# A Grid-enabled Infrastructure Testbed for research and dissemination activities in São Paulo

Rogério L. Iope

1st SPRACE Workshop

March 29, 2007

# Brief overview

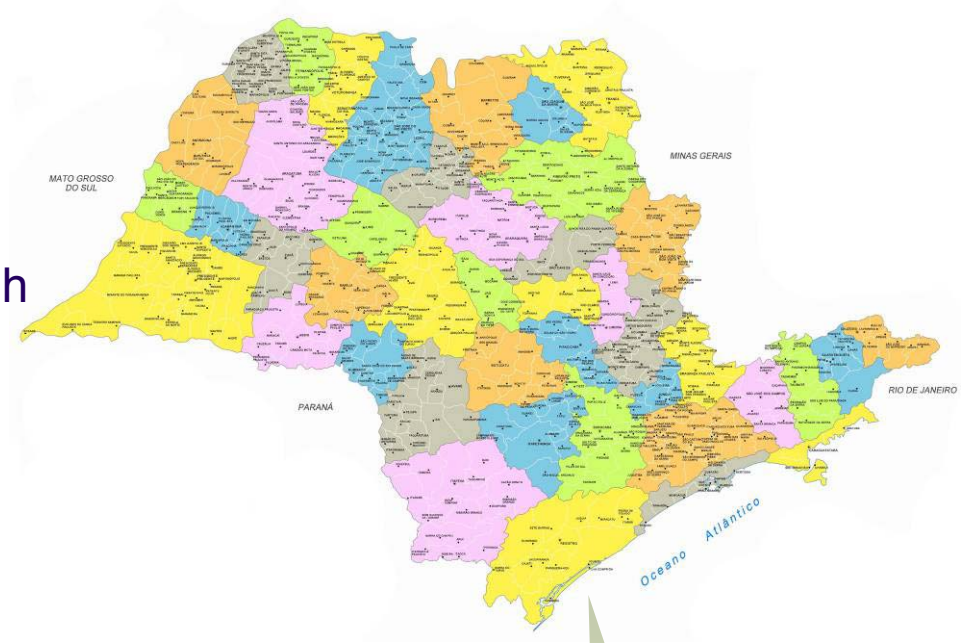
- Summary of project proposal
  - Leveraging research and dissemination activities in Grid Computing
  - Use of a geographically distributed Grid testbed as a learning tool

- Project's main characteristics
  - Not tied to a single university or research group interest
  - Strictly focused on the dissemination of Grid advanced technologies
  - Use of state-of-art VM technology to build up an extensive testbed
  - Result of a collaborative effort of leading scientists and engineers

- Target communities
  - Advanced undergraduate and graduate students from the main public universities of the state of São Paulo, Brazil
  - Systems engineers and site administrators, technicians and collaborators responsible for supporting Grid infrastructures
  - Researchers who need to learn how to use Grid technologies

# The São Paulo State

- 645 Cities
- 1/5 Brazilian Population
  - 15 million Italians
  - 7 million Portuguese and Spanish
  - 5 million Lebanese & Syrian
  - 4 million Asian
  - 3 million Germany
- 1/3 Brazilian Economy
- Size (250K km2) ≈ United Kingdom
- Population (40.5M) ≈ Spain
- GDP/PPP ($ 500B) ≈ 2 **X** Switzerland




Ref.: The World Fact Book (CIA)

