Grid Computing: **What Is It, and Why Do I Care?***

Ken MacInnis

<ken.macinnis@gmail.com>

* Or, “Mi caja es su caja!”

Outline

- Introduction and Motivation
- Examples
- Architecture, Components, Tools
- Lessons Learned and The Future
- Questions?

What is “grid computing”?

- Many different definitions:
  - Utility computing
    - Cycles for sale
  - Distributed computing
    - distributed.net RC5, SETI@Home
  - High-performance resource sharing
    - Clusters, storage, visualization, networking

“We will probably see the spread of ‘computer utilities’, which, like present electric and telephone utilities, will service individual homes and offices across the country.”

Len Kleinrock (1969)

- The word “grid” doesn’t equal Grid Computing:
  - Sun Grid Engine is a mere scheduler!

Better definitions:

- ***Common protocols*** *allowing large problems to be solved in a* ***distributed multi-resource multi-user environment****.*

"A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities."

Kesselman & Foster (1998)

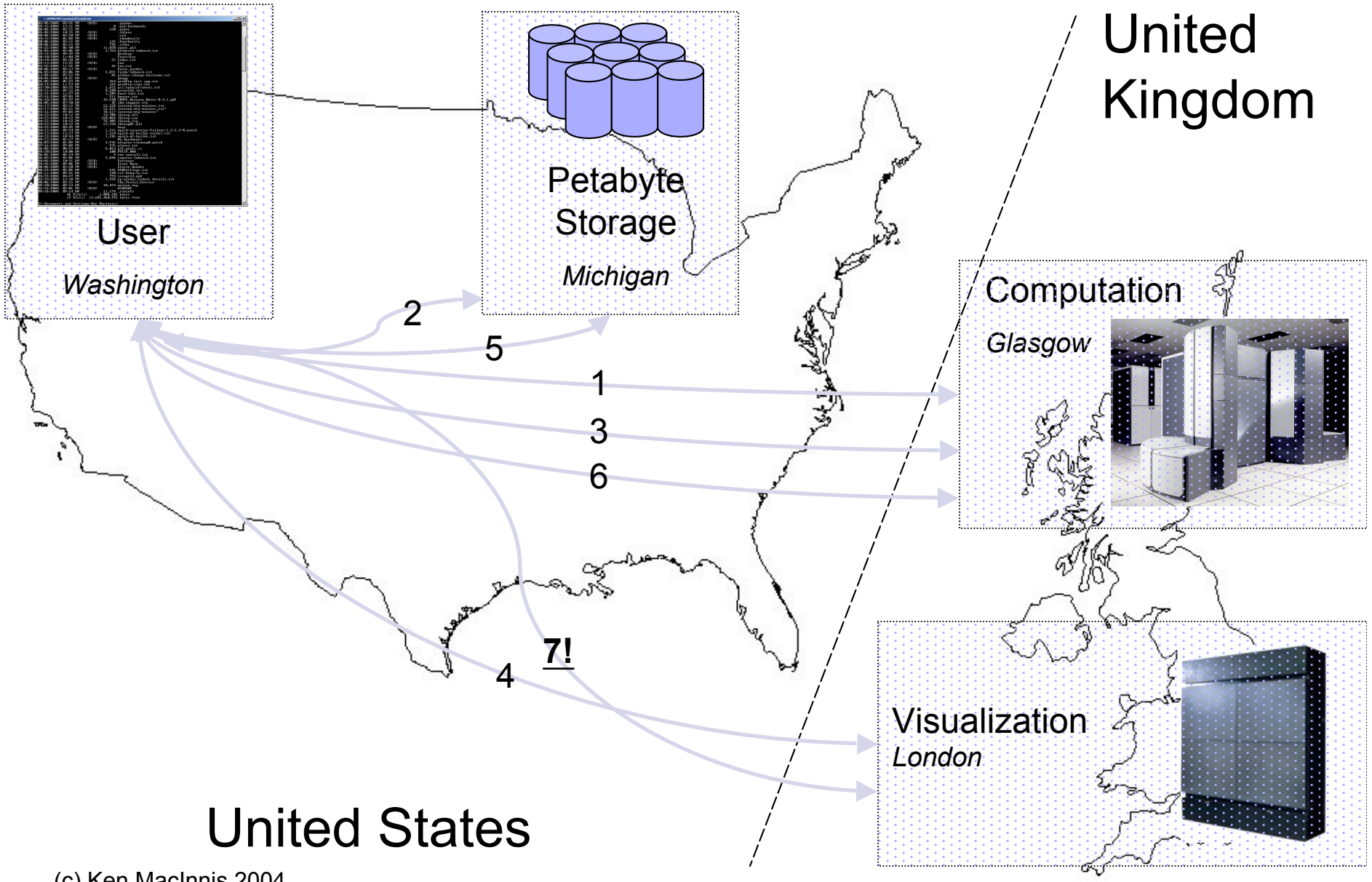
"…coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations."

