Enabling Grids for E-science in Europe

Introduction to Grid Technologies

Athanassia Assiki
(aassiki@cslab.ece.ntua.gr)
ΕΔΕΤ

Grid Technologies

Contents

Enabling Grids for E-science in Europe

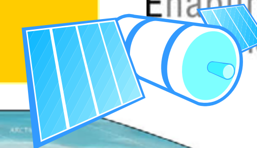
1. What is Grid ?
2. A Brief History
3. Five Big Ideas Behind Grid
4. Grid Systems' Categories
5. Grid Prospects
6. Grid Users
7. Grid Architecture

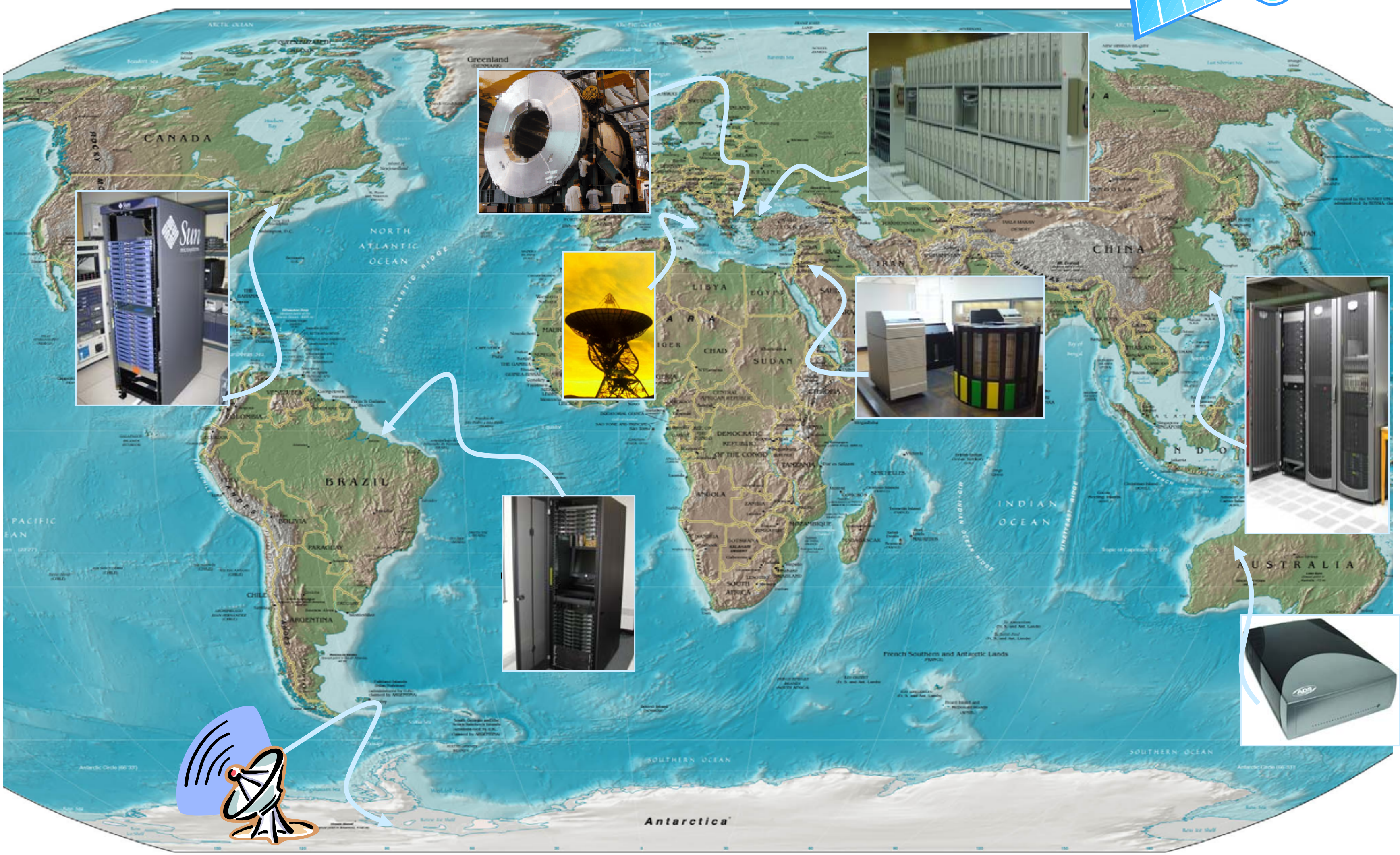
Contents

1. **What is Grid ?**
2. A Brief History
3. Five Big Ideas Behind Grid
4. Grid Systems’ Categories
5. Grid Prospects
6. Grid Users
7. Grid Architecture

Before Grid…

Enabling Grids for E-sciencE in Europe



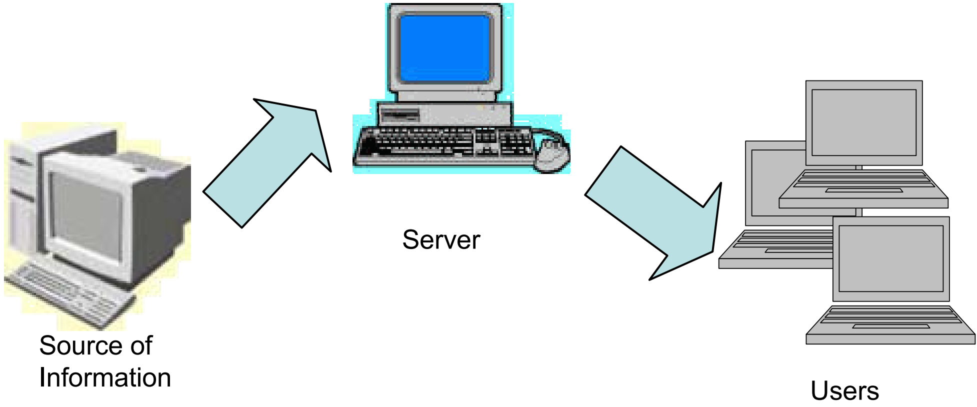


After Grid…

Grid Computing equates to the world's largest computer

What is Grid?

- The ***World Wide Web*** provides seamless access to information that is stored in many millions of different geographical locations



- The Grid is an emerging infrastructure that provides seamless access to computing power and data storage capacity distributed over the globe

What is Grid ?

Enabling Grids for E-science in Europe

- Collection of geographically distributed heterogeneous resources

  "Most generalized, globalized form of distributed computing"

- **"An infrastructure that enables flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources"**
  Ian Foster and Carl Kesselman

- Offers access to a virtual and very powerful computing system

- A user does not care, in which resource his job / jobs is going to be executed

Resource

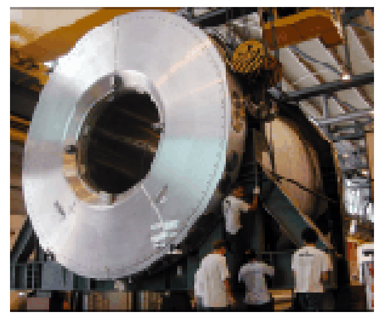
- An entity that is going to be shared

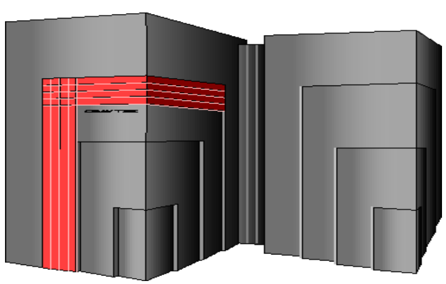
  such as:

  Computational units

  Storage units

  Software



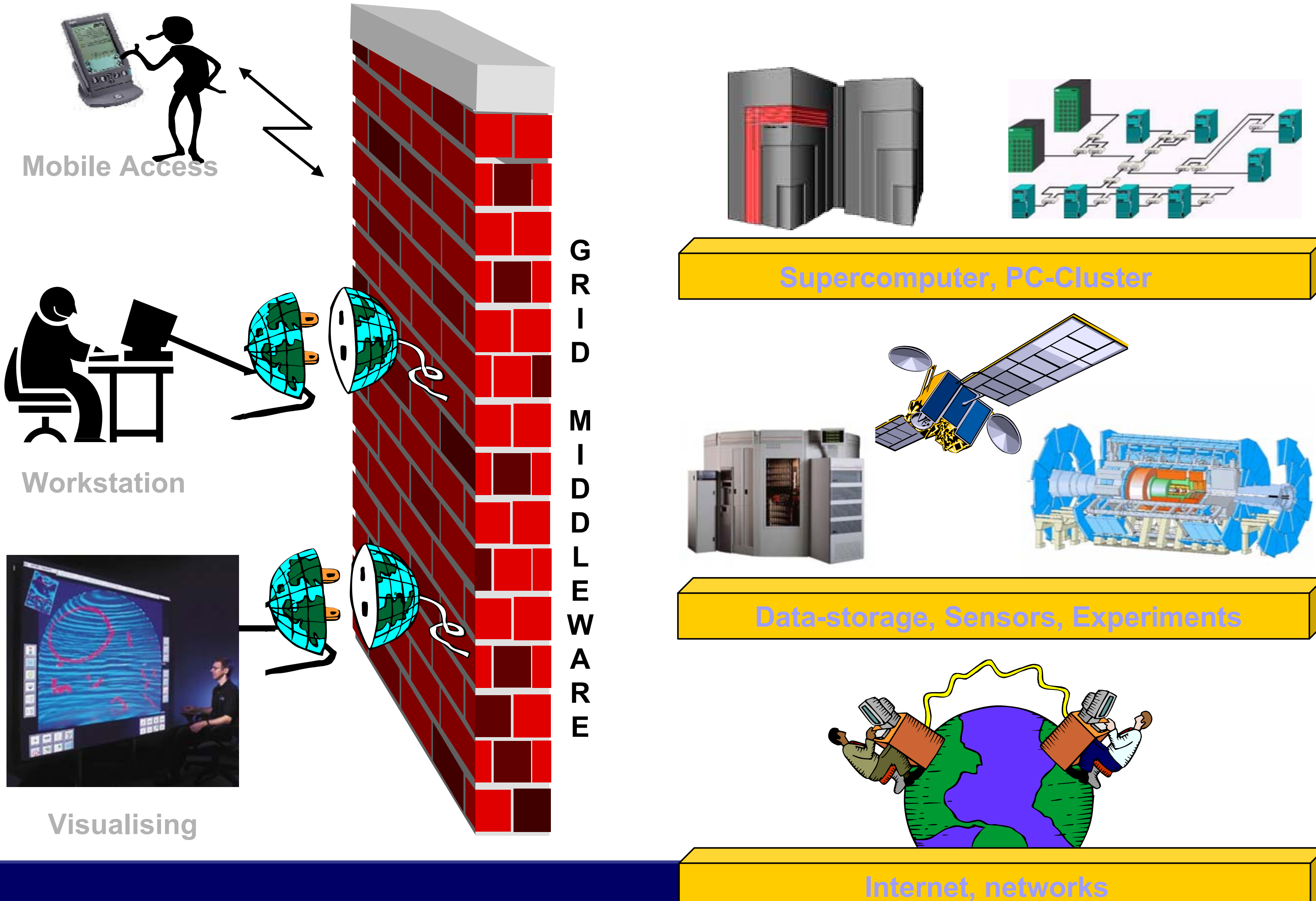


Contents

Electrical power Grid vs Grid

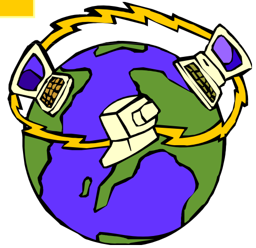
Electrical power Grid	To Grid
The use of electricity is done without worrying about where the electricity you are using comes from and how it is produced. The power grid is transparent.	The user of Grid has access to computing power and storage capacity all over the world. The Grid will be transparent.
An infrastructure, which links together power plants of many different kinds with your home through transmission stations, power stations, transformers, power lines and so forth.	An infrastructure, which links together resources such as PCs, workstations, storage elements, sensors, etc and provides the mechanism needed to access them.
The power grid is pervasive: electricity is availiable essentially everywhere and you can simply access it through a standard wall socket.	The Grid will be pervasive: remote resources will be accessible from any platform (desktops, laptops, PDAs, mobile phones) through a Web portal.
The power grid is a utility: you ask for electricity and you get it. You also pay for what you get.	The Grid is a utility: you ask for computer power or storage capacity and you get it. For the time being, it is free!

The Grid Metaphor

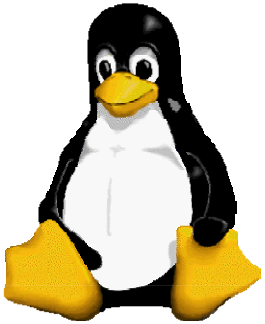


Key concepts towards Grid

Enabling Grids for E-science in Europe

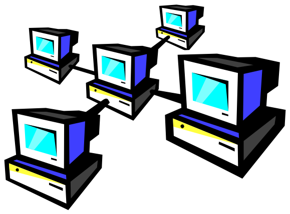


1990 : World Wide Web (CERN)



1991: The operating system Linux (University of Helsinki)

- ✓ open source operating system



1994: Cluster of PCs (NASA)

- ✓ Ethernet cards are used to provide high speed links between the computers, in order to emulate the sort of computing power previously only available from expensive supercomputers

⇨ The Grid consists of clusters connected through Internet and the operating system Linux is used mainly.

