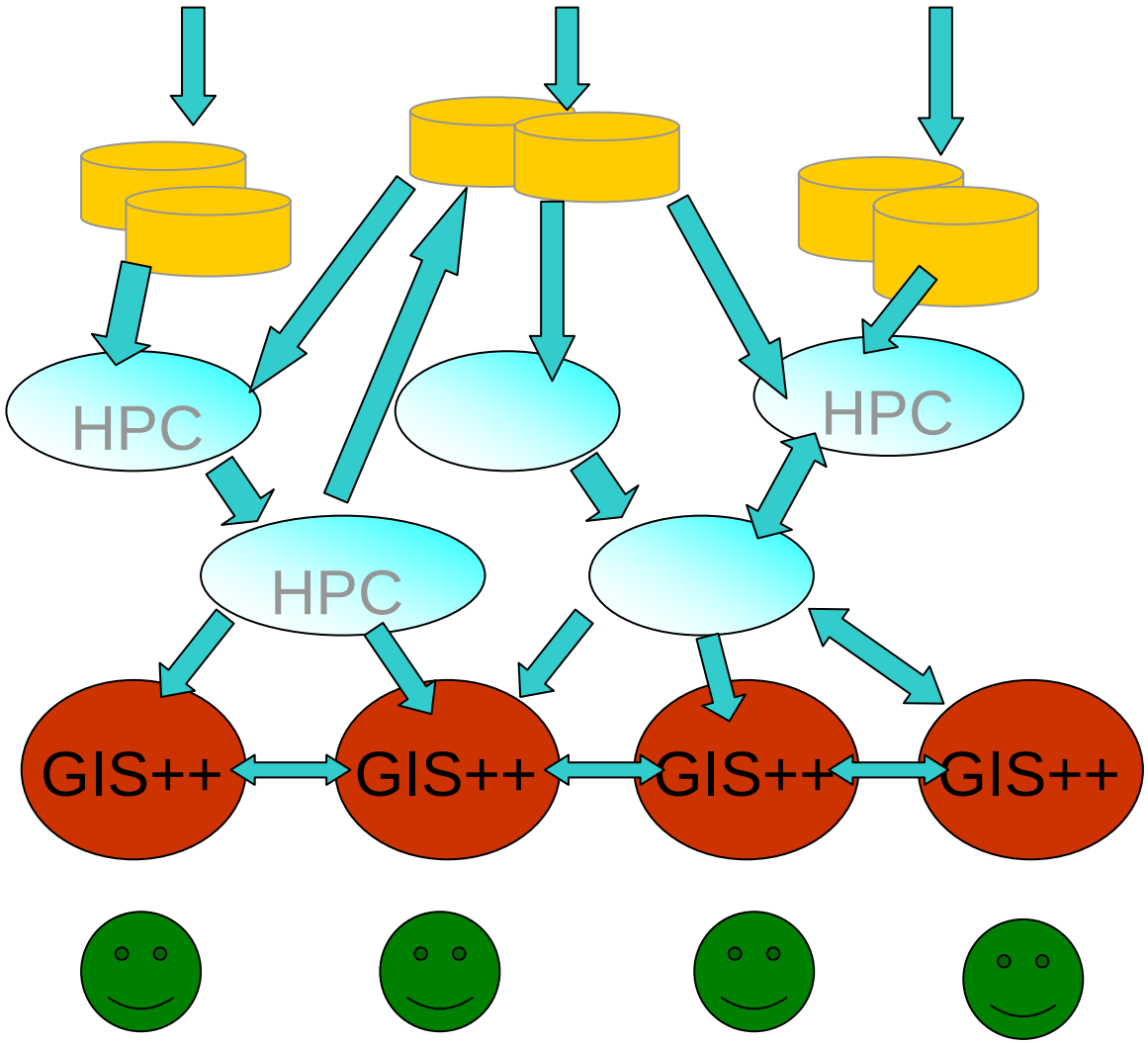
Convergence of Web Services and Grids



Towards GI Services on the Grid



GIS in Grids



Data creation

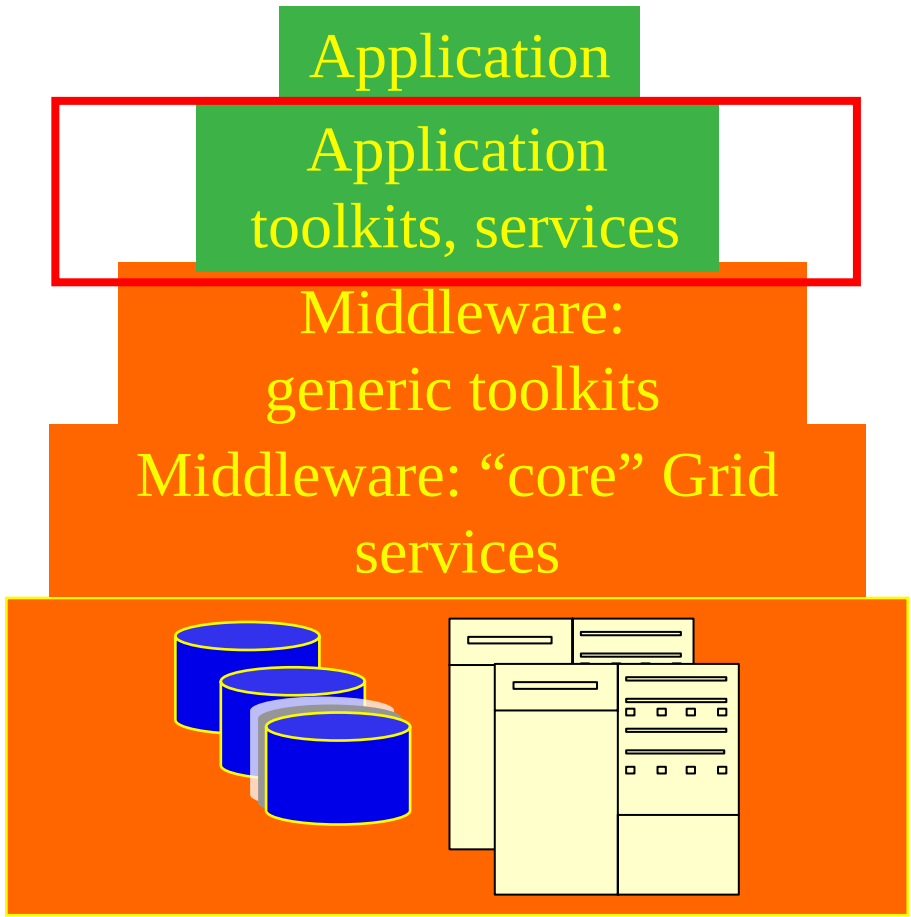
Data access & integration

Computation

Collaboration

Organisations

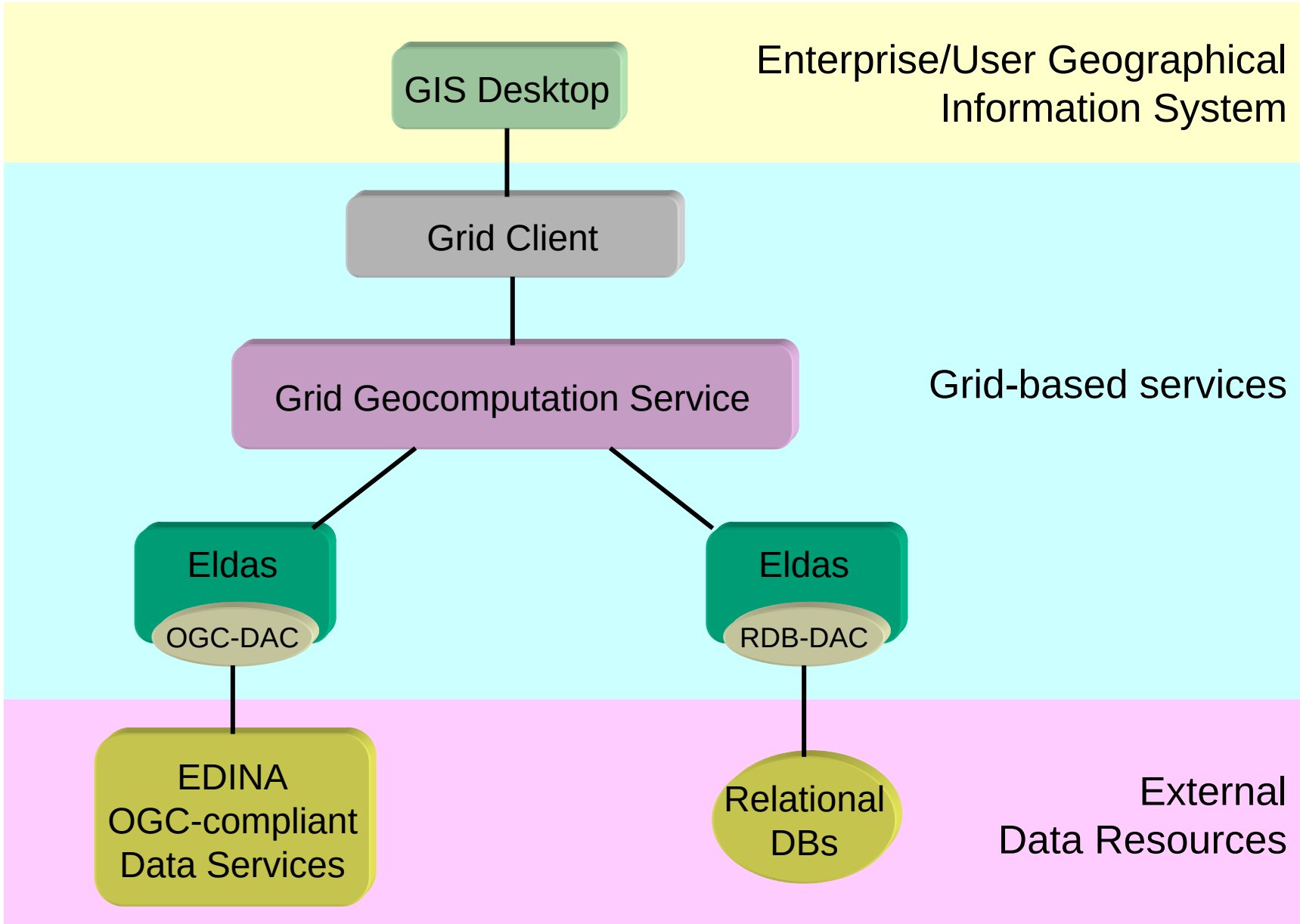
Towards Grid-aware GI services



- **Insulate applications from changing middleware**
- **Build distributed applications from components**
- **Tools/languages for application domains**
  - Data access and integration (GI data + RDBMS + legacy model results)
  - Workflow and geoprocessing services
  - Discovery of data / services

EDINA – EDIKT project

- **EDINA, in the University of Edinburgh Data Library: serves GI to the academic community**
- **EDIKT**
