

# **Planning a statewide Grid infrastructure for leveraging e-Science in São Paulo, Brazil**

**Eduardo Bach**

**São Paulo State University (UNESP)**

**Center for Scientific Computing**

OSG All-Hands Meeting  
Lincoln, Nebraska – 2012

# The SPRACE project

- São Paulo Research and Analysis Center
  - 2004 – Fermilab/DZero member
  - 2005 – OSG site: several VOs
  - 2006 – CERN/CMS member
  - 2009 – MoU with WLCG signed
- BR-SP-SPRACE is the only official WLCG Tier-2 in Latin America (so far)
- It is considered one of the most reliable Tier-2's of the world (see for instance: <http://tinyurl.com/BR-SP-SPRACE-11-02>)

# SPRACE evolution

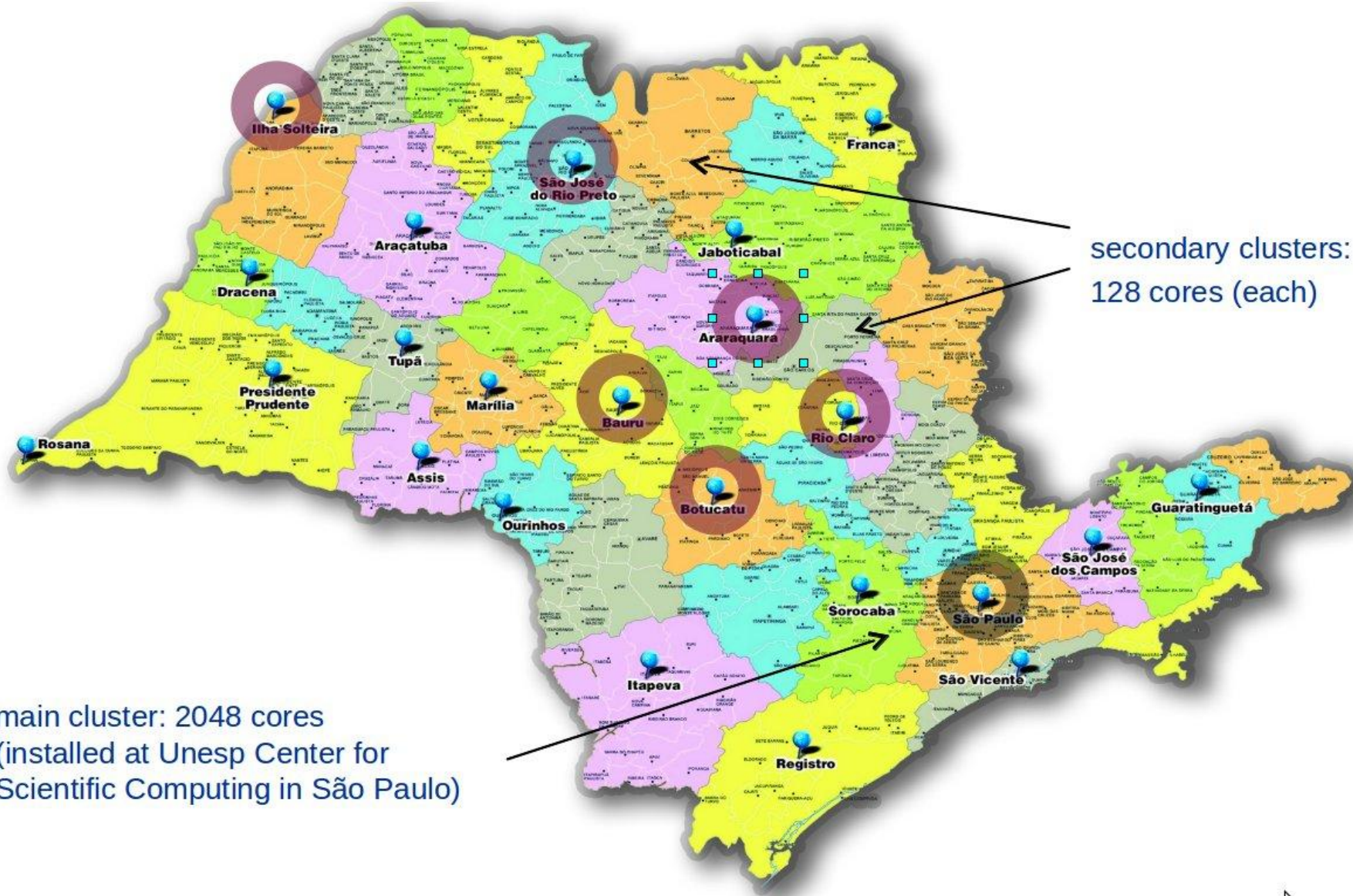
- In the beginning (2004)
  - 44 cores
  - 4 TB storage
- Last upgrade (2012)
  - 1088 cores
  - 1PB storage
- Spin-off: deployment of the first Campus Grid in Latin America (GridUNESP) that operates in close association with OSG