E-science

- Collaborative science that is made possible by the sharing across the Internet of resources (data, instruments, computation, people's expertise...)
  - ✓ Often very compute intensive
  - ✓ Often very data intensive (both creating new data and accessing very large data collections) – data deluges from new technologies
  - ✓ **Crosses organisational and administrative boundaries**

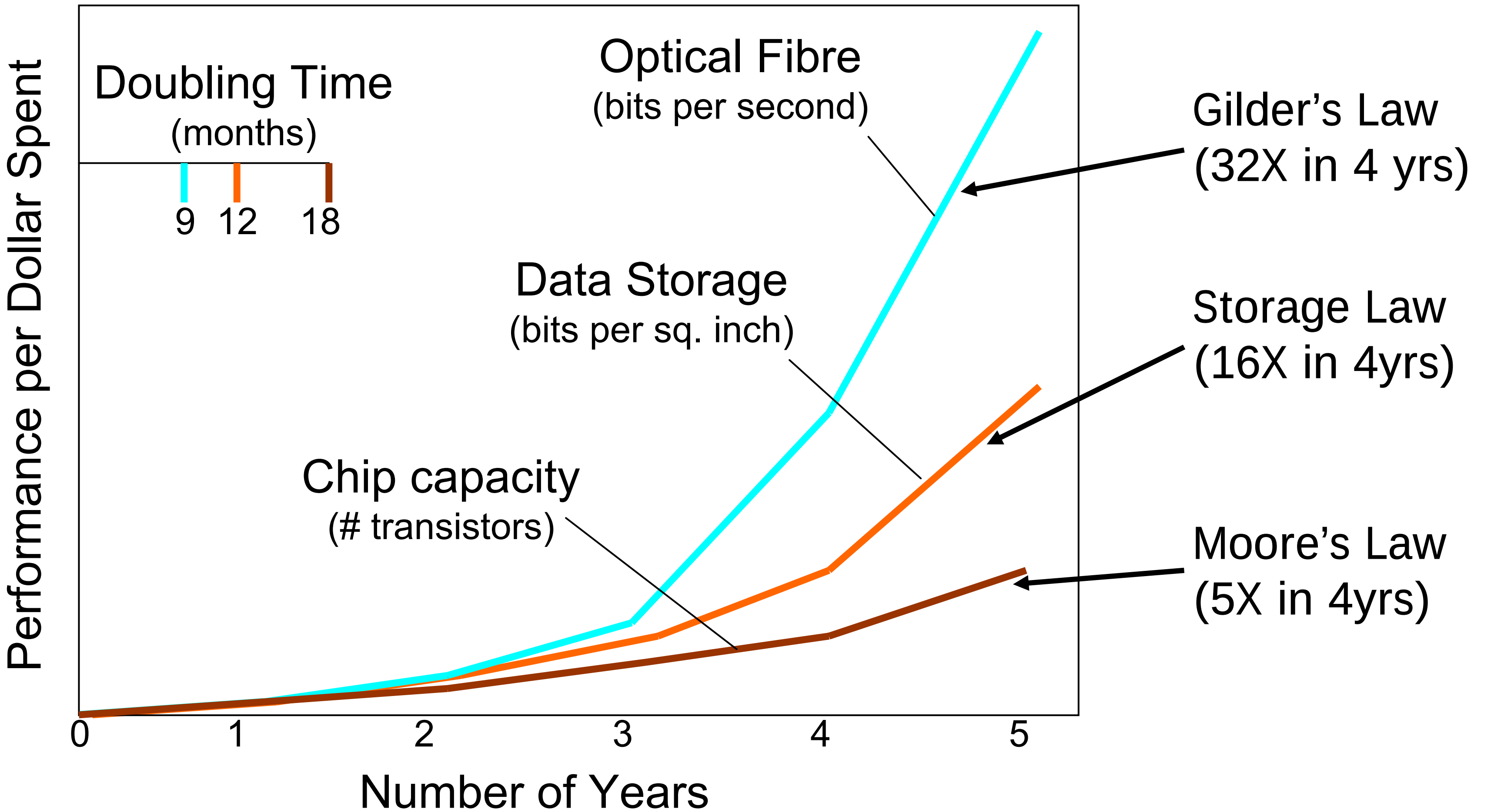
Why now ?

- Development of networking technology (doubling every nine months or so over the last years) and high-speed networks
  - ✓ widespread penetration of optical fibers
  - ✓ wireless connections
  - ✓ new Internet technologies (ADSL, WiMax)

- Moore's law everywhere
  - ✓ Instruments, detectors, sensors, scanners, …
    - ⇨ Organising their effective use is the challenge

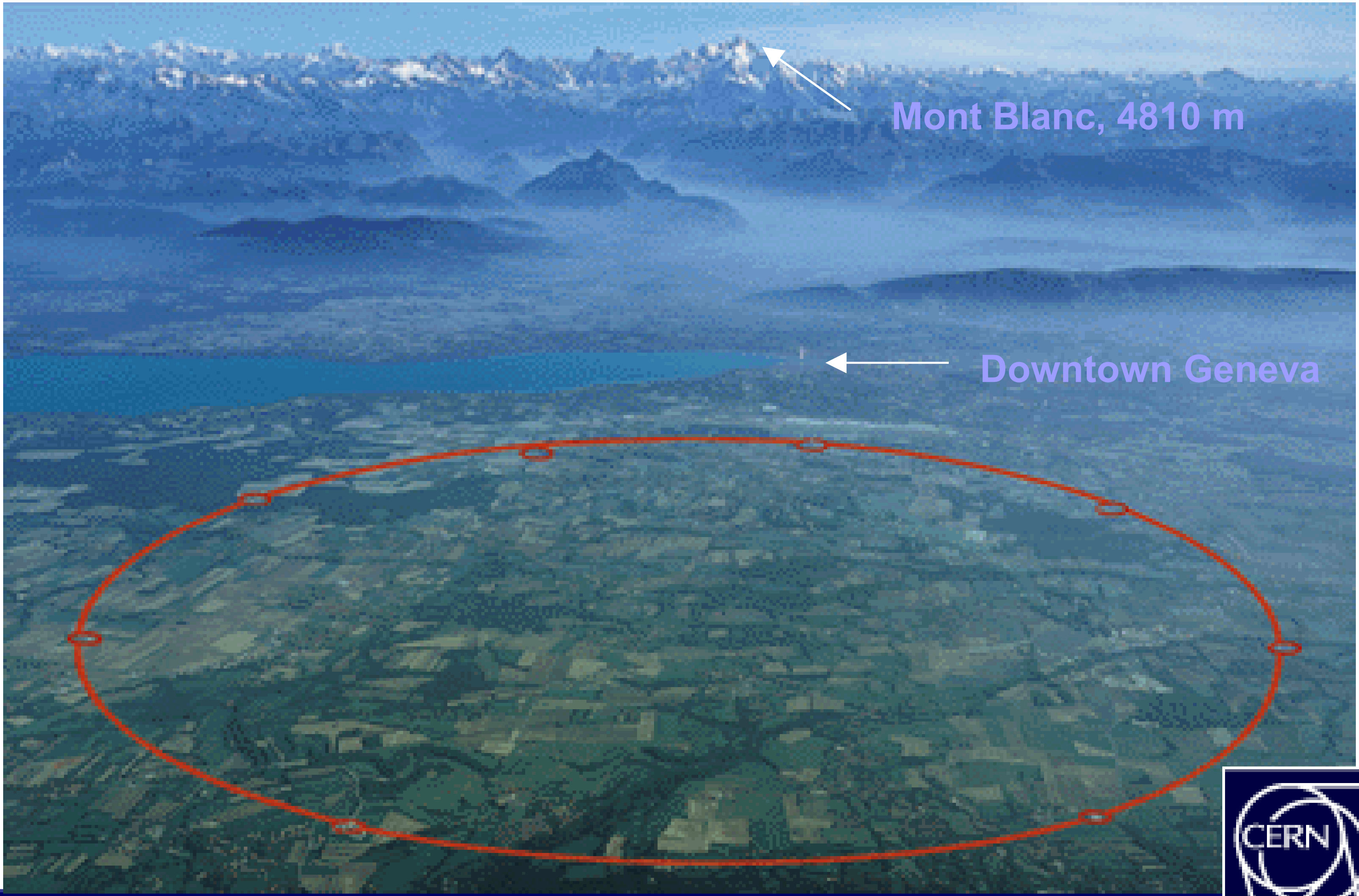
- Applications require a huge amount of computations to be executed and the collaboration among scientists

Exponential Growth



Triumph of Light – *Scientific American*. George Stix, January 2001

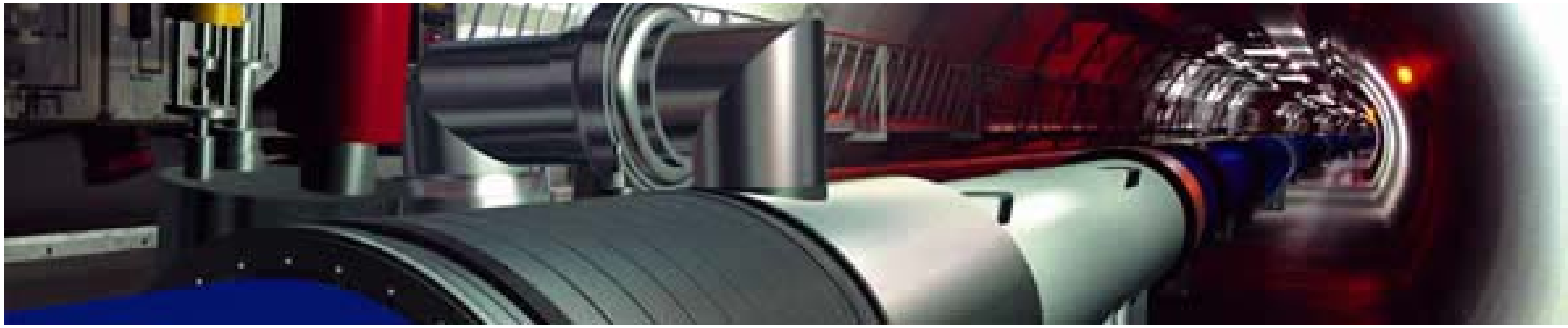
CERN



LHC (Large Hadron Collider)

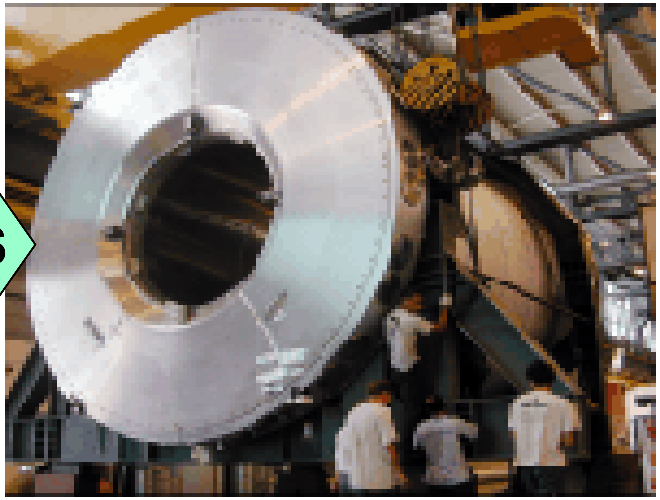
- LHC will collide beams of protons at an energy of 14 TeV
- If the Higgs boson exists, the LHC will almost certainly find it!
- *Four experiments, with detectors:*

