The Portuguese National Grid Initiative (INGRID)

Jorge Gomes

LIP Computing Division

Lisbon

e-IRG workshop, Lisbon, 11 of October 2007

- Background
  - – Portuguese participation in large infrastructure projects
- The national grid initiative
  - – Current activities
  - – Perspectives
- Future

	DataGrid	CrossGrid	LCG	EGEE-I	EELA	EGEE-II	Int.Eu.Grid
2001							
2002							
2003							
2004							
2005							
2006							
2007							

Portuguese Participation driven by LIP

- LIP is Portuguese scientific research laboratory
  - High Energy Physics (HEP)
  - Associated laboratory funded by the Portuguese public funding agencies
  - Private non-profit association
  - Created in 1986 when Portugal joined CERN (celebrated 20 years)
- LIP has three centres:
  - Lisbon, Coimbra and Faro

- LIP participation in physics experiments includes:
  - **Atlas**, **CMS**, Compass, Auger, AMS, SNO, Zeplin, Hades, n-TOF, …
- Other activities include:
  - Building DAQ systems and detectors, detectors R&D, medical physics, Geant4
  - **grid computing**, electronics, precision mechanics etc

LHC - CERN
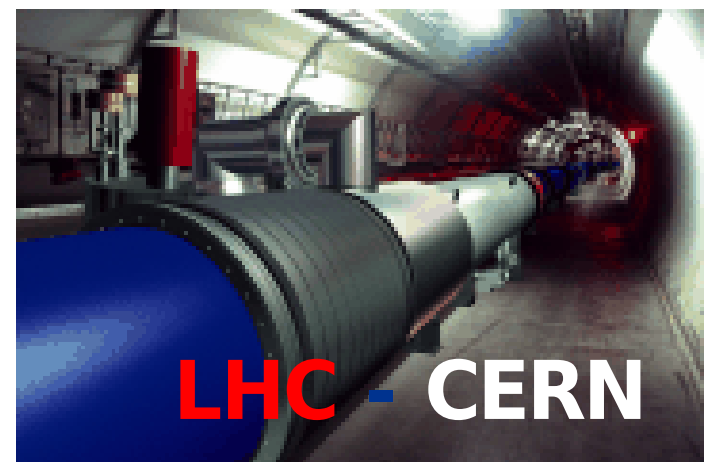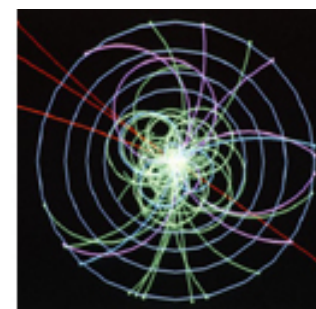
- The Large Hadron Collider (LHC) is the largest scientific instrument on earth
  – Located at CERN in the Swiss/French border
  – 27 Km of circumference
  – At 100 meters of depth (average)
  – Consuming 120 Megawatt
  – 600 million particle collisions per second
  – Reproducing the energy density that existed just a few moments after the big bang
- Objective:
  – Probe deeper into the matter structure than ever before
  – **understand unsolved questions about the universe**
    - Will allow to detect the Higgs boson
- Four experiments working in parallel
- Expected to become operational in 2008

Detector

Particle

Particle

Overall view of the LHC experiments.
Beams of protons will collide at an energy of 14 TeV
Beams of lead nuclei will collide at an energy of 1150 TeV

GENÈVE

LHC - B
Point 8

CERN
ATLAS
Point 1

ALICE
Point 2

CMS
Point 5

SPS

TI 8

TI 2

LHC - B

ATLAS

ALICE

CMS

LEP/LHC

E540 - V10/09/97
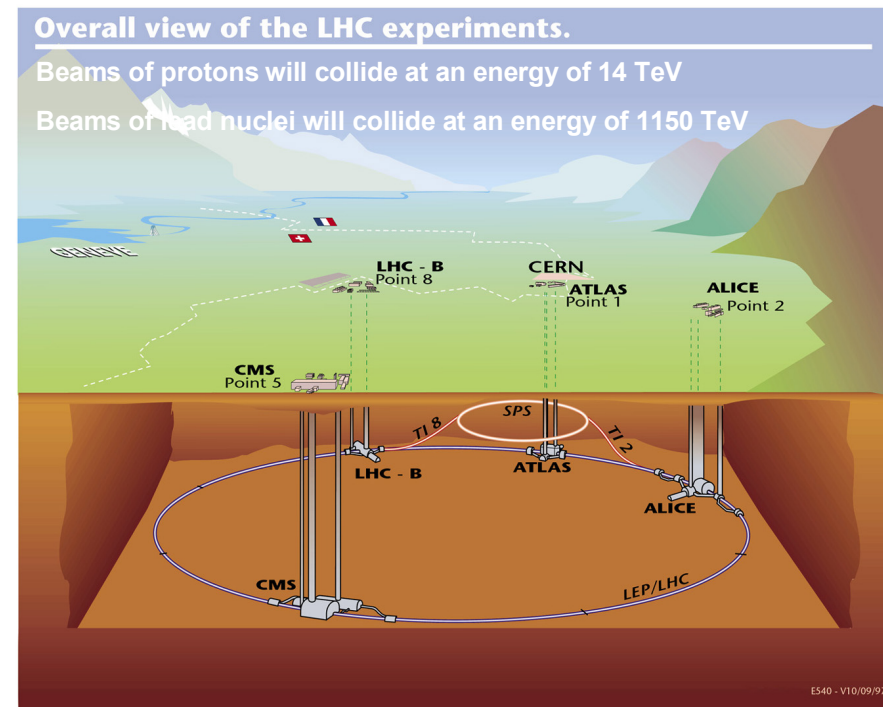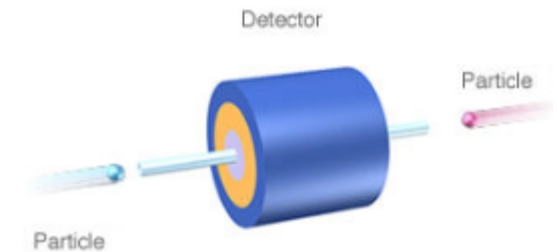
LCG aims to build/maintain a data storage and analysis infrastructure for the large LHC physics community.