# A multi-campus Grid

- GridUnesp: 8 clusters (main cluster + 7 secondary ones)
- Total Hardware
  - 2944 cores
  - 200 TB storage
- Computing and storage resources spread over the state of São Paulo



# GridUnesp sites



# GridUnesp uses KyaTera Network



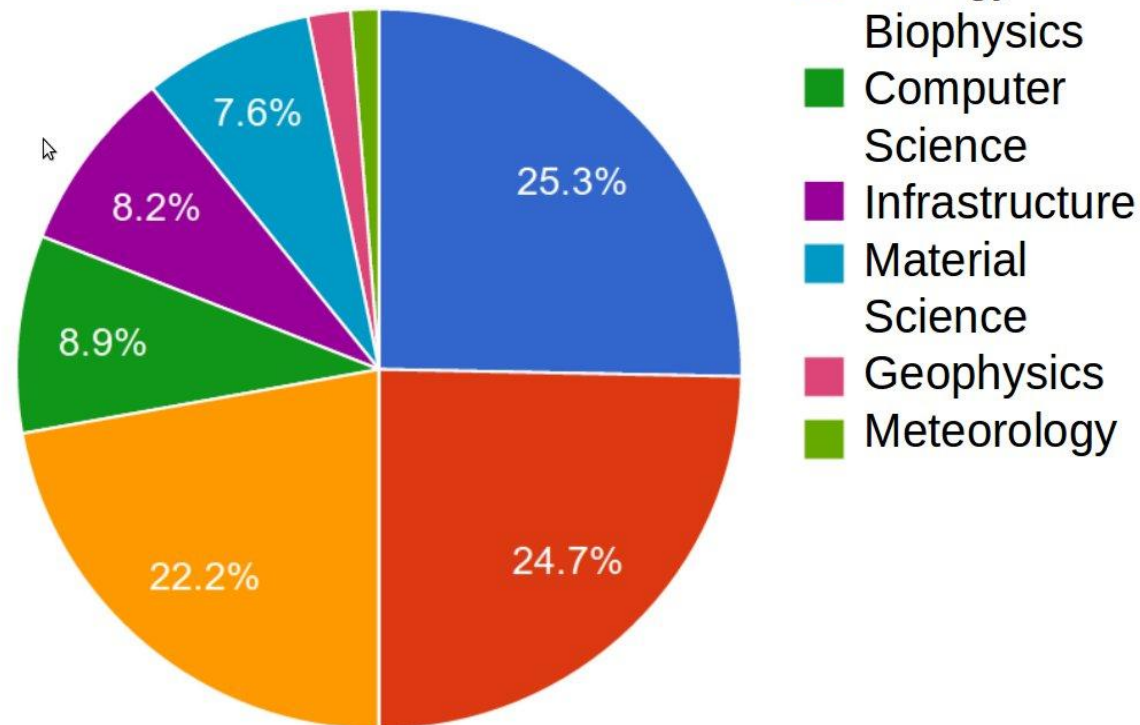
# The KyaTera Network

- A widely distributed dark-fiber-based optical network infrastructure that reaches the main research laboratories in the state of São Paulo
- Built to support all kinds of research that depend on high speed network
- DWDM network over dark fibers
- Upgrade capability (2.5G, 10G, 40G) according to the researches' needs
- GridUnesp is now the main user

