FutureGrid Overview

FutureGrid Tutorial at PPAM 2011

Torun Poland

September 11 2011

Geoffrey Fox

gcf@indiana.edu

http://www.infomall.org https://portal.futuregrid.org

Director, Digital Science Center, Pervasive Technology Institute

Associate Dean for Research and Graduate Studies, School of Informatics and Computing

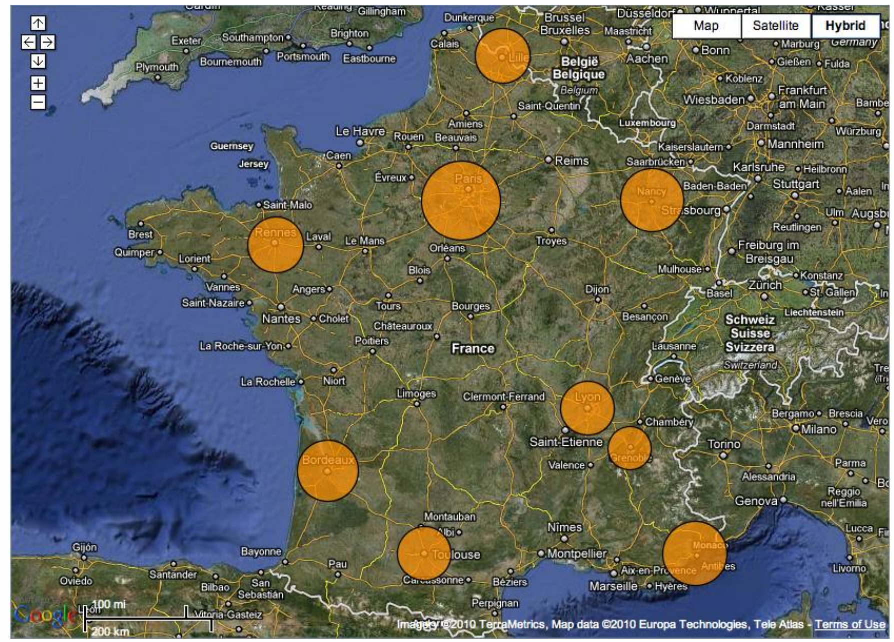
Indiana University Bloomington

FutureGrid key Concepts I

- FutureGrid is an international testbed modeled on Grid5000
- Supporting international Computer Science and Computational Science research in cloud, grid and parallel computing (HPC)
  - Industry and Academia
- The FutureGrid testbed provides to its users:
  - A flexible development and testing platform for middleware and application users looking at interoperability, functionality, performance or evaluation
  - Each use of FutureGrid is an experiment that is reproducible
  - A rich education and teaching platform for advanced cyberinfrastructure (computer science) classes

FutureGrid modeled on Grid’5000

- Experimental testbed
  - Configurable, controllable, monitorable
- Established in 2003
- 10 sites
  - 9 in France
  - Porto Allegre in Brazil
- ~5000+ cores

