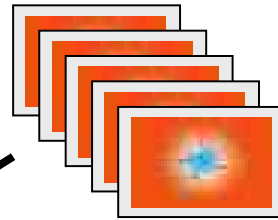




# Inside Atlas



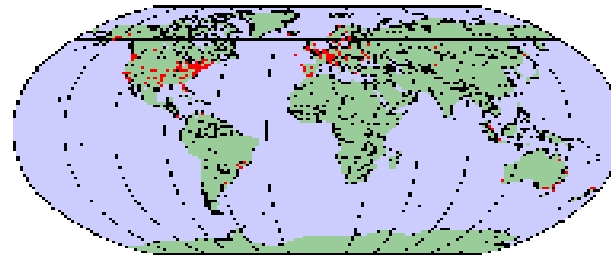
RAW data



**10 PB RAW data per year**

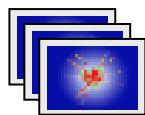
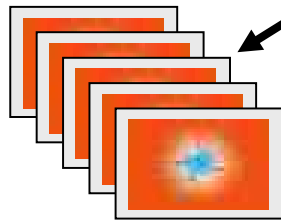
+ derived data, extracted physics data sets, filtered data sets, artificial data sets (Monte-Carlo), multiple versions, Multiple copies

**→ 20 PB of data at CERN per year**

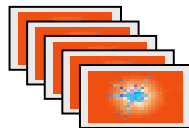


**→ 50 PB of data in addition transferred and stored world-wide per year**

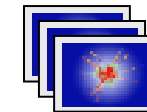
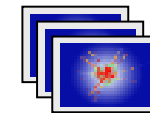
RAW data copies



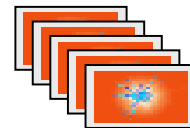
Enhanced copies



Sub-samples



Enhanced copies



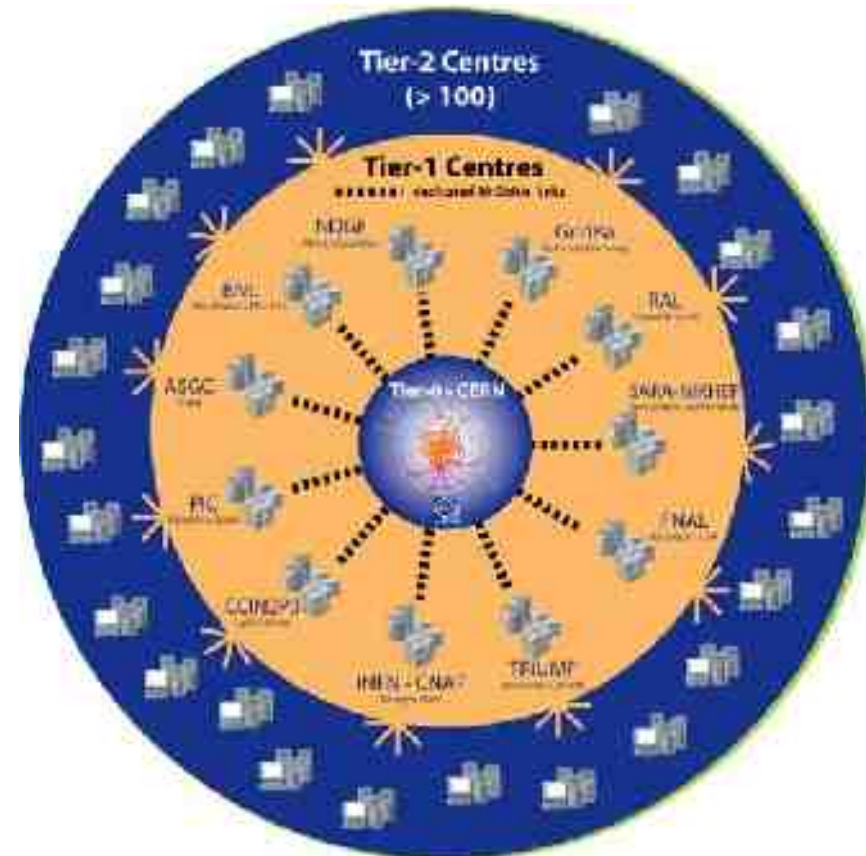
# Tier Structure

Today there are about 320 sites world-wide participating in the WLCG project.

- Tier0 center, CERN  
very large center, tape storage, first level processing and meta-data storage, quality service (24\*7)
- Tier1 center, **11** world-wide  
large capacity, tape storage, quality service
- Tier2 center, **168** world-wide  
medium size, some large, no 24\*7 service, no custodial storage
- Tier3 center, **140** world-wide  
very small to medium size, no availability guarantee, focus on end-user analysis activity

## Resource installation today:

- ~ **110 PBytes** disk space
- ~ **100 Pbytes** tape space
- ~ **150000** processor cores