# GridUnesp research areas

- 37 research groups
- 180 users

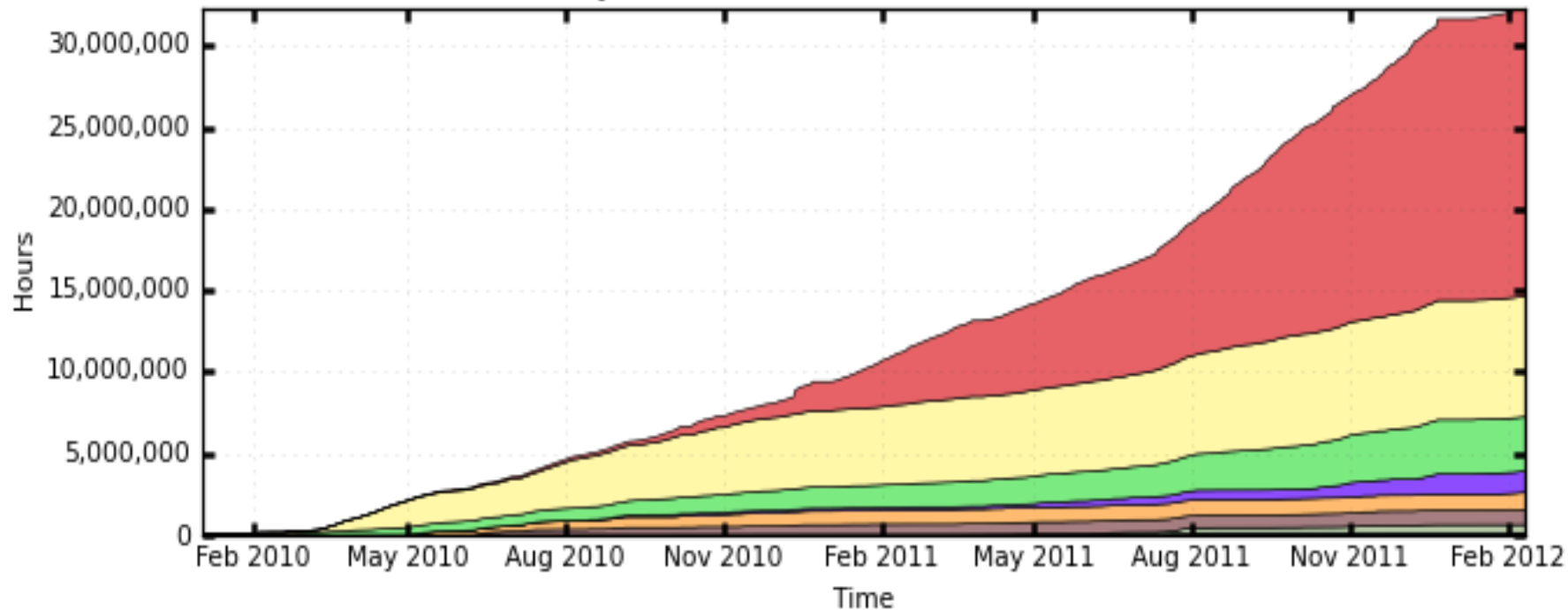
User distribution across Research areas.