- **LHC experiments are expected to produce 10 to 15 Petabytes of experimental data annually**
  - 10^9 collisions/second (1 GHz) – 100Hz (after filtering)
  - 1 collision = 1MB of data
  - 100 MB/s $\Rightarrow$ 10 PB year = 20 Km CD stack

- **Computing power > 100 000 "today's fastest" PC for analysis and simulation**

- **Needed to be available during the 15 years life time of the LHC machine**

- **Fully accessible to about 5000 scientists from more than 500 institutes around the world.**

Balloon 30 Km

CD stack with 1 year LHC data ≃ 20 Km

Concorde 15 Km

Mt Blanc 4.8 Km
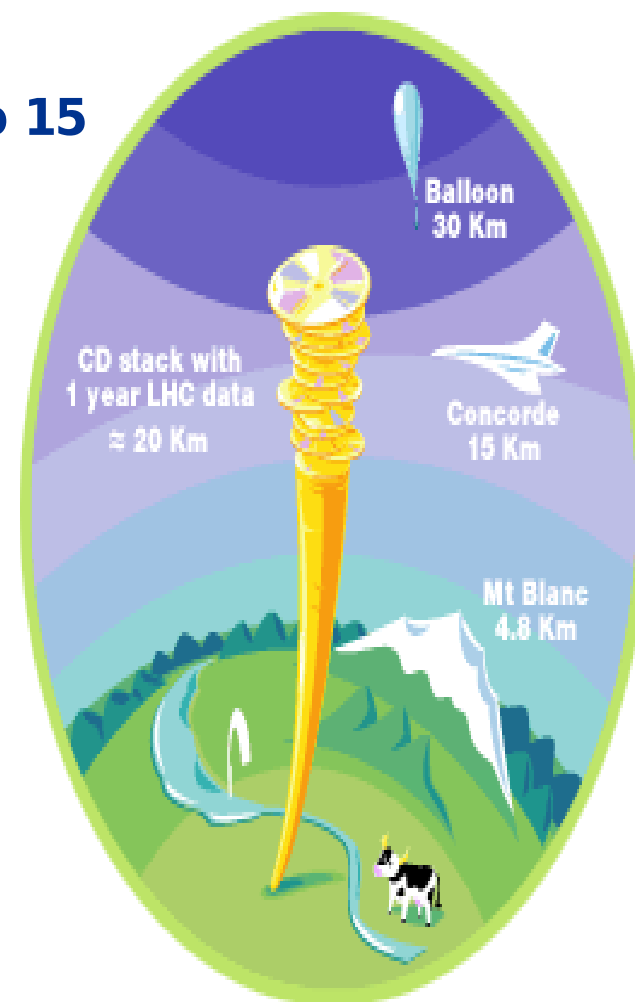
LCG depends on two major science grid infrastructures ....

BIGGEST GRID COMPUTING INFRASTRUCTURE WORLDWIDE

EGEE
Enabling Grids
for E-sciencE

240 sites
45 countries
41,000 CPUs
5 PetaBytes
>10,000 users
>150 VOs
>100,000 jobs/day

Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
...

91 partners in 32 countries
25 collaborating projects

GridPP
UK Computing for Particle Physics

- The Portuguese government has signed a MoU with CERN regarding LHC computing
- In this context an LCG Tier-2 will be established in Portugal under the responsibility of LIP
  - Computing and data storage infrastructure integrated in the LHC Computing Grid
- The Tier-2 will be federated with three sites:
  - LIP-Lisbon
    - housed at the LIP datacentre in Lisbon
  - LIP-Coimbra
    - housed in partnership with CFC
    - CFC hosts the largest HPC cluster in Portugal – "Milipeia"
  - Portuguese main node for GRID computing
    - housed at new facilities in Lisbon at the LNEC campus

- Portugal together with Spain form the EGEE Southwest federation
- LIP coordinates EGEE in Portugal
- Portuguese sites
  - LIP
    - Lisbon (core services, production and pre-prod)
    - Coimbra
  - Univ Lusiada
    - Famalicão
  - Univ Porto
    - Porto (3 clusters)
  - Univ Minho
    - Braga
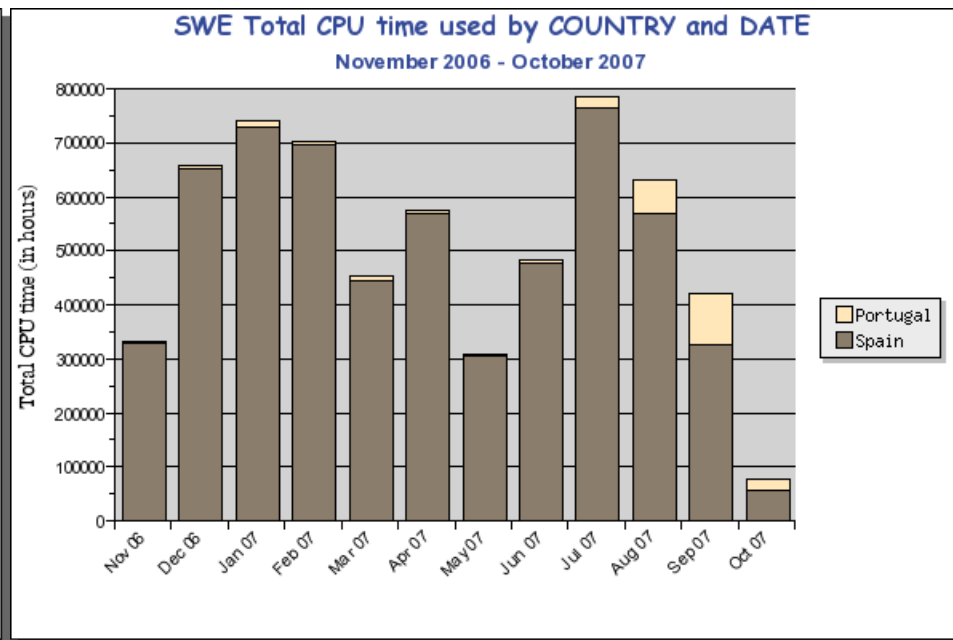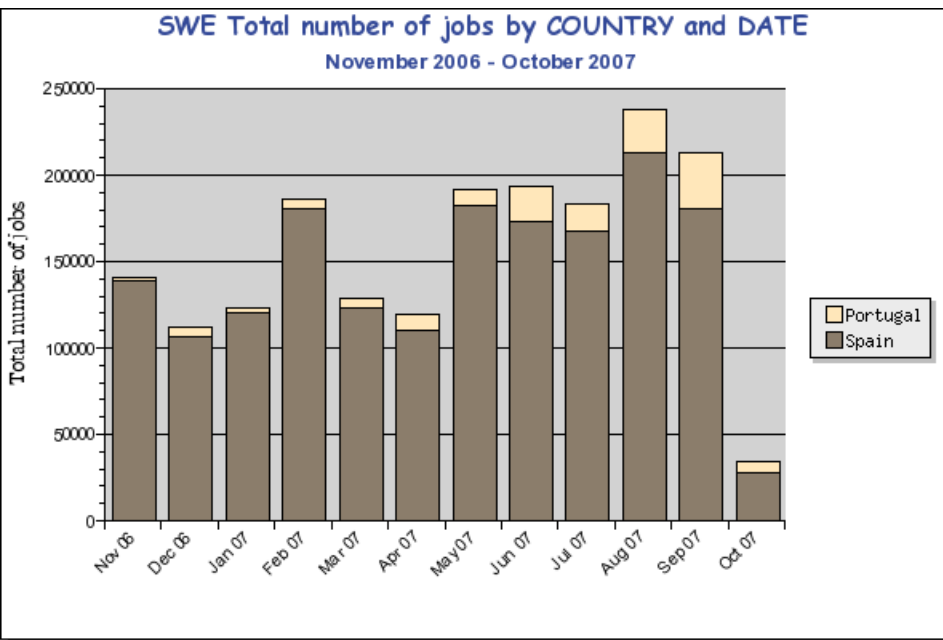 - IEETA
    - Aveiro (pre-prod)