Kesselman, Foster, Tuecke (2000)

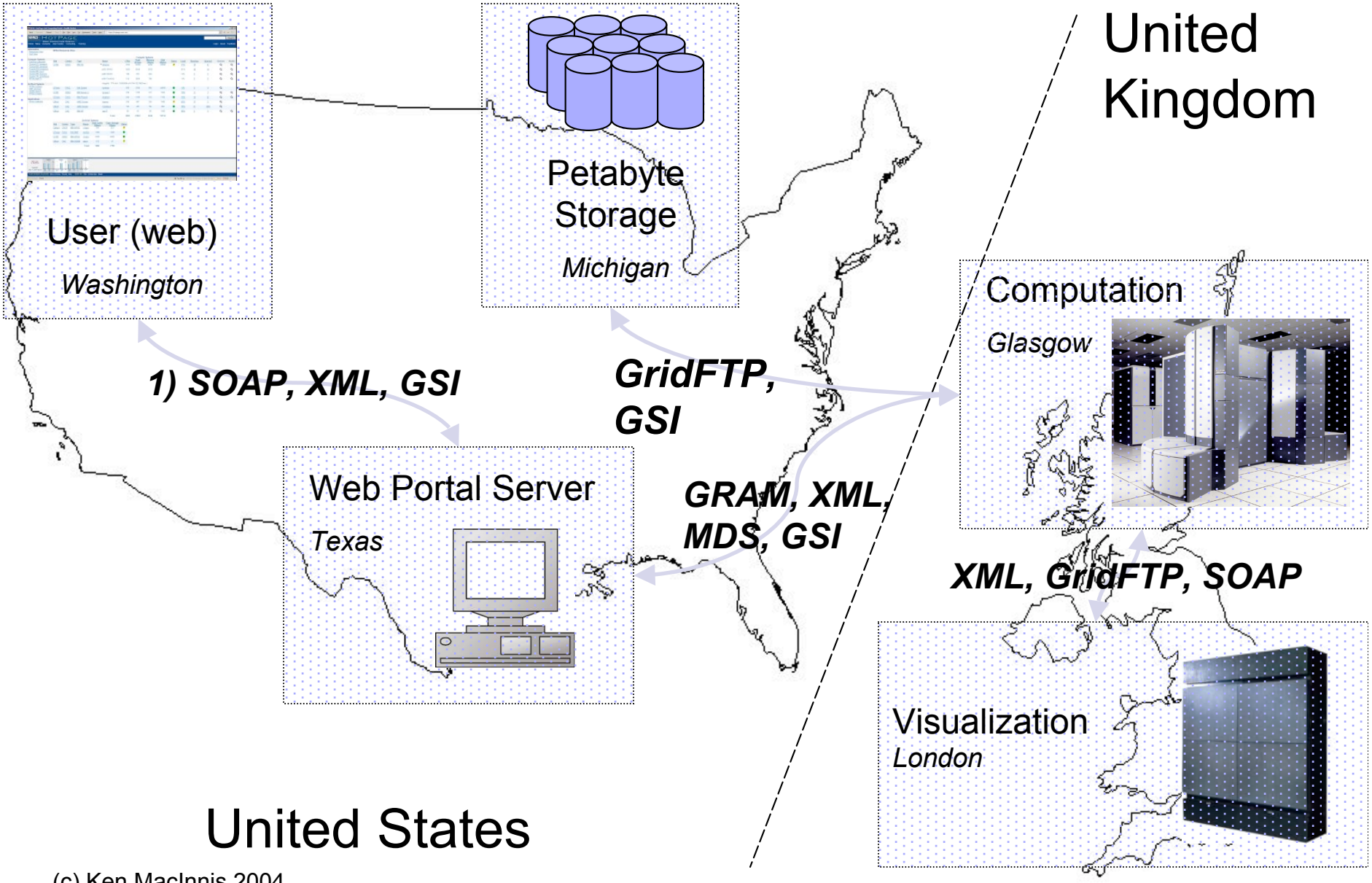
New Challenges for Computing

- Grid computing evolved out of a need to share resources
  - Flexible, ever-changing “virtual organizations”
    - High-energy physics, astronomy, more
  - Differing site policies with common needs
  - Disparate computing needs
    - Utility computing provides cycles only when needed

Traditional Architecture



Grid Architecture



Outline

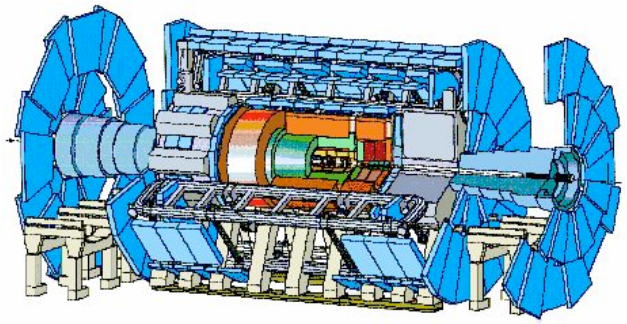
- Introduction and Motivation
- **Examples**
- Architecture, Components, Tools
- Lessons Learned and The Future
- Questions?

