

Introduction to Grid Computing

My View

Victor Penso

GSI Darmstadt / FH Darmstadt



What is Grid

all over
the world



GRID

connected to
the Internet

many
computers

1

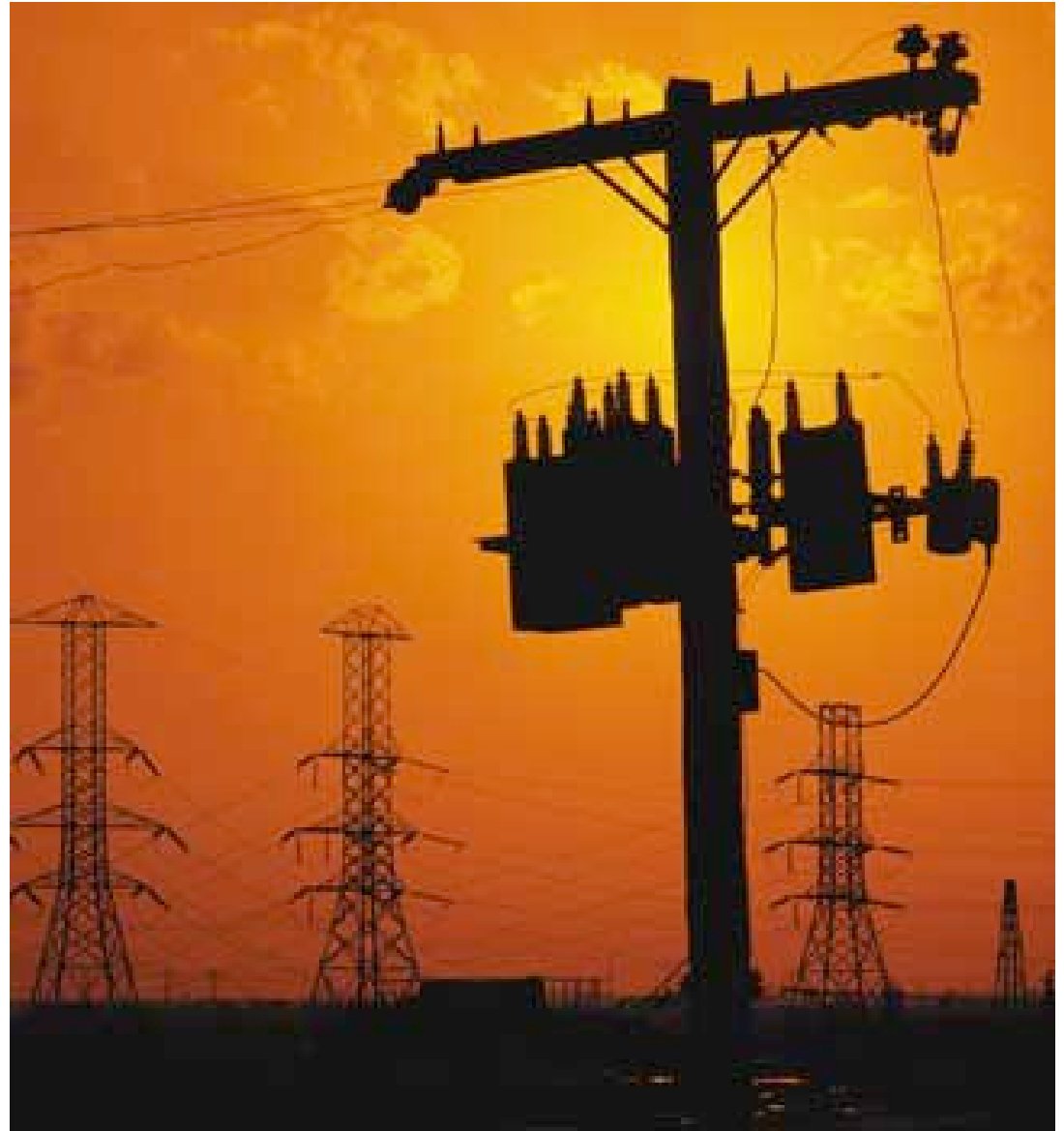
Transparency

2

Standardization

3

Reliability



Meta

Cluster

computing

computing

DISTRIBUTED

COMPUTING

Peer-to-Peer

Internet

computing

computing



- NASA Information Power Grid
- DOE Science Grid
- NSF National Virtual Observatory
- NSF GriPhyN
- DOE Particle Physics Data Grid
- NSF TeraGrid
- DOE ASCI Grid
- DOE Earth Systems Grid
- DARPA CoABS Grid
- NEESGrid
- DOH BIRN
- NSF iVDGL

- DataGrid (CERN, ...)
- EuroGrid (Unicore)
- DataTag (CERN,...)
- Astrophysical Virtual Observatory
- GRIP (Globus/Unicore)
- GRIA (Industrial applications)
- GridLab (Cactus Toolkit)
- CrossGrid (Infrastructure Components)
- EGSO (Solar Physics)

- UK e-Science Grid
- Netherlands – VLAM, PolderGrid
- Germany – UNICORE, Grid proposal
- France – Grid funding approved
- Italy – INFN Grid
- Eire – Grid proposals
- Switzerland - Network/Grid proposal
- Hungary – DemoGrid, Grid proposal
- Norway, Sweden - NorduGrid