- grid001.fc.ul.pt -> 22
- grid001.fe.ul.pt -> 22
- zephyr.up.pt -> 48
- ce.egee.di.uminho.pt -> 6
- ce02.lip.pt -> 76
- grid006.lca.uc.pt -> 87
- ce01.fam.ulusiada.pt -> 9

- **Further EGEE infrastructure improvements in the coming months**
  - **More resources at LIP**
  - **More resources at University of Porto**

- **Currently working with other organizations to deploy new sites**
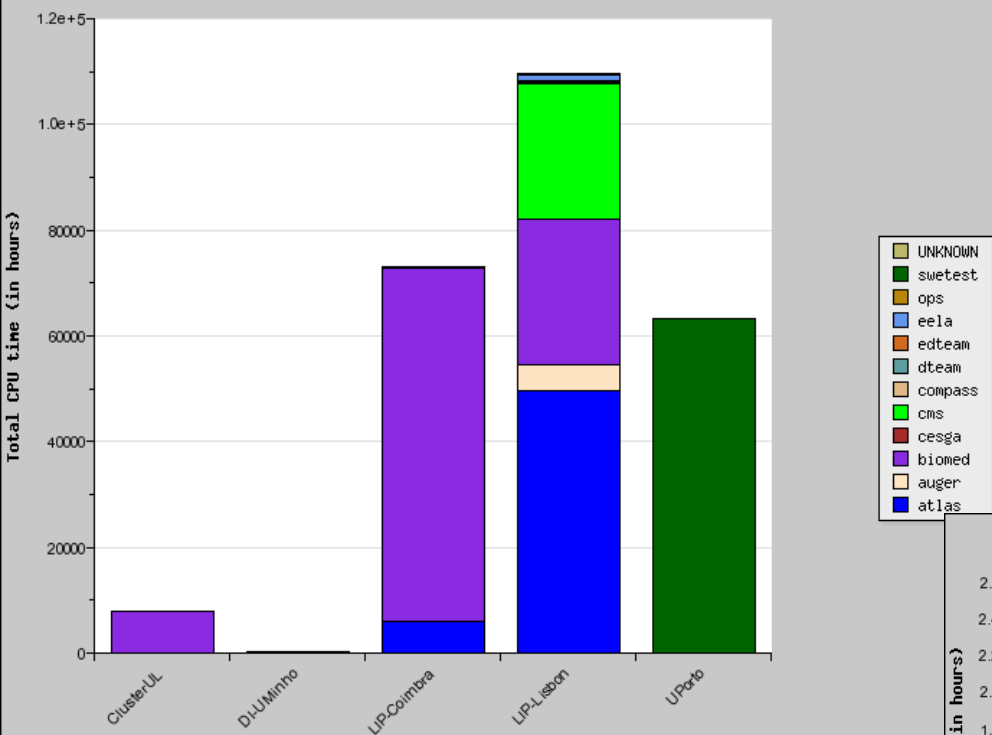 - **IST**
  - **Several other contacts**

- Total number of CPUs in the federation > 1400
- Total number of CPUs in Portugal ~ 262

- SWE federation total
  - 1.863.000 jobs during last 12 months
  - 301.400 days normalized CPU time
  - 257.000 days absolute CPU time
  - Equivalent to > 700 machines 24h/day

SWE Total number of jobs by COUNTRY and DATE
November 2006 - October 2007

SWE Total CPU time used by COUNTRY and DATE
November 2006 - October 2007

Production only

Portugal Total CPU time used by VO and SITE
ALL VOs. November 2006 - October 2007

Portugal Cumulative Total CPU time used by SITE and DATE
ALL VOs. November 2006 - October 2007

SouthWesternEurope Total number of jobs per SITE
ALL VOs. November 2006 - October 2007

Portugal Cumulative Total CPU time used by SITE and DATE
ALL VOs. November 2006 - October 2007

SouthWesternEurope Normalised CPU time per VO
ALL VOs. November 2006 - October 2007