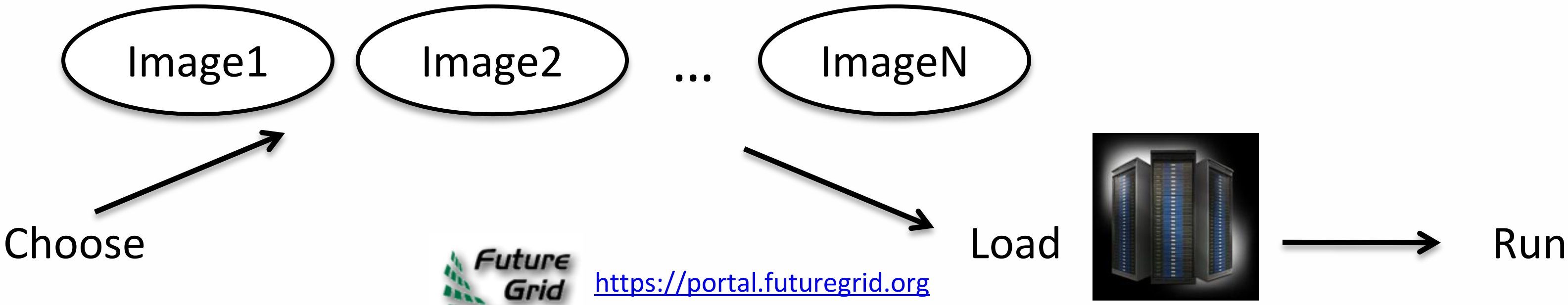


FutureGrid key Concepts II

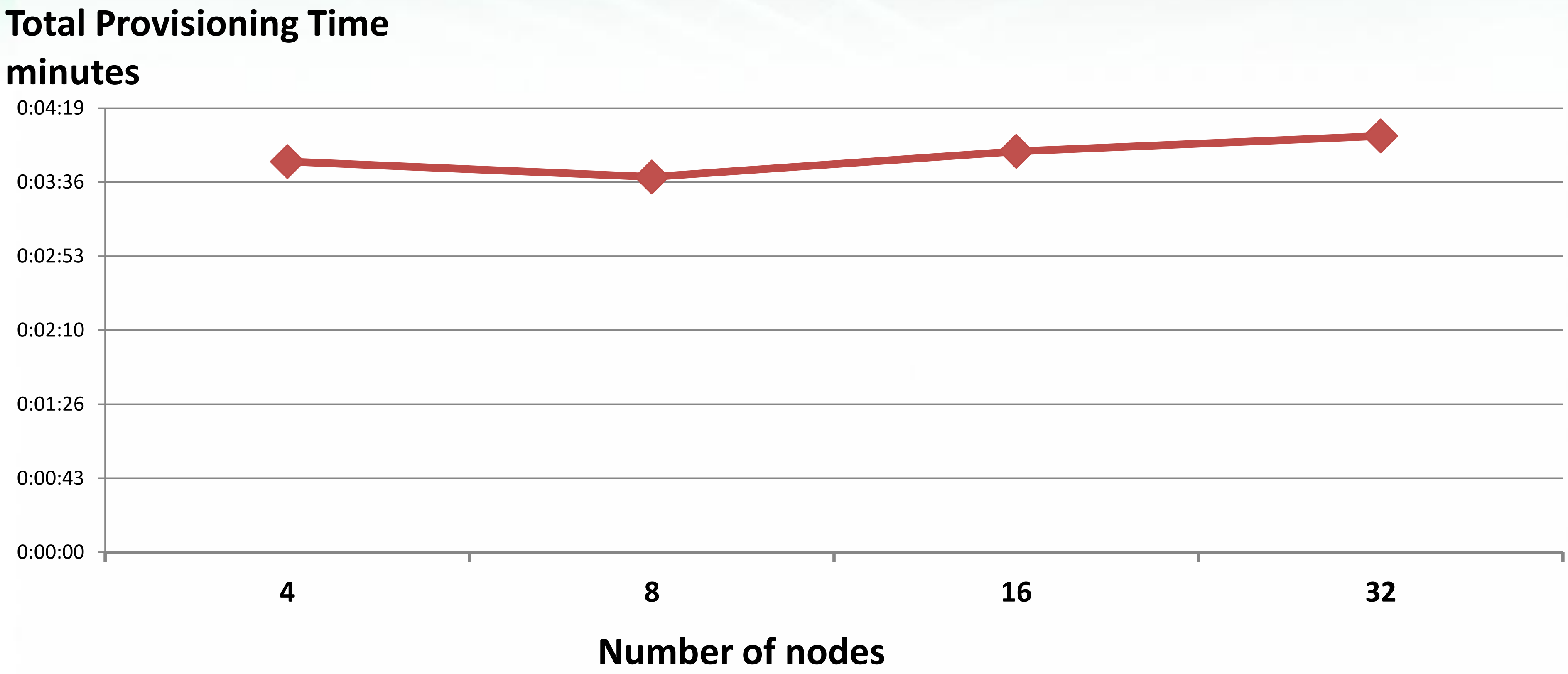
- FutureGrid has a complementary focus to both the Open Science Grid and the other parts of XSEDE (TeraGrid).
  - FutureGrid is user-customizable, accessed interactively and supports Grid, Cloud and HPC software with and without virtualization.
  - FutureGrid is an experimental platform where computer science applications can explore many facets of distributed systems
  - and where domain sciences can explore various deployment scenarios and tuning parameters and in the future possibly migrate to the large-scale national Cyberinfrastructure.
  - FutureGrid supports Interoperability Testbeds – OGF really needed!
- Note much of current use Education, Computer Science Systems and Biology/Bioinformatics

FutureGrid key Concepts III

- Rather than loading images onto VM's, FutureGrid supports Cloud, Grid and Parallel computing environments by dynamically provisioning software as needed onto "bare-metal" using Moab/xCAT
  - Image library for MPI, OpenMP, MapReduce (Hadoop, Dryad, Twister), gLite, Unicore, Globus, Xen, ScaleMP (distributed Shared Memory), Nimbus, Eucalyptus, OpenNebula, KVM, Windows …..
- Growth comes from users depositing novel images in library
- FutureGrid has ~4000 (will grow to ~5000) distributed cores with a dedicated network and a Spirent XGEM network fault and delay generator