ALICE
ATLAS
CMS
LHCB

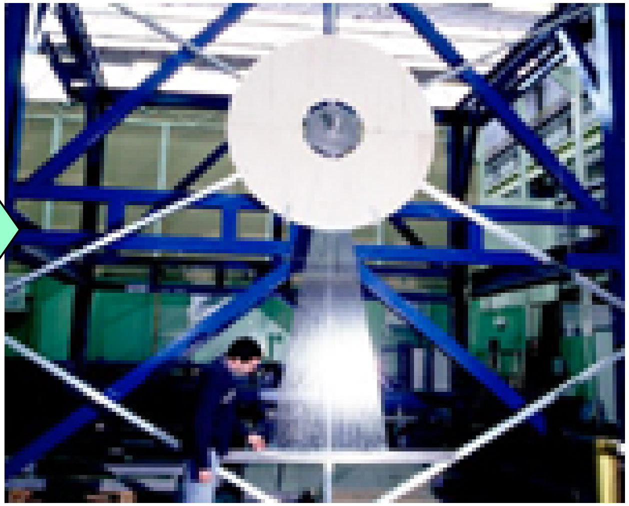


LHC detectors

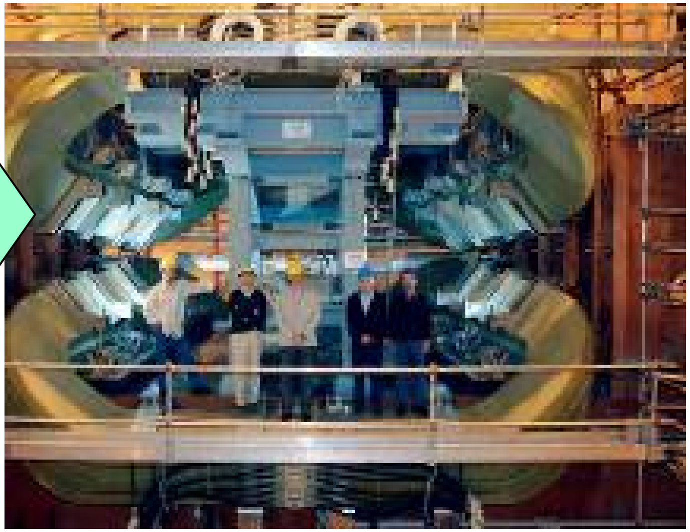
ATLAS



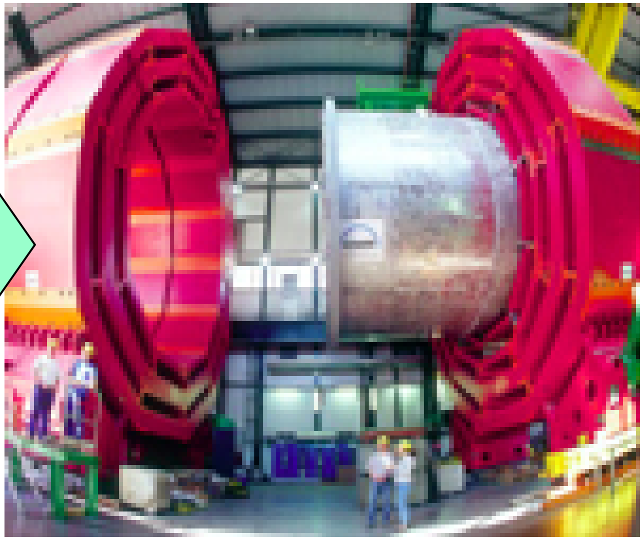
CMS



LHCb

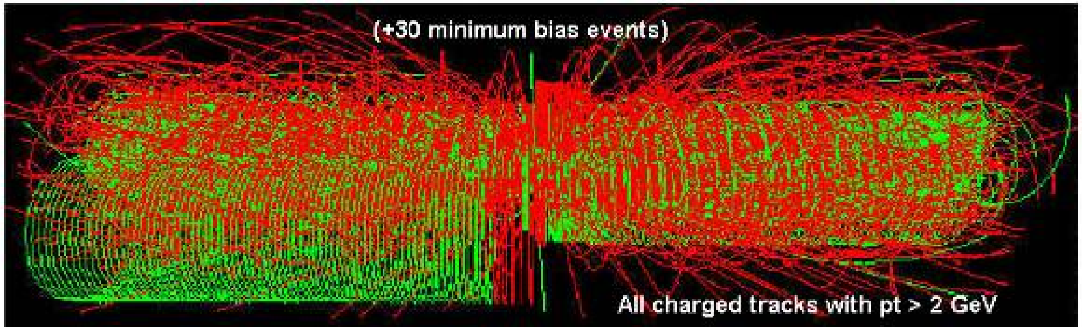


ALICE



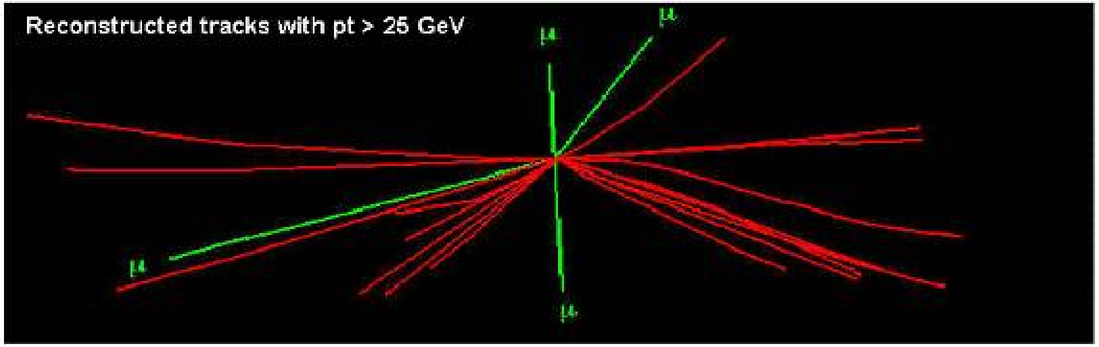
The LHC Data Challenge

Starting from this event (particle collision) …



- ✓ Data Collection
- ✓ Data Storage
- ✓ Data processing

You are looking for this “signature”…

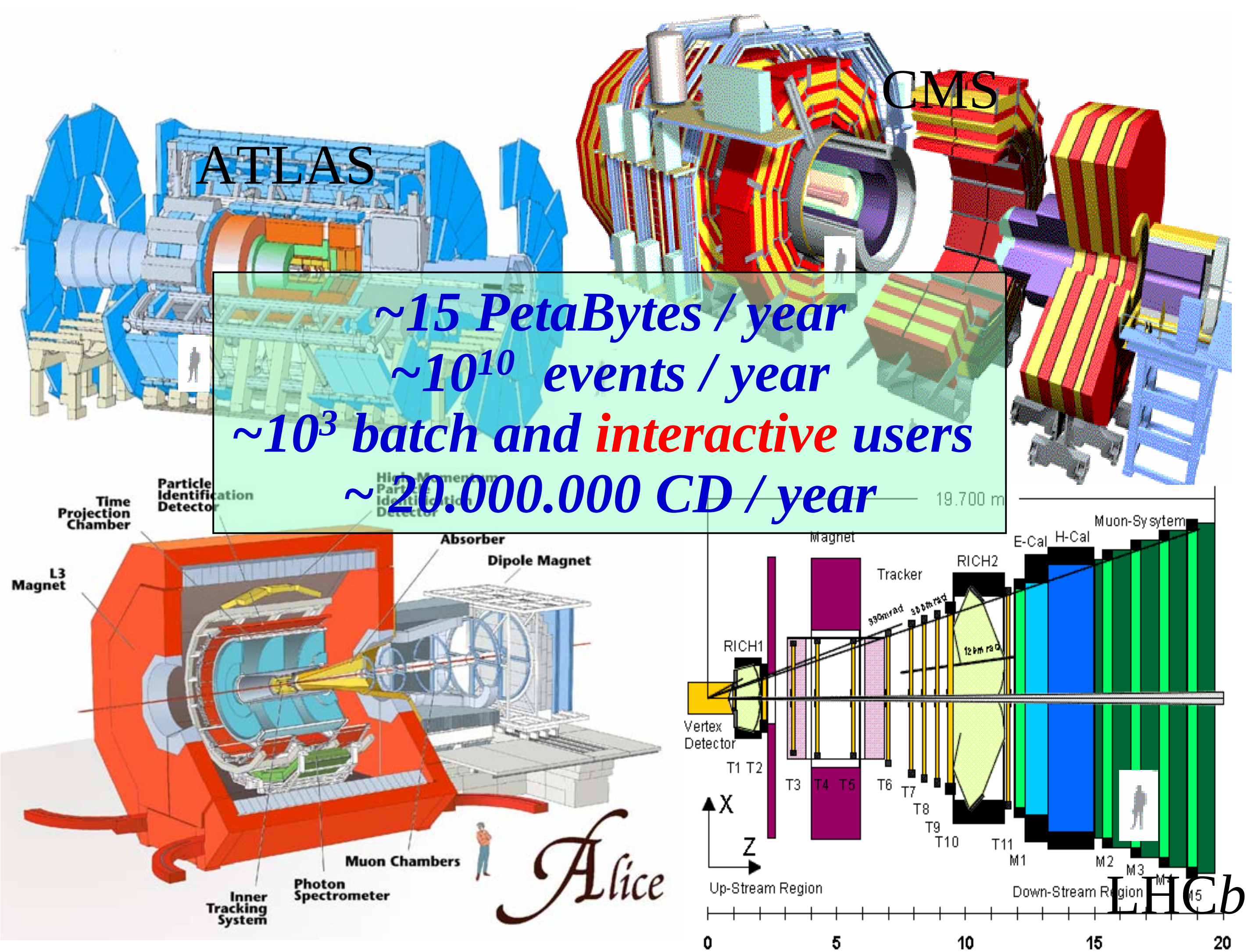


Selectivity: 1 in 10^{13}

Like looking for 1 person in a thousand world populations!

Or for a needle in 20 million haystacks!

Amount of data from the LHC detectors

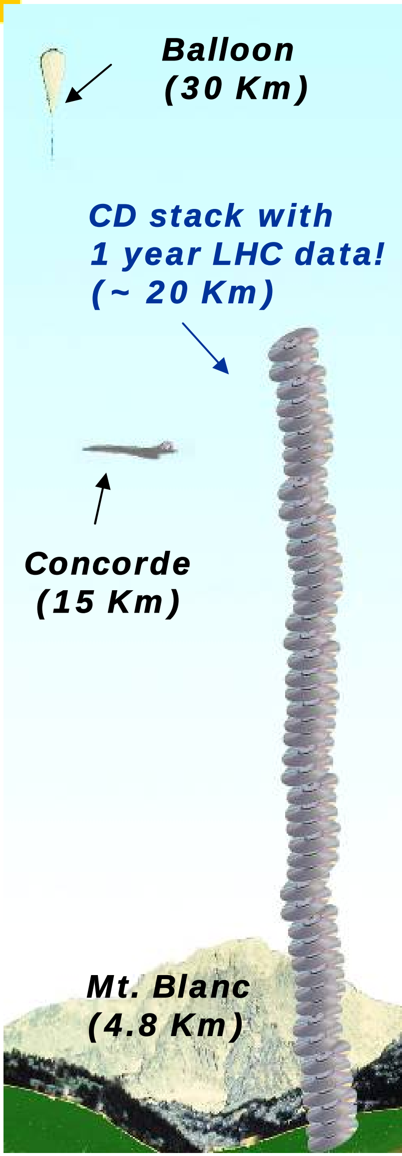


~15 PetaBytes / year

~10^{10} events / year

~10^{3} batch and interactive users

~20.000.000 CD / year



Dreams of scientists

- Nearly infinite computing power available for their institution, whenever they need it
- Nearly infinite storage so they would never have to worry where to put the data
- Reliable access and sharing to common resources, data, procedures and results
- Being able to collaborate with distant colleagues easily and efficiently

Computational problems (1)

- ***"Computer – centric" problems***

  - ✓ The user needs "teraflops"
    - ⇨ the Grid can combine large computational resources in order to tackle problems that cannot be solved on a single system, or at least to do so much more quickly

- ***"Data-centric" problems ("Data-intensive" problems)***

  - ✓ Huge amounts of data
    - ⇨ The Grid will be used to
      - ➢ collect data
      - ➢ store data
      - ➢ analyze data

    maintained in geographically distributed repositories, digital libraries and databases

Computational problems (2)

- ***“Community-centric” problems (collaborative applications)***
  - ✓ Enabling and enhancing human-to-human interactions
  - ✓ Brings people together for collaborations of various types
  - ✓ “Virtual shared space”
    - ⇨ enables the shared use of computational resources such as data archives and simulations
  - ✓ Real time requirements

High-throughput applications

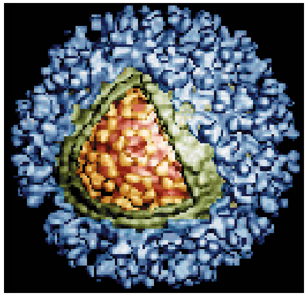
- High-throughput computing is for problems that can be divided into many different tasks, independent of each other

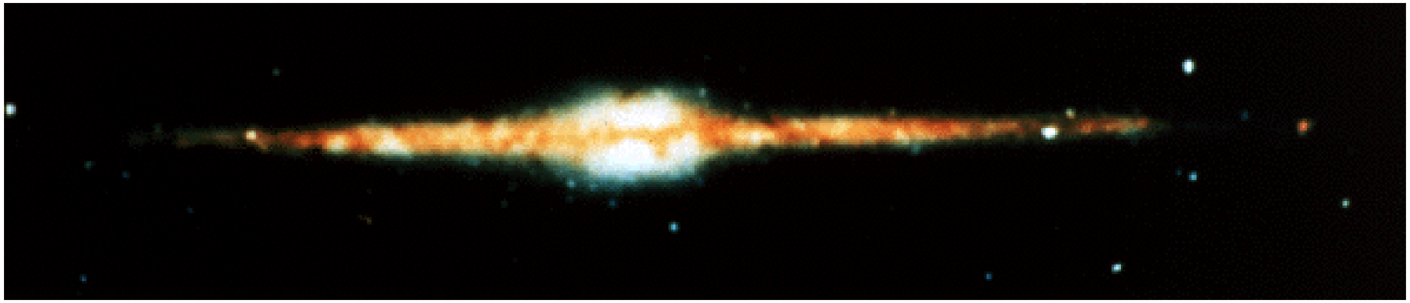
- The Grid can be used to schedule the tasks, with the goal of putting unused processor cycles (often from idle workstations) to work

- @home applications
  - SETI@HOME
  - fightAIDS@home powered by AutoDock — The Olson Laboratory — The Scripps Research Institute
  - Folding@home distributed computing
  - Be part of the adventure — evolution@home

High-performance applications

- **Supercomputing**
- "Computer – centric" problems
- Scientific applications
  - Astrophysics
  - Aerospace industry
  - Automotive industry
  - Economics
  - Climate modeling
  - Distributed Interactive Simulation





Contents

Five Big Ideas behind Grid (1)

Enabling Grids for E-science in Europe

- **Resource sharing on a global scale**

- A user enters the Grid

  ⇨ the user can use remote resources, which allow him to do things that he cannot do with the computer he owns or the computer centre he normally uses