atlas
biomed
cms
ific
lhcb
Others

36.6%
4.2%
31.9%
1.6%
2.4%
23.3%
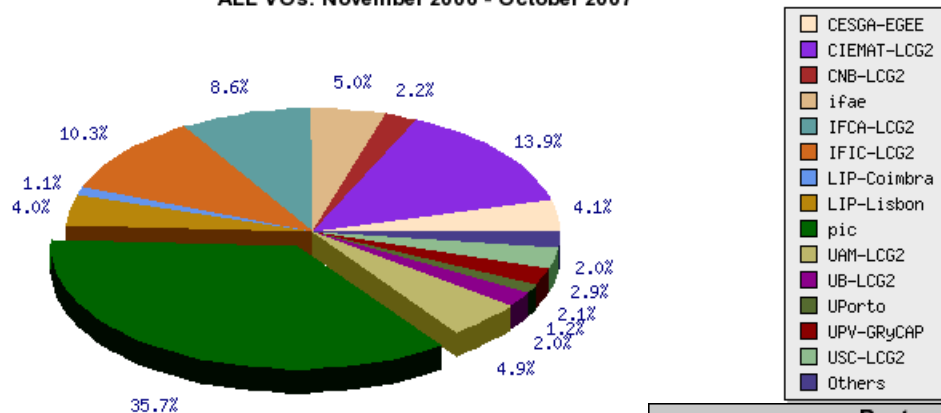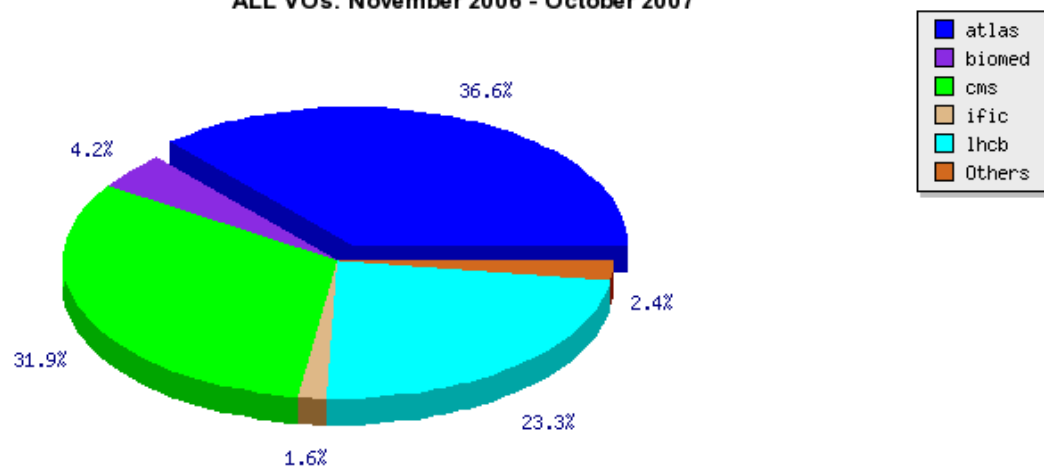
(C) CESGA 'EGEE View': SouthWesternEurope / normcpu / 2006:11-2007:10 / VO-SITE / all / ACCBAR-LIN / i
2007-10-09 15:55 UTC

Portugal Normalised CPU time per VO
ALL VOs. November 2006 - October 2007

atlas
auger
biomed
cms
swetest
Others

2.8%
30.6%
38.0%
14.9%
12.6%
1.1%
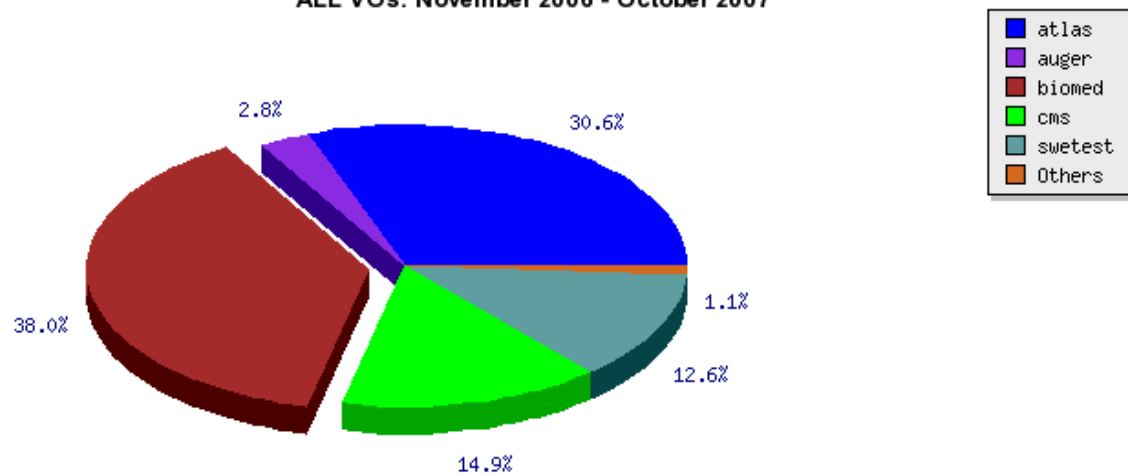
(C) CESGA 'EGEE View': Portugal / normcpu / 2006:11-2007:10 / VO-DATE / all / LINES-LIN / i
2007-10-09 15:55 UTC

