Running a successful VO

Status of the GridUnesp project

Rogério L. Iope
São Paulo State University (Unesp)
Center for Scientific Computing

Open Science Grid - 2011 All Hands Meeting
Harvard Medical School

March 07, 2011

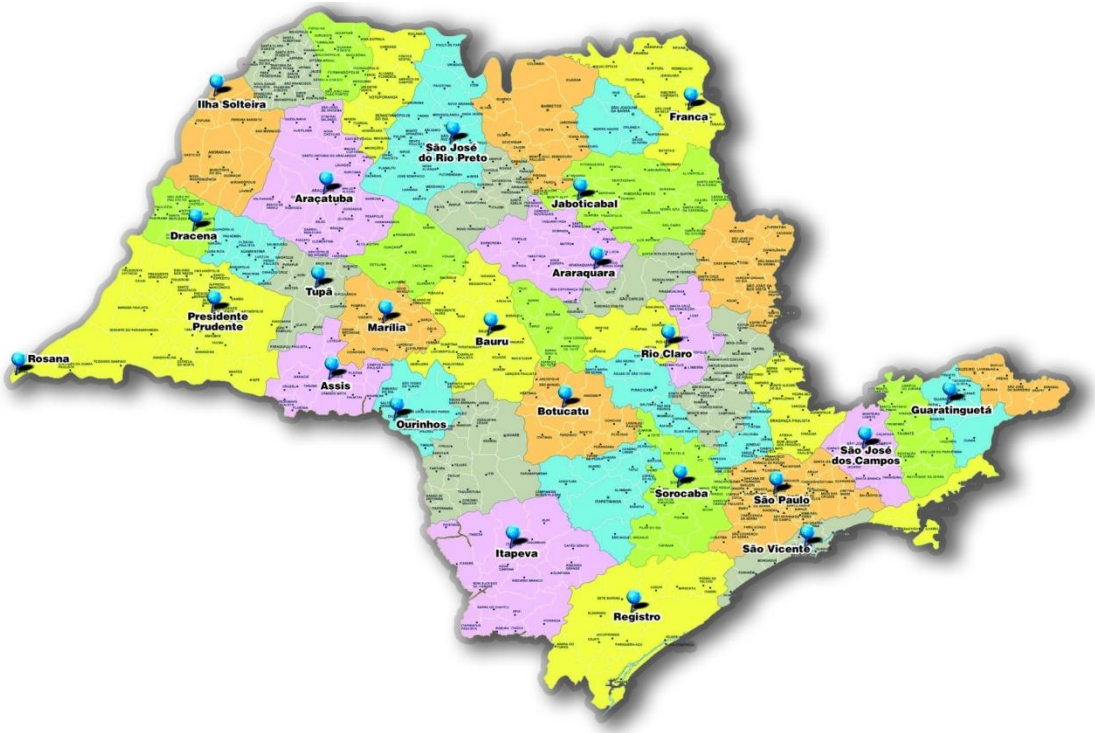
Open Science Grid

The São Paulo State University

33 institutes, colleges and schools on 23 campuses distributed throughout the State of São Paulo

Second largest university in Brazil; part of the state of São Paulo public higher education system

unesp
UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"



GridUnesp scope and evolution

A distributed computational system for scientific computing with resources dispersed over a wide geographic area

A production-quality, multi-campus Grid → one of the largest campus Grid infrastructures in Latin America

Started in 2004 with a call for Scientific Proposals sent to UNESP researchers

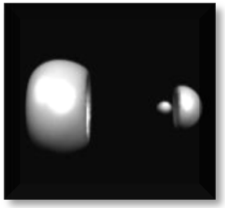
Formal proposal submitted to FINEP by the end of 2005 and approved by the middle of 2006

Hardware acquisition process finished by the end of 2008; operations started in 2009