Dynamic Provisioning Results

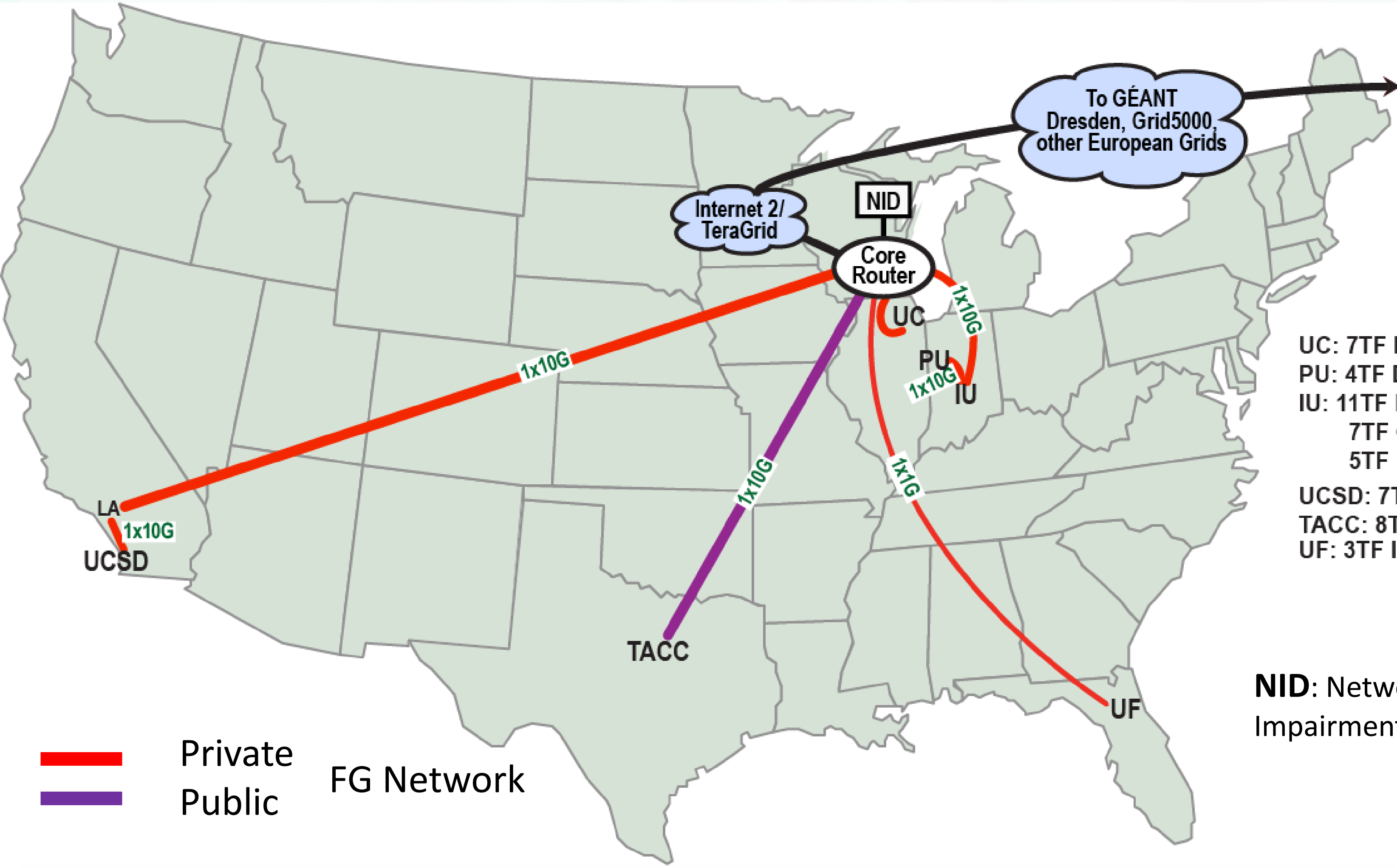


Time elapsed between requesting a job and the jobs reported start time on the provisioned node. The numbers here are an average of 2 sets of experiments.

FutureGrid Partners

- Indiana University (Architecture, core software, Support)
- Purdue University (HTC Hardware)
- San Diego Supercomputer Center at University of California San Diego (INCA, Monitoring)
- University of Chicago/Argonne National Labs (Nimbus)
- University of Florida (ViNE, Education and Outreach)
- University of Southern California Information Sciences (Pegasus to manage experiments)
- University of Tennessee Knoxville (Benchmarking)
- University of Texas at Austin/Texas Advanced Computing Center (Portal)
- University of Virginia (OGF, Advisory Board and allocation)
- Center for Information Services and GWT-TUD from Technische Universtität Dresden. (VAMPIR)
- Red institutions have FutureGrid hardware

FutureGrid: a Grid/Cloud/HPC Testbed



UC: 7TF IBM 672 cores
PU: 4TF Dell 384 cores
IU: 11TF IBM 1024 cores
7TF Cray 684 cores
5TF Disk Rich 512 cores
UCSD: 7TF IBM 672 cores
TACC: 8TF Dell 768 cores
UF: 3TF IBM 256 cores

NID: Network Impairment Device

Private
Public
FG Network

Future Grid https://portal.futuregrid.org

Compute Hardware

Name	System type	# CPUs	# Cores	TFLOPS	Total RAM (GB)	Secondary Storage (TB)	Site	Status
india	IBM iDataPlex	256	1024	11	3072	339 + 16	IU	Operational
alamo	Dell PowerEdge	192	768	8	1152	30	TACC	Operational
hotel	IBM iDataPlex	168	672	7	2016	120	UC	Operational
sierra	IBM iDataPlex	168	672	7	2688	96	SDSC	Operational
xray	Cray XT5m	168	672	6	1344	339	IU	Operational
foxtrot	IBM iDataPlex	64	256	2	768	24	UF	Operational
Bravo*	Large Disk & memory	32	128	1.5	3072 (192GB per node)	144 (12 TB per Server)	IU	Early user Aug. 1 general
Delta*	Large Disk & memory With Tesla GPU's	16 16 GPU's	96	? 3	1536 (192GB per node)	96 (12 TB per Server)	IU	~Sept 15
			TOTAL Cores 4288					

* Teasers for next machine

Future Grid https://portal.futuregrid.org

Storage Hardware

System Type	Capacity (TB)	File System	Site	Status
DDN 9550 (Data Capacitor)	339 shared with IU + 16 TB dedicated	Lustre	IU	Existing System
DDN 6620	120	GPFS	UC	New System
SunFire x4170	96	ZFS	SDSC	New System
Dell MD3000	30	NFS	TACC	New System
IBM	24	NFS	UF	New System