# Deployment strategies

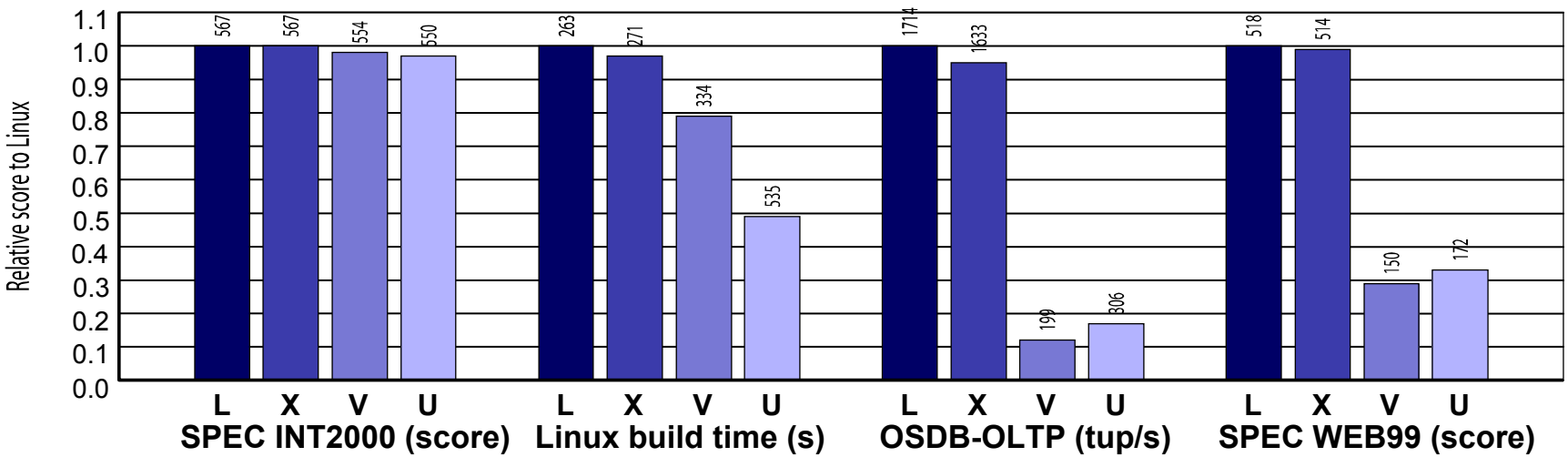
- Project deployment - main ideas
  - Servers planned to be deployed in pairs, and distributed over a wide metropolitan area
  - Each physical server is split in several virtual servers using cutting-edge virtualization technologies
  - Virtual servers house all the services needed to run Grid sites (using a “Grid-in-a-box” concept)
  - Virtualization provides server consolidation, saving hardware and deployment & management costs
  - VMs make it easy to set up and tear down new environments without time consuming installation and configuration
  - An ideal environment for education & outreach activities, as well as for porting and validating middleware and for dinamically testing configuration changes

# Virtual machines based on Xen

- Xen: a software based Virtual Machine Monitor (hypervisor)
  - Developed at the University of Cambridge
  - A low-level micro-kernel that runs on bare hardware
  - Common interface through which hosted virtual machines can access low-level services
  - Allows multiple operating systems to be hosted simultaneously on the same machine
  - Provides the protection of performance isolation between domains
  - An open-source project, published under the GPL
- Hardware support for virtualization
  - Intel VT allows Xen to support  full virtualization
  - Full virtualization enables one to run unmodified OS guests
  - Intel recent releases also include new support for I/O device virtualization
- University of Cambridge Xen page
  - http://www.cl.cam.ac.uk/Research/SRG/netos/xen/

