Future Grid Introduction

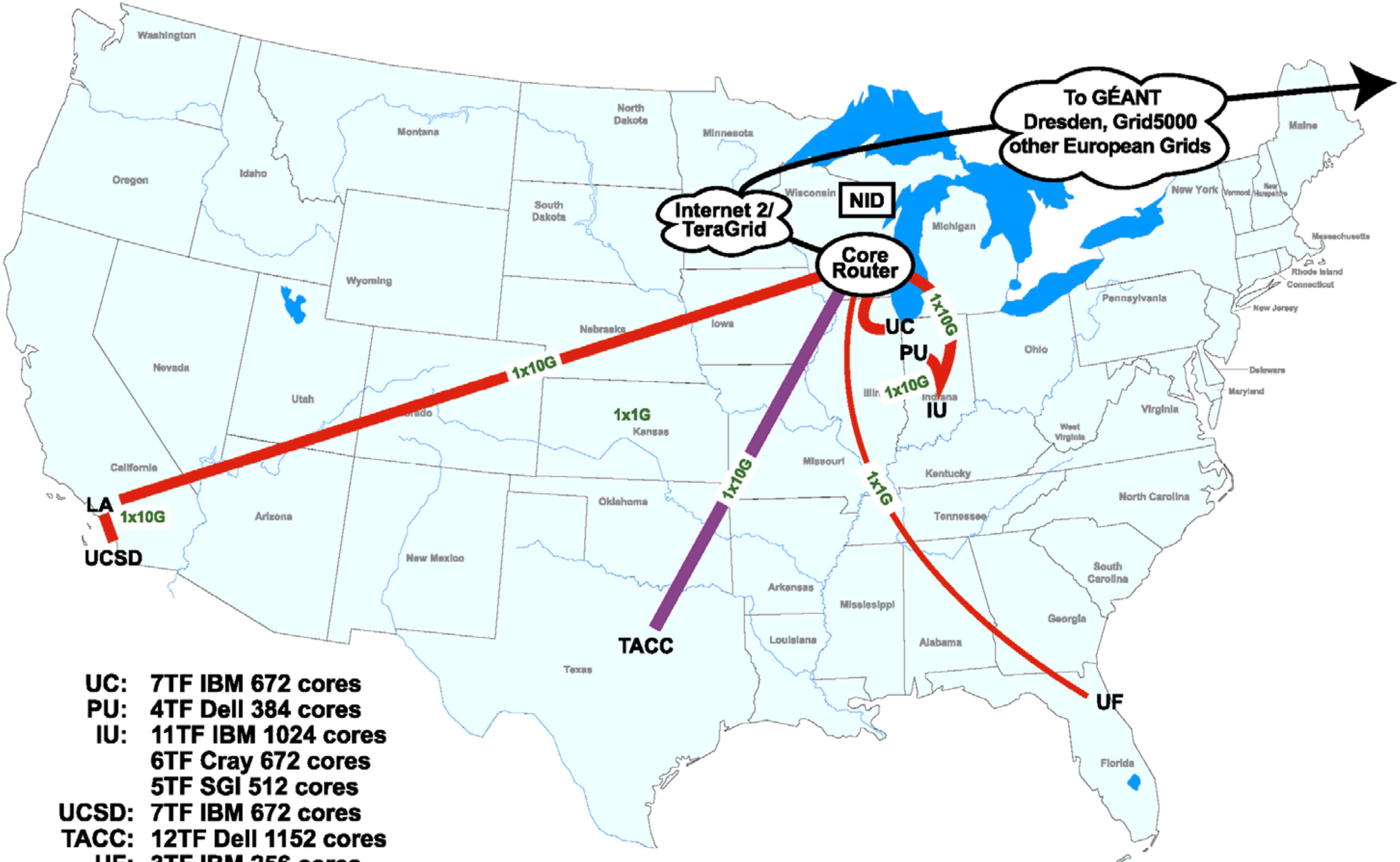
March 8 2010
www.futuregrid.org
MAGIC Meeting

Gregor von Laszewski
laszewski@gmail.com

Community Grids Laboratory,
Digital Science Center
Pervasive Technology Institute
Indiana University

FutureGrid Goals

- Support the research on the future of distributed, grid, and cloud computing.
- Build a robustly managed simulation environment or testbed to support the development and early use in science of new technologies at all levels of the software stack: from networking to middleware to scientific applications.
- Mimic TeraGrid and/or general parallel and distributed systems – FutureGrid is part of TeraGrid and one of two experimental TeraGrid systems (the other is based on GPU)
- Enable major advances in science and engineering through collaborative development of science applications and related software using FG
- FutureGrid is a (small 5600 core) Science Cloud but it is more accurately a virtual machine based simulation environment



UC: 7TF IBM 672 cores
PU: 4TF Dell 384 cores
IU: 11TF IBM 1024 cores
6TF Cray 672 cores
5TF SGI 512 cores
UCSD: 7TF IBM 672 cores
TACC: 12TF Dell 1152 cores
UF: 3TF IBM 256 cores