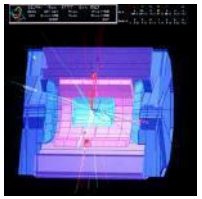
gridunesp

GridUnesp main research areas

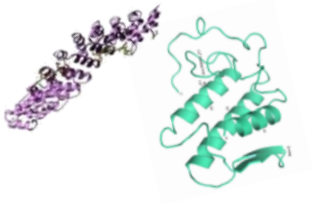
Biological Networks



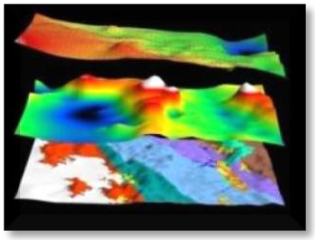
Relativistic Chemistry



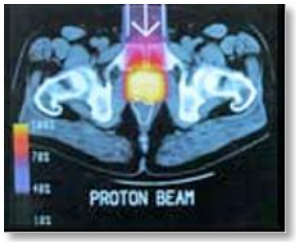
High Energy Physics



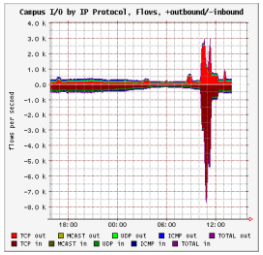
Molecular Dynamics



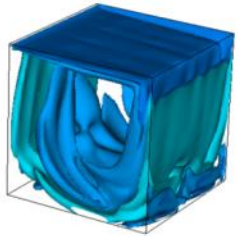
Geological Modeling



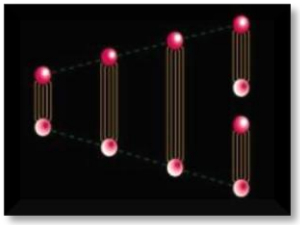
Medical Physics



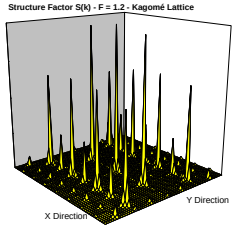
Network Security



Turbulence

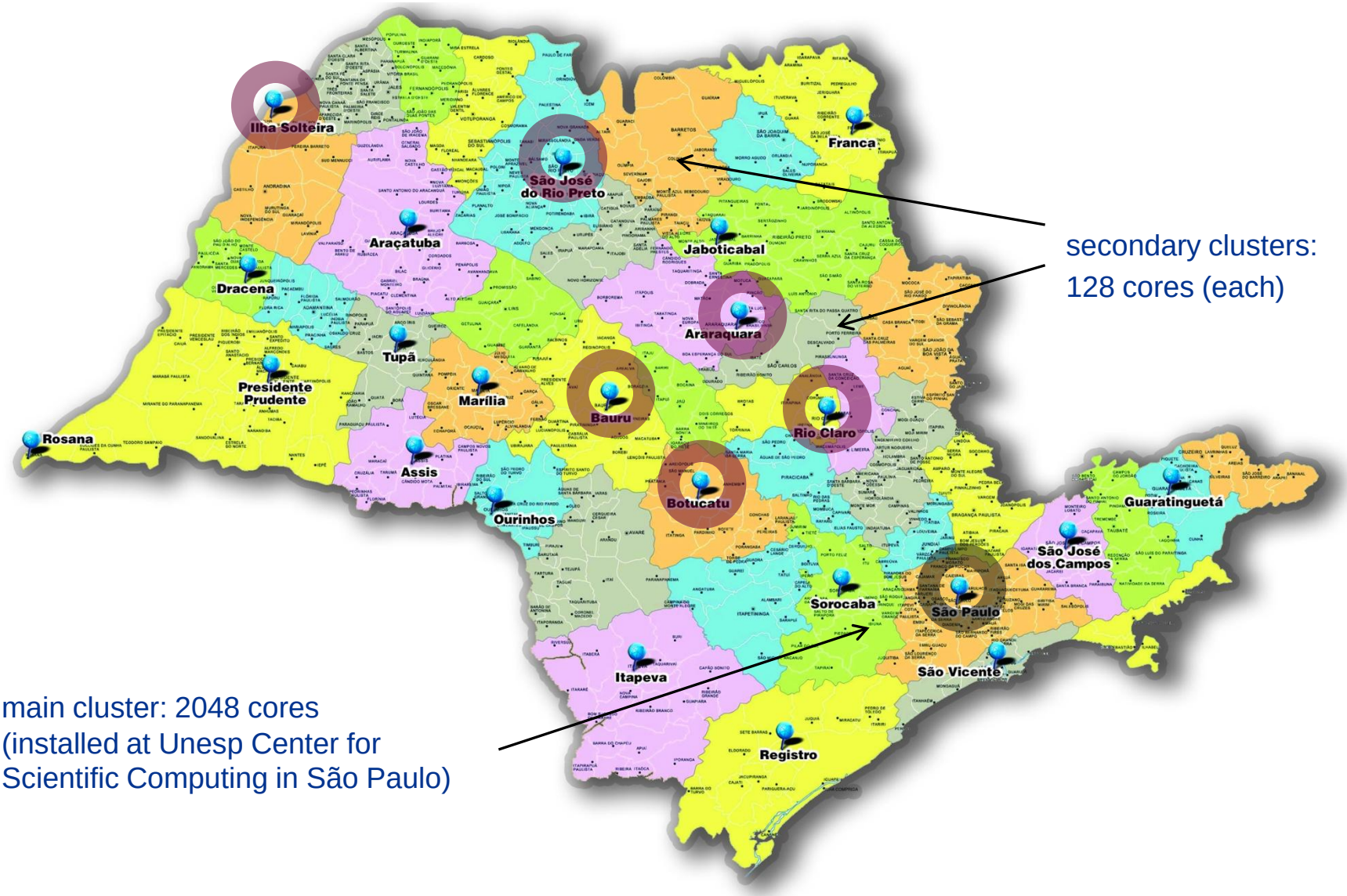


