Grid **Applications** in EELA

José Manuel Gutiérrez

Universidad de Cantabria

IBERGRID Conference

Santiago de Compostela, 14-16.05.2007

www.eu-eela.org

Information Society
and Media

E-infrastructure shared between **E**urope and **L**atin **A**merica

- The **EELA** Project

  Build a bridge between consolidated e-infrastructure initiatives in europe and emerging ones in latin america.

  Reinforce collaboration between latin america and europe

- **EELA Applications**
  - **Biomed**
  - **HEP**
  - **E-Learning**
  - **Climate**

E-infrastructure shared between Europe and Latin America

Italy: INFN
Portugal: LIP
Spain: **CIEMAT**, CSIC, Red.ES, UC, UPV
CERN

- **10 Countries**
- **21 Partners**

Argentina: UNLP
Brazil: CEDERJ, RNP, UFF, UFRJ
Chile:REUNA, UDEC, UTFSM
Cuba:CUBAENERGIA
Mexico:UNAM
Peru:SENAMHI
Venezuela:ULA
CLARA

– **15 TUTORIALS: More than 1200 participants·days delivered so far in 2006!**

- **For grid users (total # of participants: 300)**
  - *Madrid - Spain* — *(20-22 February 2006)*
  - *Merida - Venezuela* — *(27-29 April 2006)*
  - *Itacuruçá - Brazil* — *(26-28 June 2006)*
  - *Mexico City - Mexico* — *(28-30 August 2006)*
  - *Santiago de Chile* — *(6-7 September 2006)*
  - *Madrid – Spain* — *(16-18 October 2006)*
  - *Merida – Spain* — *(7-9 November 2006)*
  - *La Plata – Argentina* — *(11-12 December 2006)*
  - *Bogota – Colombia* — *(6-7 March 2007)*

- **For grid system administrators (total # of participants: 160)**
  - *Madrid - Spain* — *(20-24 February 2006)*
  - *Itacuruçá - Brazil* — *(26-30 June 2006)*
  - *Mexico City - Mexico* — *(28 August - 01 September 2006)*
  - *Madrid – Spain* — *(19-20 October 2006)*
  - *Itacuruçá - Brazil* — *(29 November-01 December 2006)*
  - *Bogota – Colombia* — *(6-9 March 2007)*

- **1st EELA GRID SCHOOL. – Itacuruçá - Brazil (4 -15 December 2006)**
  - **Training for "gridification" of new applications to be run on EELA infrastructure (23 participants): 7/8 applications gridified**
    - *Climate applications, proposed by UC-SP, SENAMHI-PE and UDEC-CL (EELA applications)*
    - *Dist Simulation of Multiple Events, proposed by UNESP-BR, UNICAMP-BR and USP-BR (non-EELA application)*
    - *EMBOSS, proposed by UNAM-MX (non-EELA application)*
    - *LEMDist, proposed by UNAM-MX (EELA application)*
    - *LMS in Grid Env, EELA application proposed by CITMATEL-CU (non-EELA partner) and CUBAENERGIA-CU (EELA partner)*
    - *SATyrus, proposed by UFRJ-BR (EELA application)*
    - *SegHidro, proposed by UFCG-BR (non-EELA application)*
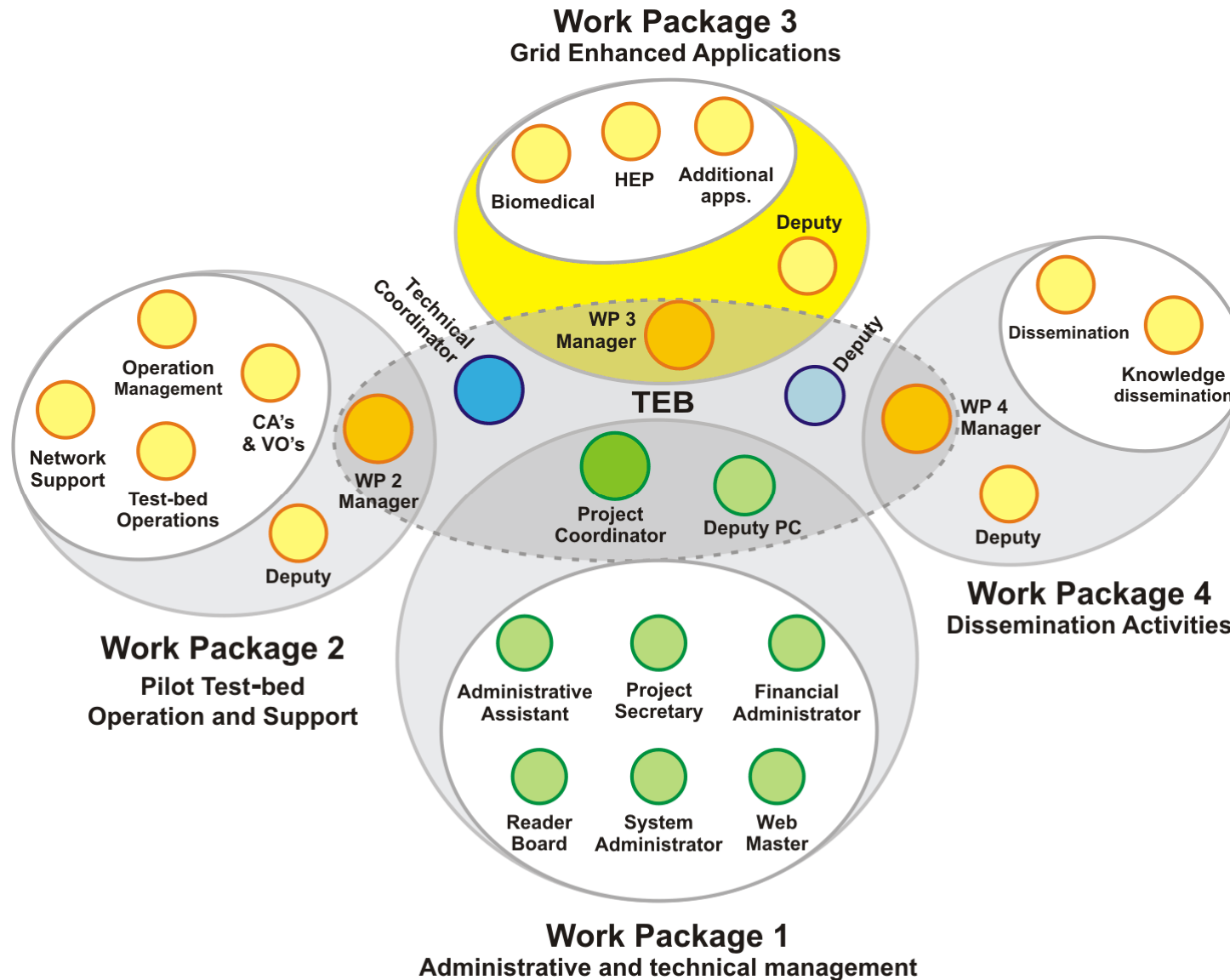   - *VoD, proposed by UFRJ-BR and CEDERJ-BR (EELA application)*