- Applications
  - ifusion
  - ienvmod
  - iusct
  - ibrain
  - ihep
  - iplanck
  - icompchem
  - open to more applications
- Project specific
  - lmain, itut, itest
- Operations specific
  - imon
- Other
  - icesga
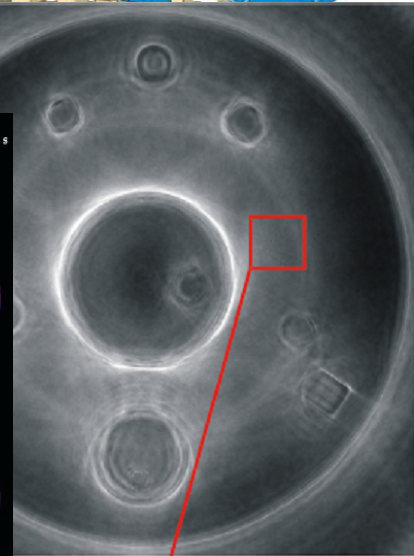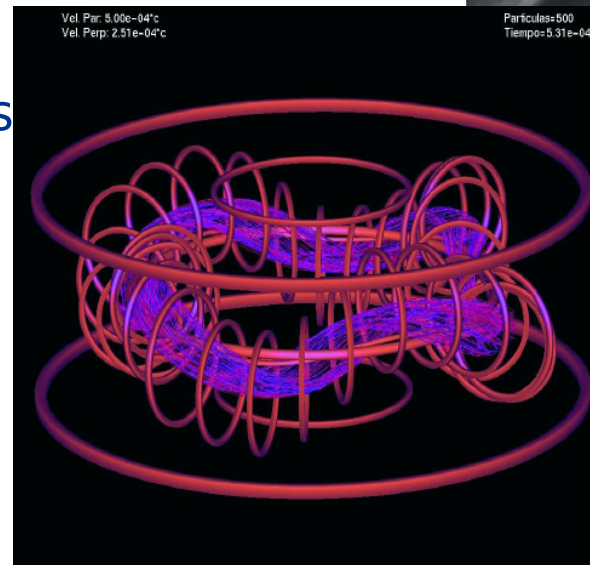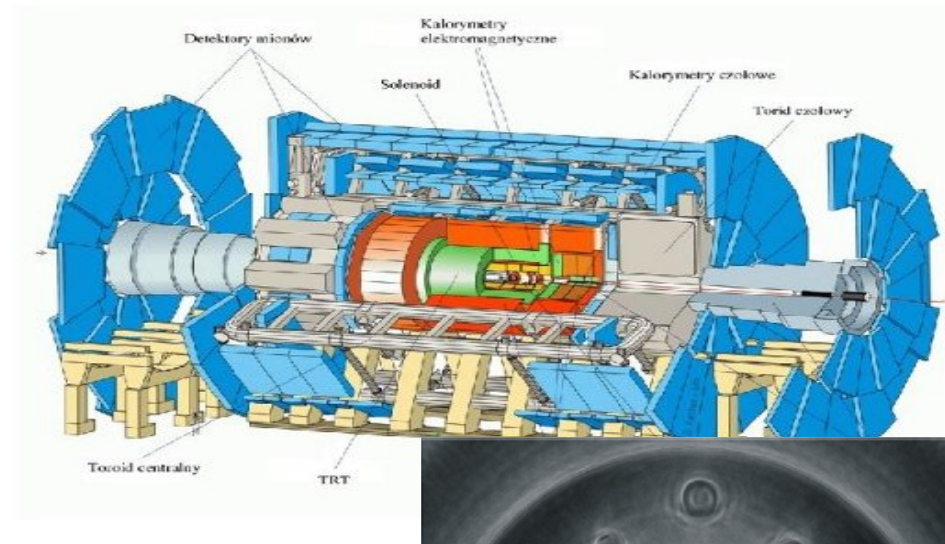
CMS PhEDEx - Transfer Rate
24 Hours from 2007-10-09 14:00 to 2007-10-10 14:00 UTC
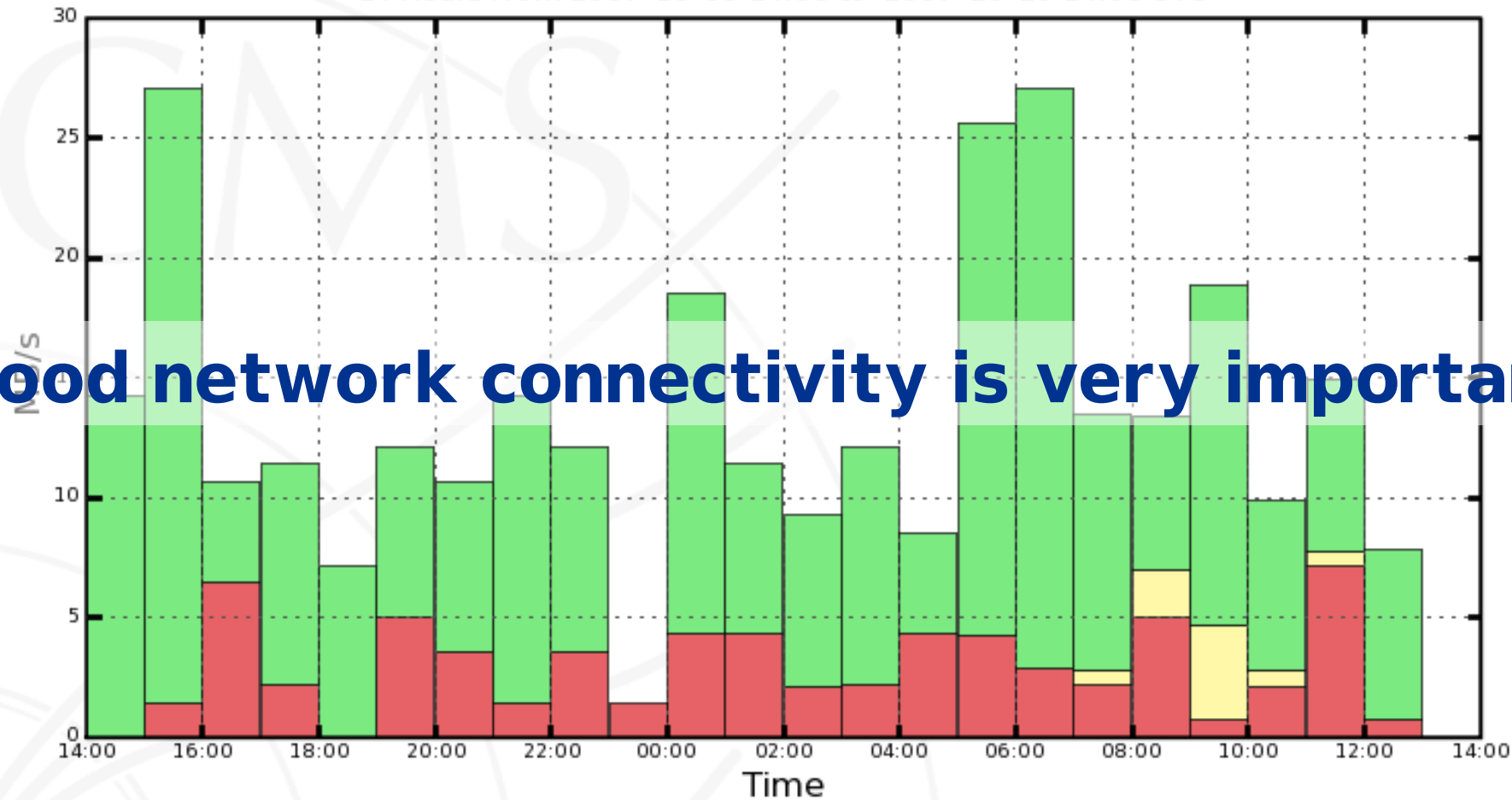
Good network connectivity is very important

T1_CERN_Buffer to T2_LIP_Lisbon   T1_IN2P3_Buffer to T2_LIP_Lisbon   T1_PIC_Buffer to T2_LIP_Lisbon

Maximum: 27.02 MB/s, Minimum: 1.42 MB/s, Average: 13.55 MB/s, Current: 7.82 MB/s

- The Portuguese NGI was officially launched in
  29 April 2006 by the Ministry of Science
- Goals:
  - Support the development of support infrastructures for distributed computing aimed at sharing resources for the resolution of complex problems with demanding computing requirements
  - Ensure the development of competence and capacity for the evolution of GRID computing in Portugal
  - Integrate Portugal in major International grid computing infrastructures
- Entities
  - Execution of the NGI programme by the Portuguese Science Foundation (FCT)
  - UMIC is closely following the initiative
- More information at the web site