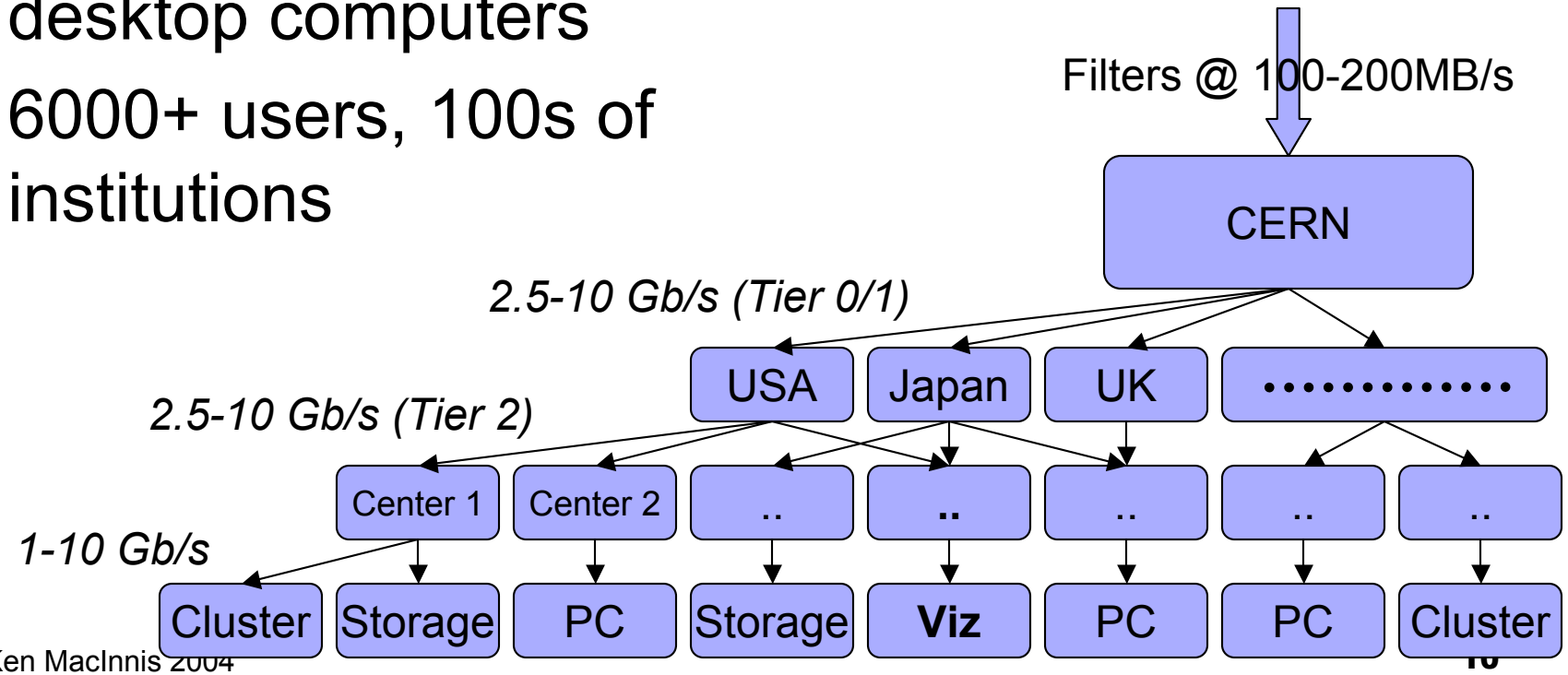
Examples

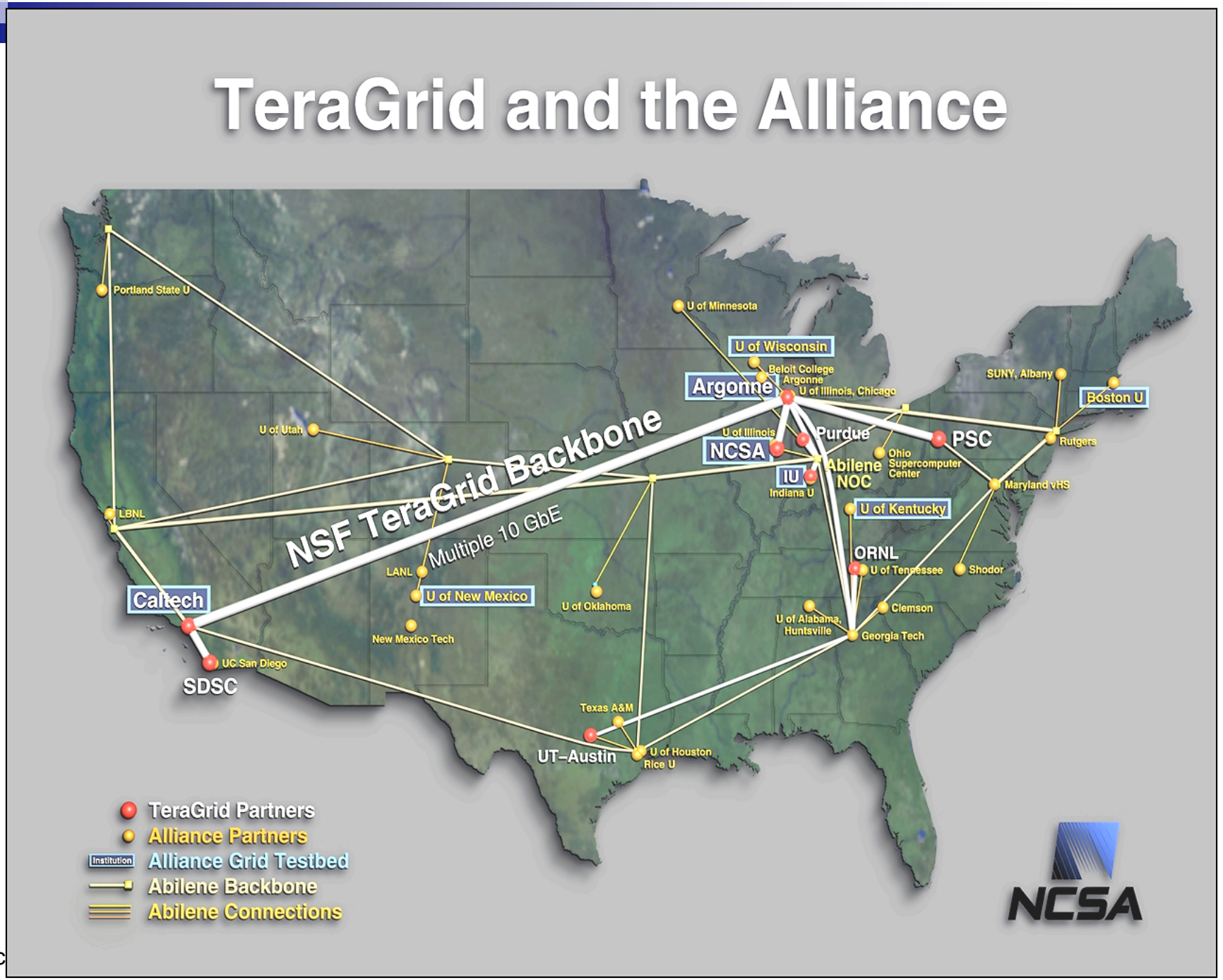
Project	***Description***
AccessGrid	Group collaboration, distance learning
DOE Science Grid	General purpose resource sharing
TeraGrid	Distributed “Terascale” computing
NEESgrid	Earthquake analysis
GridMP Global	United Devices computational projects
iVDGL	International physical science
Fraunhofer Resource Grid	German research grid projects
UK e-Science	UK collaboratory backbone
Enterprise Grid Alliance	Commercial grid development

CERN Large Hadron Collider Computing Grid

- 12-14 petabytes data per year
- Equivalent of 70,000 desktop computers
- 6000+ users, 100s of institutions






TeraGrid and the Alliance
NSF TeraGrid Backbone
Multiple 10 GbE
Portland State U
U of Minnesota
U of Wisconsin
Beloit College
Argonne
Argonne
U of Illinois, Chicago
SUNY, Albany
Boston U
U of Utah
U of Illinois
NCSA
Purdue
PSC
Rutgers
Ohio Supercomputer Center
Abilene NOC
IU
Indiana U
Maryland vHS
LBNL
U of Kentucky
ORNL
U of Tennessee
Shodor
LANL
U of New Mexico
U of Oklahoma
Caltech
Clemson
U of Alabama, Huntsville
Georgia Tech
New Mexico Tech
UC San Diego
SDSC
Texas A&M
UT–Austin
U of Houston
Rice U
TeraGrid Partners
Alliance Partners
Institution
Alliance Grid Testbed
Abilene Backbone
Abilene Connections
NCSA

(c

Example – Electronic Arts

- Electronic Arts, “Sims Online”
  - Grid supporting 250k players
  - 30,000 SQL calls per second per cluster
  - Backend is scalable with player count
  - Inherently highly available

Outline

- Introduction and Motivation
- Examples
- **Architecture, Components, Tools**
  - Security
  - Globus
  - Data movement
  - Condor, MPI
  - Web portals
- Lessons Learned and The Future
- Questions?