**Views are not
identical**

BUT

Strong links

Web-Service

*INDUSTRIES
MEETS
SCIENCE*



Grid-Service

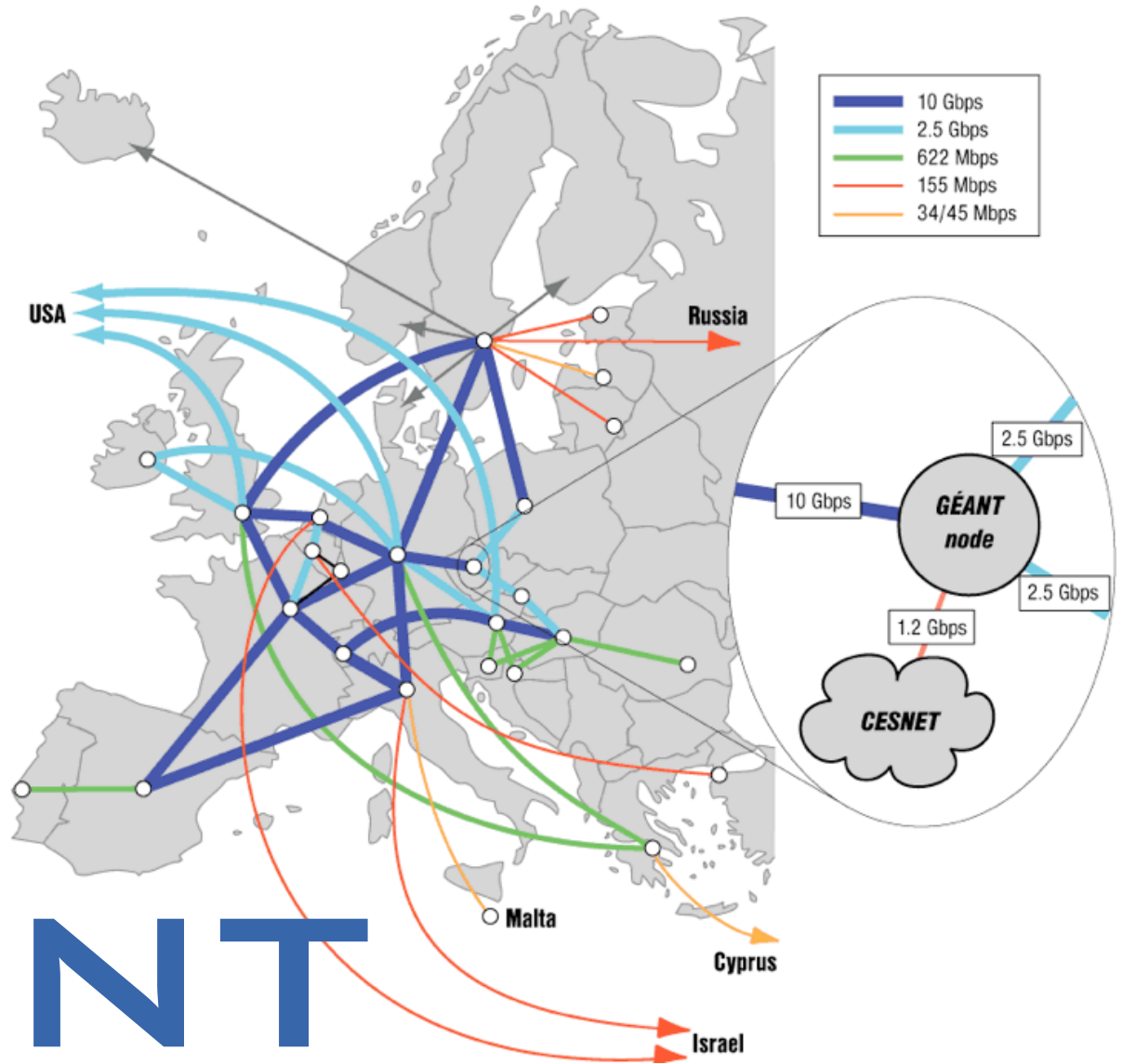
Sharing resources

Security

Balance the load

Distance not matters

Open standards



GEANT

50 Gigaflop

ACMAPS 1989

Mac G4 2003

40 Teraflop

NEC Earth
Simulator 2004

85 Teraflop

HPC2500
Fujitsu 2005

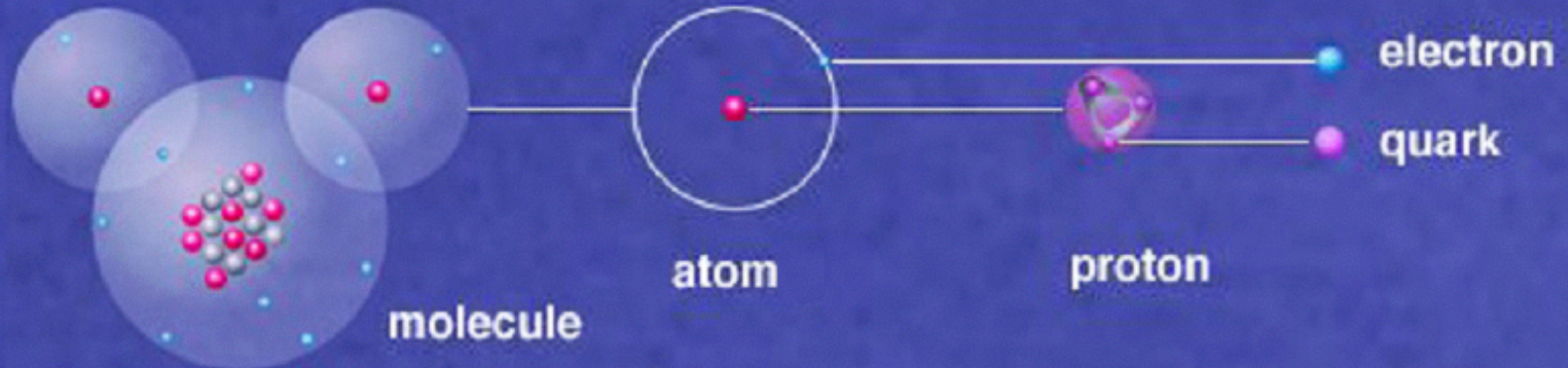
flop is a floating point operation

- CERN is leading some of the **most ambitious Grid projects** in the world.
- CERN has chosen Grid technology to solve a huge data storage and analysis challenge it faces in 2007, when the **Large Hadron Collider**, the biggest scientific instrument in the world, starts running.