  ⇨ the user can do more things than simple file exchange such as to gain access to remote software, computers, data and even he can access and control remote sensors, telescopes and other devices that belong to research centers

- *Challenges*:
  - ✓ Resources are owned by many different people and they exist within different administrative domains and they are subject to different security and access control policies
  - ✓ Heterogeneous resources (different software, different security and access control policies)

Five Big Ideas behind Grid (2)

- **Secure access**

  - ***Access policy***
    Resource providers and users must define clearly and carefully what is shared, who is allowed to share, and the conditions under which sharing occurs
  - ***Authentication***
    Mechanism for establishing the identity of a user or resource
  - ***Authorization***
    Mechanism for determining whether an operation is consistent with the defined sharing relationships

  ⇨ Reliable accounting mechanism

- *Challenges*
  - Security in resources of Grid infrastructures
  - Security in transferring data through Internet

Five Big Ideas behind Grid (3)

Enabling Grids for E-science in Europe

- **Efficient use of resources**

- Increase of users' number
  ⇨ Insufficient resources
  ⇨ Queue of people waiting to use the resources

↳ Development of algorithms for optimal allocation of resources

↳ Optimal allocation must take into account:
  - ✓ Number of jobs waiting in a queue to be executed
  - ✓ How long the execution of each previous job is going to take
  - ✓ Computational power of resources

Five Big Ideas behind Grid (4)

Enabling Grids for E-science in Europe

- **Death of distance**

↳ High-speed connections between resources make a truly global Grid possible (use of optical fibers, optimization of Transport Protocols)

↳ If the connection between resources is slow
⇨ then the use of remote but powerful resources will not be effective

- *Challenges*
  - Some scientists will need even ***higher-speed connectivity,*** so as to work with colleagues across the world to analyze large amount of data
  - Other scientists will demand ***ultra-low latency*** for their applications, so there is no delay when working with colleagues in real time on the Grid

Five Big Ideas behind Grid (5)

Enabling Grids for E-science in Europe

- **Open Standards**

✓ Software engineers developing the Grid services
✓ IT companies

⇨ So the applications, that run in one Grid site, will run in all others with the same way

↳ **Global Grid Forum**
- ❍ Defines Grid standards (Open Grid Standards Architecture)
- ❍ Develops best practices and specifications in cooperation with other leading standards organizations, software vendors, and users

↳ **Globus Toolkit**
- ❍ Developed by Globus Alliance
- ❍ Open-source infrastructure that provides many of the basic services needed to construct Grid applications

What are the challenges?

Enabling Grids for E-science in Europe

- Must share data between thousands of scientists with multiple interests
- Must link major computer centres, not just PCs
- Must ensure all data accessible anywhere, anytime
- Must grow rapidly, yet remain reliable for more than a decade
- Must cope with different management policies of different centres
- Must ensure data security: more is at stake than just money!
- Must be up and running soon ...

Contents

Enabling Grids for E-science in Europe

Grid Categories (1)

- ***Computational Grids***

- Collection of distributed computing infrastructures, which can be considered as an integrated virtual computing system

- Execution of applications with many computations
  - ✓ faster
  - ✓ efficiently
  - ✓ with low cost
  - ✓ using existing computing systems

- Applied to:
  - ✓ Science
  - ✓ Research
  - ✓ Industry

Grid Categories (2)

- **Data Grids**

  - The users and the applications administer information from databases in distributed platforms:
    - ✓ easy
    - ✓ efficiently

  - Low cost, because there is no need for transfer, copy and collection in a central point

  - More reliability during the data access

Grid Categories (3)

- **Service Grids**

- Real time processing

- Requirements:
  - ✓ collection of data from distributed laboratories
  - ✓ analysis of data
  - ✓ administration of data

Contents

Enabling Grids for E-science in Europe

1. What is Grid ?
2. A Brief History
3. Five Big Ideas Behind Grid
4. Grid Systems' Categories
5. **Grid Prospects**
6. Grid Users
7. Grid Architecture

Grid prospects (1)

- ***Better utilization of resources***

- A user's machine may use the most of its capacity

- The user can use the Grid to run his application in an idle machine

- So there is better utilization of the existing resources, with no need for obtaining new ones
- Better load balancing

- Requirements :
  - ✓ The application should be designed and written properly for the Grid
  - ✓ The remote machine must have availiable the required hardware and installed software, so as the application to be executed

Grid prospects (2)

- ***Parallel CPU Capacity***

- If an application requires computing power
  - then it can be divided in independent tasks, that can be executed in parallel

- *Challenges*

  - It is difficult an algorithm to be divided in completely independent parts

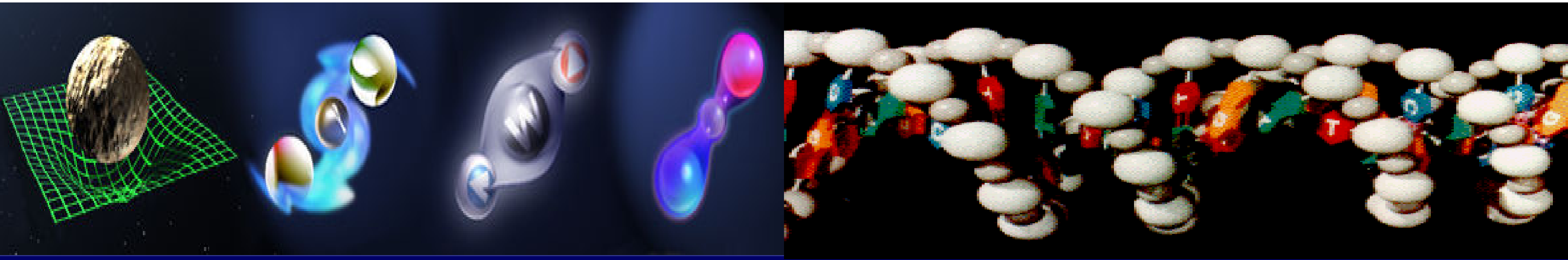
  - There may be collisions, when the various tasks ask access to common files or databases, so as to read or write data

Grid prospects (2)

Parallel CPU Capacity

- Technical problems in the communication of parallel tasks
  - ✓ limited network bandwidth
  - ✓ synchronization protocols
  - ✓ available bandwidth to storage resources

- Scientific applications such as :
  - ✓ High Energy Physics
  - ✓ Bioinformatics
  - ✓ Economic models
  - ✓ Animation
  - ✓ Video processing
  - ✓ Meteorological models



Grid prospects (3)

- ***Collaboration inside Virtual Organizations***

- "A set of individuals and / or institutions defined by highly controlled sharing rules, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share and the conditions under which sharing occurs"

  Ian Foster

- Abstract entities grouping users, institutions and resources in the same administrative domain

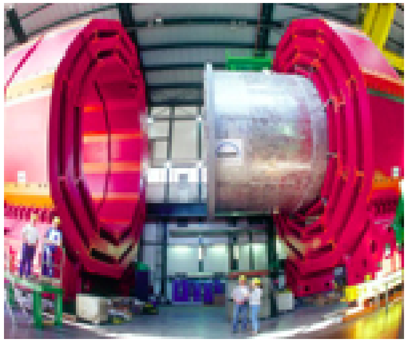
- What is going to be shared ?
  - ✓ resources
  - ✓ software
  - ✓ special equipment
  - ✓ licences
  - ✓ services
  - ✓ Internet bandwidth

Virtual Organizations (VOs)

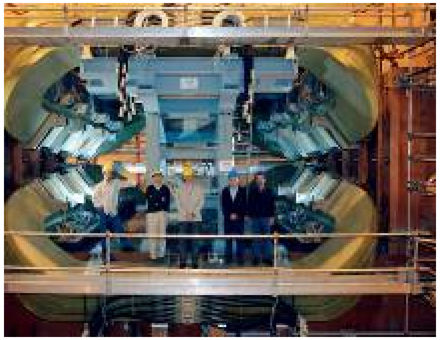
- Associated with LHC experiment

  - ALICE ⇨ ALICE experiment
  - ATLAS ⇨ ATLAS experiment
  - CMS ⇨ CMS experiment
  - DTEAM ⇨ Grid (LCG) Deployment Group
  - LHCB ⇨ LHCB experiment
  - SixTrack ⇨ Single Particle Tracking Code

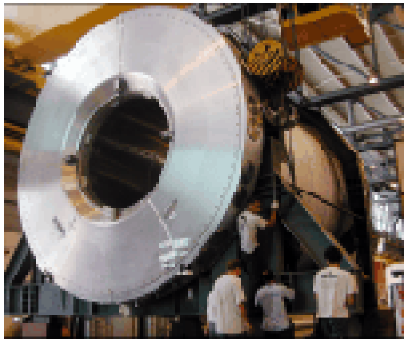
CMS



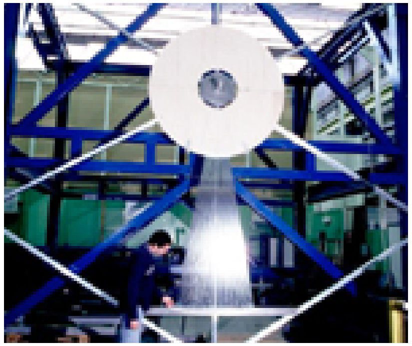
LHCb



ATLAS



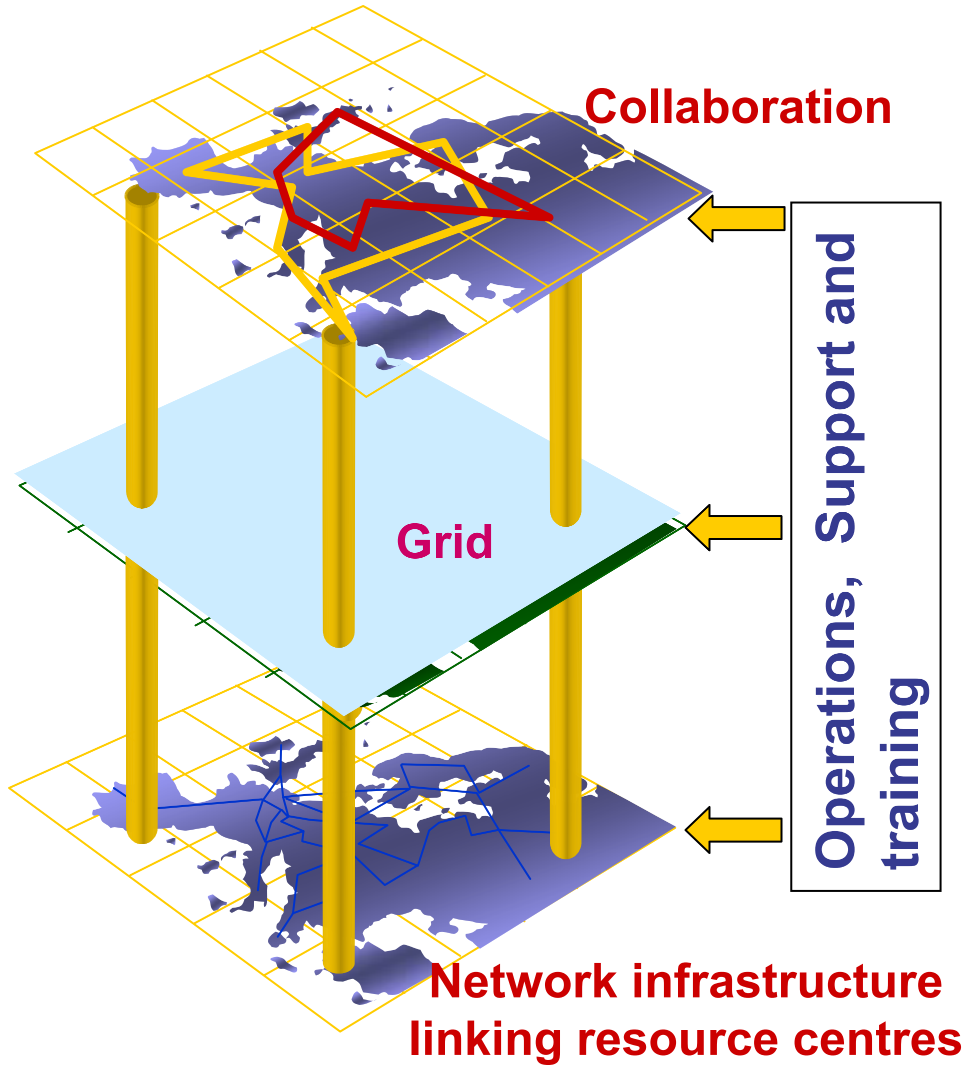
ALICE



Virtual Organizations (VOs)

- Not associated with LHC experiment
  - Babar ⇨ Babar experiment
  - D0 ⇨ D0 experiment
  - H1 ⇨ H1 experiment
  - Zeus ⇨ Zeus experiment
  - ILC ⇨ ILC Community
  - Biomed ⇨ EGEE Biomedical Activity
  - ESR ⇨ Earth Science Research
  - EGEODE ⇨ Expanding GEOsciences on Demand
  - PhenoGrid ⇨ Particle Physics Phenomenology
  - CompChem ⇨ Computational Chemistry
  - SEE-VO ⇨ South Eastern Europe VO
- More about Virtual Organizations in following presentations

The Grid


Collaboration
Grid
Network infrastructure linking resource centres
Operations, Support and training

Contents

Enabling Grids for E-science in Europe

1. What is Grid ?
2. A Brief History
3. Five Big Ideas Behind Grid
4. Grid Systems' Categories
5. Grid Prospects
6. **Grid Users**
7. Grid Architecture