# 32+ million CPU hours provided to OSG

770 Days from Week 00 of 2010 to Week 06 of 2012



gridunesp (17,671,472)	ligo (7,392,860)	dzero (3,333,229)	engage (1,300,730)	hcc (1,072,219)
sbgrid (901,252)	glow (502,327)	gluex (107,205)	cms (14,700)	osgedu (346.29)
usatlas (340.78)	alice (193.82)	osg (79.58)	fermilab (30.82)	nebiogrid (21.05)
suragrid (6.38)	geant4 (1.70)	mis (0.09)	dosar (0.07)	Other (0.00)

Total: 32,297,019 Hours, Average Rate: 0.48 Hours/s

# First conclusions

- Our team has acquired a good know-how on handling HPC / DHTC jobs
- We have already covered a broad range of research areas
- We have a good network infrastructure spread over the state of São Paulo

# Launching a new endeavor

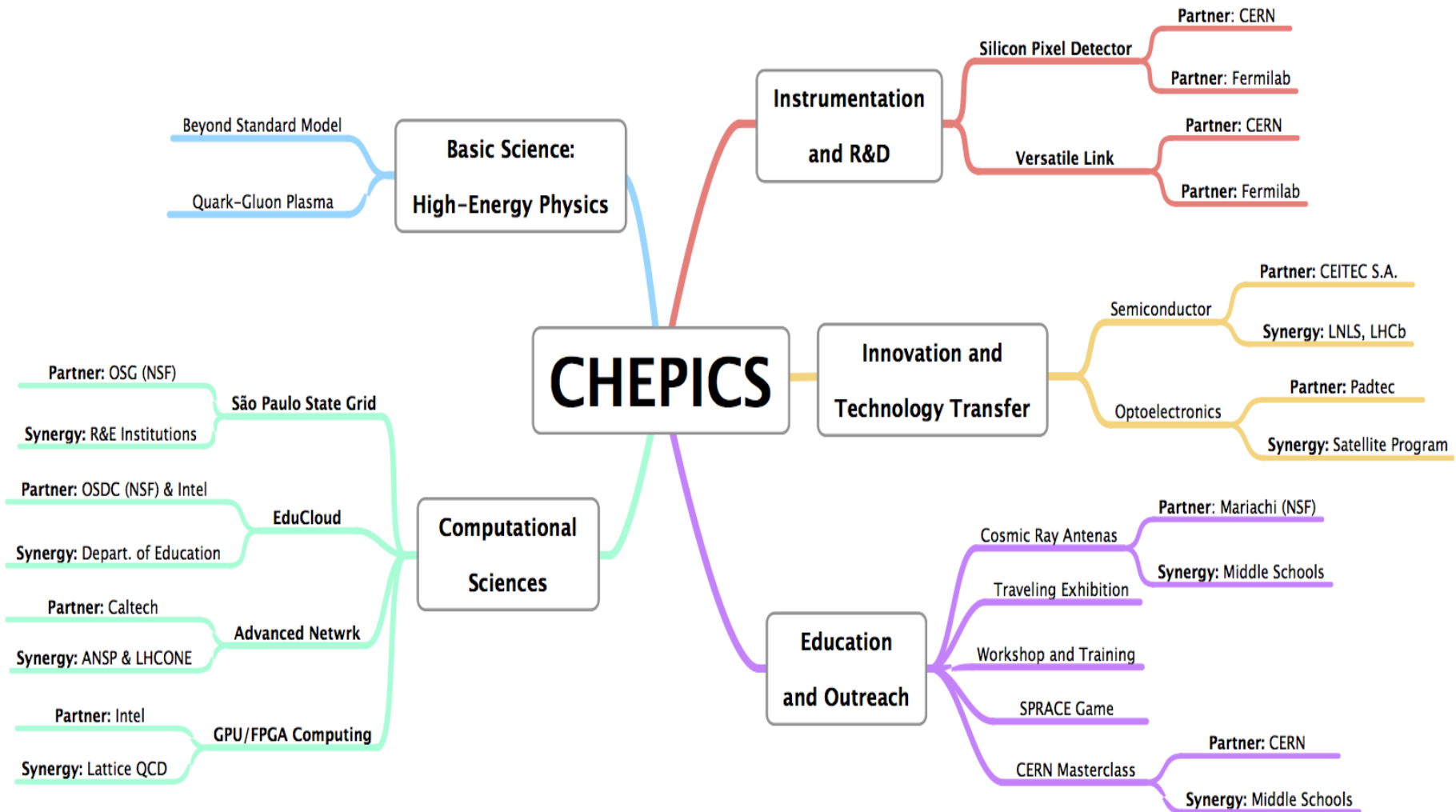


- Center for High Energy Physics, Instrumentation and Computational Sciences (CHEPICS)