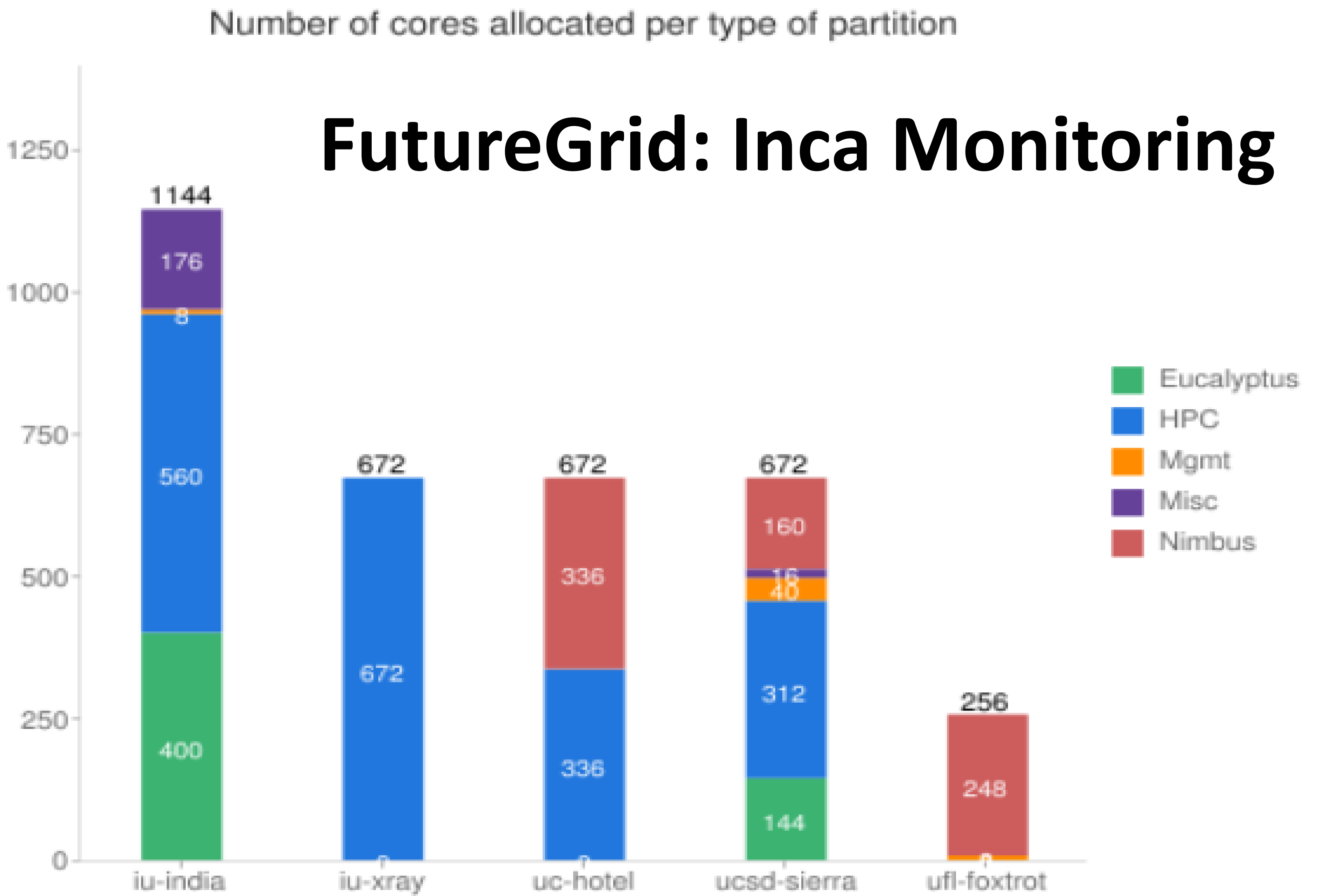
Network Impairment Device

- Spirent XGEM Network Impairments Simulator for jitter, errors, delay, etc
- Full Bidirectional 10G w/64 byte packets
- up to 15 seconds introduced delay (in 16ns increments)
- 0-100% introduced packet loss in .0001% increments
- Packet manipulation in first 2000 bytes
- up to 16k frame size
- TCL for scripting, HTML for manual configuration

FutureGrid: Online Inca Summary

Toggle aliasworks status | Toggle ping status | Toggle ssh status | Toggle nimbus-external-telnet status | Toggle eucalyptus-webpage-

FutureGrid: Inca Monitoring



5 Use Types for FutureGrid

- **122** approved projects July 17 2011 (141 September 11)
  - https://portal.futuregrid.org/projects
- **Training Education and Outreach (13)**
  - Semester and short events; promising for small universities
- **Interoperability test-beds (4)**
  - Grids and Clouds; **Standards**; from Open Grid Forum OGF
- **Domain Science applications (42)**
  - Life science highlighted (21)
- **Computer science (50)**
  - Largest current category
- **Computer Systems Evaluation (35)**
  - TeraGrid (TIS, TAS, XSEDE), OSG, EGI
- Clouds are meant to need less support than other models; FutureGrid needs more **user support** …….