Grid Users

- End Users :

  Run applications

  Exploit the computer power, storage capacity and other resources that are offered

- Application Developers :

  Design and development of applications, that can be executed in Grid infrastructures

  ✓ Virtual Organizations' developers

  - development of special purpose applications, especially for the experiments associated with

Grid Users

- **Grid Administrators :**

  Administration of Grid infrastructures and reassurance of their correct operation

  (network administrators, cluster administrators, regional administrators, virtual organization administrators)

- **Grid Developers :**

  Design, development and deployment of new services

- **Tool Developers :**

  Development of tools, compilers, libraries, that will be used by the application developers

Utilization of Grid (1)

- **Government and International Organizations**
  - The collective power of the nation's fastest computers can be used to very big or very urgent problems
  - Sharing of data archives between institutions more simply and effectively
  - Disaster response(flood, fire), urban planning

- **Training and Education**
  - E-libraries and e-learning could benefit from Grid-based tools for accessing distributed data and creating virtual classrooms with distributed students, resources and tutors

Utilization of Grid (2)

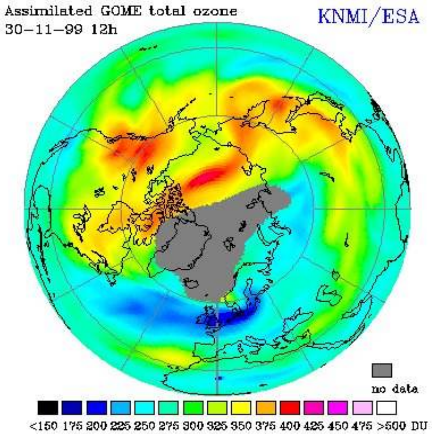
Enabling Grids for E-science in Europe

- **Science and Technology**

- Simulation of real time applications
- Execution of computations
- Connection of remote scientific instruments or sensors to computer farms, so as to be watched and managed
- Grid is suitable, when it is required:
  - ✓ Computing power
  - ✓ Storage capacity
  - ✓ Access to common databases
- High Energy Physics, Astronomy, Meteorology, Biology

Utilization of Grid (3)

- Environment
  - Simulation of environmental problems such as:
    - ✓ air and water pollution
    - ✓ greenhouse effect
    - ✓ ozone hole
  - Modelling and prediction of earthquakes
  - Weather forecast



- **Enterprises**
  - Provide application software and special services
  - Grid technologies' users for their activities

✓ The companies have geographically distributed departments

⇨ development of local Grids (**intra** – **grids**) like intranets

Grid development efforts all over the world

Enabling Grids for E-science in Europe

- Access Grid
- DISCOM
- DOE Science Grid
- Condor
- ESG (Earth System Grid)
- Fusion Collaboratory
- Globus
- GrADSoft (Grid Application Development Software)
- Grid Canada
- GRIDS (Grid Research Integration Development & Support Center)
- GriPhyN (Grid Physics Network)
- iVDGL (International Virtual Data Grid Laboratory)
- Music Grid
- NASA Information Power Grid
- NCSA Alliance Access Grid

- AstroGrid
- AVO (Astrophysical Virtual Observatory)
- Comb-e-chem
- CrossGrid
- DAME (Distributed Aircraft Maintenance Environment)
- DAMIEN (Distributed Applications and Middleware for Industrial Networks)
- DataTAG
- Discovery Net
- DutchGrid
- EDG (European DataGrid)
- EGSO (European Grid of Solar Observations)
- GEODISE (Grid Enabled Optimisation & Design Search for Engineering)

- GRIA (Grid Resources for Industrial Applications)
- Grid-Ireland
- GridLab (Grid Application Toolkit and Testbed)
- GridPP
- LCG (LHC Computing Grid)
- MyGrid
- NGIL (National Grid for Learning Scotland)
- NorduGrid (Nordic Testbed for Wide Area Computing and Data Handling)
- PIONIER Grid
- Reality Grid
- ScotGrid

- ApGrid
- ApBioNet
- Grid Forum Korea
- PRAGMA (Rim Applications and Grid Middleware Assembly)
- Grid Datafarm for Petascale Data Intensive Computing
- Gridbus Project

- The Enabling Grids for E-sciencE (EGEE) project is funded by the European Commission

- This project aims to build on recent advances in grid technology and develop a service grid infrastructure which is available to scientists 24 hours-a-day and 7 days a-week

- ***The project will primarily concentrate on three core areas:***

- ✓ The first area is to build a consistent, robust and secure Grid network that will attract additional computing resources.
- ✓ The second area is to continuously improve and maintain the middleware in order to deliver a reliable service to users.
- ✓ The third area is to attract new users from industry as well as science and ensure they receive the high standard of training and support they need.

LHC Computing Grid

- The LHC Computing Grid Project (LCG) was born to prepare the computing infrastructure for the simulation, processing and analysis of the data of the Large Hadron Collider (LHC) experiments.

⇨ The processing of the enormous amount of data, that will be generated, will require large computational and storage resources and the associated human resources for operation and support.

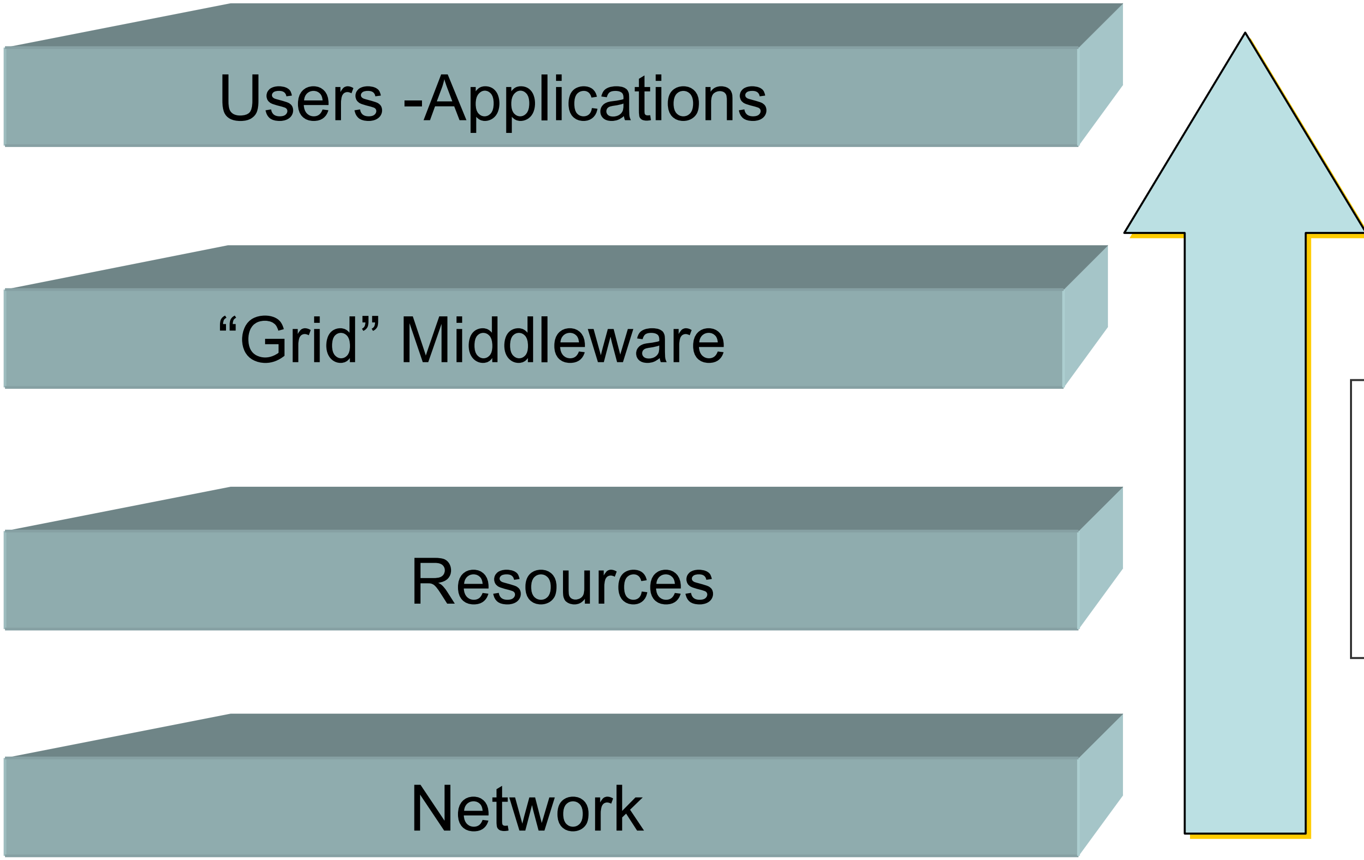
⇨ Preparation of a common infrastructure of
  - ✓ libraries
  - ✓ tools
  - ✓ frameworks

required to support the physics application software

Contents

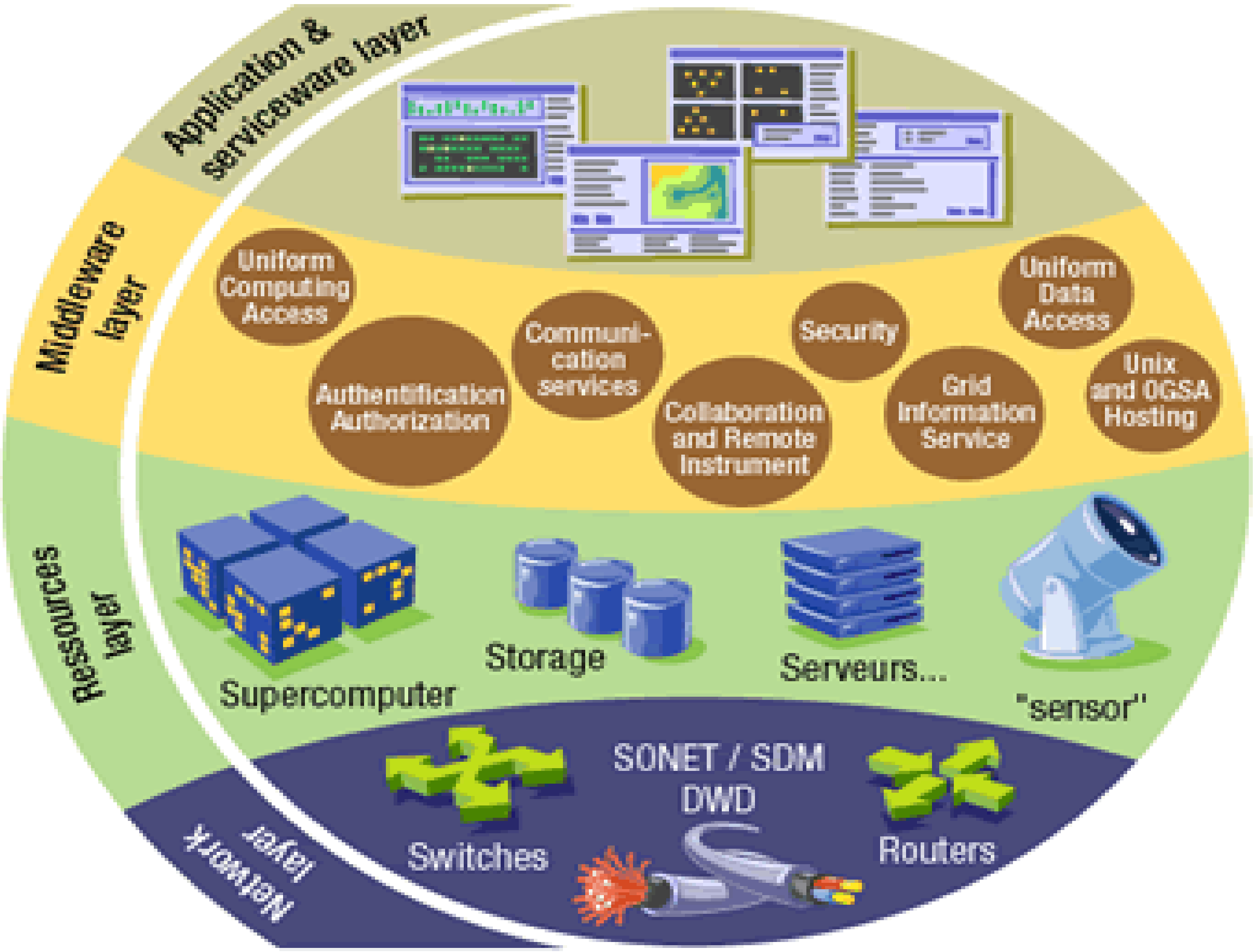
1. What is Grid ?
2. A Brief History
3. Five Big Ideas Behind Grid
4. Grid Systems' Categories
5. Grid Prospects
6. Grid Users
7. **Grid Architecture**

Grid Architecture (1)



- ***Network layer***
- ***Resource layer***
- ***Middleware layer***
- ***Application and serviceware layer***