Data flow between Tiers:  
multi GBytes/s    multi PBytes/month





## **Expansion of the CERN computing capacity**

→ 5 MW power and cooling

Several options under discussion and planning

6. Purchasing space+power in a Hosting-Company

7. Add-on based on container (the 'google-way')

8. Further expansion of the existing building

9. Remote computer center (e.g. Norway)

10. Purchase of 'cloud-computing' resources

### **TCO considerations**

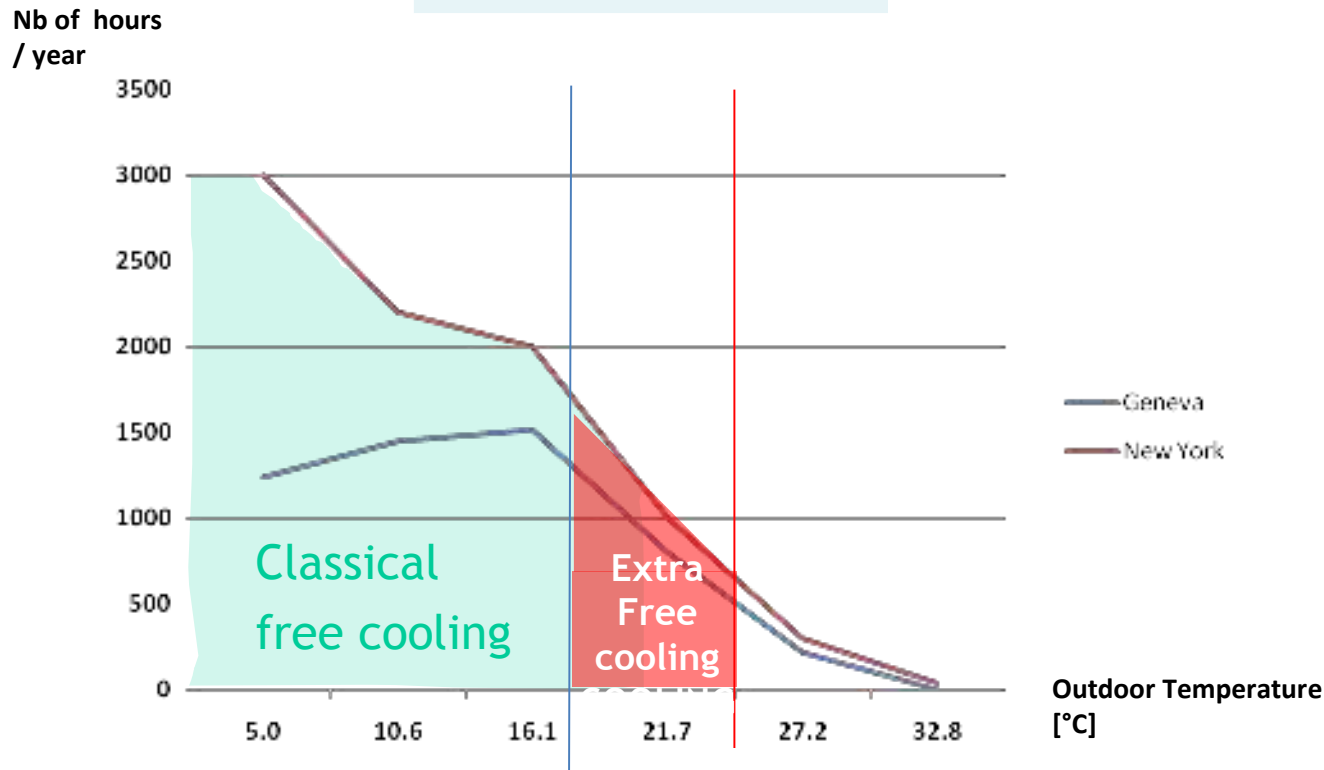
### **Implications for the computing models**

### **(data flow, work flow, operations)**



# Basic Housekeeping is essential!

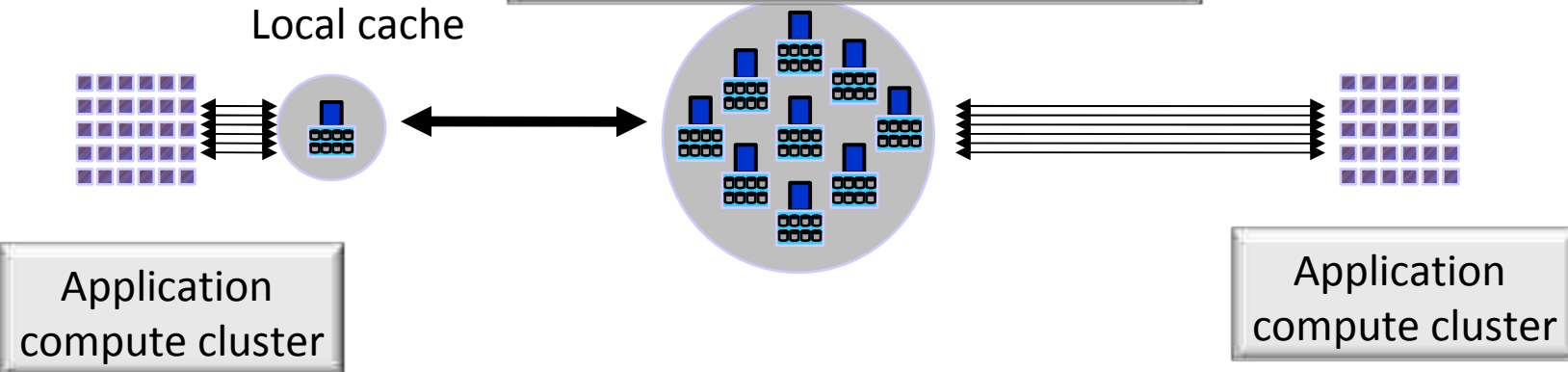
Hot is cool!