FG Projects Summary

https://portal.futuregrid.org/projects

Project Ref ID	User	Title	Institution	Date Started	Keywords
[FG-P133]	Manuel Rossetti	Supply Chain Network Simulator Using Cloud Computing	University of Arkansas, Department of Industrial Engineering	07/01/2011	discrete event simulation; supply chain; large-scale inventory
[FG-P132]	Yogesh Simmhan	Large scale data analytics	University of Southern California, Computer Engineering Division	07/01/2011	data mining, machine learning, smart grids
[FG-P131]	Judy Qiu	HBase Application and Investigation	Indiana University, School of Informatics and Computing	06/20/2011	Hadoop, HBase, Non-relational database indexing
[FG-P130]	Weiwei Chen	Optimizing Scientific Workflows on Clouds	University of Southern California, Information Sciences Institute	06/08/2011	workflow, cloud computing
[FG-P129]	Andrew Younge	Google SOC 11: Using DemoGrid and Globus Online on FutureGrid	Indiana University, Community Grids Laboratory	06/03/2011	grid, cloud, eucalyptus, nimbus, globus
[FG-P128]	Haixu Tang	De novo assembly of genomes and metagenomes from next generation sequencing data	Indiana University, School of Informatics and Computing	06/01/2011	Fragment assembly, metagenomics, bioinformatics
[FG-P127]	Cui Lin	Fresno System Architecture and Cloud Computing Class	California State University, Fresno, Computer Science	05/31/2011	system programming, system architecture, cloud computing
[FG-P126]	Srirangam Addepalli	Research Experience for Undergraduate Students in Cloud Computing	University of Arkansas, High Performance Computing Center	05/31/2011	REU
[FG-P125]	Shiyong Lu	The VIEW Project	Wayne State University, Department of Computer Science	05/31/2011	Scientific workflow, Cloud computing.
[FG-P123]	Jongwook Woo	CSULA Business Intelligence on Map/Reduce	California State University Los Angeles, CIS department	05/22/2011	Market Basket Analysis, Customer Analysis, Data Mining, Business Intelligence, MapReduce,
[FG-P122]	Massimo Canonico	Cloud computing class	University of Piemonte Orientale, Computer Science Department	05/22/2011	class, teaching, comparison of different cloud middlewares
[FG-P124]	Gregor von Laszewski	Tutorial: CCGrid2011	Indiana University, Community Grids Laboratory	05/20/2011	tutorial, future grid, nimbus, eucalyptus, hpc
[FG-P110]	Gregory Pike	FutureGrid Systems Development	Indiana University, Community Grids Lab at the Pervasive Technology Institute	05/20/2011	FutureGrid, Systems
[FG-P121]	Paul Marshall	Elastic Computing	University of Colorado at Boulder, Computer Science	05/16/2011	elastic computing, cloud computing, infrastructure-as-a-service
[FG-P120]	Jerome Mitchell	Workshop: A Cloud View on Computing	PTI, Indiana University	05/16/2011	outreach, cloud computing
[FG-P119]	Mo Zhou	Keyword-based Semantic Association Discovery	Indiana University , Computer science	05/16/2011	Keyword, Semantic Association, MapReduce, Parallel

Current Education projects

- **System Programming and Cloud Computing,** Fresno State, Teaches system programming and cloud computing in different computing environments
- **REU: Cloud Computing,** Arkansas, Offers hands-on experience with FutureGrid tools and technologies
- **Workshop: A Cloud View on Computing,** Indiana School of Informatics and Computing (SOIC), Boot camp on MapReduce for faculty and graduate students from underserved ADMI institutions
- **Topics on Systems: Distributed Systems,** Indiana SOIC, Covers core computer science distributed system curricula (for 60 students)

Current Interoperability Projects

- **SAGA,** Louisiana State, Explores use of FutureGrid components for extensive portability and interoperability testing of Simple API for Grid Applications, and scale-up and scale-out experiments
- **XSEDE/OGF** Unicore and Genesis Grid endpoints tests for new US and European grids

Current Bio Application Projects

- **Metagenomics Clustering,** North Texas, Analyzes metagenomic data from samples collected from patients
- **Genome Assembly,** Indiana SOIC, De novo assembly of genomes and metagenomes from next generation sequencing data

Current Non-Bio Application Projects

- **Physics: Higgs boson,** Virginia, Matrix Element calculations representing production and decay mechanisms for Higgs and background processes
- **Business Intelligence on MapReduce,** Cal State - L.A., Market basket and customer analysis designed to execute MapReduce on Hadoop platform