The Applications

- **WP3: Identification and support of Grid enhanced applications**

  – Coordinated by CIEMAT

  – Identifies, selects and customizes relevant applications and tools suitable for the Grid dissemination process in:

    - T3.1   Biomedicine (CUBAENERGIA, ULA, <u>UPV</u>)

    - T3.2   High Energy Physics (CERN, CIEMAT, UFRJ, <u>UNAM</u>, UNLP, UTFSM)

    - T3.3 Additional Applications

      - *e-Education (CECIERJ/CEDERJ, CIEMAT, CUBAENERGIA, <u>UFRJ</u>, UNAM)*

      - *Climate (<u>UC</u>, UDEC, SENAMHI)*

  – Aims at being the place of information exchange between already gridified applications and future ones.

E-infrastructure shared between Europe and Latin America

Work Package 3
Grid Enhanced Applications

Work Package 2
Pilot Test-bed
Operation and Support

Work Package 4
Dissemination Activities

**Work Package 1**
Administrative and technical management

Biomed

E-infrastructure shared between Europe and Latin America

- **The interest of the LA community is leaded by CUBAENERGÍA**
  - It is focused towards two main oncological problems:
    - Thyroid Cancer.
    - Treatment of Metastasis with P^{32}.
 - 9 centers in Cuba are interested (*5 Hospitals* and *4 Oncological Centers and Institutions*)

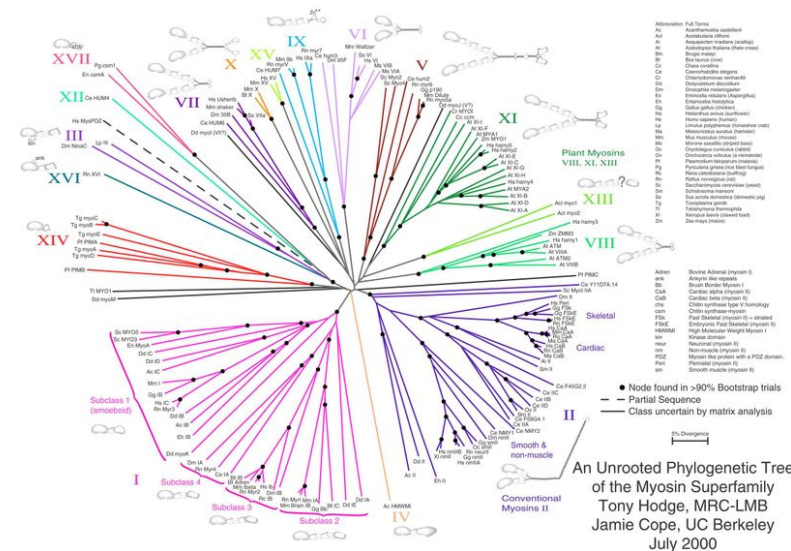- **Installed in several EELA sites, including Cuba with a local grid**
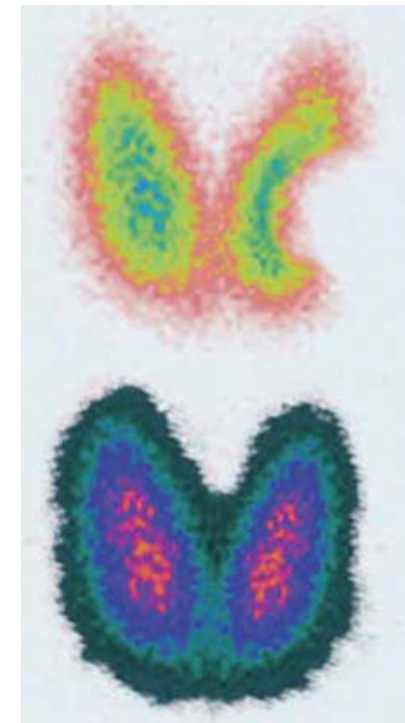
- **WISDOM can be executed on the EELA grid infrastructure.**
  - UPV already participated in previous data challenge on Malaria and Avian Flu.
- *An Experiment was prepared jointly by UPV and ULA.*
  - Two targets were accepted in _Plasmodium vivax Malaria_ in the WISDOM data challenge-II
  - One has been completely docked (100%) in the EELA infrastructure
- *Next Steps will be:*
  - Dock the second target
  - Analyze the first target
  - Inclusion of new targets in new data challenges.