Indoor conditions from 20°C to 27°C

But not too hot: Vendors now accept that their systems will run at higher temperatures, but there are reports that server energy efficiency starts to drop above a certain temperature.

# Very large scale Data Repository Global repository for everything ?!



## Data access coupling models

- direct access
- local caches

### Important metrics:

Amount of storage  
Access patterns  
Data throughput

### Dependencies:

Data integrity  
Data preservation  
Data access

### Data storage

- application agnostic
- application aware

**TCO !**



## Three cost and power scenarios for a one Exabyte space repository (rough estimate of initial investment, building+power+cooling+equipment)

1. Slow performance, 'backup' access , 10\*10 Gbit network connectivity, low number of users, tape based  
→ ~120 MEuro 0.5 MW power
5. Moderate performance, 'batch' access, 100\*10 Gbit network connectivity, medium number of users, disk based  
→ ~750 MEuro 20 MW power
9. High performance, 'interactive' access, 1000\*10 Gbit network connectivity, high number of users, solid-state-disk based  
→ ~5900 MEuro 50 MW power

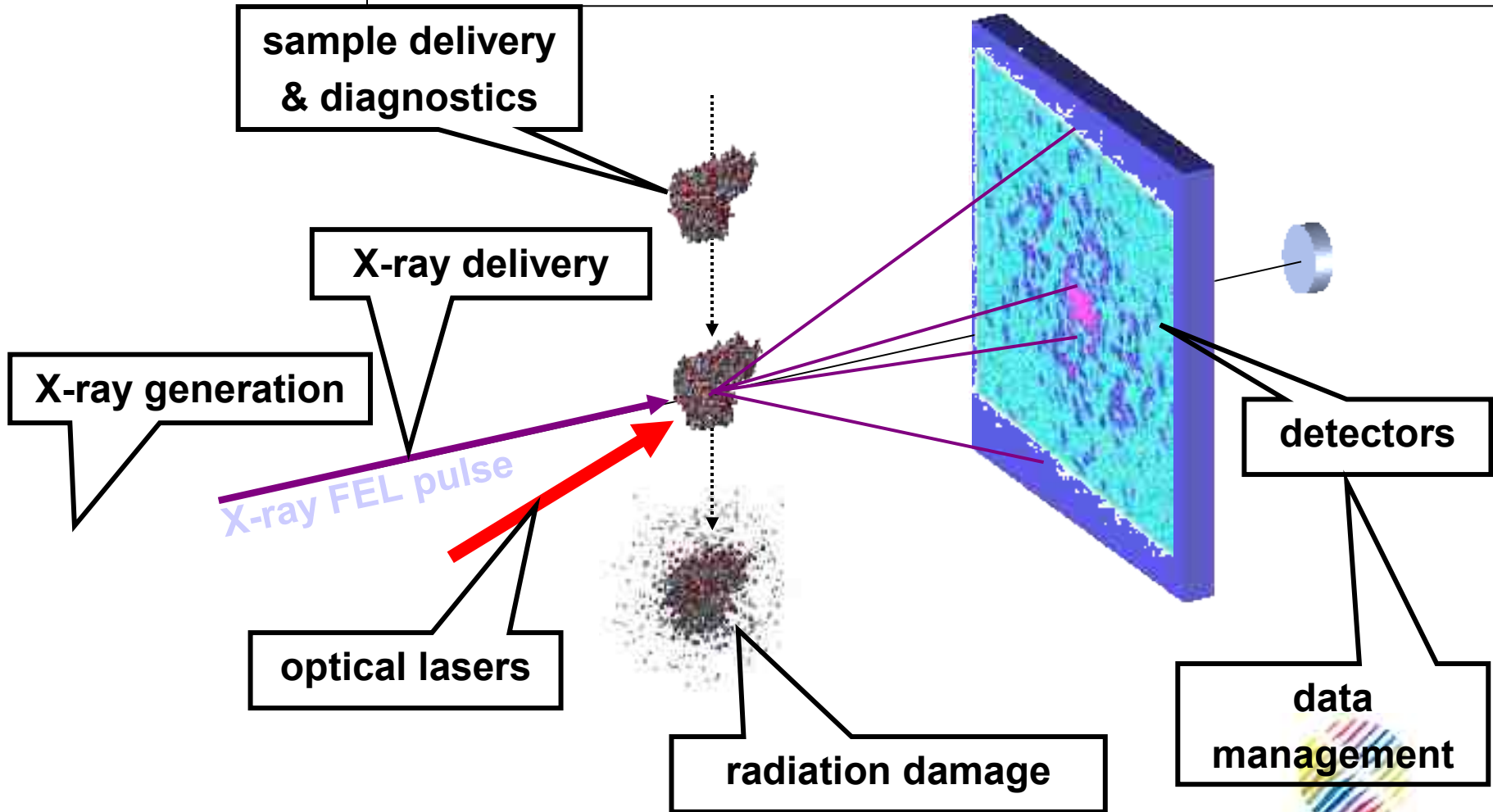
Plus running costs: ~10% of initial investment per year



## XFEL: Office and Laboratory Building



# The complete single particle experiment





# Technology Forecast – Storage at DESY

Year	Rate Capability [Gbyte/sec]	Storage Space [Petabyte]
2009	1	3
2012	5	26
2016	40	200

- not a technology problem
- money and manpower issues
- to be determined:
  - user behaviour
  - compression and accept/reject algorithms
- **potentially critical: access to data!**