  - e-Science Data, Information and Knowledge Transformation
  - Close ties to EPCC and to NeSC, Edinburgh
- **ELDAS: Enterprise Level Data Access Services, Developed by EDIKT: http://www.edikt.org/**
  - allow multiple client connections to multiple databases
  - from web services or from Grid (OGSI) services
- **Goal of project: A few months effort to**
  - Add services that can access GIS OGC-compliant data, via WS interfaces
  - Direction of effort shown in next slide – currently in design phase


GIS Desktop
Enterprise/User Geographical Information System
Grid Client
Grid Geocomputation Service
Grid-based services
Eldas
OGC-DAC
Eldas
RDB-DAC
EDINA OGC-compliant Data Services
Relational DBs
External Data Resources

Outline

- **Grids**
  - What are Grids?
  - Grids and Virtual Organisations (VO)
    - The VO in particle physics and in EGEE
- **GIS**
  - From programs to service-orientation
  - Standards: the Open GIS Consortium
- **Towards Geospatial services on Grids**
- **What are the key issues? – incomplete suggestions!**

Issues – suggestions

- **Characteristics of applications**
  - Require workflow with both loose and tight-coupled components
  - Need more than batch-oriented Grids
- **Characteristics of VOs that use GI**
- **Characteristics of GI data (increasingly diverse sources) and open standards**
- **Where do the Spatial Data Infrastructures fit in?**
- **Ask not what ‘the Grid’ can do for you, but what you can do for a VO**

On these themes:

- **EOGEO 2004 Workshop, UCL, London**
- **Sessions on Grids and GIS**
- **http://www.eogeo.org/Workshops/EOGEO2004/**