# Virtual machines advantages

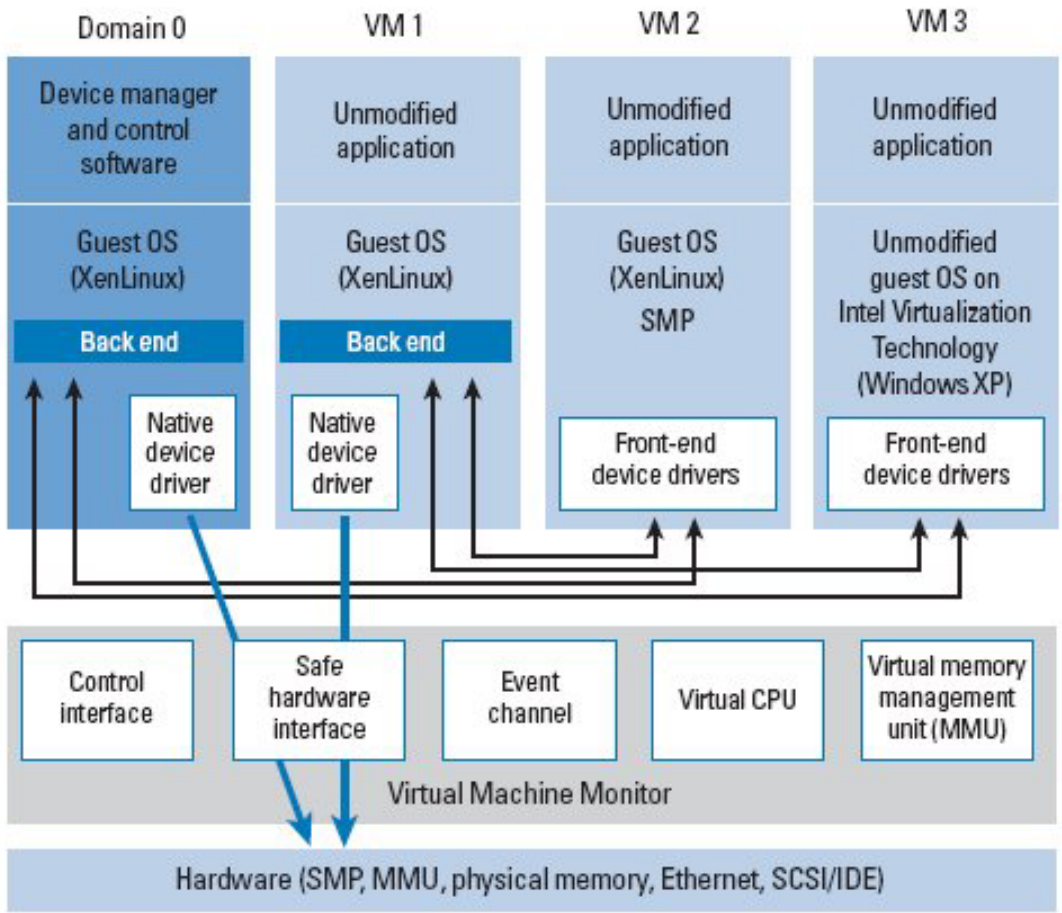
- Application isolation and security
- Improved productivity on deploying and updating software
- More control to each site & great recoverability
- Flexible configuration and management
- The ability to serialize and migrate entire sites
- Consolidation and reduced TCO