- Physicists smash particles into each other to:
 - identify their **components**
 - **create** new particles
 - reveal the nature of the **interactions** between them
 - create an environment similar to the one present at the origin of our Universe

- **What for?** To answer fundamental questions like:
*How did the Universe begin? What is the origin of mass?
 What is the nature of antimatter?*

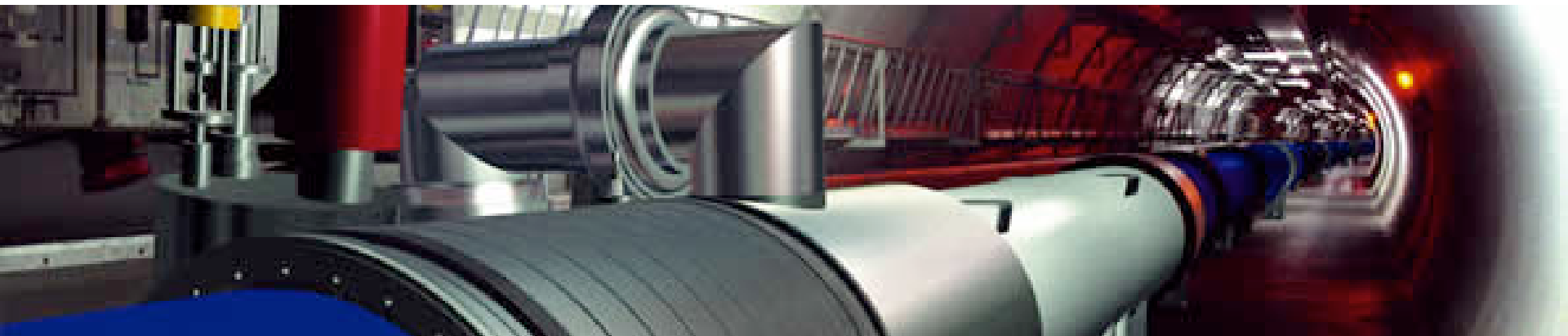


LHC is due to switch on in 2007

Four experiments, with detectors as big as cathedrals :
ALICE
ATLAS
CMS
LHCb



- LHC will collide beams of protons at an energy of **14 TeV**
- Using the latest super-conducting technologies, it will operate at about – **300°C**, just above absolute zero of temperature.
- With its **27 km circumference**, the accelerator will be the largest superconducting installation in the world.



- 40 million collisions per second
- After filtering, 100 collisions of interest per second
- A Megabyte of data digitised for each collision
= recording rate of **0.1 Gigabytes/sec**
- 10^{10} collisions recorded each year
= **10 Petabytes/year of data**