Architecture – Software
Tools and Components

- Security: X.509
- Toolkits:
  - Globus Toolkit 2.x
  - Globus Toolkit 3.x+
- Job Scheduling: Condor, Condor-G
- Parallel Programming: MPICH-G2

X.509 Security

- Public/private certificate system
  1. User creates private certificate, and public certificate request
  2. Certificate Authority verifies user's identity
  3. Certificate Authority provides a signed response (public certificate) to request
  4. Any site with CA's public key can now verify user's ID (chain of trust)
  - Does not require pre-staging a public key, i.e. SSH
  - Grid Security Infrastructure (GSI) proxy can then be used as a "placeholder"
    - Single sign on
    - Impersonation / delegation

X.509 Security *cont.*

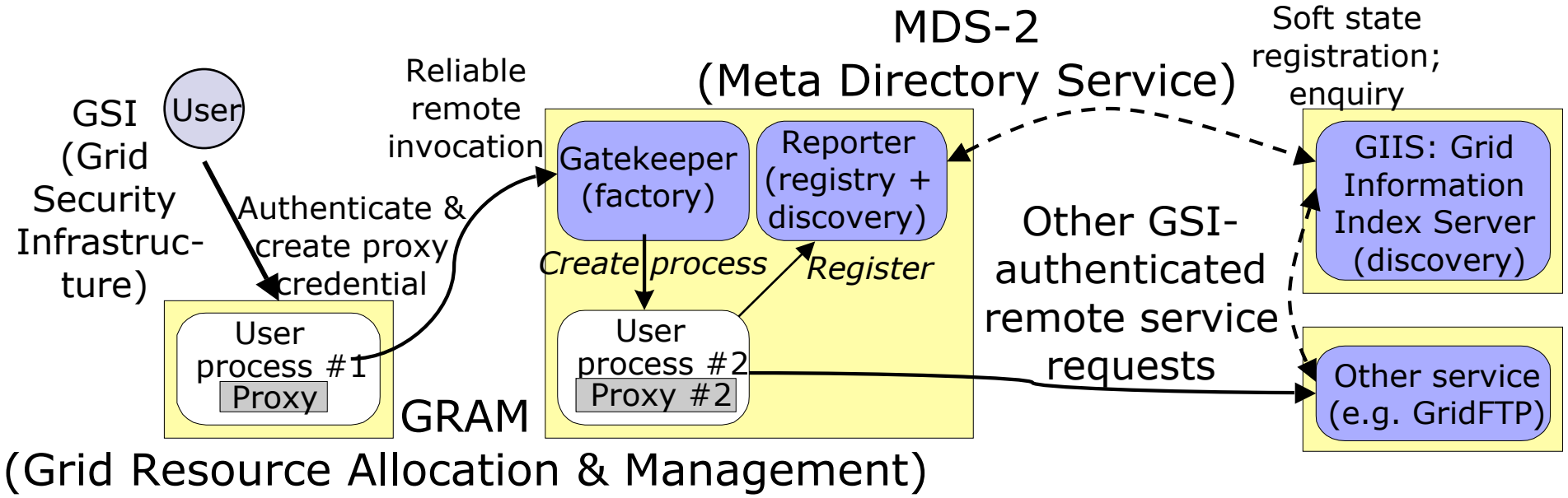
- *Certificate Authority* is the atom of trust
  - Each Virtual Organization may have >=1 CA
  - No single CA can compromise all certificates
  - Each CA can belong to multiple Virtual Organizations
- Your *Distinguished Name* [DN] is your Grid ID
  - Remote site performs DN-UID mapping at will
- Trust revocation
  - Each CA publishes Certificate Revocation List [CRL]
  - Online Certificate Status Protocol (OCSP) responders

Globus 2.x - *Motivation*

- Globus is today's de-facto grid toolkit
  - Also ***only*** toolkit implementation (well, almost)
- Provides many of the "Grid" protocols
- Open source, open framework!
- Builds on Linux, FreeBSD, Mac OS X, even Tru64!
  - Windows? Portability is a problem.

The Globus Toolkit 2.x in One Slide

- Grid protocols (GSI, GRAM, …) enable resource sharing within virtual orgs; toolkit provides reference implementation (= Globus Toolkit services)



Globus 2.x – *What's Right*

- Strong security implementation
  - GSI provides good delegation support
- Traditional supercomputing works well
  - Built for "jobs" and "clusters"
  - High-speed data transfer
- Excellent base for further community development
  - PyGlobus, GridFTP clients, etc.

Globus 2.x – *What's Wrong*

- First pass at a "Field of Dreams":
  - Not everything is clusters and jobs: sensors, instruments, networks, databases
  - The "whole system" is left out of the picture
- Large and complex to install and support
  - End users cannot configure on their own
  - Lots of memory leaks
  - C and Perl

Globus 3.x

- Globus Toolkit 3 introduced Grid Services
  - Grid services == Web services with state
  - **Open Grid Services Architecture (OGSA)**
    - Protocols and formats
  - **Open Grid Service Infrastructure (OGSI)**
    - WSDL specifications defining standard interfaces, behaviors, and schema implementing OGSA
  - Java, .NET (as well as C/C++)

Globus 4.x

- Globus Toolkit 4 introduces **WSRF**
  - **Web Services Resource Framework**
    - "Web services for grid computing"
  - Announced Jan 2004
    - Collaboration of Globus Alliance, IBM, HP, Akamai, SAP, others.
  - Industry backing shows the Grid is here to stay!