- Workshop INGRID'06 took place in Braga 11 November 2006
  - 170 attendees

- NGI has establish cooperation on grid computing with Spain (IBERGRID)

- NGI has been following closely the:
  - Grid computing activities in Europe
  - e-IRG activities (EU: e-infrastructures reflection group)
  - European Grid Initiative (EGI) related activities

- First call for pilot applications launched in November of 2006

  - Organized by FCT
  - Evaluation by international board
  - 37 projects submitted
  - 13 projects approved
  - Total funding ~ 1.500.000 €

- GRID for ATLAS/LHC data simulation and analysis

- G-Cast: Application of GRID-computing in a coastal morphodynamics nowcast-forecast system

- GridClass - Learning Classifiers Systems for Grid Data Mining

- PoliGrid - distributed policies for resource management in Grids

- Collaborative Resources Online to Support Simulations on Forest Fires (CROSS-Fire): a Grid Platform to Integrate Geo-referenced Web Services for Real-Time Management

- P-found: GRID computing and distributed data warehousing of protein folding and unfolding simulations

- GERES-med: Grid-Enabled REpositorieS for medical applications

- BING –Brain Imaging Network Grid

- GRITO – A Grid for preservation

- PM#GRID - GRID Platform Development for European Scale Satellite Based Air Pollution Mapping

- AspectGrid: Pluggable Grid Aspects for Scientific Applications

- Call to fund equipments to support the approved projects
  - Possibly until the end of the 2007 ...

- Open a new call for applications in 2008

- Improvement of network connectivity for GRID computing (FCCN)

- Deployment of a grid infrastructure (INGRID+) in coordination with IBERGRID and other international initiatives

- International connectivity through Spain using proprietary fibre
  - North: Minho – Galicia
  - South: going through Spanish Estremadura …
- Providing
  - better Geant connectivity
  - a loop providing redundancy for both countries
  - grid computing support

- Improved connectivity for grid clusters in Portugal
  - Dedicated bandwidth for grid computing

- Main node for GRID computing
  - First step towards INGRID+
  - Project started in summer 2007
  - Expected to become operational in September 2008
  - Consortium of: LNEC, FCCN and LIP
- Dedicated facility for GRID computing
  - **Provide core services for the integration of Portuguese GRID resources into a coherent infrastructure**
  - Provide a large set of GRID computing and storage resources
  - Housing for GRID resources from other organizations
- Users
  - National grid initiative projects
  - The Portuguese Tier-2 for the LHC
  - Projects in the context of IBERGRID
  - Portuguese researchers with demanding applications
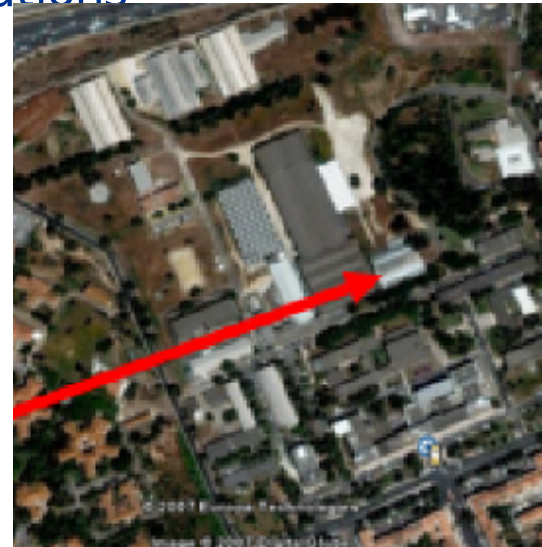
- GRID Data Centre
  - To be built at the LNEC's campus in Lisbon
  - Located near the FCCN NOC
  - High bandwidth network connectivity to the FCCN backbone
  - Adequate cooling and power infrastructure for GRID computing

- Will house:
  - Core GRID services for INGRID+
  - GRID computing cluster
    - \> 500 CPU COREs and ~ 200TB of storage already in 2008
    - Later deployments for > 2000 CPU COREs and ~ 500TB of storage and look forward …
    - The cluster will be managed by LIP
  - FCCN robotic storage
  - LNEC GRID cluster
  - Other GRID resources …
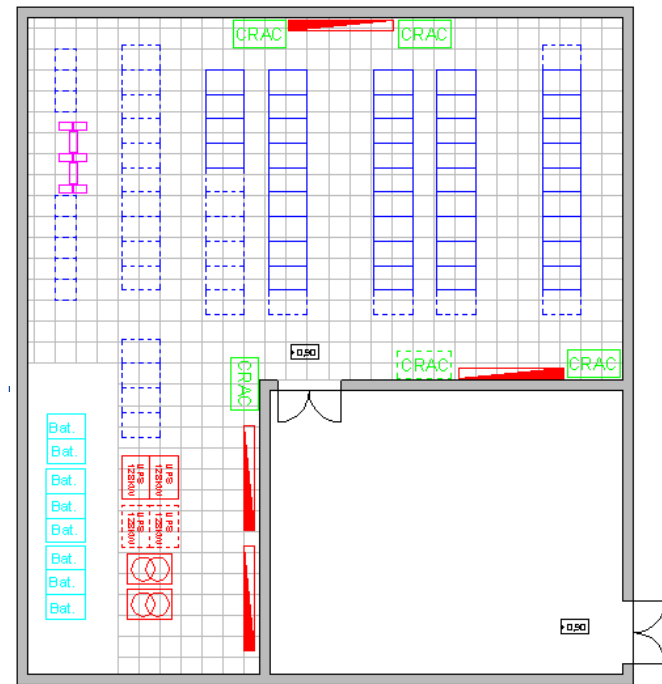
- In addition of the Main node for grid computing the Portuguese LCG Tier-2 will be composed of 2 other LIP clusters