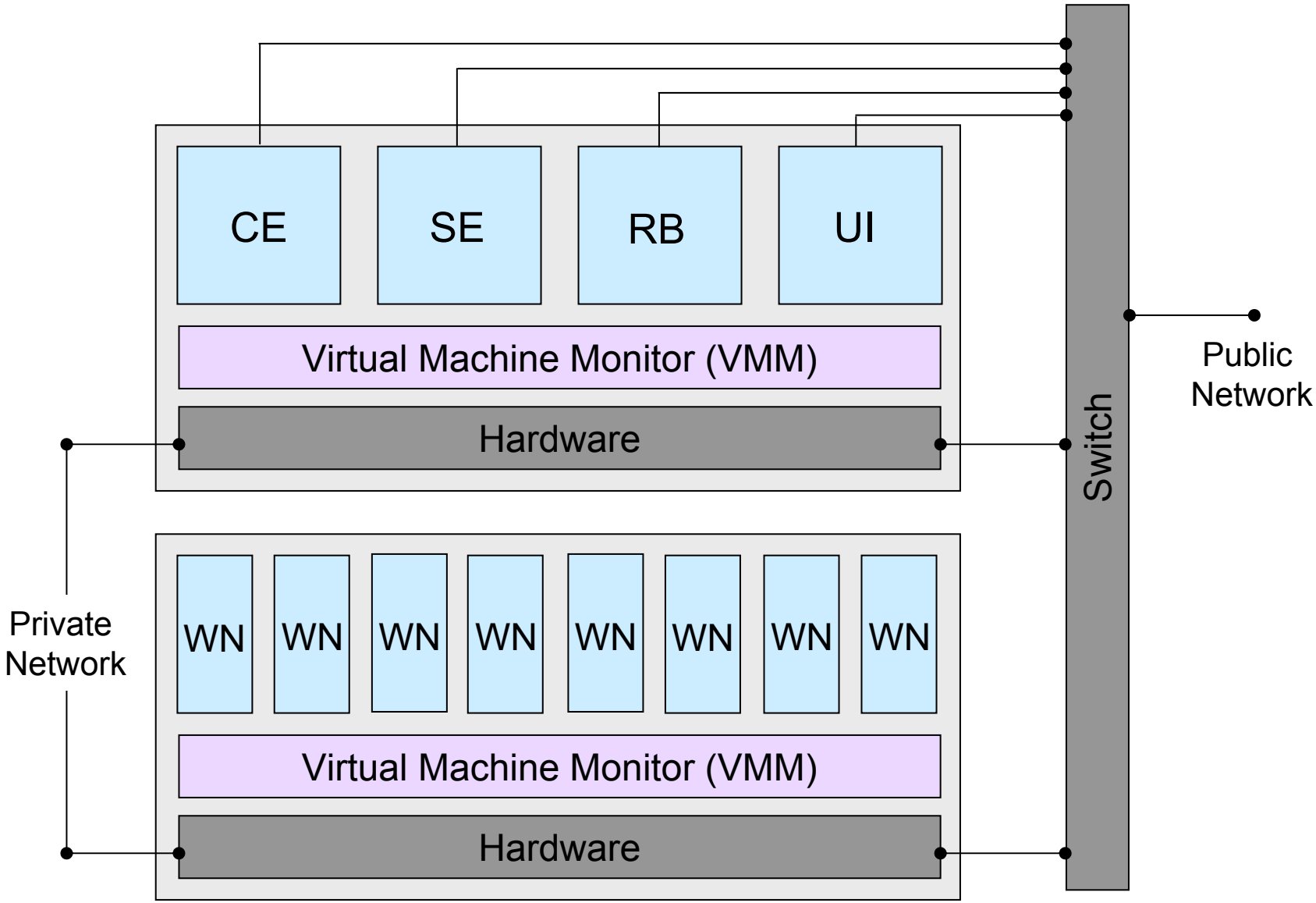

**Benchmark suite running on Linux (L), Xen (X), VMware Workstation (V), and UML (U)**

# Xen 3.0 architecture

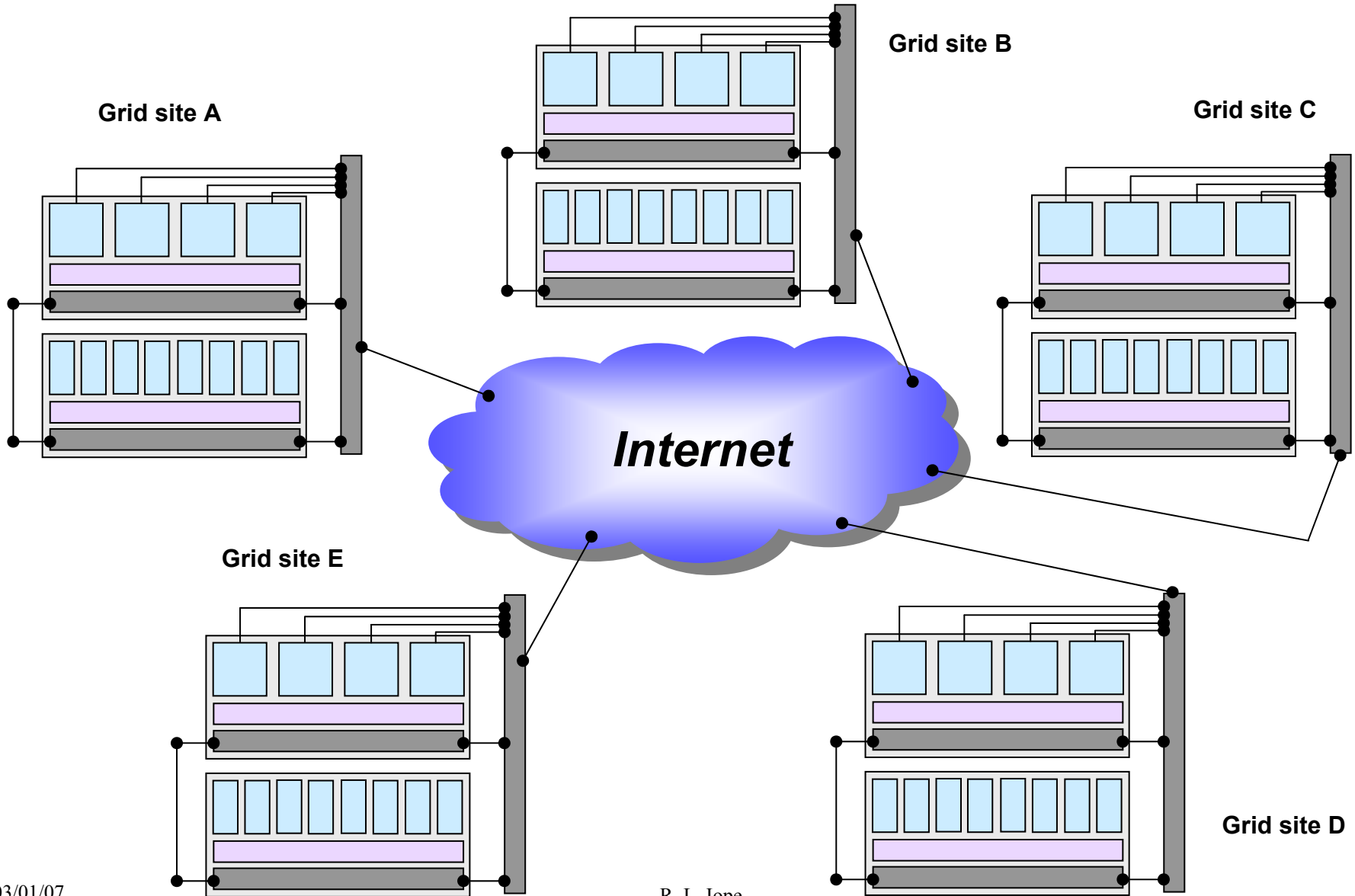
- A machine running Xen hypervisor can host a number of different guest Operating Systems
- The Xen VMM abstracts the underlying hardware and provides hardware access for the different virtual machines
- Domain 0 has special priviledges and runs the management and control software to create, manage, and destroy VMs