- Users access the service through a Web Portal (CeCalcULA)
- Access to the EELA Grid is performed through the *Gate-to-Grid*.
- *Gate-to-Grid* is an EELA Grid Node which provides a WSRF-Based Web Interface.
- Security with MyProxy server
- Study on the complete genome of the *Plasmodium falciparum* for the identification of DHFR antigenic proteins

E-infrastructure shared between **E**urope and **L**atin **A**merica
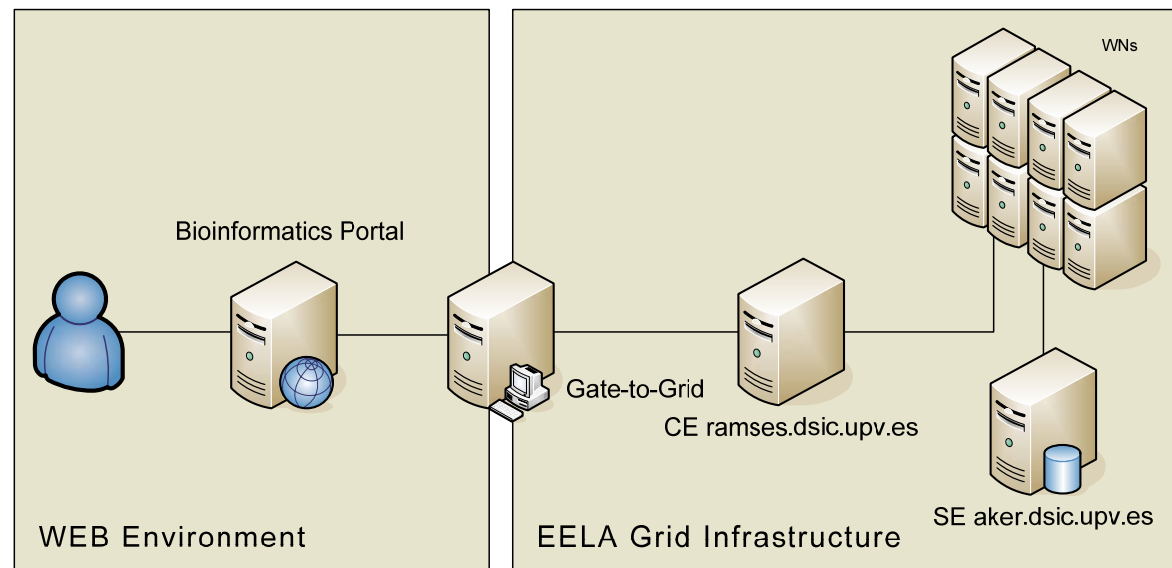
- **A Phylogeny is a Reconstruction of the Evolutionary History of a Group of Organisms.**

- *Applications:* **Gene Function Prediction, Drug Discovery and Conservation Biology.**

- **Interest of the LA Community**
    - ULA is Leading the Interest in the LA.
    - A Grid Service is being Developed to Run a Parallel Version of MrBayes from the Bioinformatics Portal.
- **Successful tests with a paralell MrBayes vesion in a Grid-service**

An Unrooted Phylogenetic Tree
of the Myosin Superfamily
Tony Hodge, MRC-LMB
Jamie Cope, UC Berkeley
July 2000

HEP

– Initial applications

- **ALICE**

  To study the physics of strongly interacting matter at extreme energy densities, where the formation of a new phase of matter, the quark-gluon plasma, is expected.

E-infrastructure shared between **E**urope and **L**atin **A**merica

MonALISA
MONitoring Agents using a Large
Integrated Services Architecture

Usage info ❓

		Job parameters			Application software				Details
Run# ▾	PID	Owner	Events	ROOT	ALIROOT	GEANT	Date	Output dir	
		podesta (15) ▾					last year ▾		
7	1711466	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/007/	test pro
6	1711465	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/006/	test pro
5	1711443	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/005/	test pro
4	1711433	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/004/	test pro
3	1711432	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/003/	test pro
3	1797427	podesta					25.01.2007 20:06	/alice/cern.ch/user/p/podesta/pp_900GeV/03/	MC simu
2	1711421	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/002/	test pro
2	1794911	podesta					25.01.2007 20:06	/alice/cern.ch/user/p/podesta/pp_900GeV/02/	MC simu
1	673395	podesta					06.09.2006 21:08	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/001/	test pro
1	1711410	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/001/	test pro
1	1794908	podesta		v5-13-02	v4-04-Rev-08	v1-6-1	25.01.2007 20:06	/alice/cern.ch/user/p/podesta/pp_900GeV/1/	MC simu
	1769912	podesta					23.01.2007 20:37	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/test/	test eve
	1771607	podesta					23.01.2007 21:36	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/test/	test eve
	1771617	podesta					23.01.2007 21:36	/alice/cern.ch/user/p/podesta/demo/	test eve
	1771691	podesta					23.01.2007 23:39	/alice/cern.ch/user/p/podesta/demo/	test eve
TOTAL : 15 jobs			0					**Export folders**	

Submission of ALICE Grid jobs from Latin America

– Initial applications

  ▪ **LHCb**

  To full investigate the CP violation in the Bd and Bs systems,
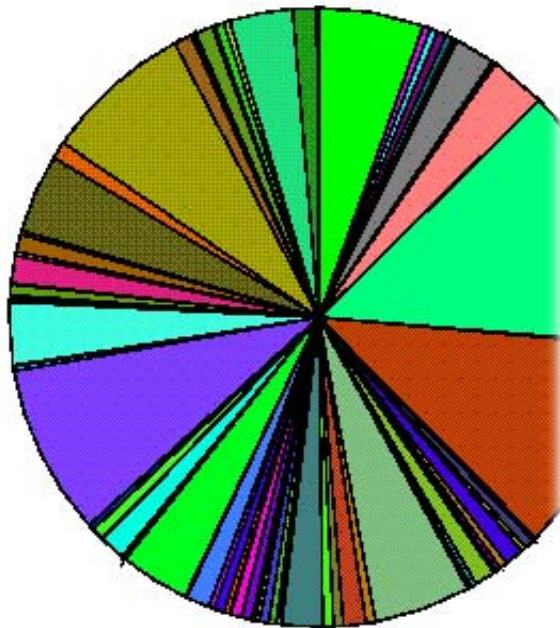 *to possibly renew the new physics beyond the standard model.*
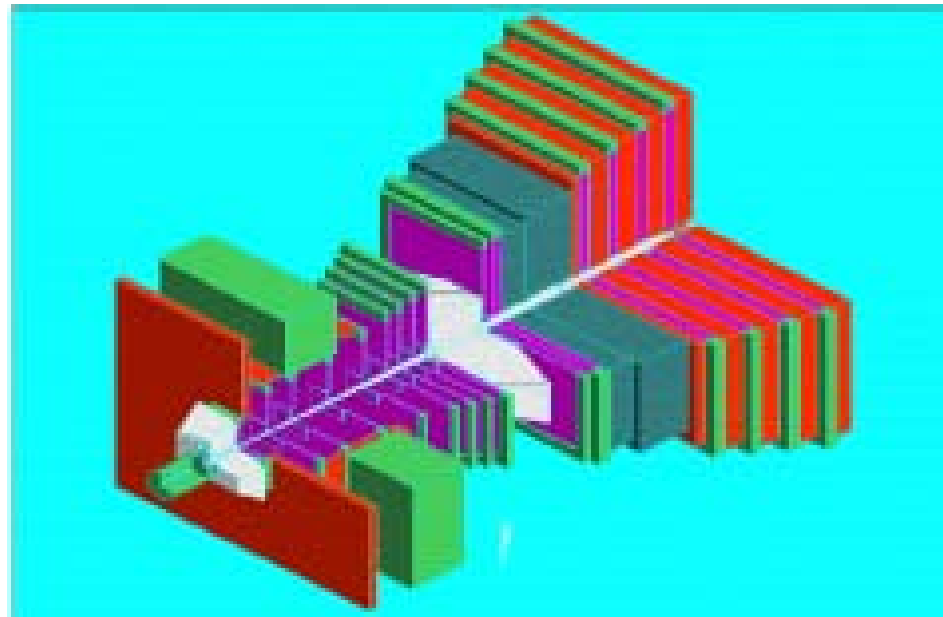
E-infrastructure shared between **E**urope and **L**atin **A**merica