Add Network Fault Generator and other systems running FutureGri

FutureGrid Hardware

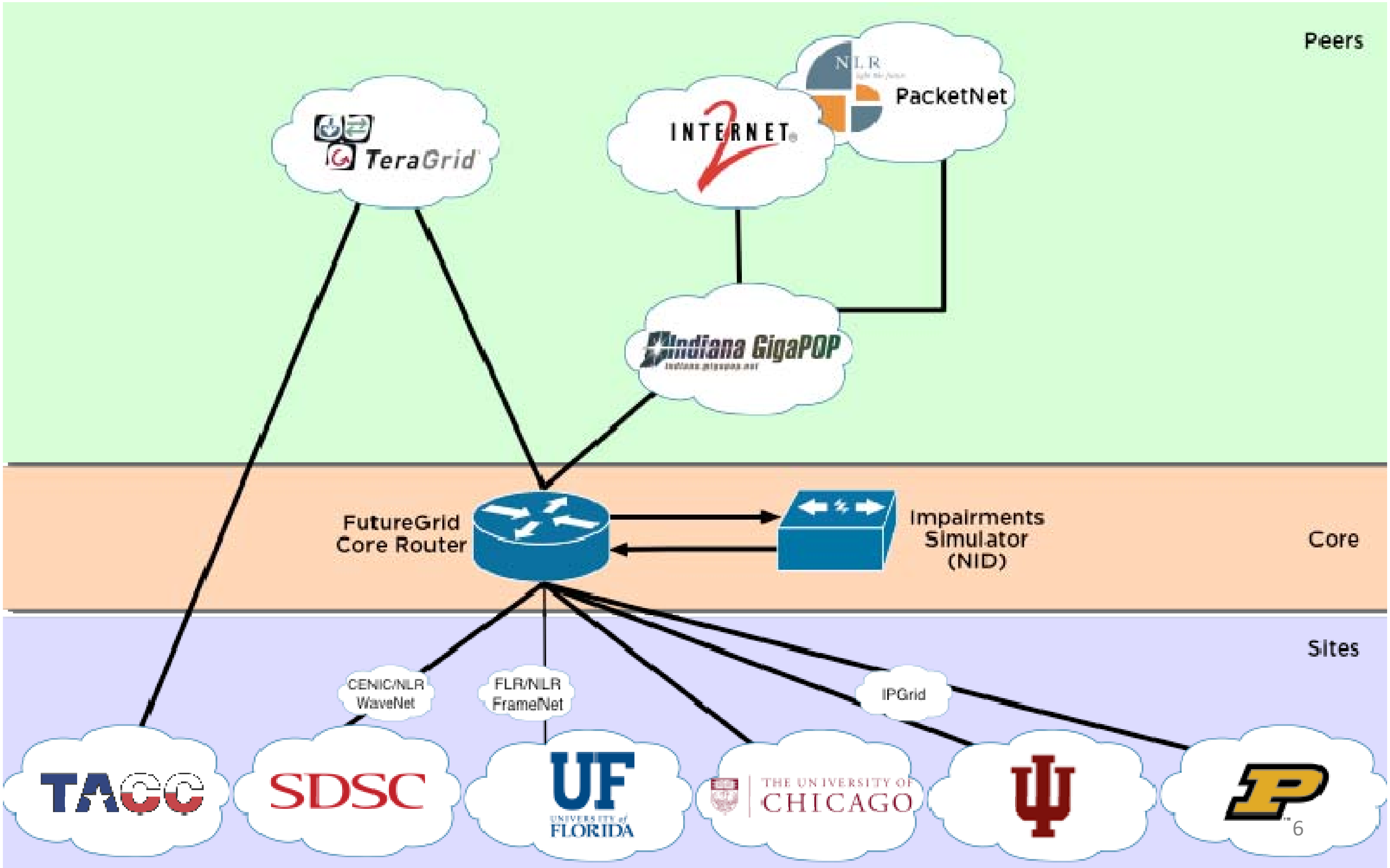
System type	# CPUs	# Cores	TFLOPS	RAM (GB)	Secondary storage (TB)	Default local file system	Site
Dynamically configurable systems							
IBM iDataPlex	256	1024	11	3072	335*	Lustre	IU
Dell PowerEdge	192	1152	12	1152	15	NFS	TACC
IBM iDataPlex	168	672	7	2016	120	GPFS	UC
IBM iDataPlex	168	672	7	2688	72	Lustre/PVFS	UCSD
Subtotal	784	3520	37	8928	542		
Systems not dynamically configurable							
Cray XT5m	168	672	6	1344	335*	Lustre	IU
Shared memory system TBD	40**	480**	4**	640**	335*	Lustre	IU
Cell BE Cluster	4						
IBM iDataPlex	64	256	2	768	5	NFS	UF
High Throughput Cluster	192	384	4	192			PU
Subtotal	552	2080	21	3328	10		
Total	**1336**	**5600**	**58**	**10560**	**552**		

- FutureGrid has dedicated network (except to TACC) and a network fault and delay generator
- Can isolate experiments on request; IU runs Network for NLR/Internet2
- (Many) additional partner machines will run FutureGrid software and be supported (but allocated in specialized ways)

Storage Hardware

System Type	Capacity (TB)	File System	Site	Status
DDN 9550 (Data Capacitor)	339	Lustre	IU	Existing System
DDN 6620	120	GPFS	UC	New System
SunFire x4170	72	Lustre/PVFS	SDSC	New System
Dell MD3000	30	NFS	TACC	New System