Data Movement – *GridFTP*

- High performance, secure, reliable data transfer built on FTP
- Features built for distributed systems:
  - **GSI security on control and data channels**
  - Multiple data channels for parallel transfers
  - Partial file transfers
  - **Third-party (direct server-to-server) transfers**
  - Authenticated data channels
  - Reusable data channels
  - Command pipelining

Data Movement - *GridFTP*

- Existing FTP clients
  - Normal FTP clients simply add GSI authentication
  - Full-featured clients include third-party transfers, additional transfer modes, extra features
- Upcoming features
  - Java rewrite (server was based on wuftpd)
  - Striped data transfers

Data Movement – *SRB/MCAT*

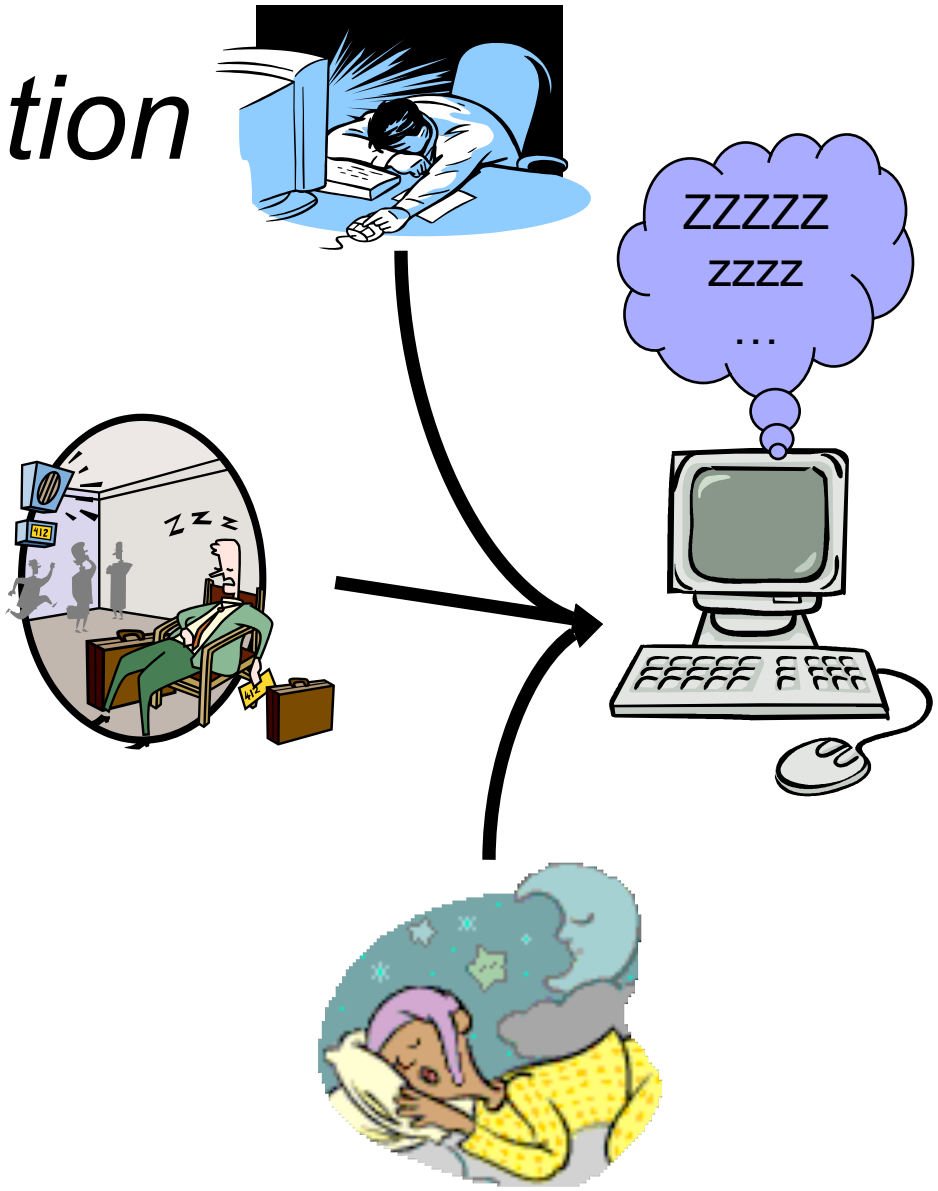
- Storage Resource Broker, Metadata Catalog
  - “A uniform interface for connecting to heterogeneous data resources over a network and accessing replicated data sets.”
- Provides additional tools for:
  - Collaboration
  - High performance movement
  - Backing storage awareness
  - C, C++, Java, Perl, Python bindings
- **What is SRB?**
  - SRB is a Distributed File System
  - SRB is a Data Grid Management System
  - SRB is a Digital Library
  - SRB is a Semantic Web