CPU time 15773343 h

	%
DIRAC.LHCBONLINE.ch	0.00%
DIRAC.Lyon.fr	5.49%
DIRAC.NIPNE.ro	0.00%
DIRAC.UCD.ie	0.00%
DIRAC.Zurich-MH.ch	0.38%
DIRAC.Zurich.ch	0.41%
DIRAC.joel.ch	0.00%
LCG.ACAD.bg	0.47%
LCG.AUVER.fr	0.36%
LCG.BHAM-HEP.uk	0.15%
LCG.BIFI.es	0.11%
LCG.Barcelona.es	2.13%
LCG.Bari.it	0.19%
LCG.Bologna.it	0.01%
LCG.Brunel.uk	2.93%
LCG.CERN.ch	13.67%
LCG.CESGA.es	0.00%
LCG.CGG.fr	0.04%
LCG.CNAF-GRIDIT.it	0.06%
LCG.CNAF.it	11.21%
LCG.CNB.es	0.20%
LCG.CPPM.fr	0.52%
LCG.CSCS.ch	0.26%
LCG.CY01.cy	0.03%
LCG.Cagliari.it	0.00%
LCG.Cambridge.uk	0.00%
LCG.Catania.it	0.01%
LCG.Dortmund.de	0.98%
LCG.Durham.uk	0.40%
LCG.EELA-CIEMAT.es	0.25%
LCG.EELA-UFRJ.br	0.01%
LCG.ETF-RTH.lv	0.11%
LCG.Edinburgh.uk	0.06%
LCG.FESB.hr	0.00%

	%
LCG.FORTH.gr	1.06%
LCG.Ferrara.it	0.02%
LCG.Firenze.it	0.26%
LCG.GOG.sg	0.05%
LCG.GR-01.gr	0.04%
LCG.GR-03.gr	0.07%
LCG.GR-04.gr	0.03%
LCG.GR-05.gr	0.02%
LCG.GRIDKA.de	5.06%
LCG.GRNET.gr	0.50%
LCG.Glasgow.uk	1.17%
LCG.HG-02.gr	0.51%
LCG.HG-04.gr	0.56%
LCG.HG-06.gr	2.13%
LCG.HPC2N.se	0.05%
LCG.HellasGrid.gr	0.19%
LCG.ICI.ro	0.05%
LCG.IHEP.su	0.38%
LCG.IN2P3.fr	0.47%
LCG.INR.ru	0.22%
LCG.IPP.bg	0.06%
LCG.IPSL-IPGP.fr	0.04%
LCG.IRB.hr	0.10%
LCG.ITEP.ru	0.51%
LCG.ITWM.de	0.03%
LCG.Imperial.uk	0.55%
LCG.Iowa.us	0.00%
LCG.JINR.ru	0.09%
LCG.KFKI.hu	0.29%
LCG.KIAE.ru	0.52%
LCG.KIAM.ru	0.31%
LCG.Krakow.pl	1.22%
LCG.LAL.fr	0.08%
LCG.LAPP.fr	0.01%
LCG.LISA.nl	0.00%
LCG.LPC.fr	3.64%
LCG.LPN-fails.fr	0.03%
LCG.LPN.fr	0.19%
LCG.Lancashire.uk	1.34%

	%
LCG.LeSC.uk	0.04%
LCG.Legnaro.it	0.66%
LCG.Liverpool.uk	0.29%
LCG.Manchester.uk	8.93%
LCG.Milano.it	0.24%
LCG.NCP.pk	0.01%
LCG.NCU.tw	0.01%
LCG.NIKHEF-save.nl	0.04%
LCG.NIKHEF.nl	3.28%
LCG.NIPNE.ro	0.13%
LCG.Napoli-Atlas.it	0.07%
LCG.Napoli.it	0.08%
LCG.OU.il	0.06%
LCG.Oxford.uk	0.56%
LCG.PAKGRID.pk	0.00%
LCG.PDC.se	0.06%
LCG.PIC.es	1.43%
LCG.PNPI.ru	0.21%
LCG.Padova.it	0.90%
LCG.Pisa.it	0.20%
LCG.QMUL.uk	4.11%
LCG.RAL-HEP.uk	0.91%
LCG.RAL.uk	7.75%
LCG.RHUL.uk	0.94%
LCG.SARA.nl	0.11%
LCG.SINP.ru	0.16%
LCG.SRCE.hr	0.00%
LCG.Sheffield.uk	0.90%
LCG.Sofia.bg	0.02%
LCG.TAU.il	0.10%
LCG.TCD.ie	0.02%
LCG.Torino.it	0.49%
LCG.UCL-CCC.uk	0.00%
LCG.ULAKBIM.tr	0.32%
LCG.USC.es	3.30%
LCG.WARSAW.pl	1.25%
LCG.WCSS.pl	0.04%
LCG.WEIZMANN.il	0.05%