Logical Diagram



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)

- Blue institutions have FutureGrid hardware

Other Important Collaborators

- NSF
- Early users from an application and computer science perspective and from both research and education
- Grid5000/Aladin and D-Grid in Europe
- Commercial partners such as
  - Eucalyptus ….
  - Microsoft (Dryad + Azure) – Note current Azure external to FutureGrid as are GPU systems
  - Application partners
- TeraGrid
- Open Grid Forum
- Open Nebula, Open Cirrus Testbed, Open Cloud Consortium, Cloud Computing Interoperability Forum. IBM-Google-NSF Cloud, UIUC Cloud

FutureGrid Usage Scenarios

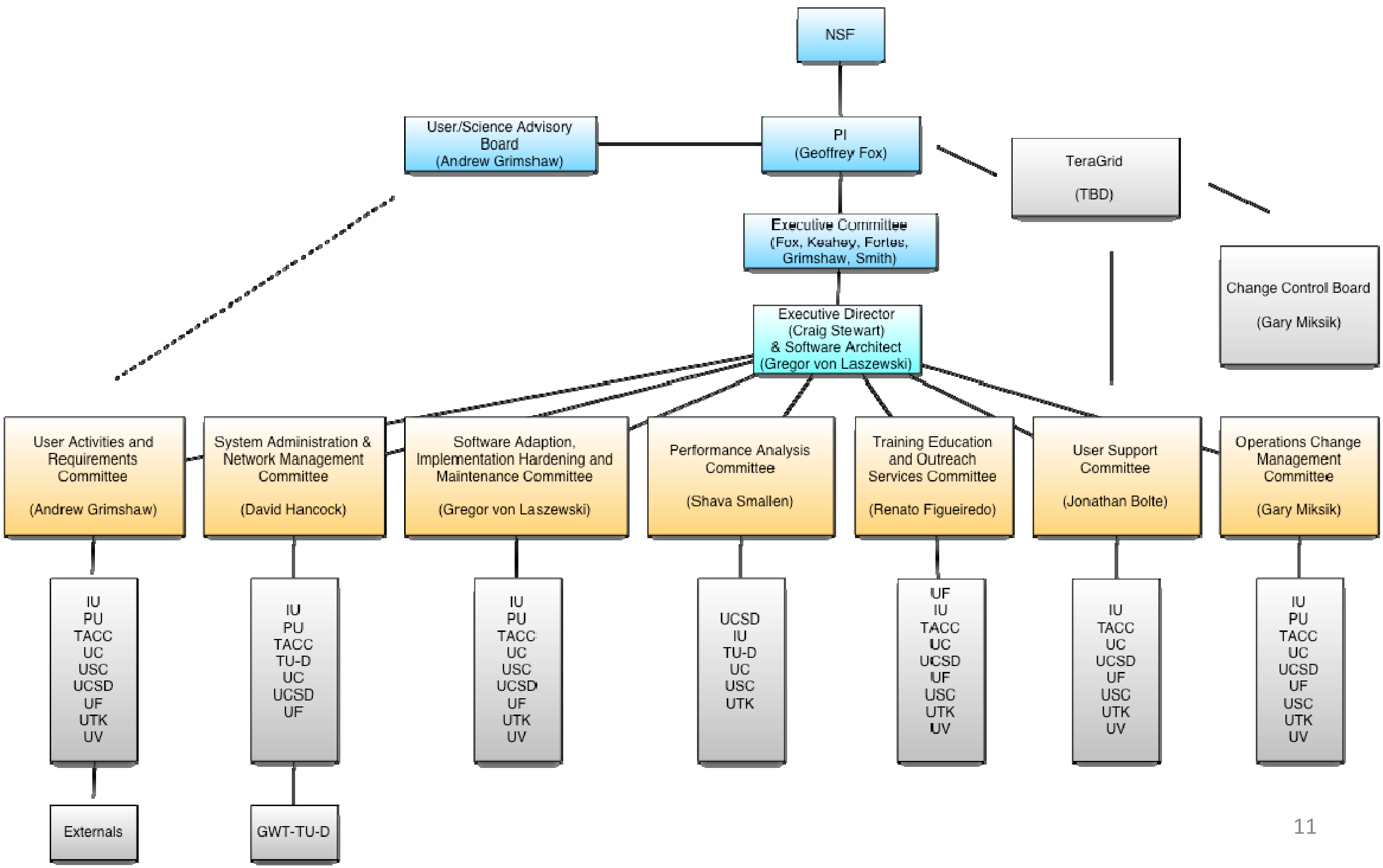
- Developers of end-user applications who want to develop new applications in cloud or grid environments, including analogs of commercial cloud environments such as Amazon or Google.
  - Is a Science Cloud for me?
- Developers of end-user applications who want to experiment with multiple hardware environments.
- Grid/Cloud middleware developers who want to evaluate new versions of middleware or new systems.
- Networking researchers who want to test and compare different networking solutions in support of grid and cloud applications and middleware. (Some types of networking research will likely best be done via through the GENI program.)
- Education as well as research
- Interest in performance requires close to OS support

Future Grid Users

- Application/Scientific users
- System administrators
- Software developers
- Testbed users
- Performance modelers
- Educators
- Students