Data Movement – *Globus RLS*

- ***R**eplica **L**ocation **S**ervice*
- Provides access to logical names for data items to target names
- Example:
  - Distributed databases can have multiple repositories
  - RLS provides mapping to closest location

Condor - *Motivation*

- Modern computing is highly distributed
- Most desktops are lightly loaded
  - Little use: typing documents, reading email
  - No use: at a meeting, overnight, vacation
- We should harness these wasted cycles!
- ***Cycle Scavenging (Cycle Sucking)***



Condor – *The Fundamentals*

- Condor is a distributed computing job manager
  - Set of daemons on each workstation allows for job management
- *Metrics:* Keyboard use, load, time of day
  - User-definable, highly expandable
- *Checkpointing:* Suspend, restart, move
  - User returns to keyboard, job fails, etc.
- Condor-G (Grid):
  - Submission front-end to Grid resources
    - Meta-scheduling!
  - Maintains credentials (auto-renewal, etc.)
  - Crash-resistant: Persistent state allows for graceful error recovery

Condor – *Summary*

- Condor's benefits:
  - Flexible cycle sucking to utilize all resources
  - Integrates with non-Condor Grid resources
  - Advanced workflow and check pointing tools
  - Portable: Windows, UNIX, Java support
- Condor's drawbacks
  - Configuration is mystifying to the neophyte
  - No source (yet) -- promised

MPICH-G2

- MPI – *Message Passing Interface*
- MPICH-G2 is grid-enabled MPI v1.1
  - GSI security
  - Distributed job startup
  - Heterogeneous architectures
- In other words:
  - Grid-enabled resources can compute as one MPI group
- Except:
  - High-latency interconnect now!

Web Portals

- Web portal definition:
  - *Loosely integrated features bundled into an efficient, common interface presented through an accessible medium (the WWW)*
- Examples:
  - Yahoo is more than just a search engine: games, finance, mail, more!