1 Megabyte (1MB)
A digital photo

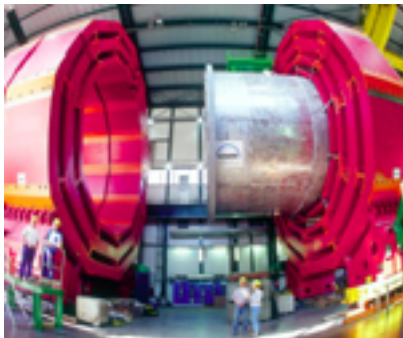
1 Gigabyte (1GB)
= 1000MB
A DVD movie

1 Terabyte (1TB)
= 1000GB
World annual book production

1 Petabyte (1PB)
= 1000TB
Annual production of one LHC experiment

1 Exabyte (1EB)
= 1000 PB
World annual information production

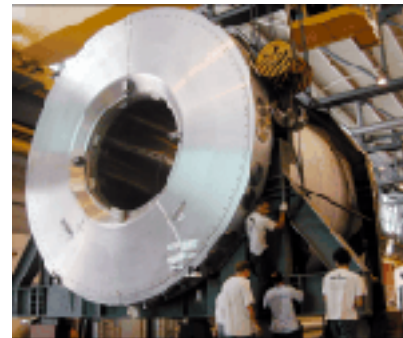
CMS



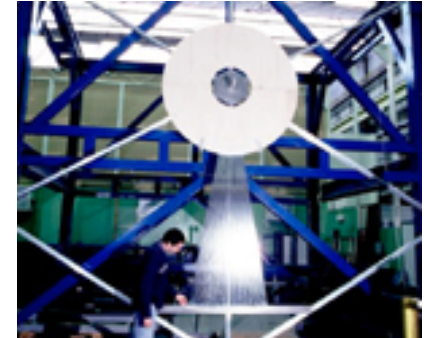
LHCb



ATLAS



ALICE



Share data, accessible

anywhere, anytime

therefore

Link computer centers

solve

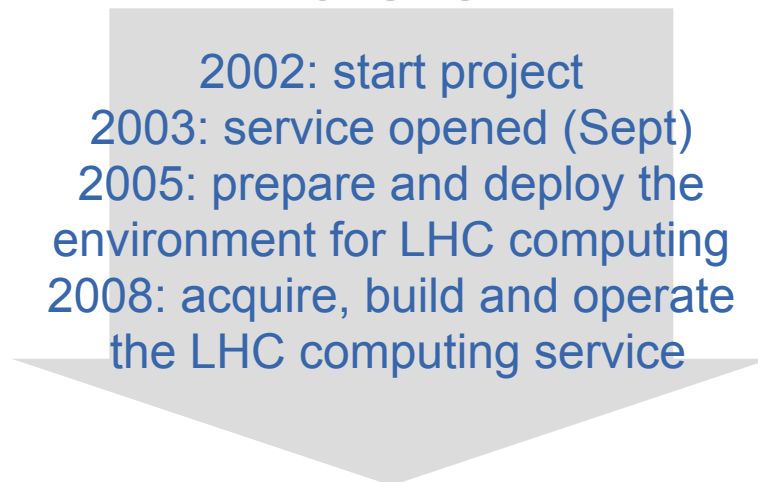
Security & Reliability

before

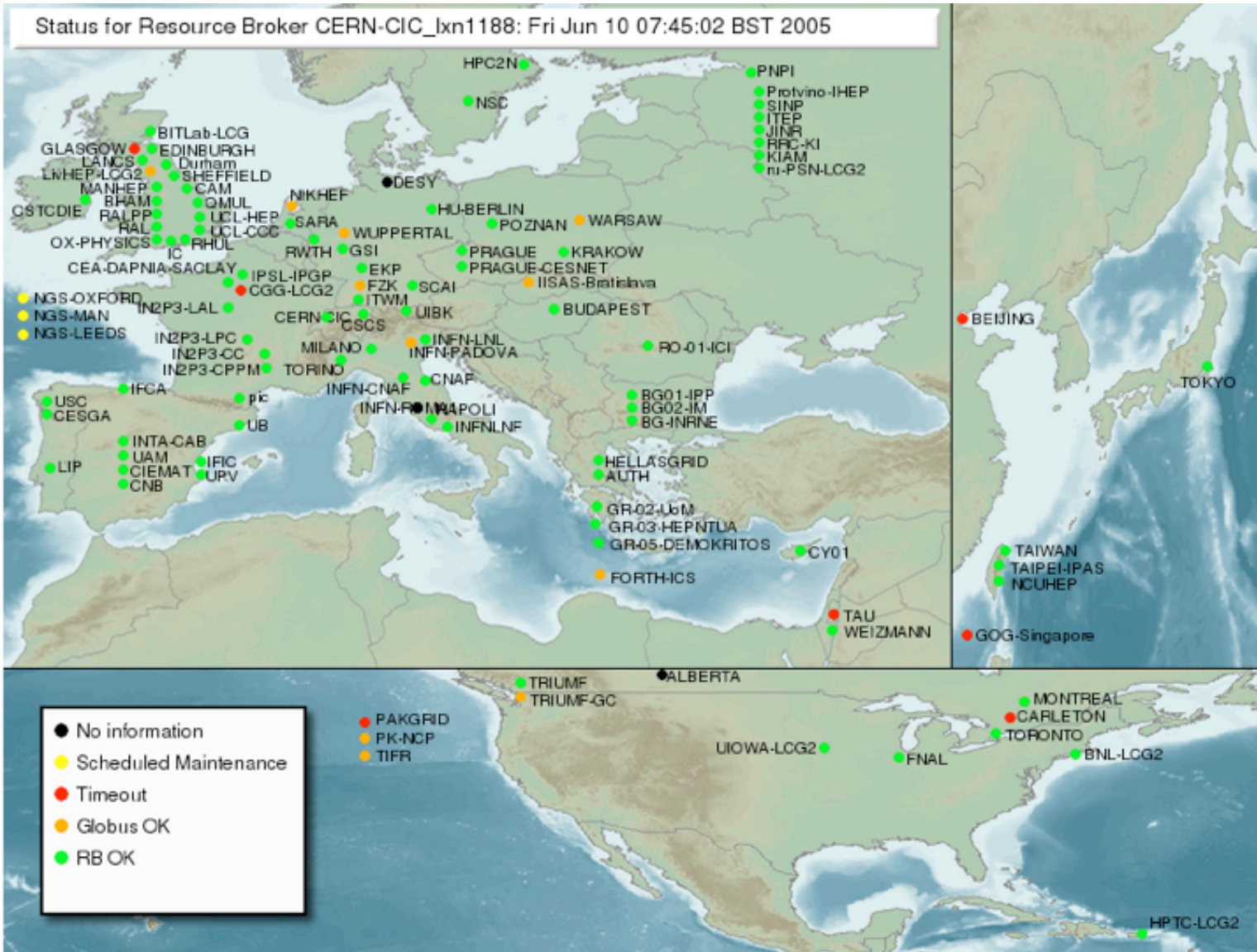
2007

- **Mission: Grid deployment** project aimed at installing a functioning Grid to help the LHC experiments collect and analyze the data coming from the detectors.
- **Strategy:**
 - Integrate thousands of computers at dozens of participating institutes worldwide into a global computing resource.
 - Rely on software being developed in advanced grid technology projects, both in Europe and in the USA.