- **Applications of interest to EELA partners and other communities in Latin America**
  - Other LHC applications
    - ATLAS
      - *CIEMAT, UFRJ, UNLP and UTFSM*
    - CMS
      - *CIEMAT and LA participants in EELA member countries, but no EELA partner from LA…*
  - New projects
    - Pierre Auger Observatory
      - *INFN-Catania, LIP, UFRJ, UNAM, and UNLP*

e-Learning

E-infrastructure shared between **E**urope and **L**atin **A**merica

- **Storage servers are running on the storage elements**
- **Multimedia server continues to manage requests**
- **Video streams continues to be sent to clients (UDP)**
- **Backup of classes stored on the grid**

- **Planed Services**
  – Access to distributed computer enhanced instrumentation
  – Remote access to simulation and modelation capabilities with high performance computing support
  – Interactive visualization
  – Distributed data analysis with access to data base systems
  – Experiment repository system.

| Food Engineering | NMR |

| Electrophoresis | cns_solve |

-**Based on Moodle**

-**Course management system to produce web-based courses that support a social constructionist framework of education**
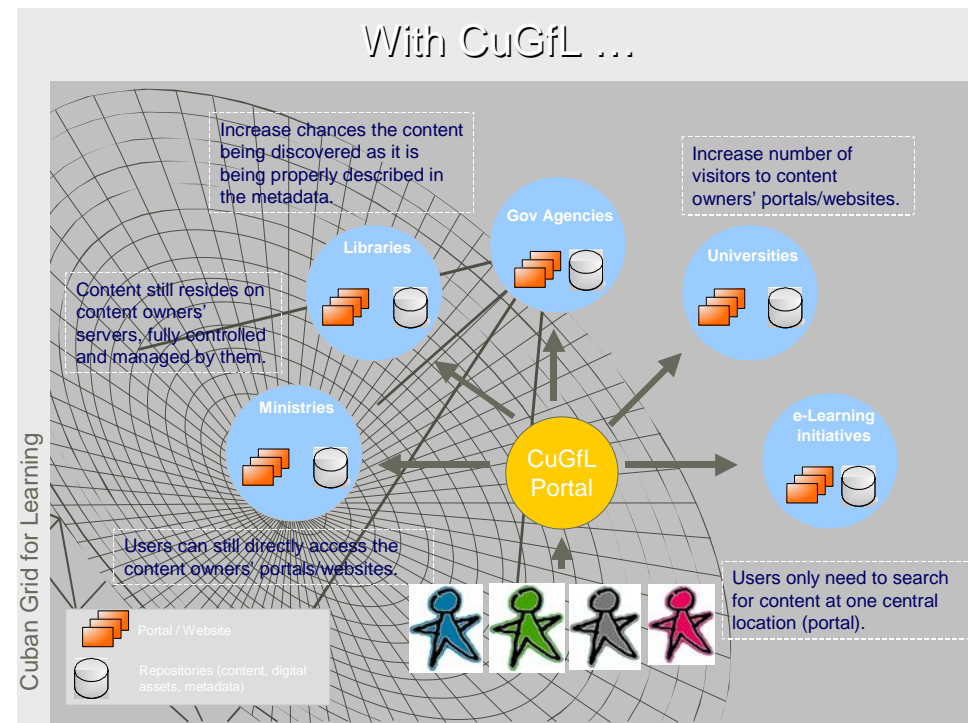
SATyrus:

- SATisfiability-based, neuro-symbolic architecture;

- Exact formulation synthesizer.

- Airport Management (surface)

Target Problem, e.g., colouring
Constraints Modeling
SATyrus compiler
Energy Function
Optimizer, e.g., stochastic HONNs
Global/Local optima

E-infrastructure shared between **E**urope and **L**atin **A**merica
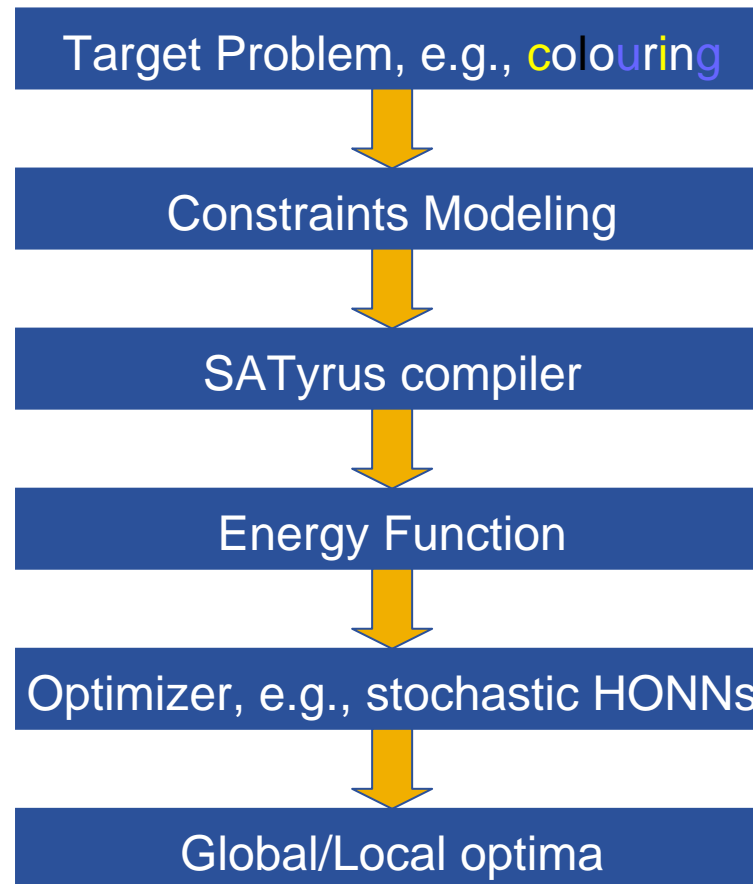
- **Use of Inductive Logic Programming to extract relevant knowledge from structured data**