Lattice QCD



High Tc Superconductivity

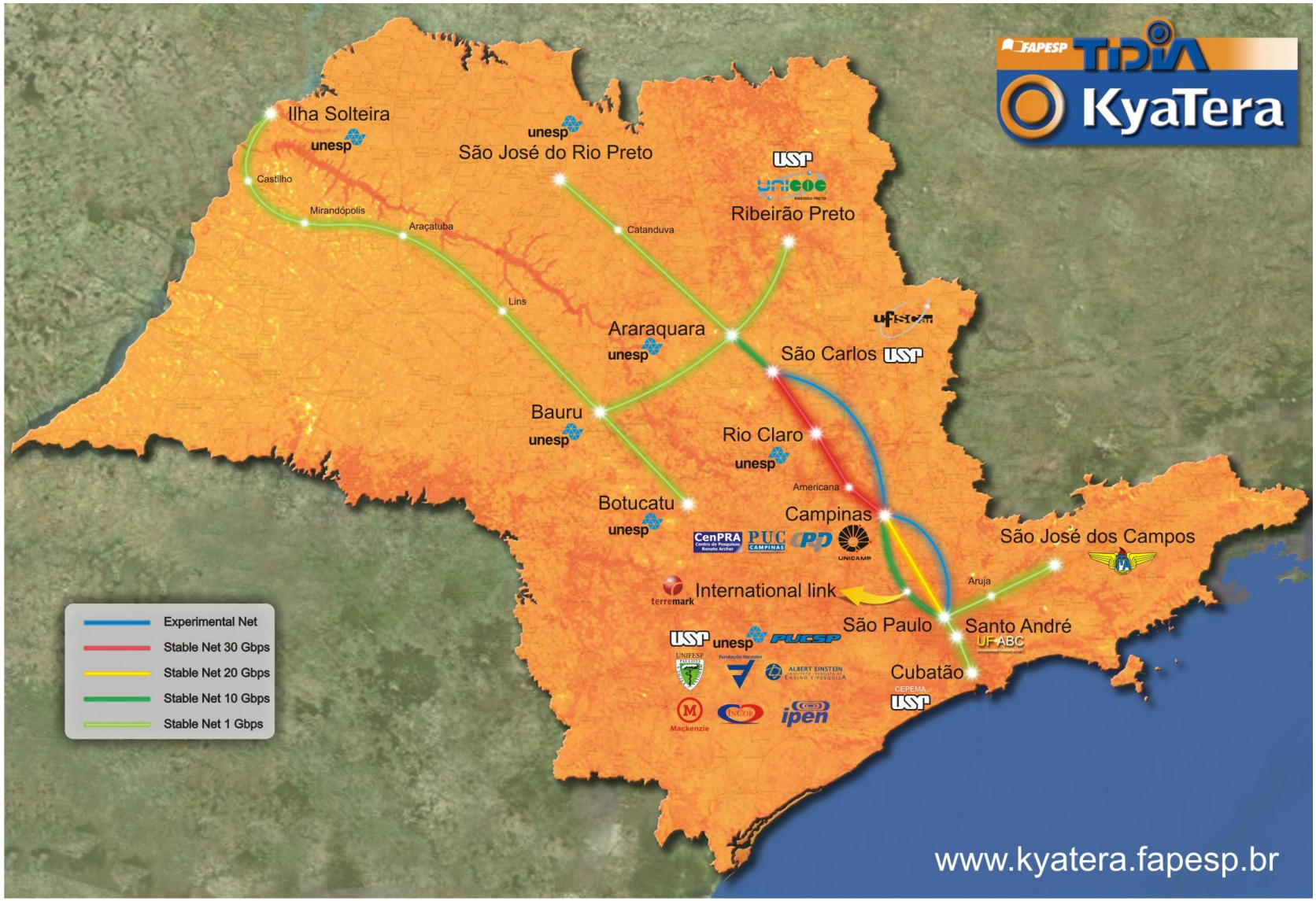
GridUNESP sites



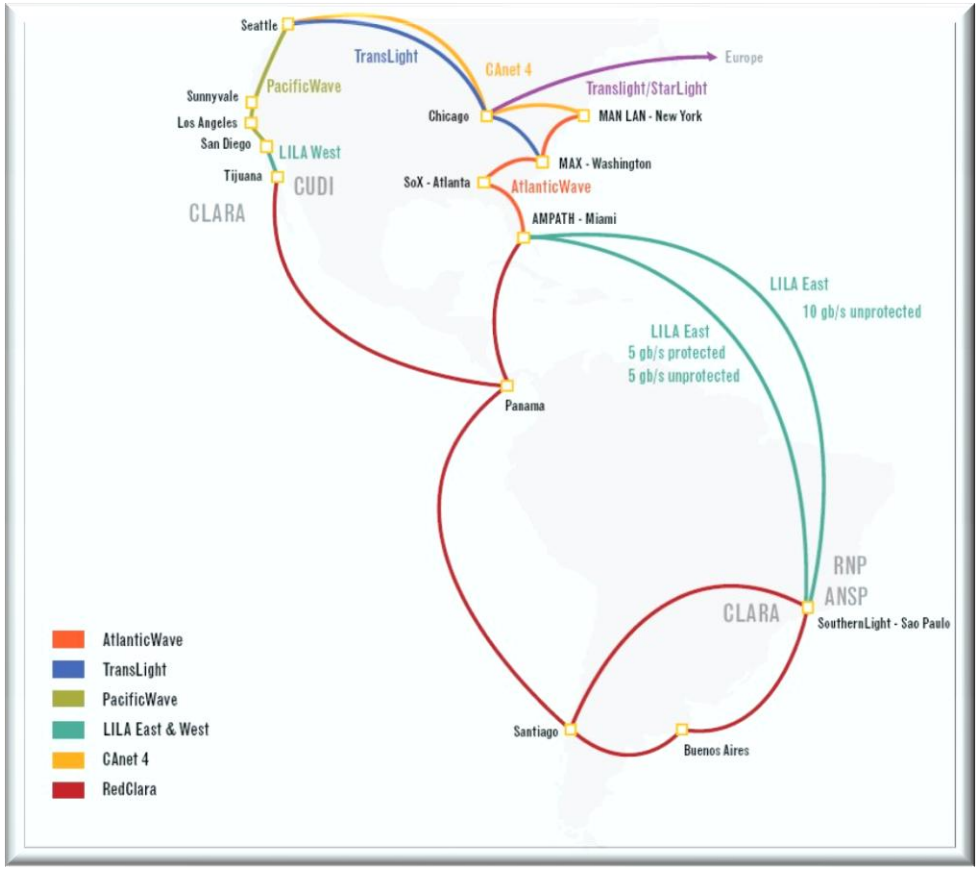
secondary clusters:
128 cores (each)

main cluster: 2048 cores
(installed at Unesp Center for
Scientific Computing in São Paulo)