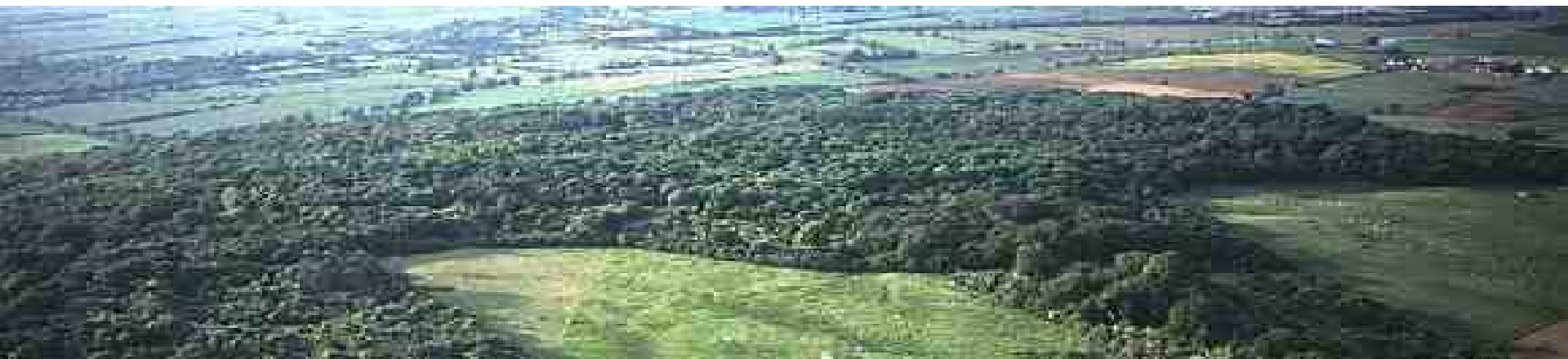




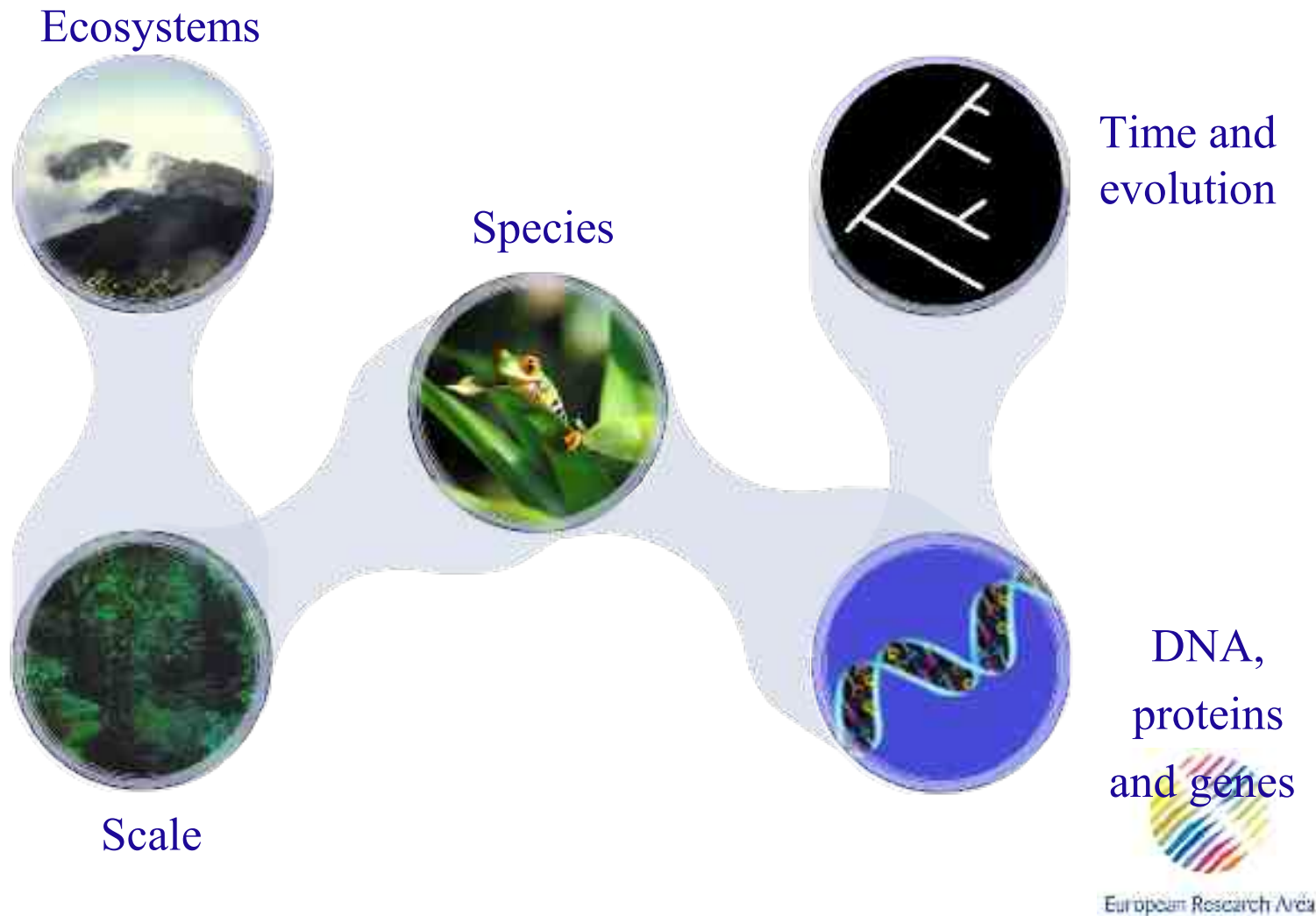
European  
Research Area  
Board (ERA)