Observations

Model: first order clauses

Background knowledge

Language bias

Constraints
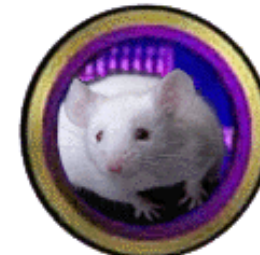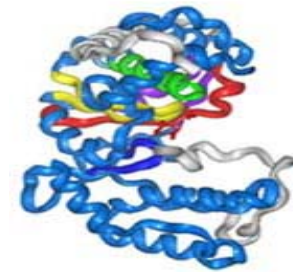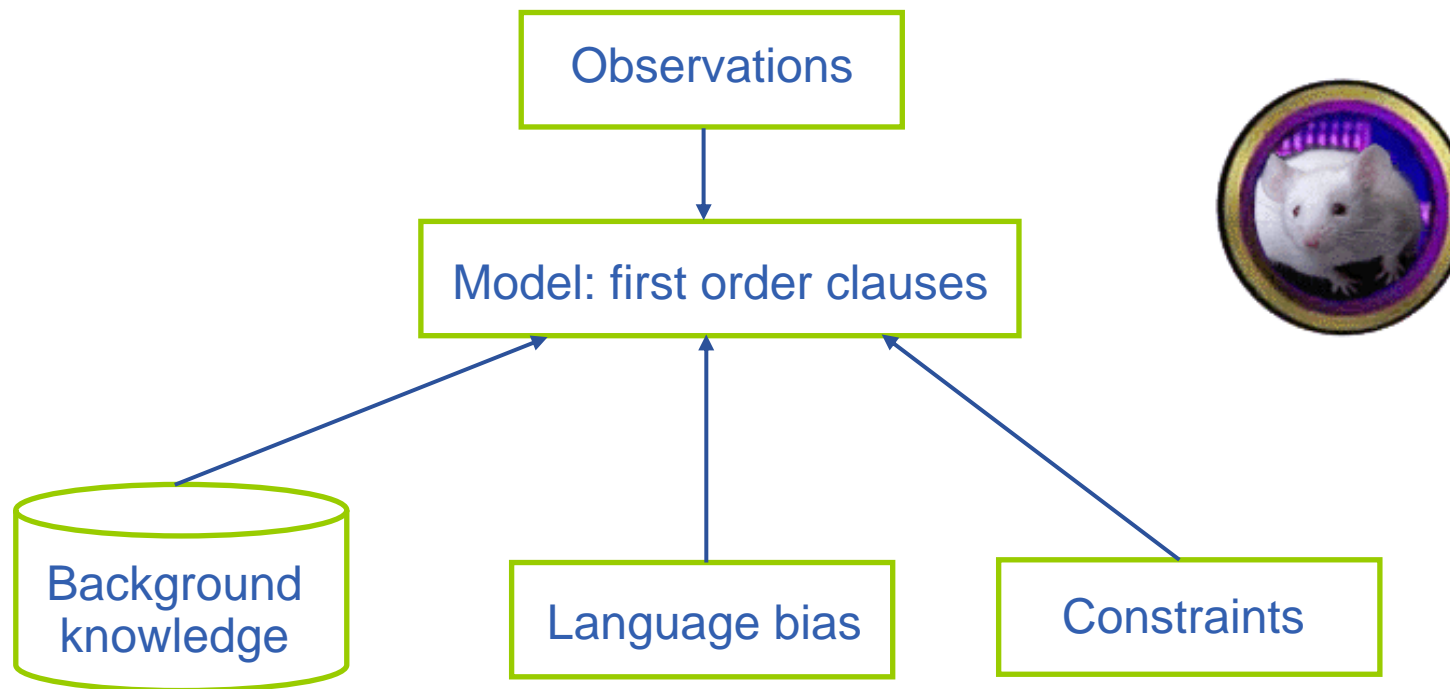
Climate

- **GOAL: Predict local impacts of "El Niño" in Latin America**

 A challenging problem for the climate community, with **huge** socio-economical impact in Latin America.
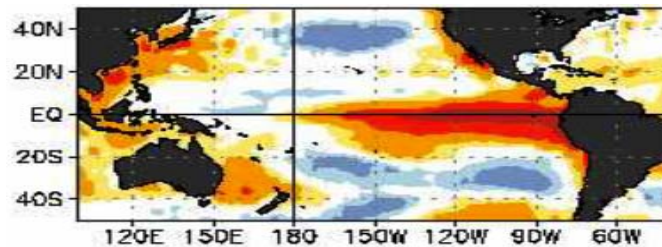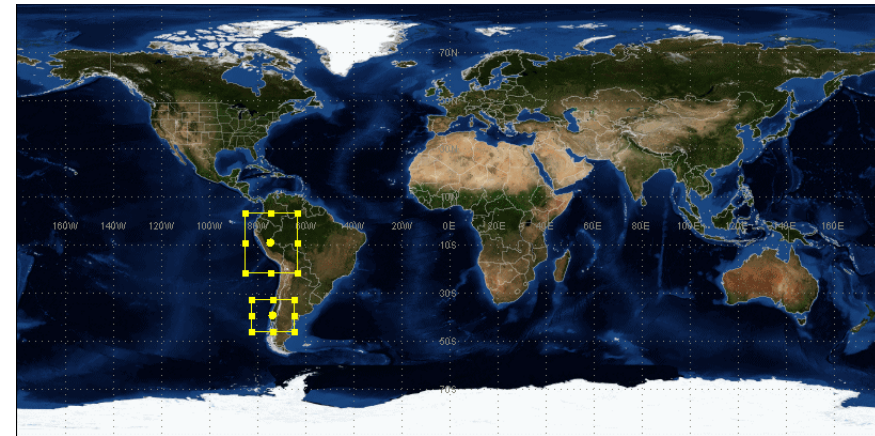
  Anomalous heating   Anomalous cooling
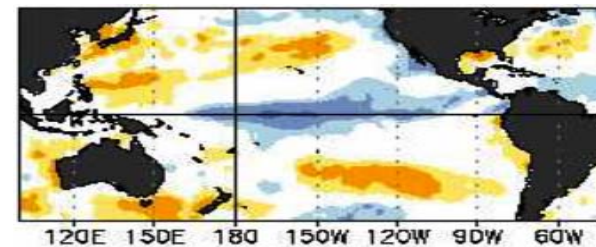
- **GRID helps to** share computing resources, heterogeneous data, as well as know-how in a user-friendly form.