Connectivity between sites: the KyaTera network



International connectivity



Hardware & software infrastructure

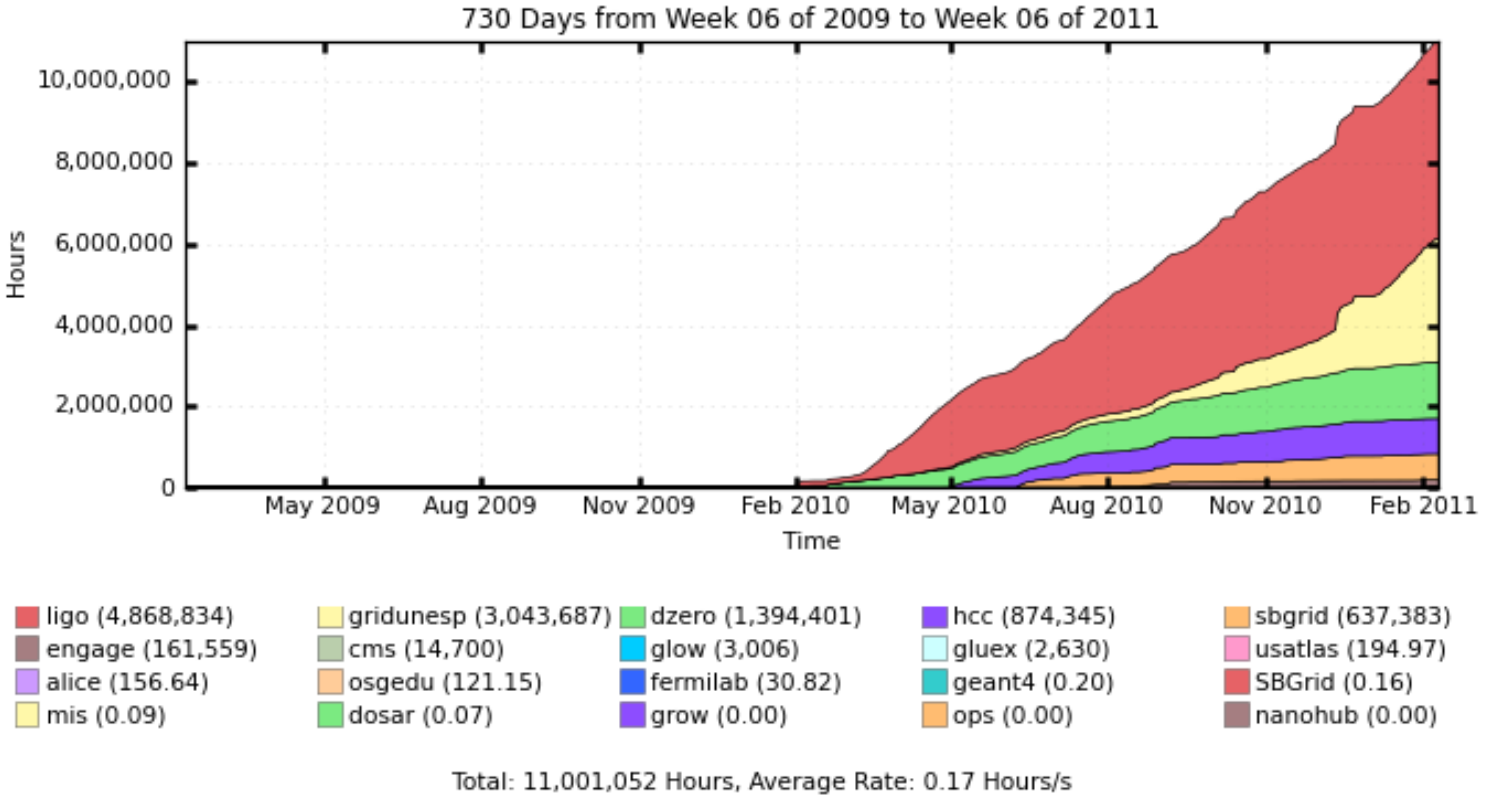
- **Main cluster**
  - 2048 processing cores, 4 data servers (24 TB each), SAN storage system (36 TB), InfiniBand and gigabit switches

- **Secondary clusters (7)**
  - 128 processing cores, SAN storage system (6 TB), gigabit switch

- **Software / middleware**
  - Intel software: C/C++/Fortran compilers, VTune, MKL, MPI, Cluster Toolkit
  - Grid middleware: Open Science Grid

→ formal agreement w/ OSG Executive Board
→ establishment of the ´GridUNESP VO´
→ GridUnesp is the first OSG VO outside U.S.