Current Computer Science Projects

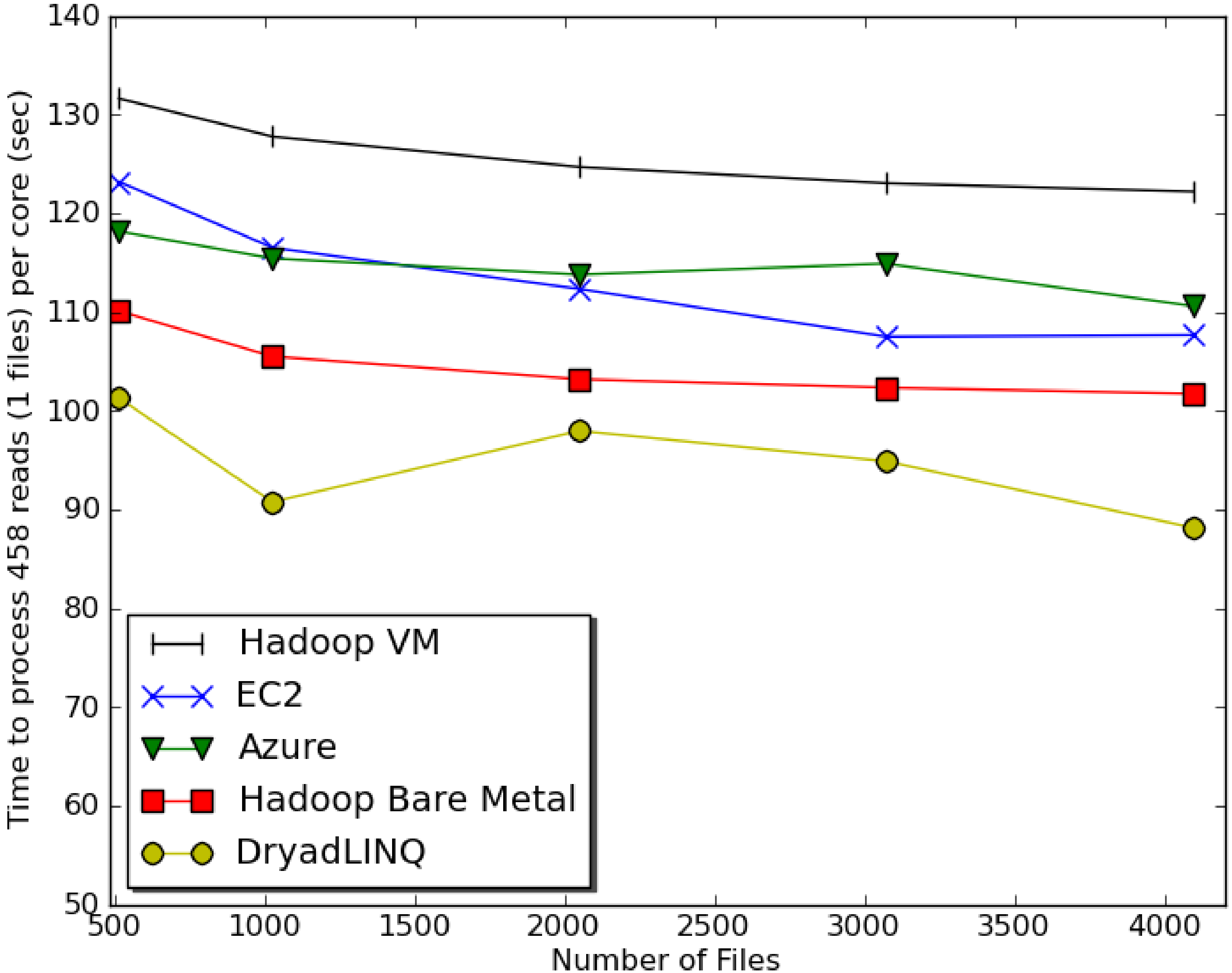
- **Data Transfer Throughput,** Buffalo, End-to-end optimization of data transfer throughput over wide-area, high-speed networks
- **Elastic Computing,** Colorado, Tools and technologies to create elastic computing environments using IaaS clouds that adjust to changes in demand automatically and transparently
- **The VIEW Project,** Wayne State, Investigates Nimbus and Eucalyptus as cloud platforms for elastic workflow scheduling and resource provisioning

Current Technology Projects

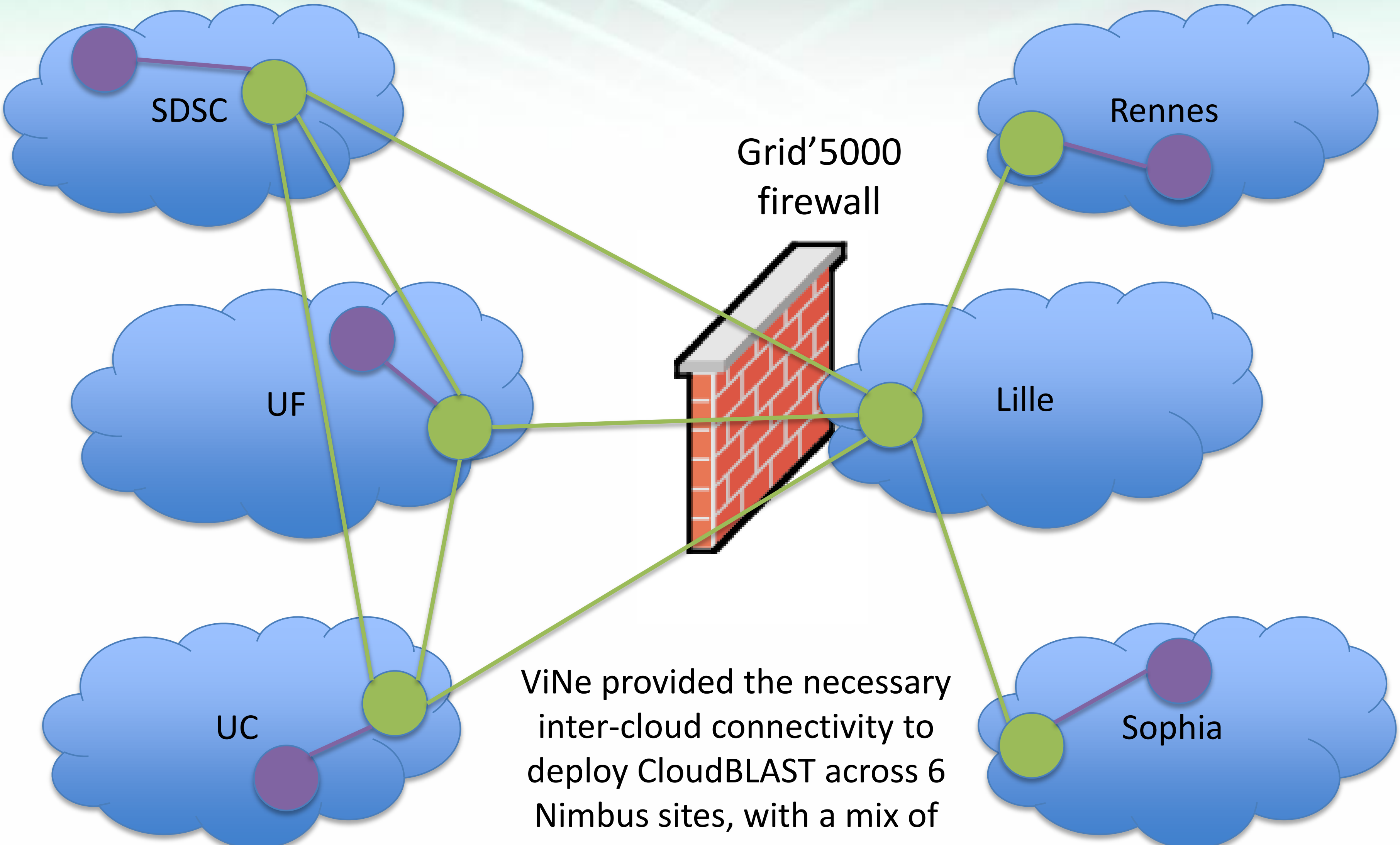
- **ScaleMP for Gene Assembly,** Indiana Pervasive Technology Institute (PTI) and Biology, Investigates distributed shared memory over 16 nodes for SOAPdenovo assembly of Daphnia genomes
- **XSEDE,** Virginia, Uses FutureGrid resources as a testbed for XSEDE software development
- **Globus Online,** Indiana PTI, Chicago, Investigates the feasibility of providing DemoGrid and its Globus services on FutureGrid IaaS clouds