- **A new integrated climate application** developed in EELA from the scratch, with no similar counterpart in any other Earth Science/Climate EU Project.
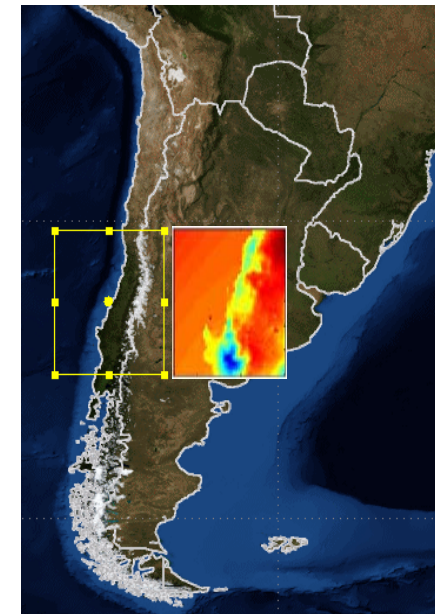
WRF. Regional Simulation

It runs in any site of the production testbed:
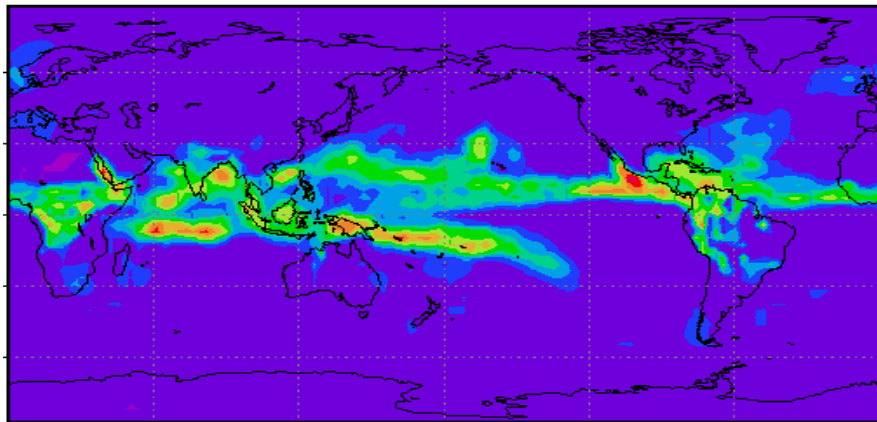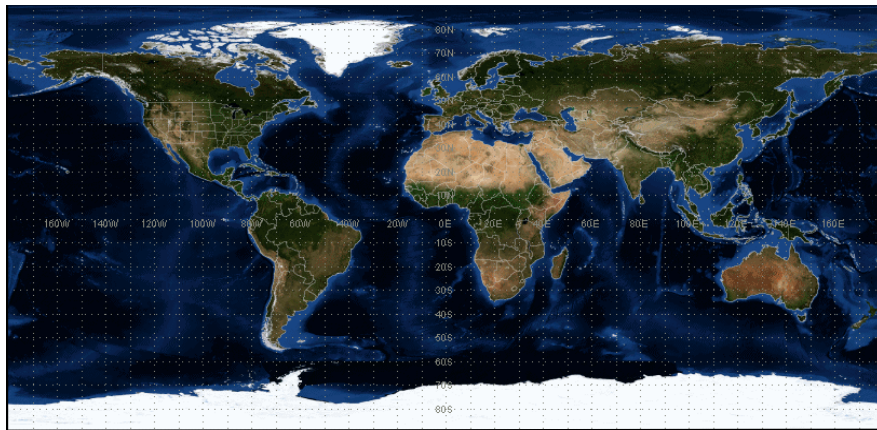
UFRJ

CAM. Global simulation

Requires metadata catalog so it runs

In: UC and CIEMAT (EU), UFRJ (LA).

**E**-infrastructure shared between **E**urope and **L**atin **A**merica

Three applications have been identified (climate sequence):

- Global atmospheric circulation model (**CAM**) .... *Deployed !*
- Regional weather model (**WRF**) .... *Deployed !*
- Datamining clustering tools (SOM) .... *Work in progress*

This sequence poses several computational challenges

Nontrivial dependent relationships among the applications.
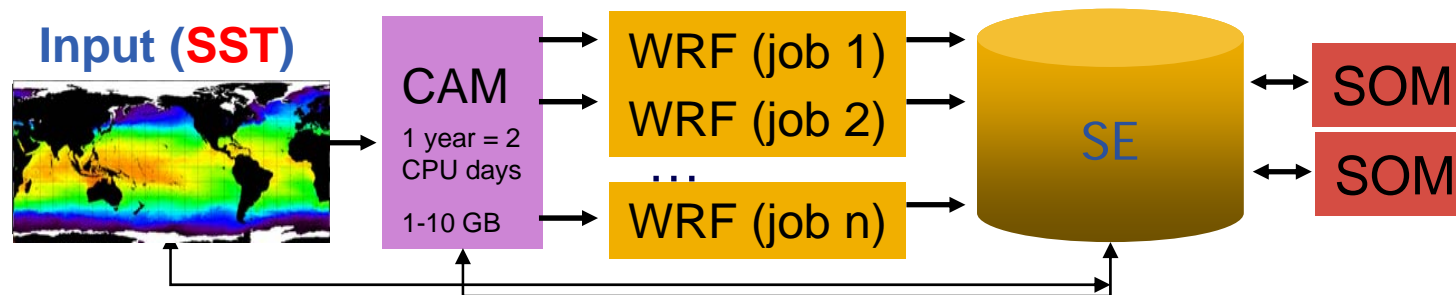
This sequence of jobs demands middleware solutions for:

- Preparing and submitting dependent jobs / data sharing (workflow).
- Restarting interrupted experiments.
- Manage metadata (for datasets and application status).