● **LIFEWATCH**

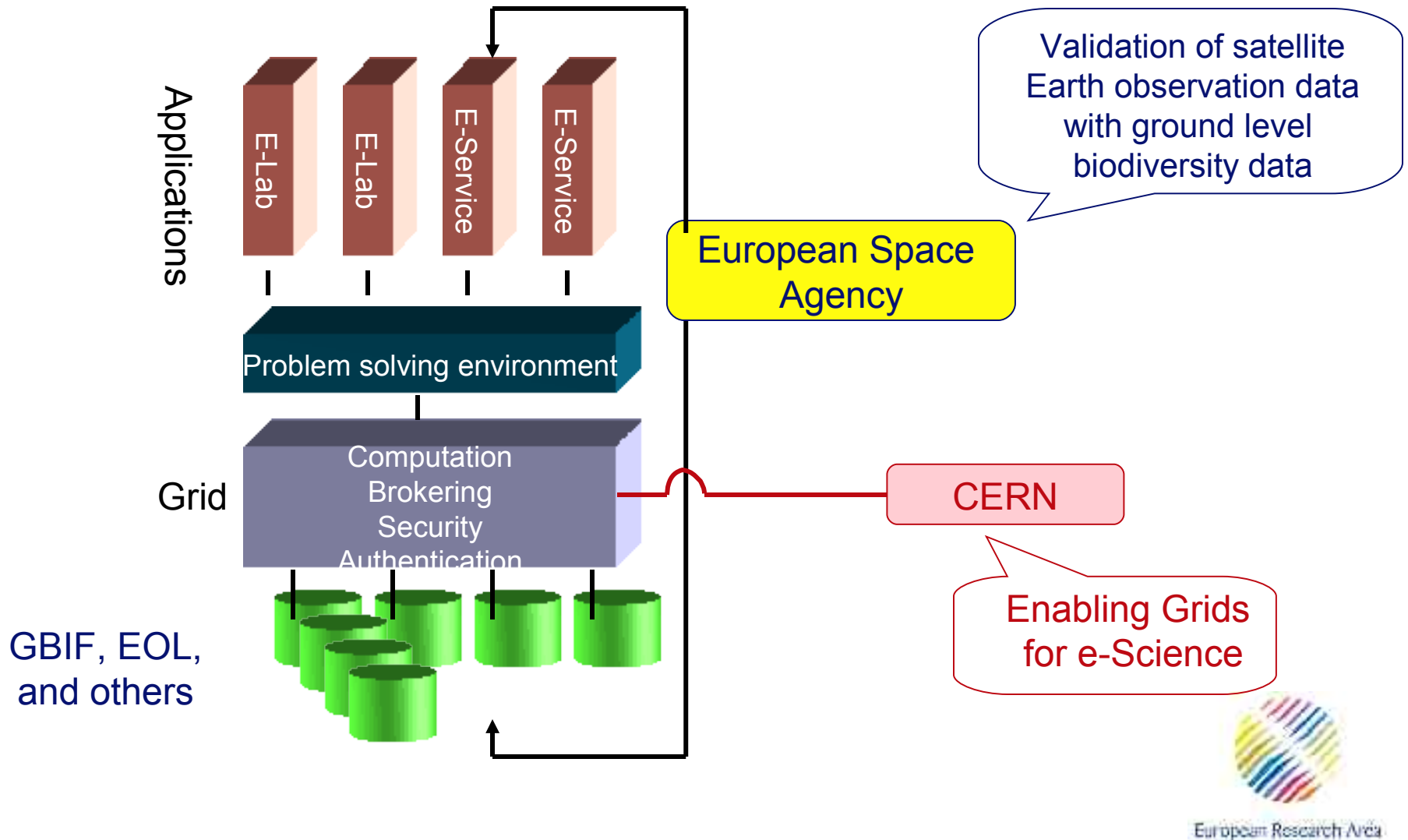
# Large-scale e-Infrastructures for Biodiversity Research