Typical FutureGrid Performance Study

Linux, Linux on VM, Windows, Azure, Amazon Bioinformatics



OGF 2010 Demo from Rennes



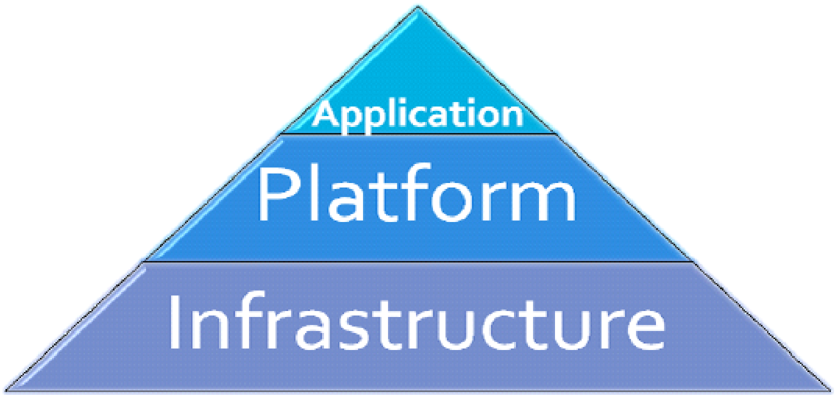
ViNe provided the necessary inter-cloud connectivity to deploy CloudBLAST across 6 Nimbus sites, with a mix of public and private subnets.

Home

Term Project Final Report

TERM PROJECT
Aparna Tiwari
Pradnya Kakodkar
Ratul Bhawal
Sumaya Alwaris

In this project we have built a complete distributed system by going through the layers in figure1 in a number of phases. Also, we have done performance analysis and some bench marking on performance of C-MPI v/s Java-MPI for PageRank application.



Phase#1
Pagerank using Java MPI:
This project constituted the first phase of our term project
In this project we implemented a page rank algorithm which is one of the essential algorithms used by search engines to calculate the value of each page. This project implemented the algorithm in two phases: first phase was a sequential implementation and the second phase was a parallel implementation. The sequential version was more efficient and faster for a small dataset of pages while the parallel version was mostly used for larger sets of data.
The parallel version was implemented using Java MPI.The command line arguments provided by the user were validated and then different computations were performed depending on whether the process is a master process or slave process. Master process will do some extra computation as compared to slave nodes. It sorts the final page rank values received and then writes the top 10 Page Ranks to the output file. It also computes the Job turnaround time and file IO times.

Phase#2
Pagerank Performance analysis on academic cloud
This project constituted the second phase of our term project. In this phase we ran performance tests on Eucalyptus and bare metal.
We analyzed the performance of the Eucalyptus cloud and bare metal clusters running on FutureGrid for MPI. In order to calculate the performance of these systems, we choose the page rank algorithm. We used the Java implementation of the MPI page rank algorithm which was completed earlier in

ADMI Cloudy View on Computing Workshop June 2011

Concept and Delivery by Jerome Mitchell: Undergraduate ECSU, Masters Kansas, PhD Indiana

- Jerome took two courses from IU in this area Fall 2010 and Spring 2011 on FutureGrid
- ADMI: Association of Computer and Information Science/Engineering Departments at Minority Institutions
- Offered on FutureGrid
- 10 Faculty and Graduate Students from ADMI Universities
- The workshop provided information from cloud programming models to case studies of scientific applications on FutureGrid.
- At the conclusion of the workshop, the participants indicated that they would incorporate cloud computing into their courses and/or research.