timeline



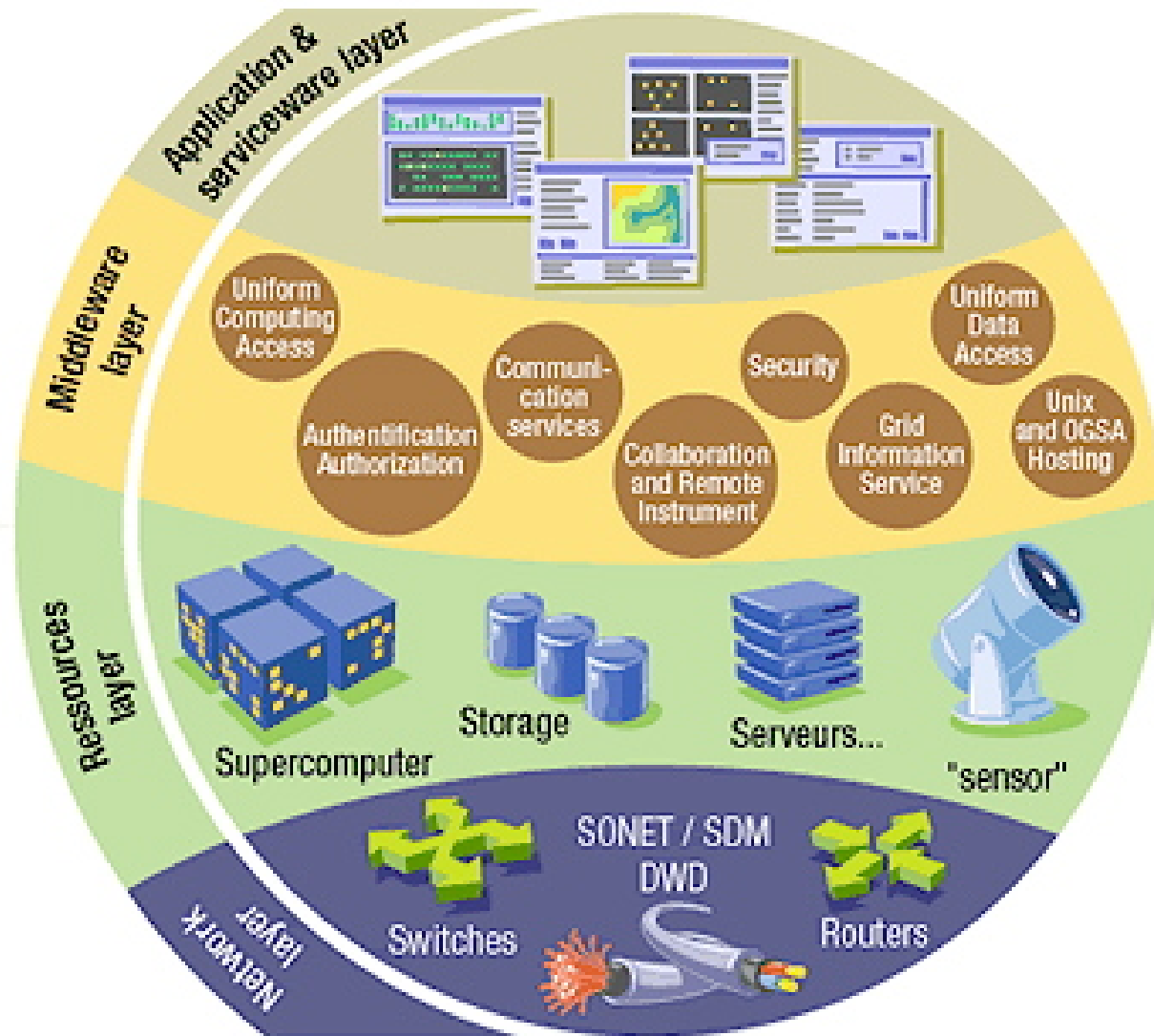
Status for Resource Broker CERN-CIC_1xn1188: Fri Jun 10 07:45:02 BST 2005