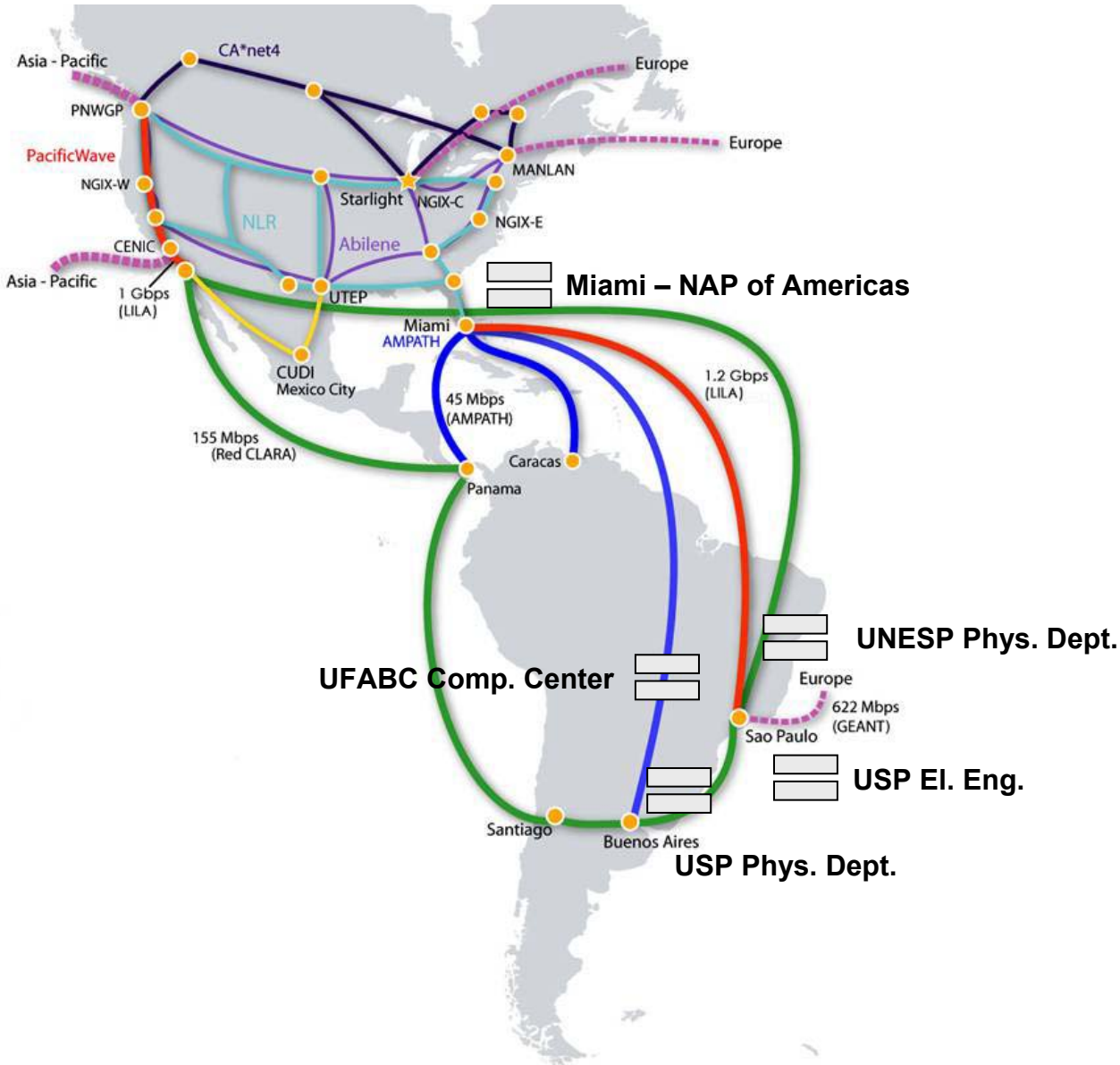
# Planned Grid site architecture



# Set of sites spread over a MAN/WAN



# Sites deployed in a wide area



# Grid testbed use cases

- Education & Outreach testbed for hands-on sessions
  - Pre-configured images stored on a central storage can be instantiated to act as a pool of machines for use by the trainees
  - Extra machines can be quickly created
  - Faulting services can be promply restored
- Deployment of a testing infrastructure
  - Grid system administrators need to test new middleware releases in an environment that colsely resembles the real production infrastructure
- Specialized testbed for middleware development
  - Research middleware instances can be safely tested
  - New middleware install procedures can be verified

[S. Childs, B. Coghlan, J. McCandless, “GridBuilder: A tool for creating virtual Grid testbeds”]

# Closing remarks

- “Practical experience through lab work has long been recognized as an essential part of training education on computer systems.”
- “Users of computing resources tend to stop trying to use a new resource if they get frustrated while attempting to understand and use that resource.”
- Grid computing concepts are not yet easy to grasp, and Grid infrastructures will not turn to be simple to use in a foreseeable future
- A clear need exists: systematic training activities are required to develop new user communities fluent in the use of such new and complex environments
- A Grid infrastructure testbed can ultimately act as a fundamental tool for disseminating essential concepts and practical skills needed by students, pratictioners, and researchers in the fields of natural and applied sciences, engineering and computer sciences, for helping them conduct and support scientific analyses and new technological developments using emerging Grid computing technologies