ADMI Cloudy View on Computing Workshop Participants

DeShea Simon
Hampton University

Timothy Holston
Mississippi Valley State University

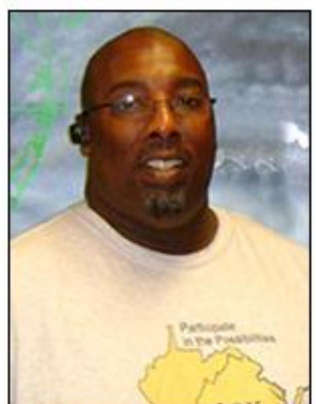
Mohammad Hasan
Elizabeth City State University

Constance Bland
Mississippi Valley State University

Candace Adams
Auburn University

Felicia Doswell
Norfolk State University

Yenhung Hu
Hampton University



Willie Fuller
Norfolk State University

Natarajan Meghanathan
Jackson State University

Darnell Johnson
Elizabeth City State University

Workshop Purpose

- Introduce ADMI to the basics of the emerging Cloud Computing paradigm
  - Learn how it came about
  - Understand its enabling technologies
  - Understand the computer systems constraints, tradeoffs, and techniques of setting up and using cloud
- Teach ADMI how to implement algorithms in the Cloud
  - Gain competence in cloud programming models for distributed processing of large datasets.
  - Understand how different algorithms can be implemented and executed on cloud frameworks
  - Evaluating the performance and identifying bottlenecks when mapping applications to the clouds

FutureGrid Tutorials

- **Tutorial topic 1: Cloud Provisioning Platforms**
- Tutorial NM1: Using Nimbus on FutureGrid
- Tutorial NM2: Nimbus One-click Cluster Guide
- Tutorial GA6: Using the Grid Appliances to run FutureGrid Cloud Clients
- Tutorial EU1: Using Eucalyptus on FutureGrid
- **Tutorial topic 2: Cloud Run-time Platforms**
- Tutorial HA1: Introduction to Hadoop using the Grid Appliance
- Tutorial HA2: Running Hadoop on FG using Eucalyptus (.ppt)
- Tutorial HA2: Running Hadoop on Eualyptus

- **Tutorial topic 3: Educational Virtual Appliances**
- Tutorial GA1: Introduction to the Grid Appliance
- Tutorial GA2: Creating Grid Appliance Clusters
- Tutorial GA3: Building an educational appliance from Ubuntu 10.04
- Tutorial GA4: Deploying Grid Appliances using Nimbus
- Tutorial GA5: Deploying Grid Appliances using Eucalyptus
- Tutorial GA7: Customizing and registering Grid Appliance images using Eucalyptus
- Tutorial MP1: MPI Virtual Clusters with the Grid Appliances and MPICH2
- **Tutorial topic 4: High Performance Computing**
- Tutorial VA1: Performance Analysis with Vampir
- Tutorial VT1: Instrumentation and tracing with VampirTrace

FutureGrid Viral Growth Model

- Users apply for a project
- Users improve/develop some software in project
- This project leads to new images which are placed in FutureGrid repository
- Project report and other web pages document use of new images
- Images are used by other users
- And so on ad infinitum ....
- **Please bring your nifty software up on FutureGrid!!**

Software Components

- Portals including “Support” “use FutureGrid” “Outreach”
- Monitoring – INCA, Power (GreenIT)
- **Experiment Manager**: specify/workflow
- Image Generation and Repository
- Intercloud Networking ViNE
- Virtual Clusters built with virtual networks
- Performance library
- **Rain** or **R**untime **A**daptable **I**nsertio**N** Service for images
- Security Authentication, Authorization,
- Note Software integrated across institutions and between middleware and systems Management (Google docs, Jira, Mediawiki)
- Note many software groups are also FG users

“Research”

Above and below

Nimbus OpenStack Eucalyptus

FutureGrid Software Architecture

- **Note on Authentication and Authorization**
- We have different environments and requirements from TeraGrid
- Non trivial to integrate/align security model with TeraGrid

Detailed Software Architecture

Access Services

- **IaaS** — *Nimbus, Eucalyptus, OpenStack, OpenNebula, ViNe, …*
- **PaaS** — *Hadoop, Dryad, Twister, Virtual Clusters, …*
- **HPC User Tools & Services** — *Queuing System, MPI, Vampir, PAPI, …*
- **Additional Tools & Services** — *Unicore, Genesis II, gLite, …*
- **User and Support Services** — *Portal, Tickets, Backup, Storage,*

Management Services

- **Image Management** — *FG Image Repository, FG Image Creation*
- **Experiment Management** — *Registry, Repository Harness, Pegasus Exper. Workflows, …*
- **Dynamic Provisioning** — *RAIN: Provisioning of IaaS, PaaS, HPC, …*
- **Monitoring and Information Service** — *Inca, Grid Benchmark Challange, Netlogger, PerfSONAR Nagios, …*

FutureGrid Operations Services

- **Security & Accounting Services** — *Authentication Authorization Accounting*
- **Development Services** — *Wiki, Task Management, Document Repository*

Base Software and Services
OS, Queuing Systems, XCAT, MPI, …

FutureGrid Fabric
Compute, Storage & Network Resources

Development & Support Resources
Portal Server, …

https://portal.futuregrid.org