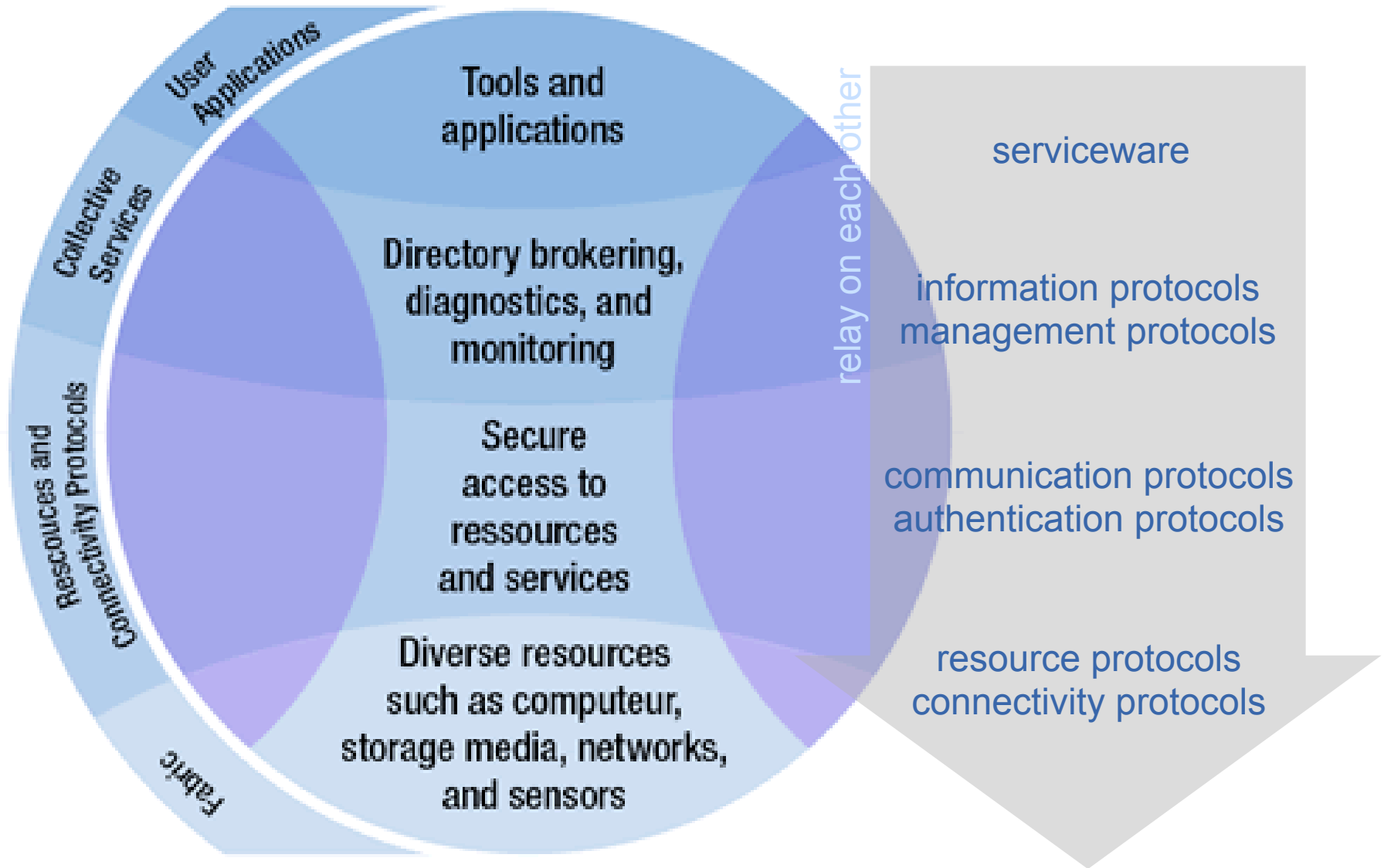


Total Sites:
149

Total CPU:
13758

Storage:
9,8 PB

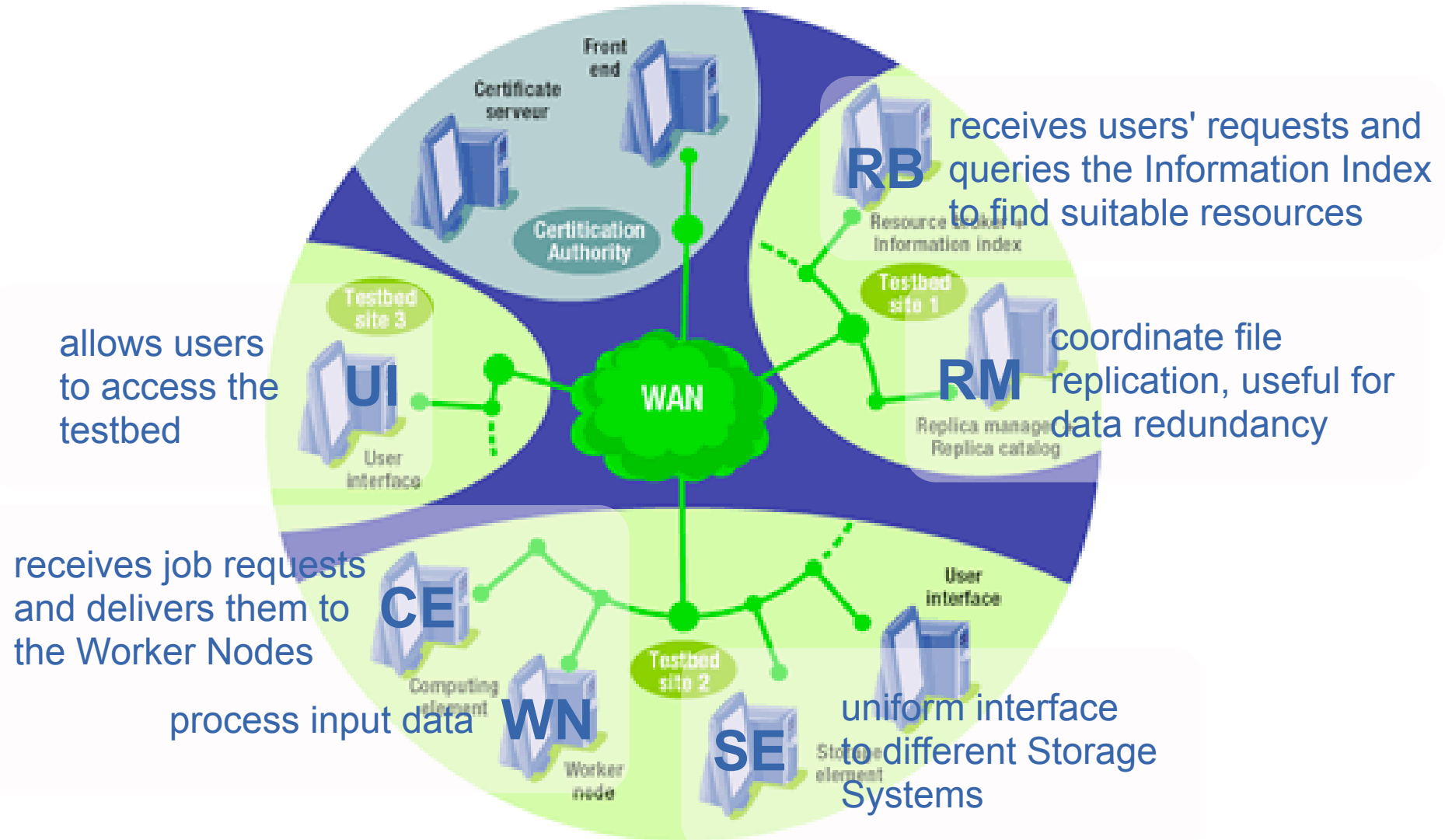




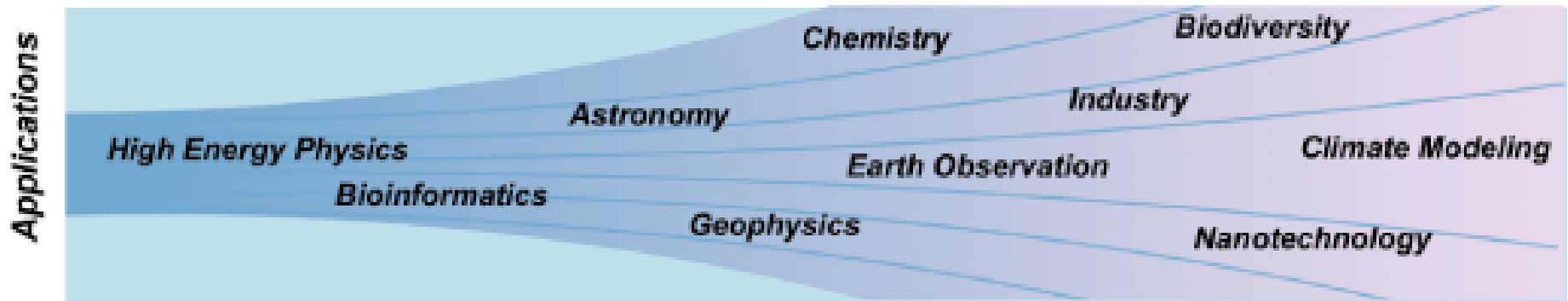
Group of people:

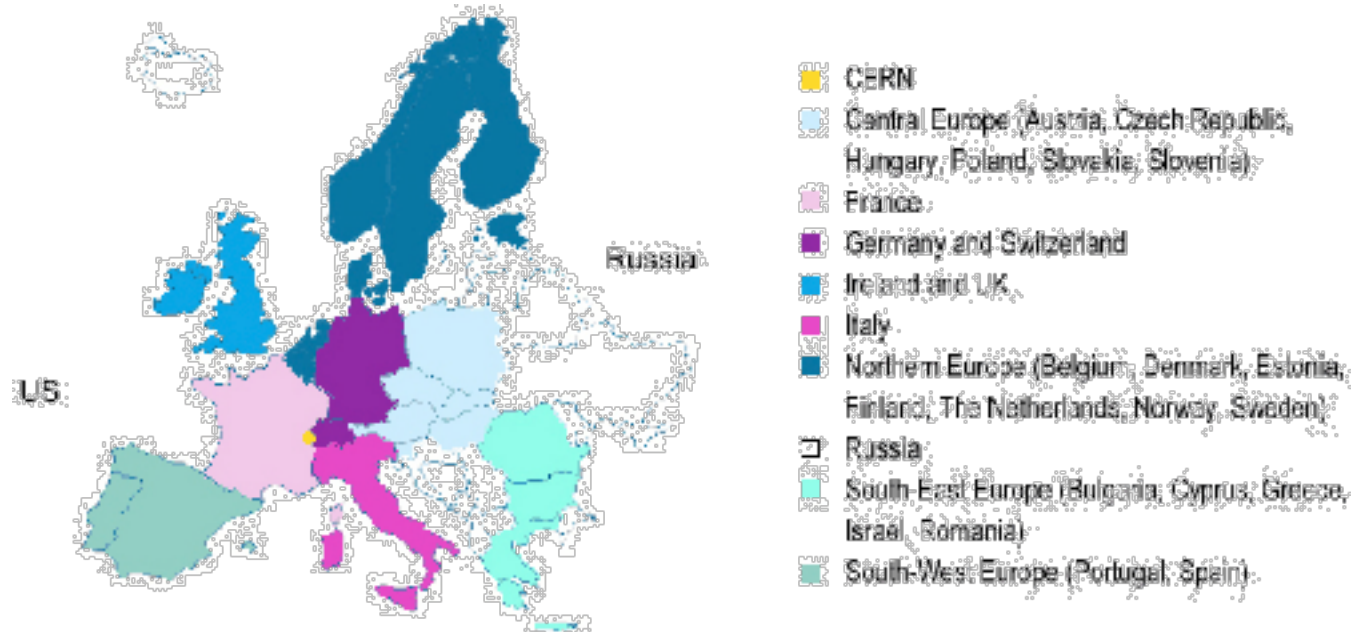
industrial consortium
crisis-management Team
group of scientists

need to share their
resources to achieve
a common goal.



- **Mission:** Deliver 24/7 Grid service to European science; re-engineer and “harden” Grid middleware for production; “market” Grid solutions to different scientific communities
- Be the first **international multi-science production Grid facility**