- Supported by FutureGrid Infrastructure & Software offerings

Management Structure



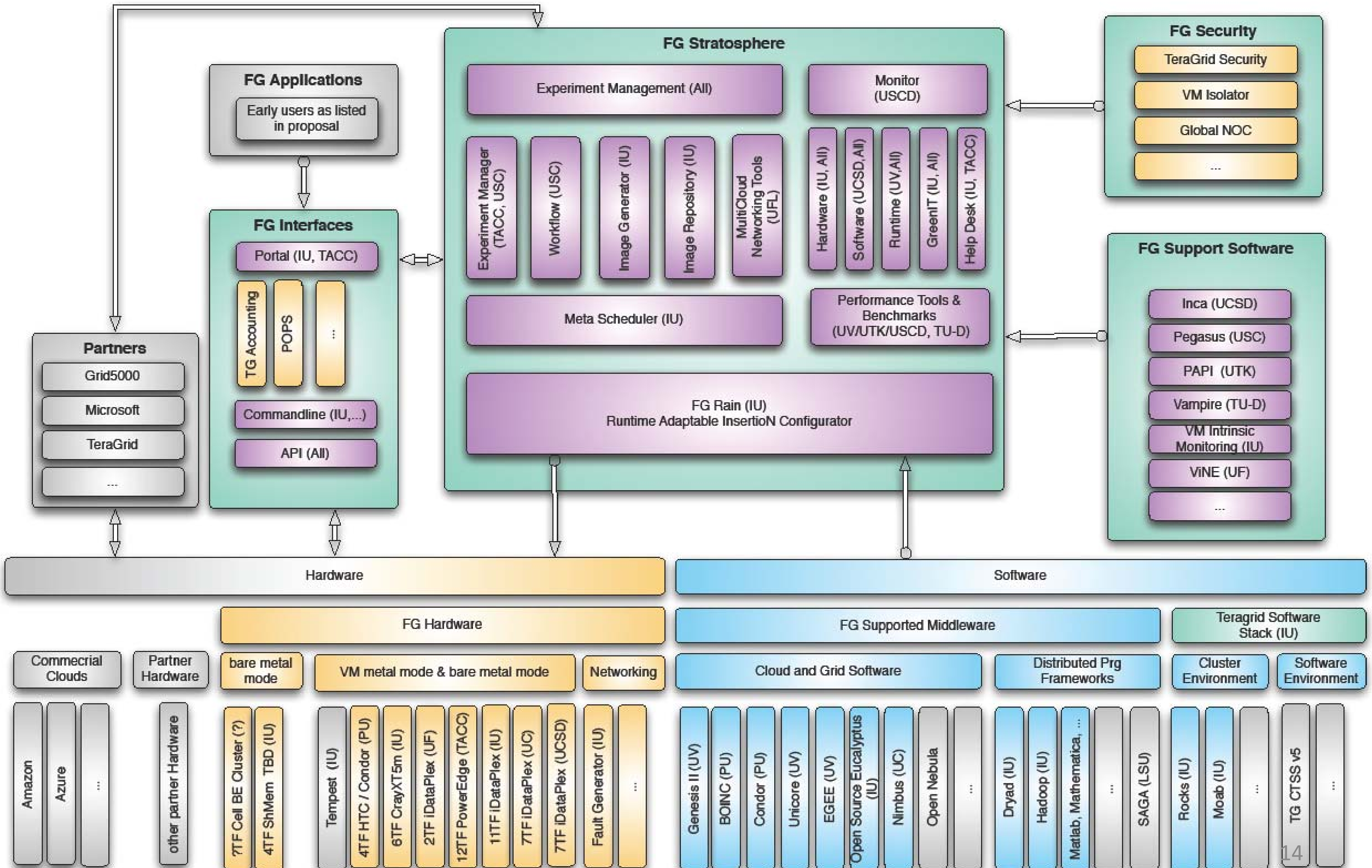
FutureGrid Working Groups

- **Systems Administration and Network Management Committee**: This committee will be responsible for all matters related to systems administration, network management, and security. David Hancock of IU will be the inaugural chair of this committee.
- **Software Adaptation, Implementation, Hardening, and Maintenance Committee**: This committee will be responsible for all aspects of software creation and management. It should interface with TAIS in TeraGrid. Gregor von Laszewski from IU will chair this committee.
- **Performance Analysis Committee**: This committee will be responsible for coordination of performance analysis activities. Shava Smallen of UCSD will be the inaugural chair of this committee.
- **Training, Education, and Outreach Services Committee**: This committee will coordinate Training, Education, and Outreach Service activities and will be chaired by Renato Figueiredo.
- **User Support Committee:** This committee will coordinate the management of online help information, telephone support, and advanced user support. Jonathan Bolte of IU will chair this committee.
- **Operations and Change Management Committee** (including CCB): This committee will be responsible for operational management of FutureGrid, and is the one committee that will always include at least one member from every participating institution, including those participating without funding. This is led by Craig Stewart

FutureGrid Architecture

- Open Architecture allows to configure resources based on images
- Managed images allows to create similar experiment environments
- Experiment management allows reproducible activities
- Through our modular design we allow different clouds and images to be “rained” upon hardware.
- Note will be supported 24x7 at “TeraGrid Production Quality”
- Will support deployment of “important” middleware including TeraGrid stack, Condor, BOINC, gLite, Unicore, Genesis II