# CHEPICS' goals

- To establish a research, innovation and dissemination center focused on high-energy physics, scientific instrumentation and computational sciences in order to leverage Brazil's participation in the CERN experiments, empowering the Brazilian groups to meet the challenges of the LHC upgrade

# Research branches



# Computational Sciences

- **São Paulo State Grid**
- Cloud Computing - EduCloud
- Advanced networking
- Hybrid (GPU/FPGA) computing

# Research at São Paulo

- We estimate that ~1800 projects bought computing power last year (from small to big power)
- Only a few research centers have appropriate physical infrastructure and human resources
- Very few large projects: lack of know-how
- Conclusion: very bad rate of power usage x budget ☐ Coordination needed

# SPSGrid

- SPSGrid (São Paulo State Grid)
- The idea was born before GridUnesp (2003)
- Only in 2011 all conditions existed.  
Launched in December 2010 during the São Paulo OSG School
- Based on the NYSGrid - New York State Grid (and others)



# Grid Community

- The SPSGrid will rely on a local support community, which will be built based on the successfully support model adopted by the U.S. Open Science Grid
- Consortium model between funding agencies, research institutions and researchers
- Based on the OSG software stack

# Advantages for everyone

- To the research institutes
  - Better use of power and network
  - Higher fault tolerance
- To the funding Agencies
  - More scientific production w/ less money
- To the researchers
  - Less system administration done by researchers
  - They can do what they do best: research

# Support services

- Grid Operations Center for the SPSGrid and other grid initiatives
  - Engineering services
  - Monitoring services
  - Consulting services
  
- But not only technical support:
  - Education and Training - EduGrid

# Education and Training

- Ongoing project
- Main company partners: Intel, SGI
- Main public university partners: UFABC, USP, UFSCar
- First focus: Grid over Virtual machines (VMs)
- Current focus: course management based on Moodle. Using grid in a box concept
- Future: VMs started from a Moodle context

# Initial players

- MoU assignments with the biggest research groups (status: ongoing)
  - UNESP, USP and UNICAMP
  - UNIFESP, UFABC and UFSCar
  - LNLS (National Laboratory for Synchrotron Light)
  - CTBE (Brazilian Bioethanol Science and Technology Laboratory)
  - INPE (National Institute for Space Research)
- Next step: the establishment of a coordination council so the definitions of the policies can be started

# Network is ready (KyaTera)



# Vegetative growth model

- After big players at the beginning, a few other institutions will be chosen for initiating the vegetative growth model
- At each new institution, a “seed” cluster will be placed, so the researchers can attach their own computational power under their institutions’ administrative domain
- After a testing period, other institutions will then be selected and inserted into the SPSgrid

# Conclusions

- We have the necessary infrastructure and know-how to implement a statewide Grid infrastructure
- CHEPICS can provide financial support
- We will face great challenges at the time of the policies definitions



# **Thank you for your attention!**

- Questions?

# GridUNESP research areas (detailed)

- **Biology and Biophysics**

- Molecular Dynamics
- Computational Biophysics
- EEG & Apnea
- Proteomics
- Genomics and Phylogenetics
- Amphibians at high elevations

- **Chemistry**

- Modeling for new materials
- Quantum chemistry
- Intermetallic phases
- Coordination compounds
- Vibrational circular dichroism

- **Computer Science**

- Data mining and IPFIX
- Grid Algorithm optimization
- Numerical methods

- **Geosciences**

- Terrestrial deformation
- Platform modeling

- **Materials Science**

- Superconductor vortices
- Electronic Structure
- Photo-dissociation of polymers
- Strong correlated electrons

- **Meteorology**

- Historic precipitation in S. Paulo
- Multi-scale interaction

- **Physics**

- Chaos and phase transition
- Lattice QCD
- Dark Energy Survey
- Few-body systems
- High-energy physics