100 million €

>400 software engineers

12 federation 70 partners

"The world will only need five computers"
attributed to Thomas J. Watson, IBM

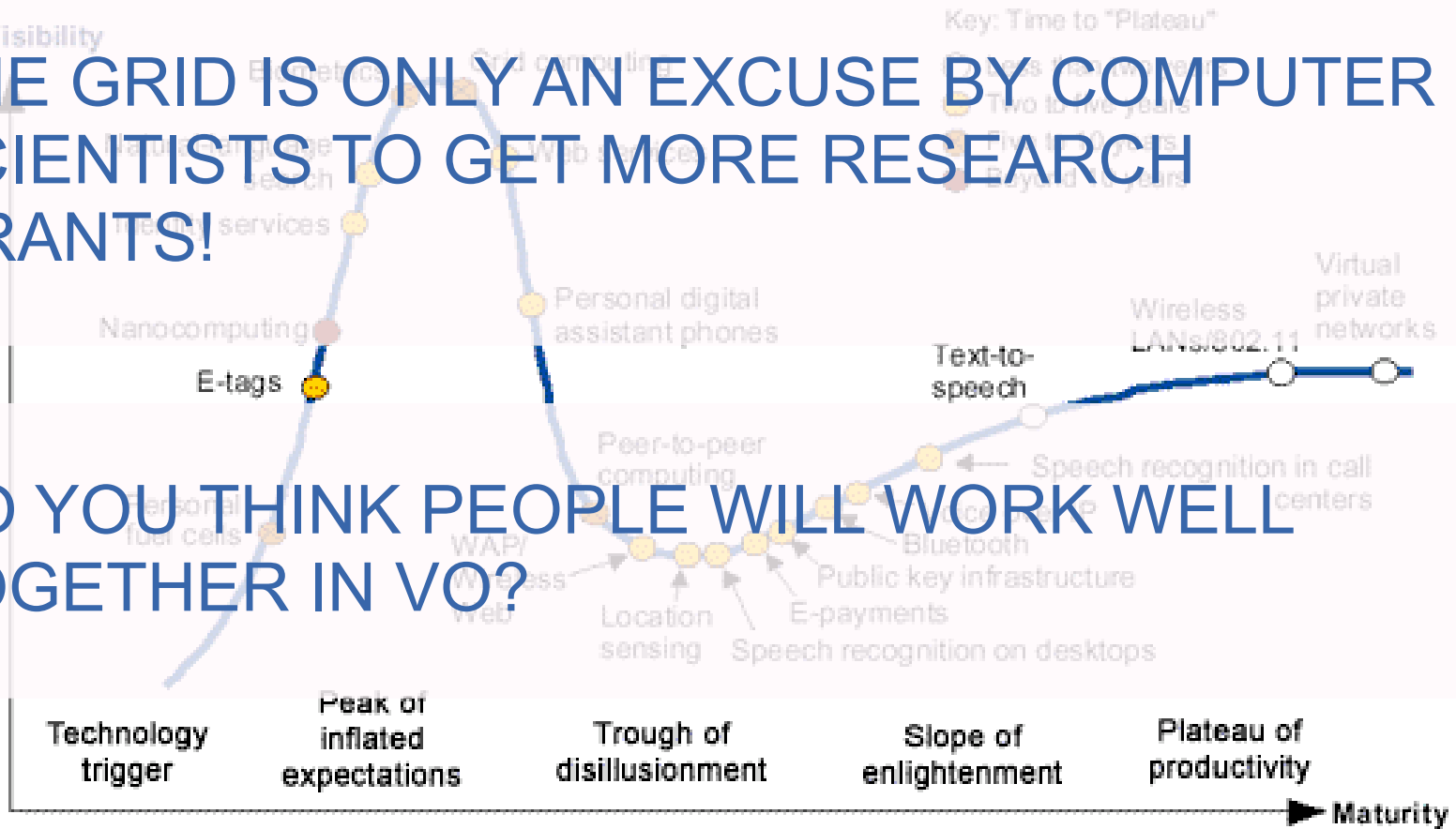
"640 kilobytes is all the memory you will ever need"
attributed to Bill Gates, Microsoft

"There is absolutely no need for a computer in the home"
attributed to Ken Olsen, DEC (once a leading minicomputer manufacturer)

The 2002 Hype Cycle of Emerging Technologies

THE GRID IS ONLY AN EXCUSE BY COMPUTER SCIENTISTS TO GET MORE RESEARCH GRANTS!

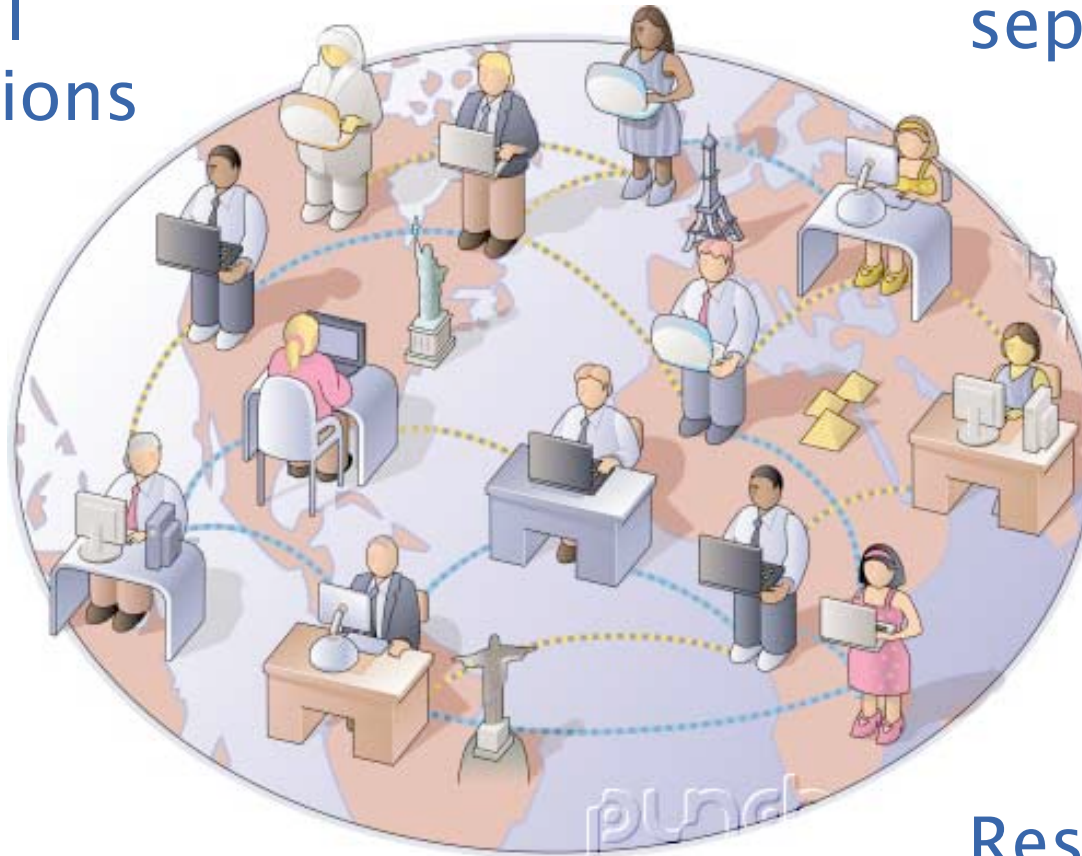
DO YOU THINK PEOPLE WILL WORK WELL TOGETHER IN VO?



Source: Gartner Research

Mutually distrustful organizations

Geographically separated



Different security requirements

Resources are heterogeneous