FutureGrid Architecture



Development Phases

- Phase 0: Get Hardware to run
- Phase I: Get early users to use the system
- Phase II: Implement Dynamic Provisioning
- Phase III: Integrate with TeraGgrid

Objectives: Software

- Extensions to existing software
- Existing open-source software
- Open-source, integrated suite of software to
  - instantiate and execute grid and cloud experiments.
  - perform an experiment
  - collect the results
  - tools for instantiating a test environment,
    - Torque, MOAB, xCAT, bcfg, and Pegasus, Inca, ViNE, a number of other tools from our partners and the open source community
    - Portal to interact
  - Benchmarking

FG Stratosphere

- Objective
  - Higher than a particular cloud
  - Provides all mechanisms to provision a cloud on a given FG hardware
  - Allows the management of reproducible experiments
  - Allows monitoring of the environment and the results
- Risks
  - Lots of software
  - Possible multiple path to do the same thing
- Good news
  - We know about different solutions and have identified a very good plan with risk mitigation plans

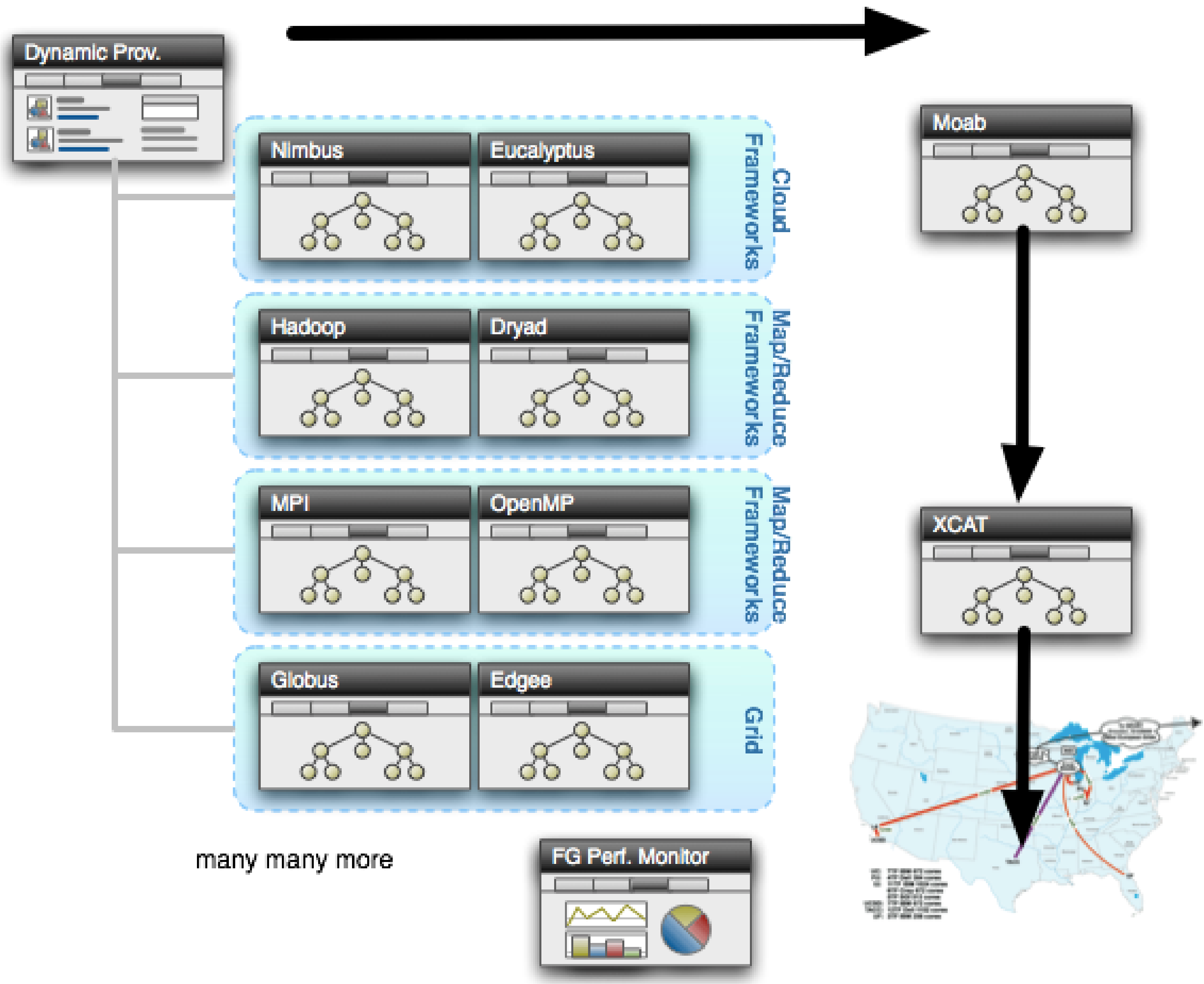
Rain
Runtime Adaptable InsertioN Service

- Objective
  - Provide dynamic provisioning
  - Running outside virtualization
  - Cloud neutral
    - Nimbus, Eucalyptus, …
  - Future oriented
    - Dryad
    - …
- Risks
  - Some frameworks (e.g. MS) are more complex to provision