GridUnesp site usage

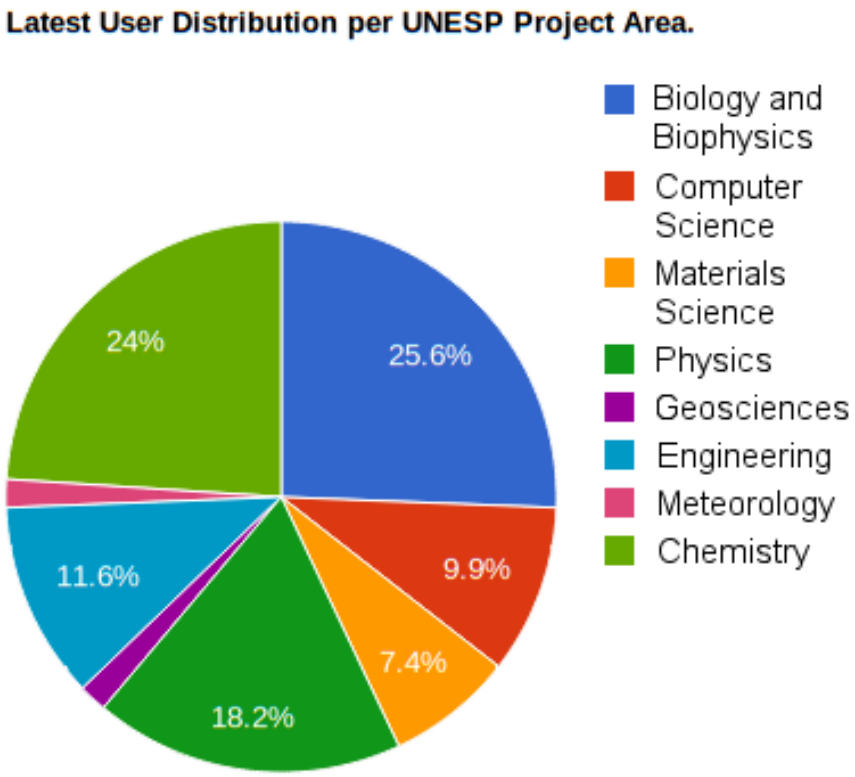


Main cluster usage approaching 12 million CPU hours
(usage is coming closer to its maximum capacity - 90% in Feb)

Projects and users registration progress

GridUnesp is a multipurpose VO: 29 projects from different scientific disciplines and around 170 users (researchers and students)

User distribution per scientific disciplines



Extensive work has been done on assisting users on porting and adapting cluster applications to the Grid environment

GridUnesp and the São Paulo State Grid initiative

GridUnesp - Complex user requirements

Most of GridUnesp users come from the HPC community

→ efforts centered at transforming their needs into grid-friendly workflows
→ we successfully run typical HPC applications like Gromacs, Gaussian, Gamess, MrBayes, etc, with MPI support

Local applications have complex and challenging requirements

→ support for MPI, multithreaded (SMP) jobs, large memory jobs and long run jobs (> 20 days)
→ main cluster batch scheduler switched to Torque/Maui to better support long run queues and complex job requirements

Conclusions and perspectives

One vision for the future of HPC: a Grid of high-performance clusters interconnected by advanced optical metro/wide area networks - GridUNESP project designed with this vision in mind

→ it is providing the means for researchers to work on scientific projects that demand a great deal of computing power

Hardware and network infrastructure fully under our control

Building a multidisciplinary infrastructure is challenging

Next steps include:

→ integration of the inter-cluster WAN infrastructure
→ deployment of a Certificate Authority for the state of São Paulo
→ training courses and workshops for researchers & developers

References

GridUnesp and OSG Agreement

http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=827

R. L. Iope, S. F. Novaes, *GridUnesp and the São Paulo State Grid initiative*

http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1020

J. Caballero, H. Severini, R. Gardner, *Employing Open Science Grid to support National Grid Initiatives in South America and South Africa*, CHEP 2010 Conference Proceedings

http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1022

References

R. L. Iope, *The São Paulo State University Campus Grid Initi*ative, in Proceedings of the Second EELA-2 Conference, eds. R. Mayo et al (CIEMAT 2009), 25-27 November 2009, Choroní, Venezuela, pp. 257-266.

https://twiki.grid.iu.edu/twiki/pub/VirtualOrganizations/DOSAR_GridUNESP_OSG/GridUNESP_1st_paper.pdf

Conference Proceedings available at

http://www.ciemat.es/recursos/doc/Redes_Cientifico_Tecnicas/422103760_17112009151715.pdf

R. L. Iope, N. Lemke, G. A. von Winckler, *GridUNESP: a multi-campus Grid infrastructure for scientific computing*, in Proceedings of the 3rd Latin American Conference on High Performance Computing (CLCAR 2010), 25-28 August 2010, Gramado, Brazil, pp. 76-84.

Slides available at

http://gppd.inf.ufrgs.br/CLCAR2010files/slides/32.pdf