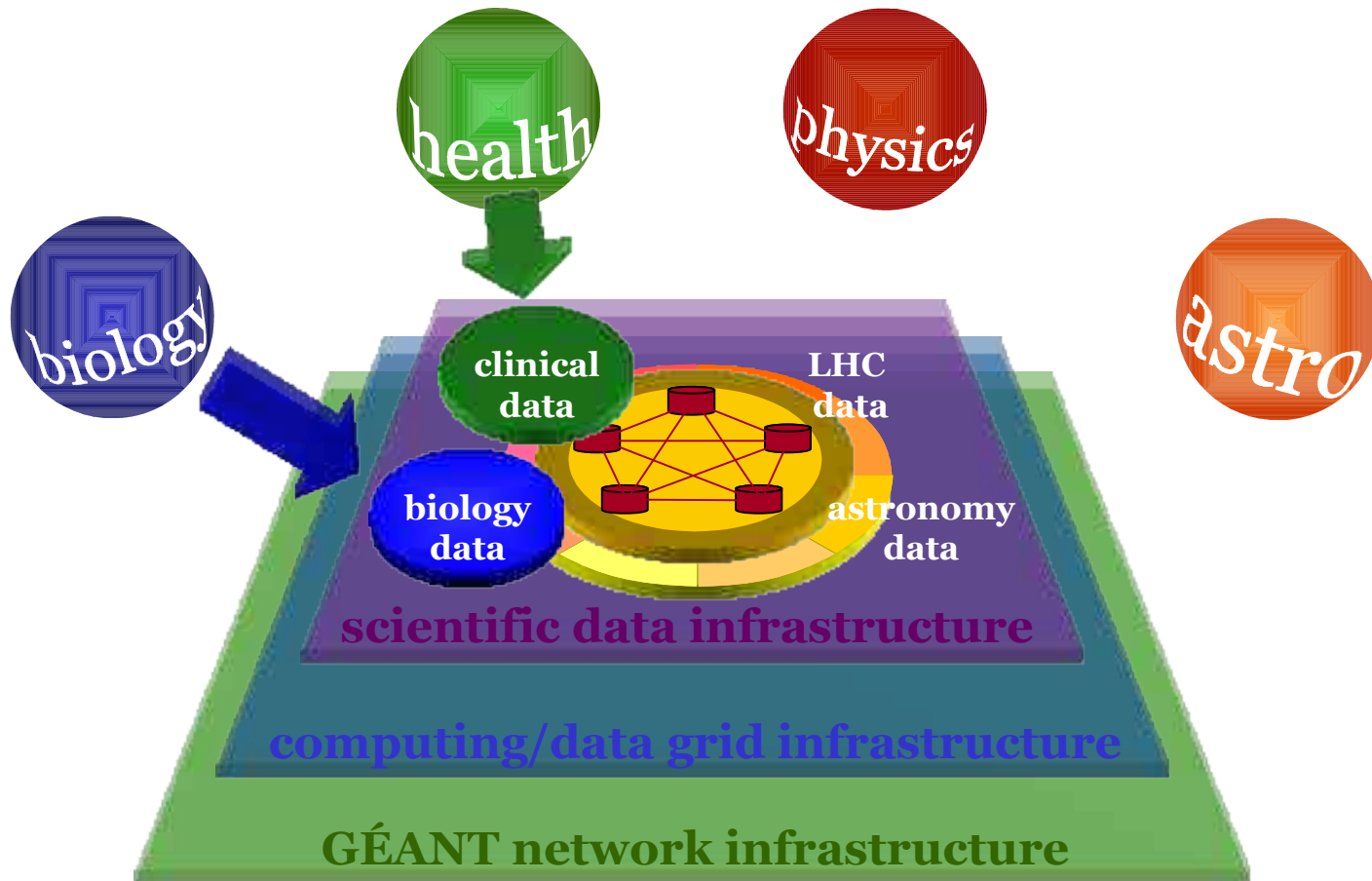
# Big questions in biodiversity research