- The LHC starts in 2008

- The LIP clusters in LCG/EGEE will be upgraded until the end of the year
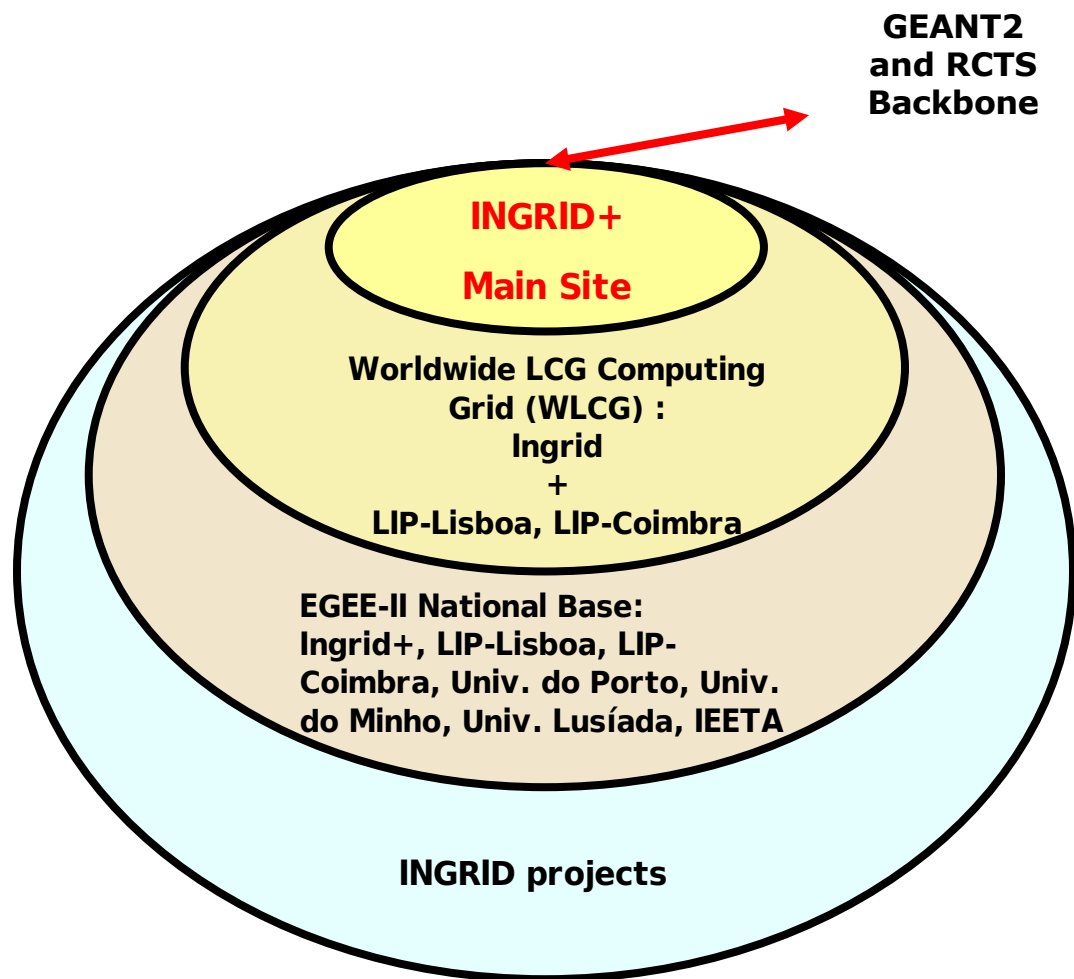
GEANT2 and RCTS Backbone

INGRID+

Main Site

Worldwide LCG Computing Grid (WLCG) :
Ingrid
+
LIP-Lisboa, LIP-Coimbra

EGEE-II National Base:
Ingrid+, LIP-Lisboa, LIP-Coimbra, Univ. do Porto, Univ. do Minho, Univ. Lusíada, IEETA

INGRID projects
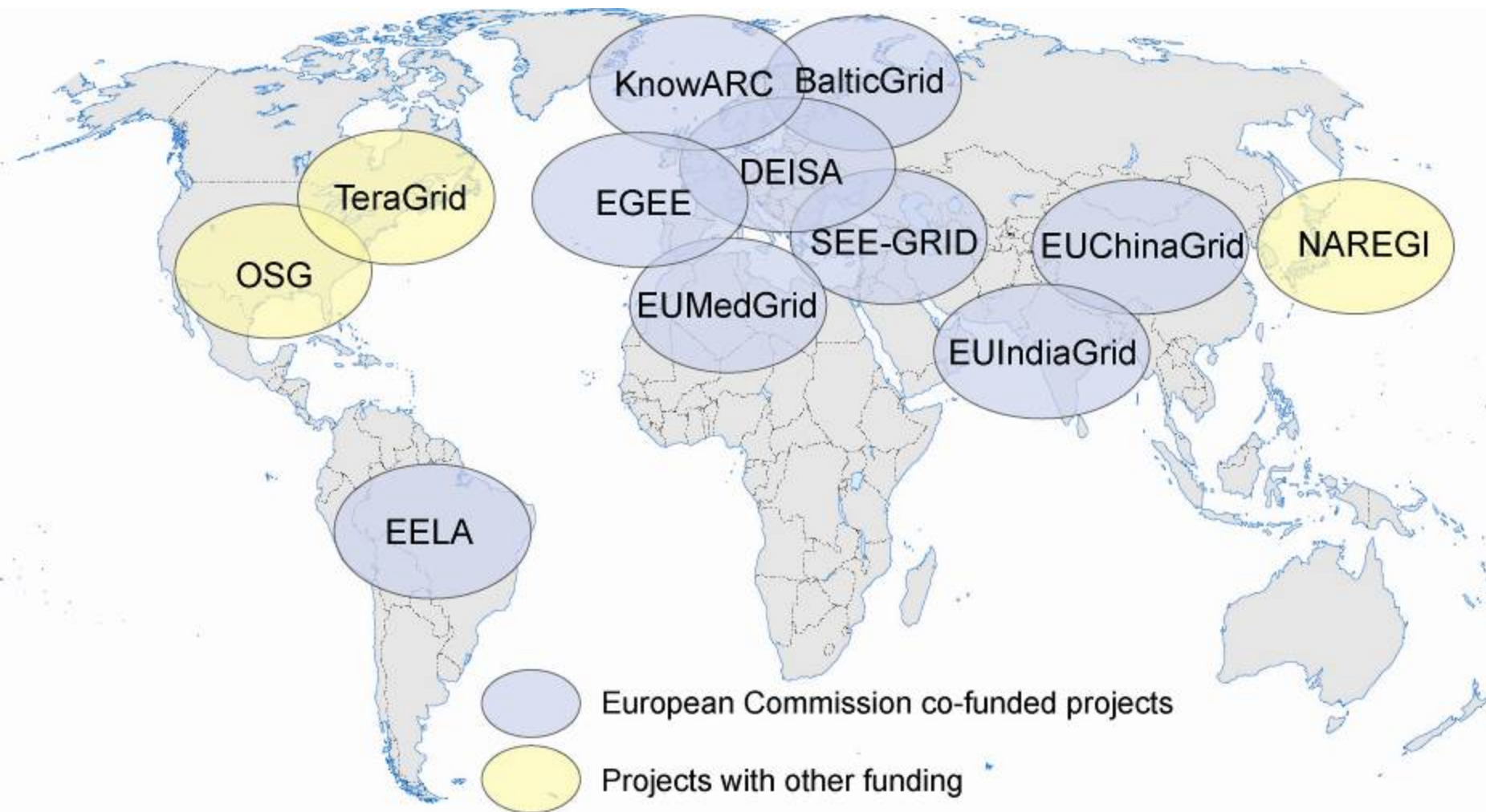
1. Main node @ LNEC
   - Core services
   - Resources
- Additional resources from Portuguese WLCG sites
- Resources from Portuguese EGEE sites
- INGRID projects
- Virtual organizations