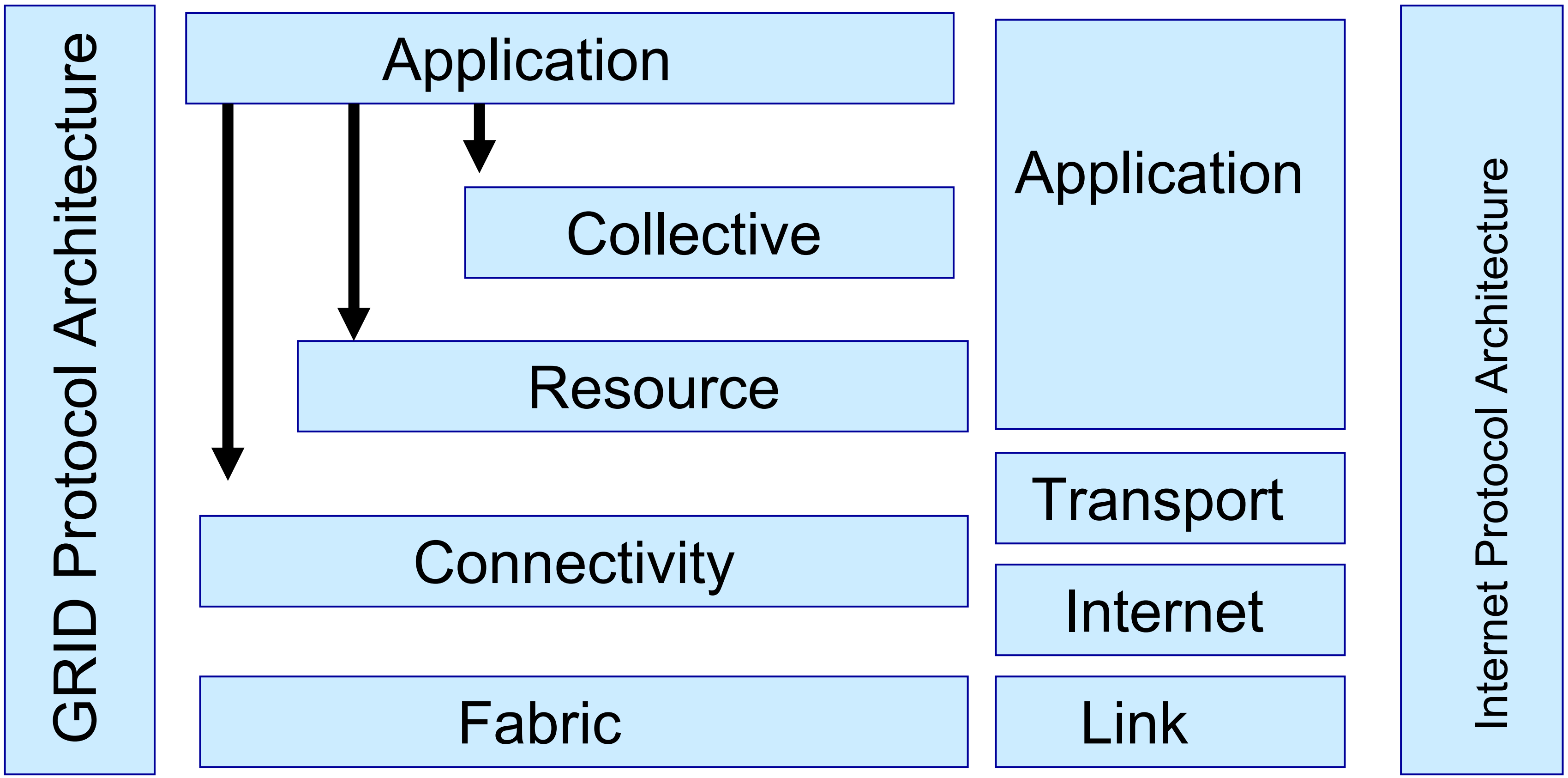
Grid Architecture (2)



Grid Architecture (3)

- **A layered grid architecture and its relationship to the Internet protocol architecture**

  (Foster, Kesselman, & Tuecke)



Middleware

- The Grid Middleware ensures seamless communication between different computers and different parts of the world

⇨ The Grid middleware hide much of the complexity of the Grid environment from the user, giving the impression that all of these resources are availiable in a coherent virtual computer centre

- LCG, Globus, Condor

Middleware

- Optimises use of the widely dispersed resources
- Organises efficient access to scientific data
- Deals with authentication to the different sites that the scientists will be using
- Finds convenient places for the “job” to be run
- Interfaces to local site authorisation and resource allocation policies
- Runs the jobs
- Monitors progress
- Recovers from problems

⇨ Tells you when the work is complete and transfers the result back!

Globus Toolkit

- Grid project
  - Protocols
  - Basic services

- Object-oriented approach
  - ➪ Provides a bag of services from which developers of specific applications can choose what best suits them to meet their own particular needs

- Developed by Globus Alliance

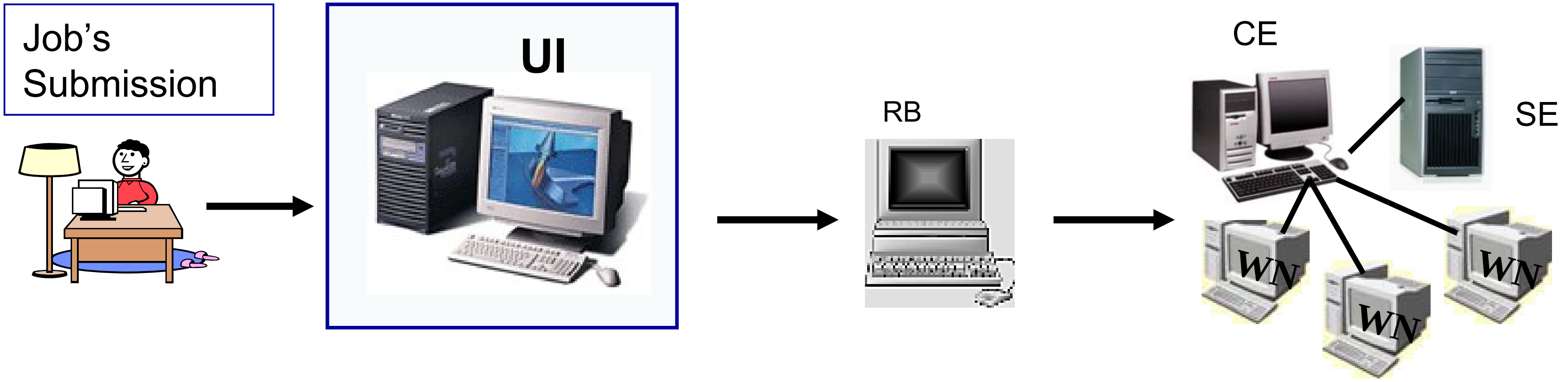
- The toolkit provides a set of software tools needed to construct a computational Grid, such as security, resource location, resource management and communications

- The creators of the Globus Toolkit are making the software availiable under an “***open source***” licensing agreement

- GRAM, GSI, MDS, GRIS, GIIS, GridFTP, Replica Catalog, Replica Management System

User Interface (1)

- Allows users to access Grid functionalities
- A machine where users have a personal account and where the user certificate is installed
- Gateway to Grid Services



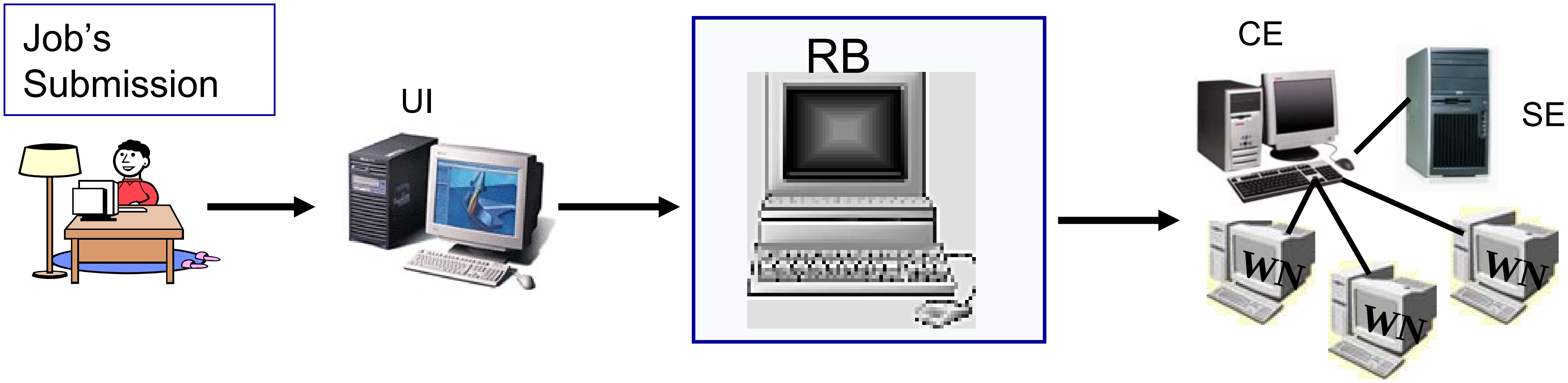
User Interface (2)

- *It provides a Command Line Interface to perform some basic Grid operations such as:*

  - List all the resources suitable to execute a given job
  - Submit jobs for execution
  - Show the status of submitted jobs
  - Cancel one or more jobs
  - Retrieve the logging and bookkeeping information of jobs
  - Retrieve the output of finished jobs
  - Copy, replicate and delete files from Grid

Resource Broker

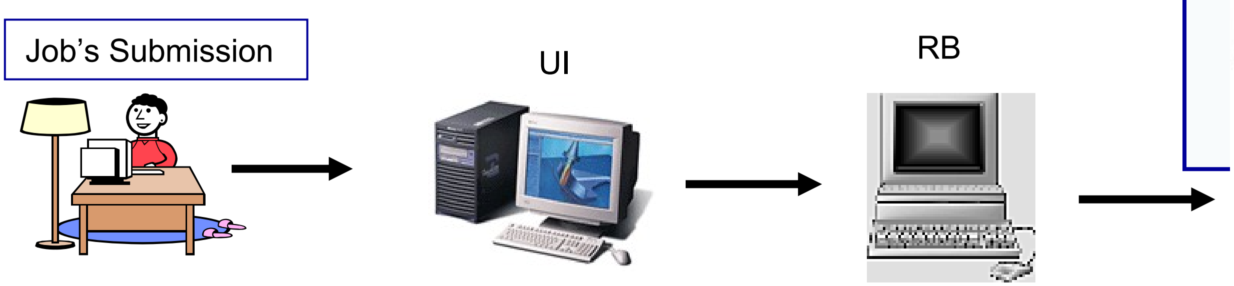
- The resource broker is responsible for the acceptance of submitted jobs and for sending those jobs to the appropriate Computing Element
- Retrieves information from Information Catalogues so as to find the proper available resources depending on the job requirements



Computing Element

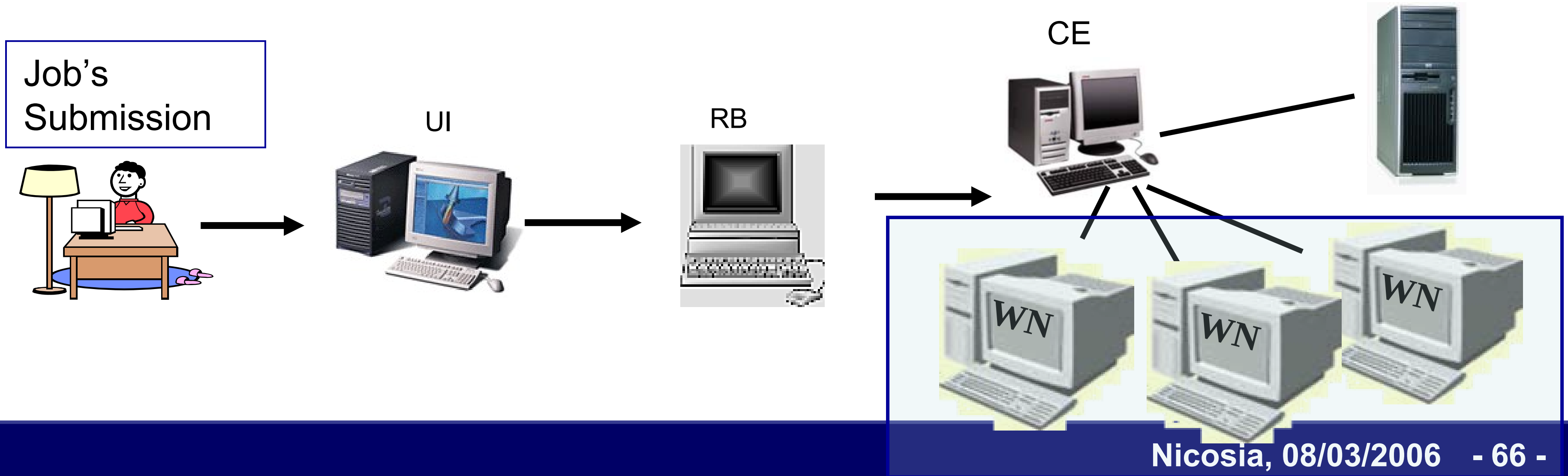
- "Grid interface"
- It is built on a farm of a computing nodes called Worker Nodes (WNs)
- Executes the basic queues functions
- In the Computing Element, a process is being executed that accepts jobs and dispatch them for execution to the Worker nodes (WNs)
- The state of an executing job is being watched by the Computing Element