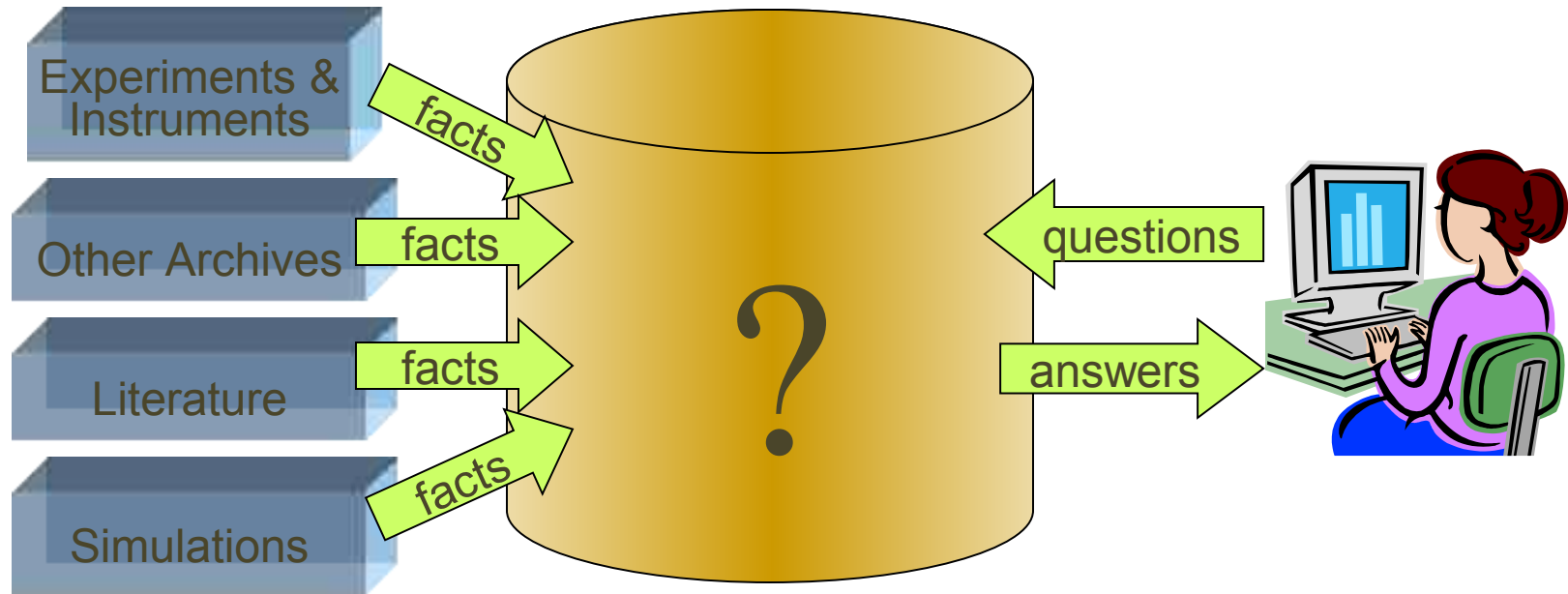


# A cyber community of infrastructures



# scientific data as an infrastructure



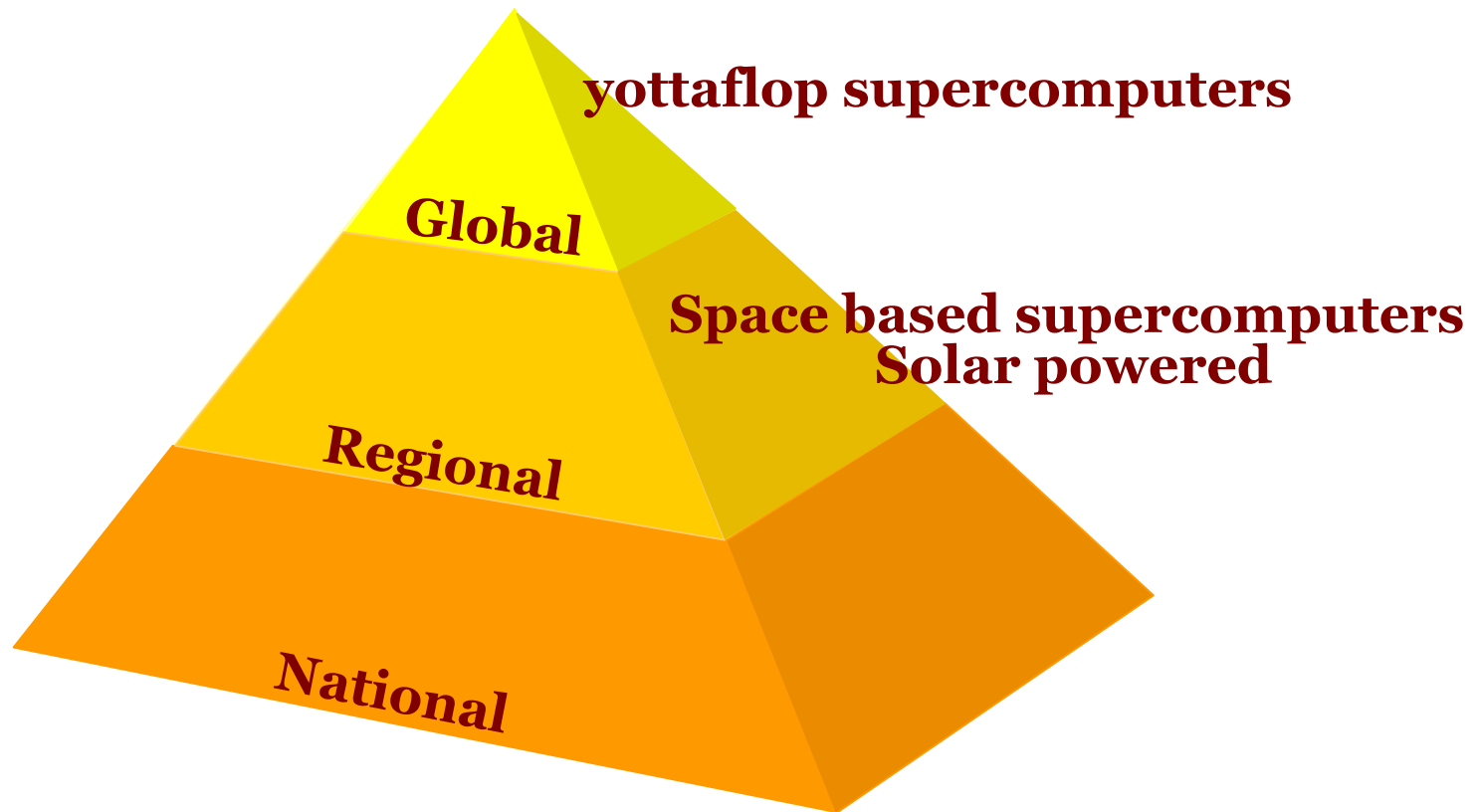


- ◆ Data ingest ◆
- ◆ Managing petabytes+ ◆
- ◆ Common schema(s) ◆
- ◆ How to organize? ◆
- ◆ How to re-organize? ◆

- How to coexist & cooperate with other scientists and researchers?
- Data query and visualization tools
- Support/training
- Performance
  - Execute queries in a minute
  - Batch (big) query scheduling