UI

Resource
Broker

1 **supersubmiter.sh**

Generate
CAM.jdl **2**

3

Submit
CAM.jdl

Insert entry in
**CAM collection
AMGA**

WN

upload_info_CAM.sh

Insert entry in
**WRF collection
AMGA**

change_status_CAM.sh

CAM

Update status
Insert runon
**CAM collection
AMGA**

checkpoint_CAM.sh

update_history_CAM.sh

Insert entry in
**CHECKPOINT
collection
AMGA**

Insert entry in
**HISTORYCAM
collection
AMGA**

- Using GENIUS to interact with the applications (CAM+WRF)
- In the future a climate specific portal will be developed (JSR168) to run and track scientific experiments.

The second year will focus on a scientific challenge:

High resolution regional simulations for El Niño and La Niña 1982-1983 and 1997-98 strong events. Comparison with historical local data, including sensitivity studies of SST.
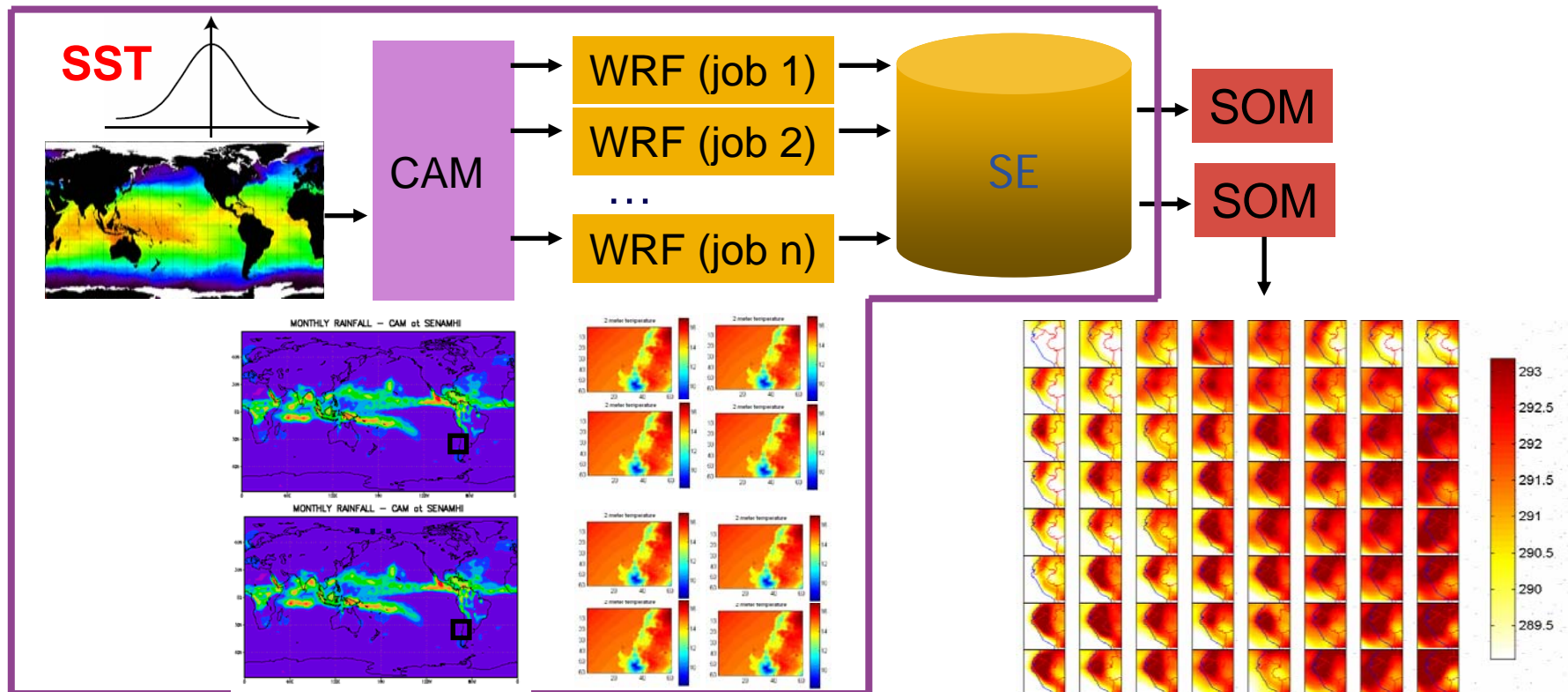
Other Applications

Volcano Sonifications

- **Currently no definitive method to predict the eruption of a volcano has been discovered or implemented (yet)**

- **Some of the calculations have been performed in the EELA e-Infrastructure.**

E-infrastructure shared between **E**urope and **L**atin **A**merica

T3.3 Applications have been ported to the Grid in

EGRIS-1 Itacuruça (Brazil) 04-16/12/2006

But also new ones:

EMBOSS

SegHidro

Distributed Simulation of Multiple Failure Events on Optical Networks

GENECODIS

EGRIS-2 Mérida (Venezuela)

30/07-10/08/2007

THE CALL IS OPEN!

http://indico.eu-eela.org/conferenceDisplay.py?confId=90

Or just look for "egris-2" in Google…

E-infrastructure shared between Europe and Latin America

WP3 web page:

http://www.eu-eela.org/eela_wp3.php

WP3 documents:

http://documents.eu-eela.org

E-infrastructure shared between Europe and Latin America

WP3 contacts

Vicente Hernández (Biomed)

    vhernand@dsic.upv.es

Lukas Nellen (HEP)

    lukas@nucleares.unam.mx

Inês Dutra (e-Learning)

    ines@dcc.fc.up.pt

José Manuel Gutiérrez (Climate)

    manuel.gutierrez@unican.es

Rafael Mayo

    rafael.mayo@ciemat.es

Thanks for your attention!