⇨ Each LCG-2 site is consisted of one or more CEs and a farm of WNs behind them



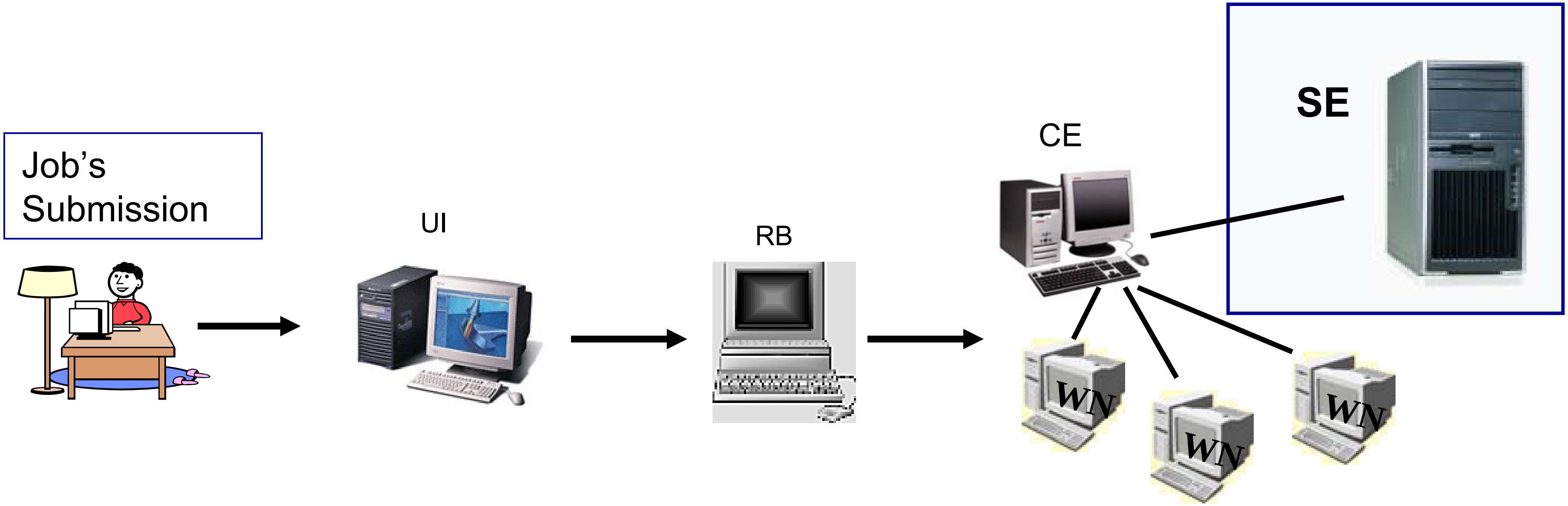
Worker Node

- The submitted jobs are being executed in the Worker noeds
- Need only outbound connectivity
- Only basic services of middleware are required in the Worker nodes
  such as Application libraries
  Application Programming Interfaces (API)
  Commands for performing actions on Grid resources and Grid data



Storage Element

- It provides uniform access to storage resources
  (it may control simple disk servers, large disk arrays or Mass Storage Systems (MSS)
- Each LCG-2 site provides one or more SEs



Information System (IS)

- It provides information about the Grid resources and their status
  - ⇨ This information is essential for the operation of the whole Grid
- Location of availiable Computing Elements to run jobs
- Finding of SEs that holding replicas of Grid files and the catalogs keeping the information on these files
- The information is stored in databases
- The published information is used for
  - ✓ monitoring purposes ⇨ for analyzing usage and performance of the Grid, detecting fault situations and any other interesting events
  - ✓ accounting purposes ⇨ for creating statistics of the applications run by the users in the resources

Monitoring systems

- ***GridICE***

✓ Distributed monitoring tool designed for Grid systems

✓ GridICE is a proper monitoring system to
  - collect
  - store
  - monitor

  data about the resources' situation

↳ Low level data like CPU load, available memory, storage utilization, etc

↳ Data about services

↳ Data about Grid like the number of each site's CPUs, the number of submitted jobs, the number of executing jobs, the number of waiting jobs, the number of each site's free CPU, the available storage space, etc

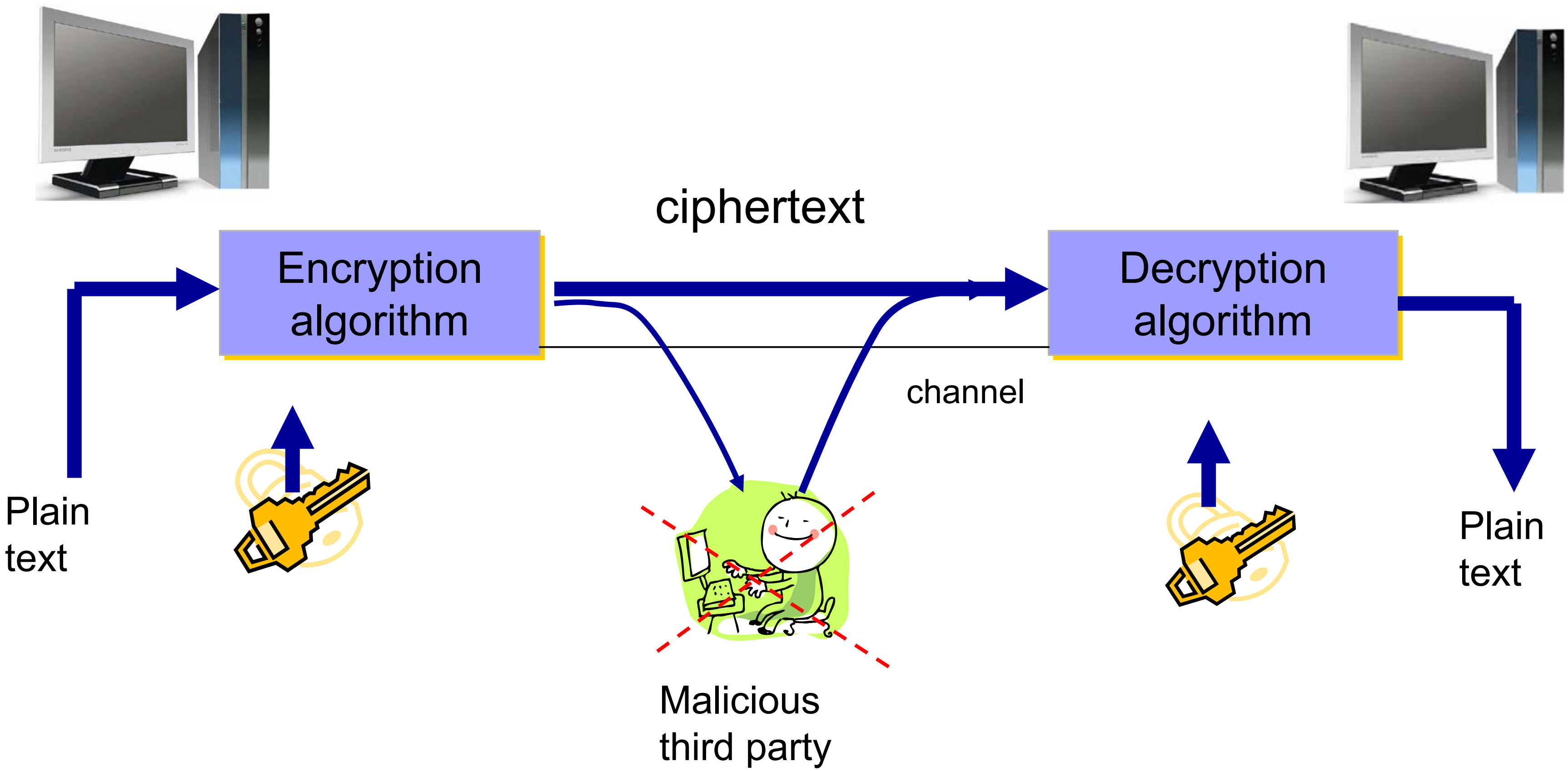
GridICE

(http://gridice2.cnaf.infn.it:50080/gridice/site/site.php)

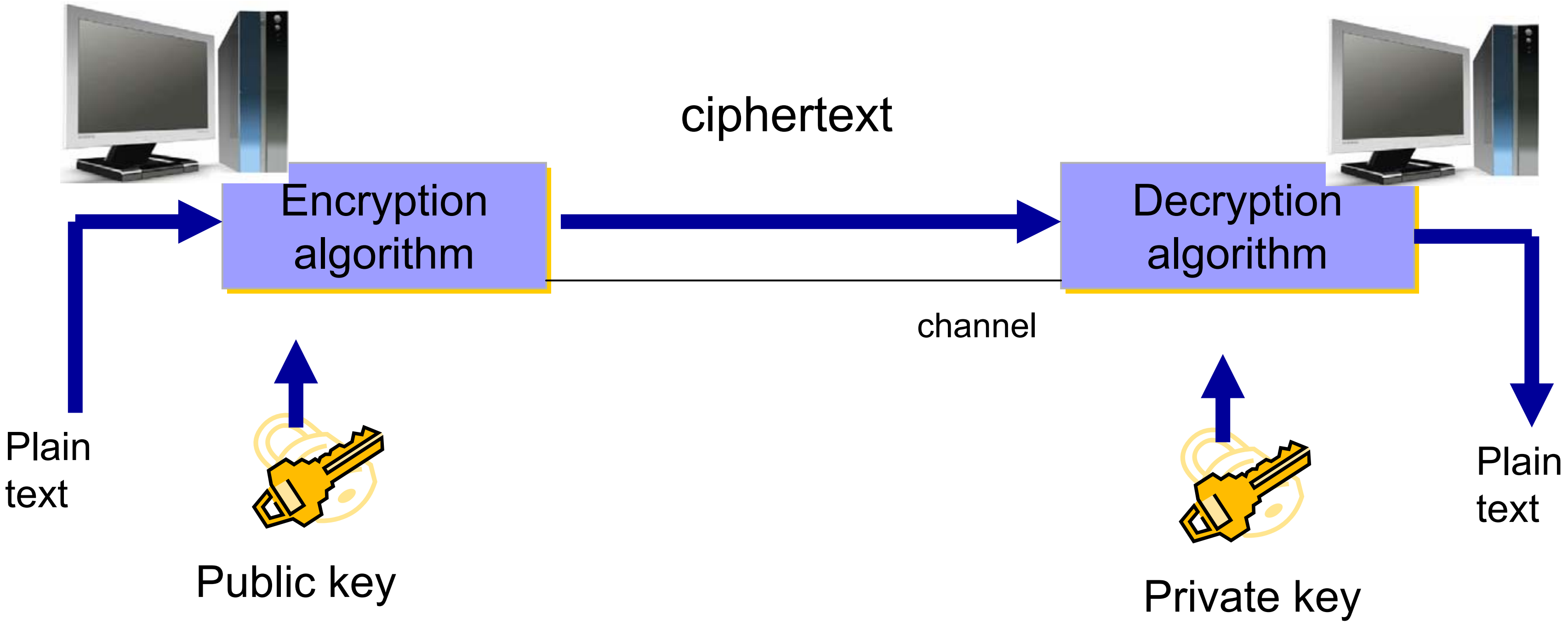
INFN - GridICE - Grid Monitoring Service - Microsoft Internet Explorer

	Computing Resources										Storage Resources		
Site ▼	GK#	Q#	RunJob	WaitJob	SlotLoad	MH#	Power	WN#	CPU#	CPULoad	Available	Total	%
CIEMAT-LCG2	1	5	2	0	3%	40	613K	24	96	0%	2.4 TB	4.1 TB	[illegible]0%
CINES	1	2	0	0	0%	2	-	-	-	-	4.5 GB	7.5 GB	[illegible]9%
CNB-LCG2	1	4	8	0	5[illegible]%	11	126K	8	32	25%	63.5 GB	67.7 GB	6%
CNR-ILC-PISA	1	6	2	1	5[illegible]%	4	18K	2	4	25%	381.3 GB	381.6 GB	0%
CSCS-LCG2	1	5	0	0	0%	3	-	-	-	-	5.8 TB	6.8 TB	14%
CY01-LCG2	1	7	2	1	8%	5	-	-	-	-	629 GB	655.2 GB	4%
CYFRONET-IA64	1	6	0	0	6%	-	-	-	-	-	-	-	100%
CYFRONET-LCG2	1	10	2	0	2%	1	-	-	-	-	103.5 GB	2 TB	95%
DESY-HH	4	41	190	17	87%	12	-	-	-	-	4.1 TB	6.9 TB	[illegible]0%
Durham	1	8	4	1	5%	-	-	-	-	-	1.2 TB	1.8 TB	29%
ENEA-INFO	1	3	0	0	7%	17	67K	14	14	0%	14.8 GB	18.2 GB	19%
ESA-ESRIN	1	3	0	0	0%	8	12K	6	6	0%	389.1 GB	845.1 GB	54[illegible]%
FMPhI-UNIBA	1	3	0	2	0%	1	-	-	-	-	2.6 TB	2.7 TB	6%
FZK-LCG2	2	29	518	4	97%	-	-	-	-	-	2.3 TB	10.3 TB	78%
GOG-Singapore	1	6	0	0	14%	2	-	-	-	-	36.2 GB	59.6 GB	[illegible]9%
GR-01-AUTH	-	-	-	-	-	9	-	-	-	-	465.2 GB	478.7 GB	3%
GR-02-UoM	1	8	3	5	[illegible]8%	3	-	-	-	-	856.1 GB	873.6 GB	2%
GR-03-HEPNTUA	2	8	3	0	23%	2	-	-	-	-	846.5 GB	873.6 GB	3%
GR-04-FORTH-ICS	1	9	3	2	60[illegible]%	6	41K	3	8	28%	0.9 GB	7.6 GB	89%
GR-05-DEMOKRITOS	1	2	4	1	10%	4	-	-	-	-	1.1 TB	1.1 TB	2%
GRIF	2	5	4	0	22%	10	-	-	-	-	995.4 GB	1.1 TB	11%
GSI-LCG2	4	6	0	0	0%	-	-	-	-	-	65.3 GB	67.7 GB	4%
HEPHY-UIBK	1	2	1	0	4%	2	-	-	-	-	1.1 TB	1.6 TB	32%
HG-01-GRNET	1	15	14	0	22%	28	512K	23	92	14%	34 TB	35.4 TB	4%
HPC2N	1	1	33	864	34%	2	-	-	-	-	-	-	-