Dynamic Provisioning

- Change underlying system to support current user demands
  - Linux, Windows, Xen, Nimbus, Eucalyptus, Hadoop, Dryad
  - Switching between Linux and Windows possible!
- Stateless images
  - Shorter boot times
  - Easier to maintain
- Stateful installs
  - Windows
- Use Moab to trigger changes and xCAT to manage installs

Dynamic Provisioning



Dynamic provisioning Examples

- Give me a virtual cluster with 30 nodes
- Give me a Eucalyptus environment with 10 nodes
- Give me a hadoop environment with x nodes

- Run my application on hadoop, dryad, … and compare the performance

Command line

- fg-deploy-image
  - host name
  - image name
  - start time
  - end time
  - label name
- fg-add
  - label name
  - framework hadoop
  - version 1.0

- Deploys an image on a host
- Adds a feature to a deployed image

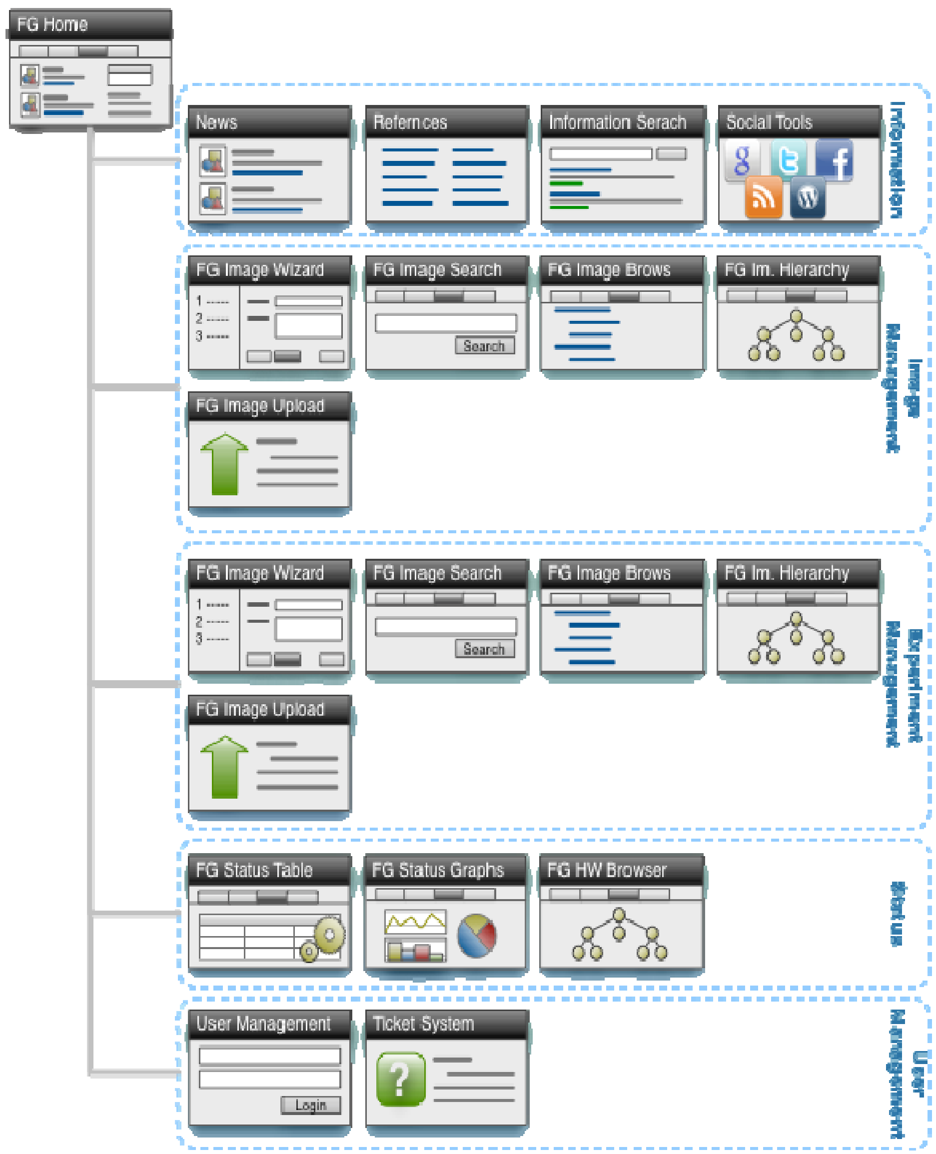
xCAT and Moab

- xCAT
  - uses installation infrastructure to perform installs
  - creates stateless Linux images
  - changes the boot configuration of the nodes
  - remote power control and console (IPMI)
- Moab
  - meta-schedules over resource managers
    - TORQUE and Windows HPC
  - control nodes through xCAT
    - changing the OS
    - remote power control