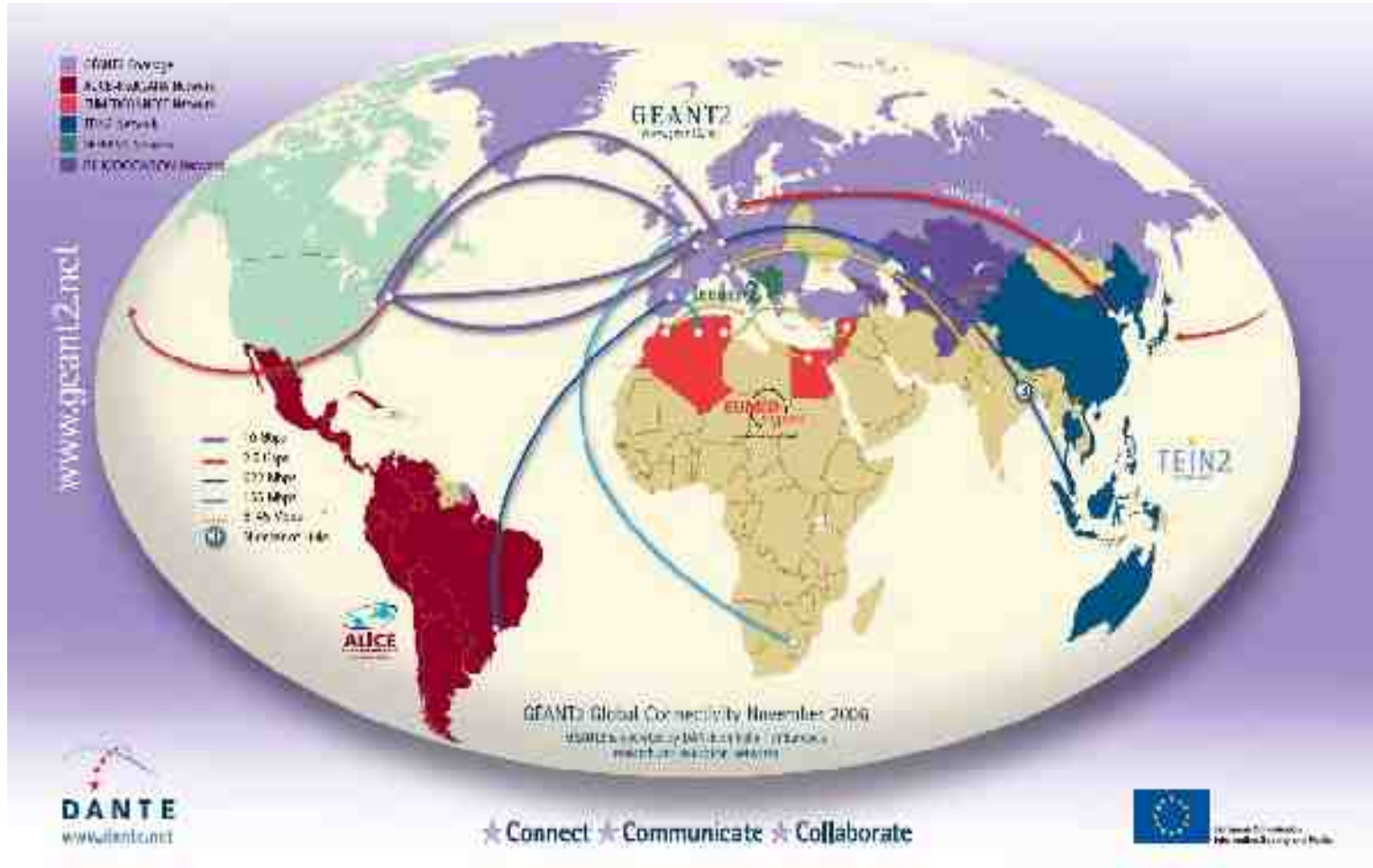
# Personal computing operates at the exa scale





European Research Area Board (ERA)

# GÉANT: global reach is now all pervasive even to small villages





# Looking back from 2070 on SKA if all goes well

## People

- 10s of thousands of students will have benefited from their training but more importantly for their ability to work across cultures. They will develop new ways of working, new technologies and inspire others to follow them
- Virtual Research Communities the norm – Ph.D. Programmes are global
- People from currently marginalised areas will be integrated into the knowledge based society through education, connectivity and local employment prospects.
- All world citizens will have access to data and have sufficient own computing power to analyse – the 5<sup>th</sup> Freedom will be a global reality.



# Technologies

- Data transfer rates will be orders of magnitude more than in 2010
- Secure world data storage solved
- Whole Galaxy simulations can be undertaken
- Energy requirements for supercomputers must be less than in 2010 or radical solutions such as off world locations will be needed. How to “feed a biological supercomputer?”
- Electronic and photonic requirements of SKA will have driven many other areas of application in mobile computing and communication
- Energy efficient cooling systems developed
- Materials and manufacturing systems for mass production of complex objects economically result (e.g. Dish construction technologies could be transferred to transport)
- New materials currently unknown are found stable in space and can be synthesised on earth

# Wider Aspects

- An intergovernmental organisation exists to foster all international S and T projects covering political and finance issues – join ERIC to ARIC and URIC?!”!
- Stabilisation of certain areas of the world that are disadvantaged leading to political stability
- Science communication between world leaders breeds trust and leads to further unblocking of world trade
- State aid rules are uniform allowing all countries to compete on equal terms.
- All school students can look at the data derived in real time and are inspired to follow S and T careers
- Further global agreement on standards for communication and sub nano scales



# Whimsical Thoughts

- Is there anyone out there or does the “Eerie Silence” continue?
- Is the SKA the new “Tower of Babel?”
- SKA dish graffiti art becomes a collector’s item and traded by international auction houses
- What will be the top news item resulting from the SKA – who will be the first SKA Nobel Laureate?

**Remember, the future is  
not what it used to be!**

**Make the most of it.**