Cryptography components



Public key infrastructure

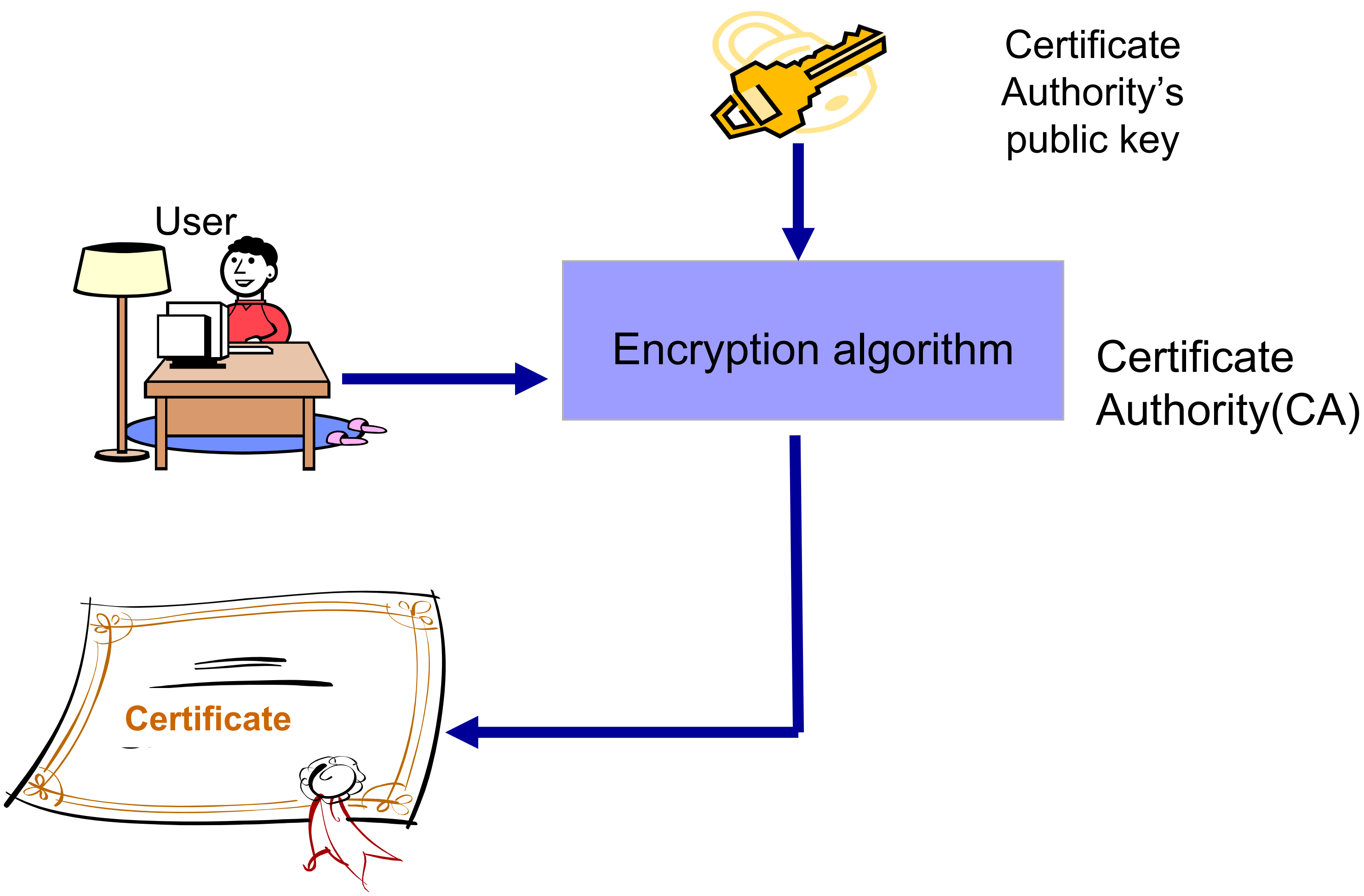


Digital certificate X.509

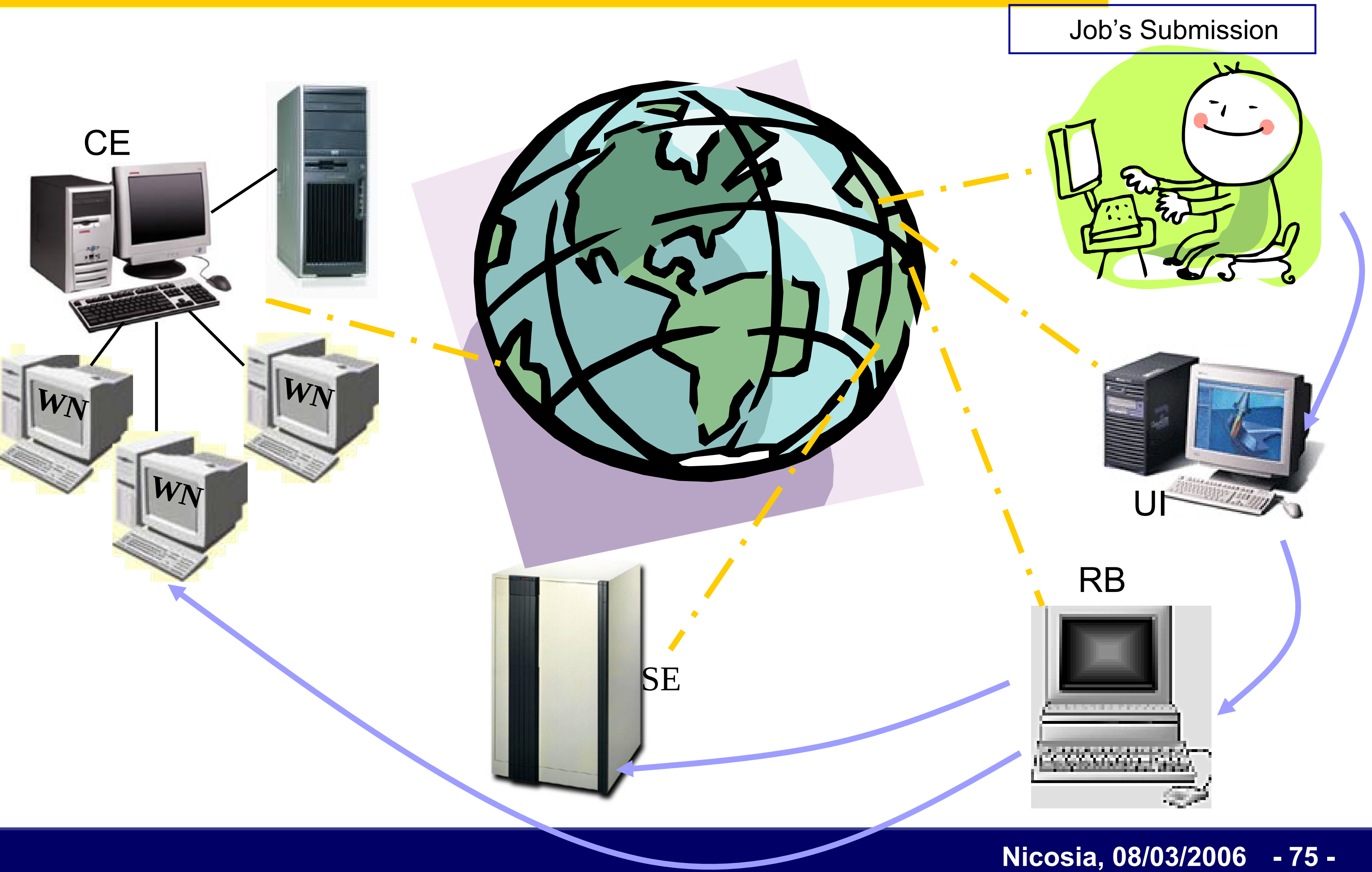
- Each entity (user, resource) must obtain a certificate
- The certificate includes information, such as the expiration date, the Certification Authority that signed it, the owner's public key and a DN
- The DN defines uniquely the owner and has the following fields:

C = Owner's country
O = Owner's organization
OU = Owner's group
CN = Owner's name

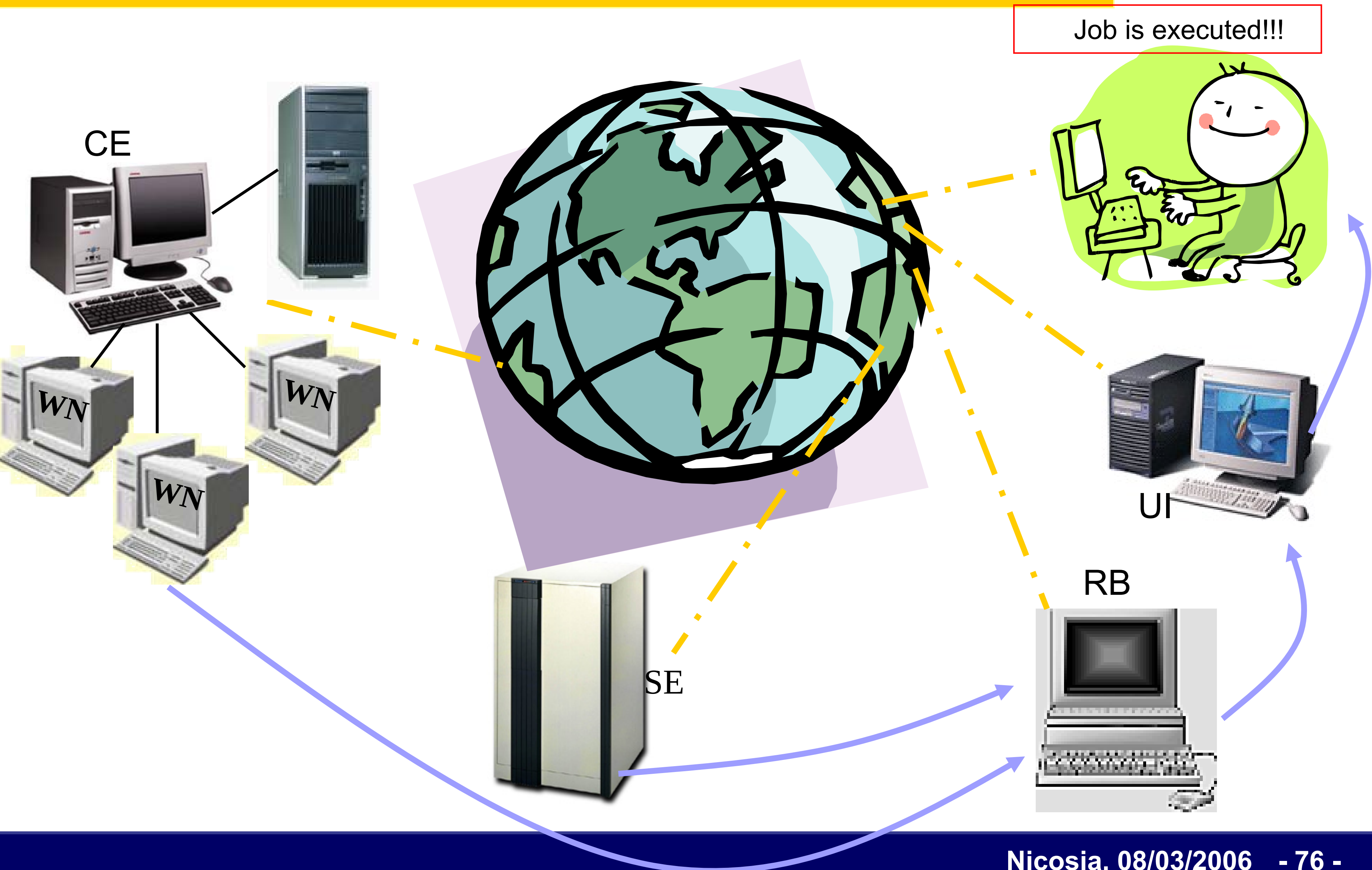
Certificate authority



Υποβολή εργασίας στο Grid



Υποβολή εργασίας στο Grid



References

- ✓ Grid café:
  http://gridcafe.web.cern.ch/gridcafe
- ✓ Global Grid Forum:
  http://www.gridforum.org/
- ✓ Gridtoday:
  http://www.gridtoday.com/gridtoday.html
- ✓ Grid Computing
  http://en.wikipedia.org/wiki/Grid_computing
- ✓ Distributed Computing
  http://en.wikipedia.org/wiki/Distributed_systems
- ✓ Supercomputing
  http://en.wikipedia.org/wiki/Supercomputing
- ✓ LCG-2 User Guide
  http://egee.itep.ru/User_Guide.html
- ✓ EGEE (Enabling Grids for E-science)
  http://public.eu-egee.org/intro/