Start with an infrastructure fully interoperable with other gLite based infrastructures

Potential for linking ~80 countries by 2008

- Start with the gLite infrastructure already in place in Portugal
  - Actually three gLite infrastructures
    - EGEE, int.eu.grid, EELA
- Core services deployed at the FCCN NOC and LIP computer centre in Lisbon
  - Services and infrastructure managed by LIP
- Certification authority LIPCA accredited by the International Grid Trust Federation (IGTF)
  - Fully supported by IGTF relying parties worldwide
- Deploy and reconfigure services to obtain higher autonomy as individual infrastructure
- Deploy support for national virtual organizations

- Welcome new users
  - Provide adequate training
  - Establish new virtual organizations
  - Integrate new resources
- Deploy and support new middleware services targeted at the user needs
  - Enhance the infrastructure
  - Follow the technology
- Coordination with other European infrastructures extremely important
  - IBERGRID
  - Major grid infrastructures such as EGEE

- Dissemination
  - Attract users and communities
- Training
  - For end-users, site managers and developers
- Operation
  - Middleware and site deployment coordination
  - Accounting and monitoring

**Well know structure used by other infrastructures
Much know how available from other infrastructures**

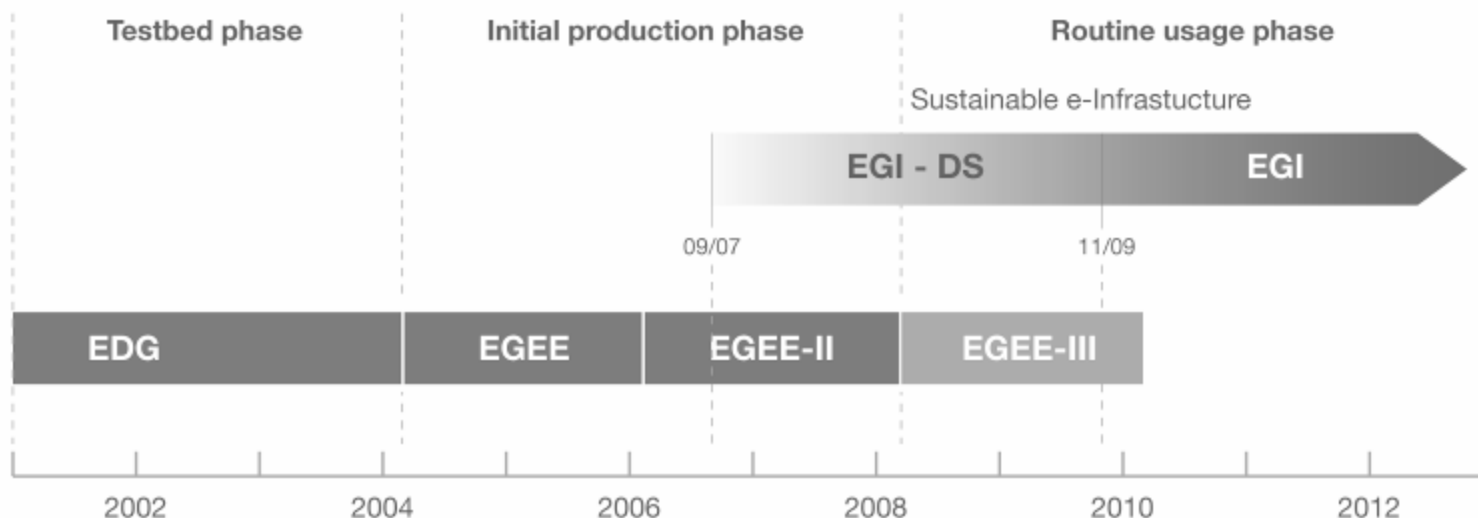**Should profit from it and not reinvent the wheel**

  - Integrate additional middleware functionalities
- Middleware validation
  - Certify the middleware
- Security
  - Incident handling
  - Policies, procedures and best practices
  - Auditing

- A common grid infrastructure between Portugal and Spain is being planned
- Profit from the collaboration in EU projects
- <u>More on next talk by Ramon Gavela</u>
- Important step for sharing resources across borders
- Potential positive impact in the Portuguese NGI
  - Foster collaboration in many areas
  - Develop a common sustainable e-infrastructure

- EGEE-III proposal submitted on 20 September to EU
  - call infra-2007-1.2.3

- Main goals:
  - Expand EGEE infrastructure
    - More resources
    - More users and communities
  - Prepare migration to a sustainable infrastructure based on NGIs

- Will start immediately after EGEE-II
- Two years project from 2008 to 2010

- Need to prepare permanent, common Grid infrastructure
- Ensure the long-term sustainability of the European e-Infrastructure independent of short project funding cycles
- Coordinate the integration and interaction between National Grid Infrastructures (NGIs)
- Operate the production Grid infrastructure on a European level for a wide range of scientific disciplines

- **EGI Design Study proposal approved to the European Commission (started 1st September'07)**

- **Supported by 31 National Grid Initiatives (NGIs)**

- **2 year project to prepare the setup and operation of a new organisation... able pan-E...**

- **Feder... uild a European organisation**

- **Well defined, complimentary responsibilities between NGIs and EGI**

EGI workshop held during EGEE'07 conference last week in Budapest

Prof. Gaspar Barreira appointed spokesman of the advisory committee

http://www.eu-egi.org

Thank you