Experiment Manager

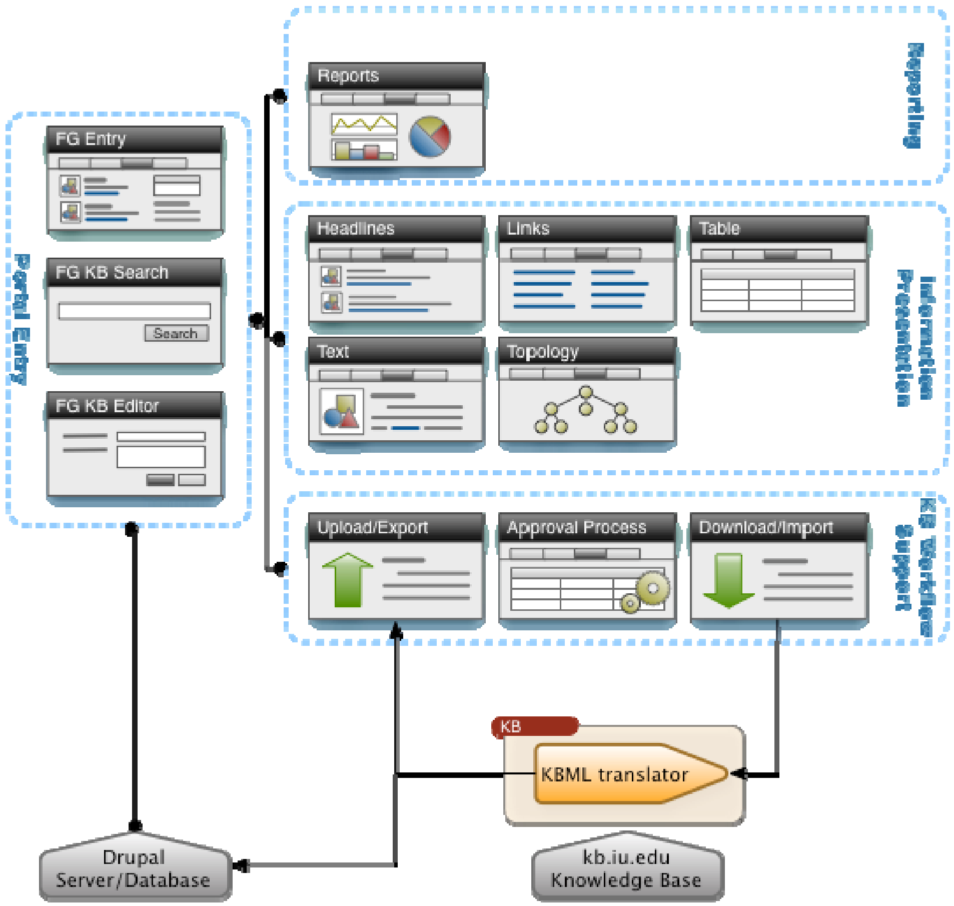
- Objective
  - Manage the provisioning for reproducible experiments
  - Coordinate workflow of experiments
  - Share workflow and experiment images
  - Minimize space through reuse
- Risk
  - Images are large
  - Users have different requirements and need different images

User Portal

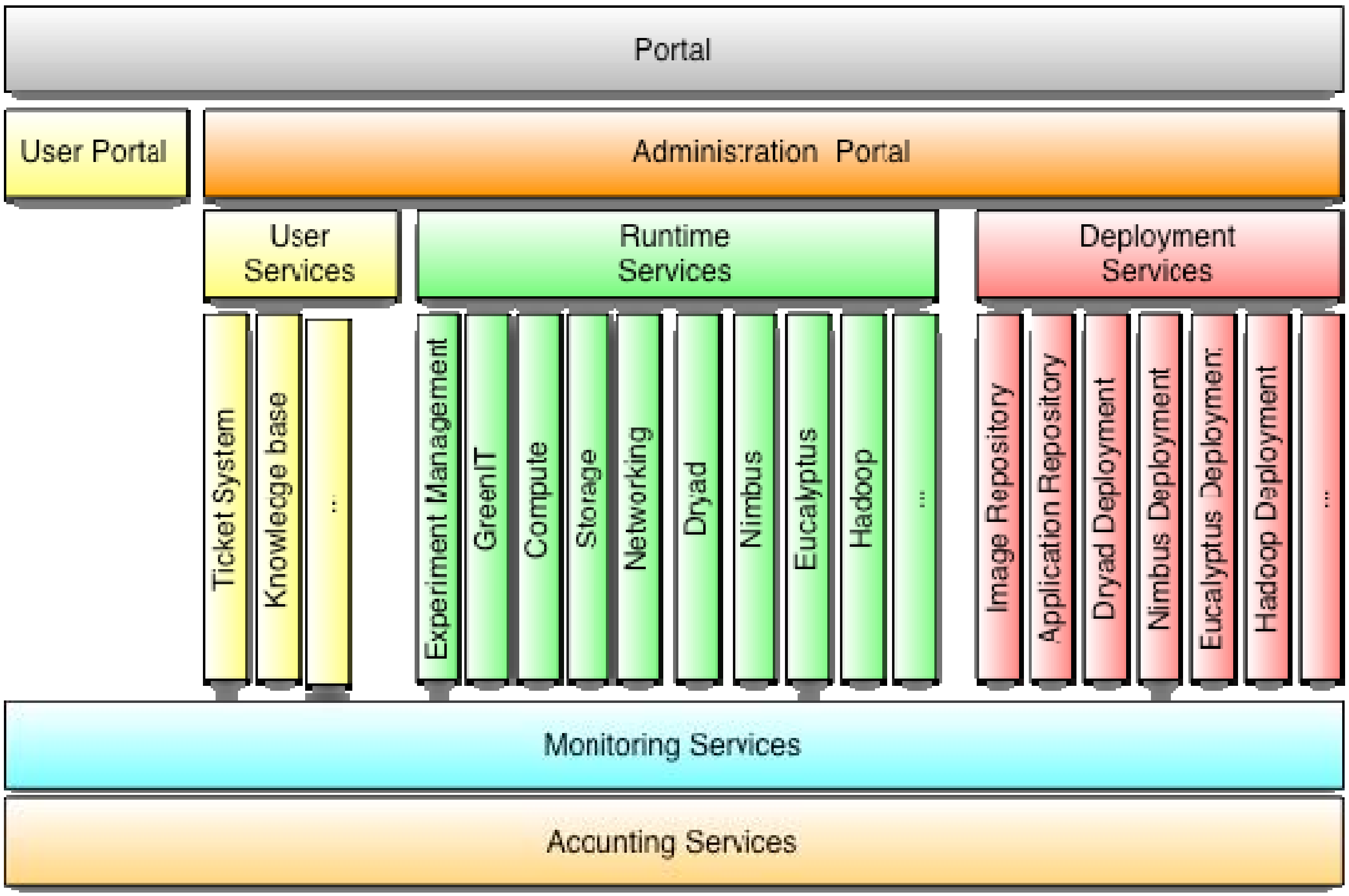


FG Home
News
Refernces
Information Serach
Social Tools
Information
FG Image Wizard
FG Image Search
FG Image Brows
FG Im. Hierarchy
Search
FG Image Upload
Image Management
FG Image Wizard
FG Image Search
FG Image Brows
FG Im. Hierarchy
Search
FG Image Upload
Experiment Management
FG Status Table
FG Status Graphs
FG HW Browser
Status
User Management
Ticket System
Login
User Management

FG Information Portal



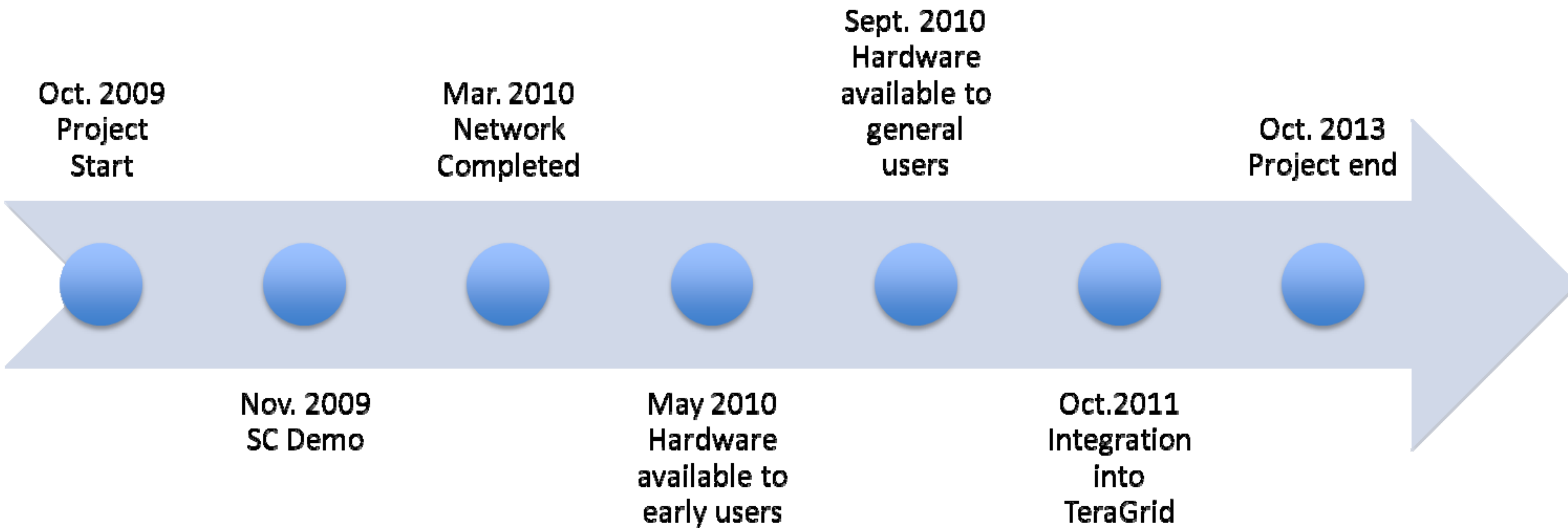
FG Administration Portal



Integration within TeraGrid / TeraGrid XD

- Sure it's part of TeraGrid
- Allocation: separate from TG processes for two years
- It is a very exciting project, it will teak effort
  - TG may change, good that we can wait
- We are looking for early adopters!

Milestones



Oct. 2009 Project Start
Nov. 2009 SC Demo
Mar. 2010 Network Completed
May 2010 Hardware available to early users
Sept. 2010 Hardware available to general users
Oct.2011 Integration into TeraGrid
Oct. 2013 Project end

Questions

- Please contact Gregor via mail
  - laszewski@gmail.com
- We will update our web page soon
  - http://futuregrid.org