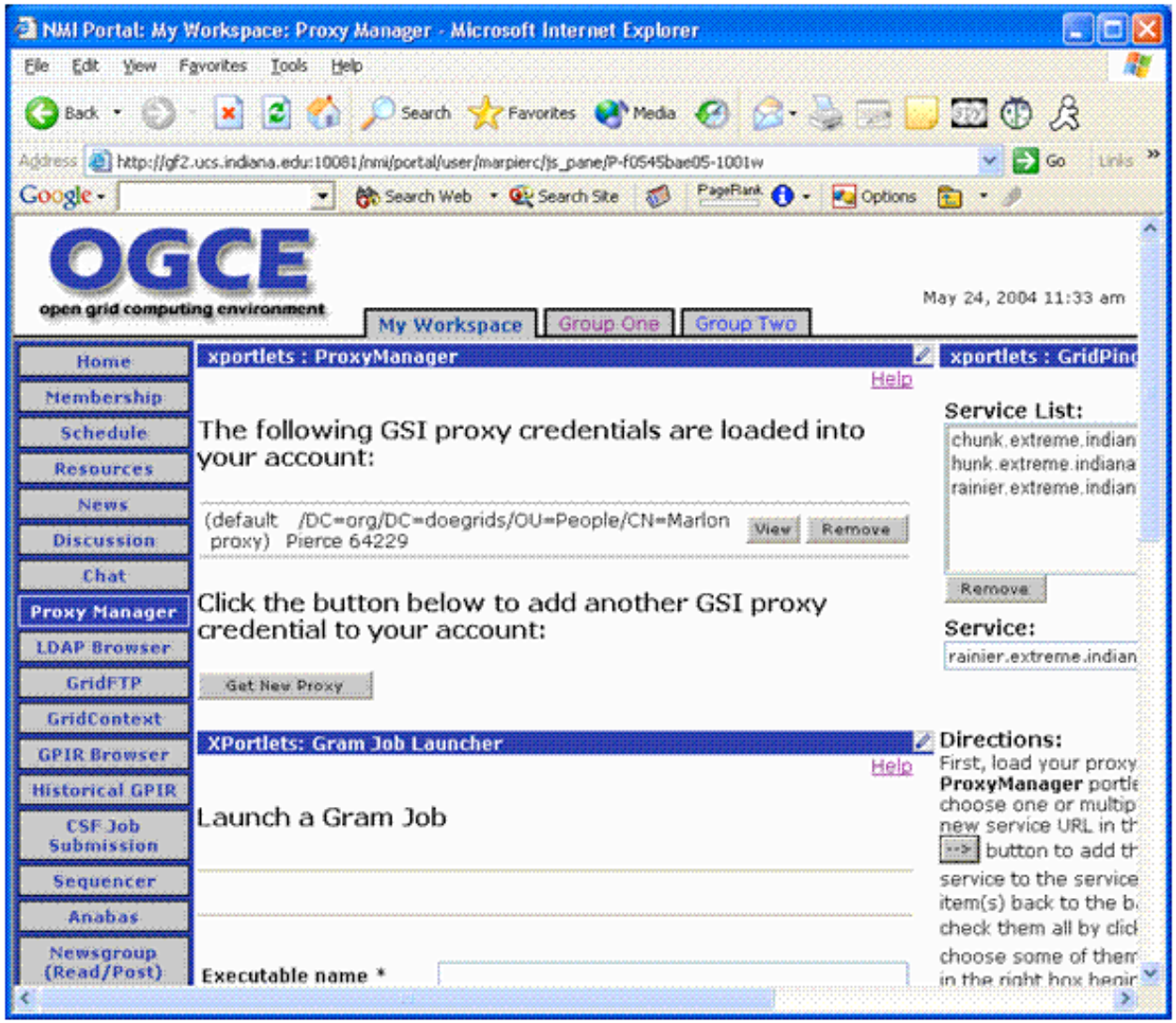
Web Portals - *Motivation*

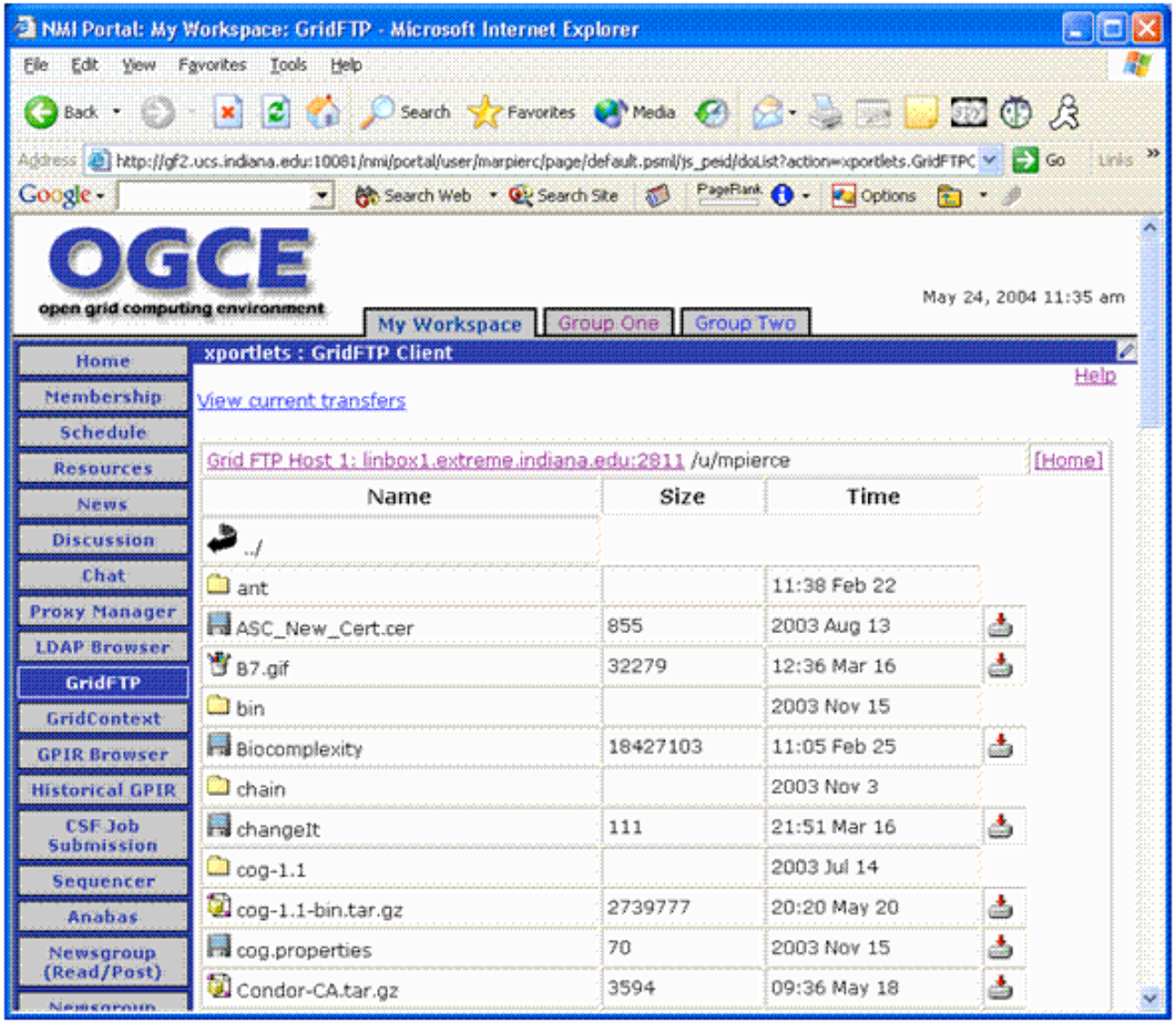
- Grid tools are too difficult to maintain:
  - Installation, maintenance, upgrades
  - Globus is truly one of Dante's nine rings
- Also want
  - Consistent interface
  - No prerequisites past a web browser
  - One stop shopping
  - Access anywhere

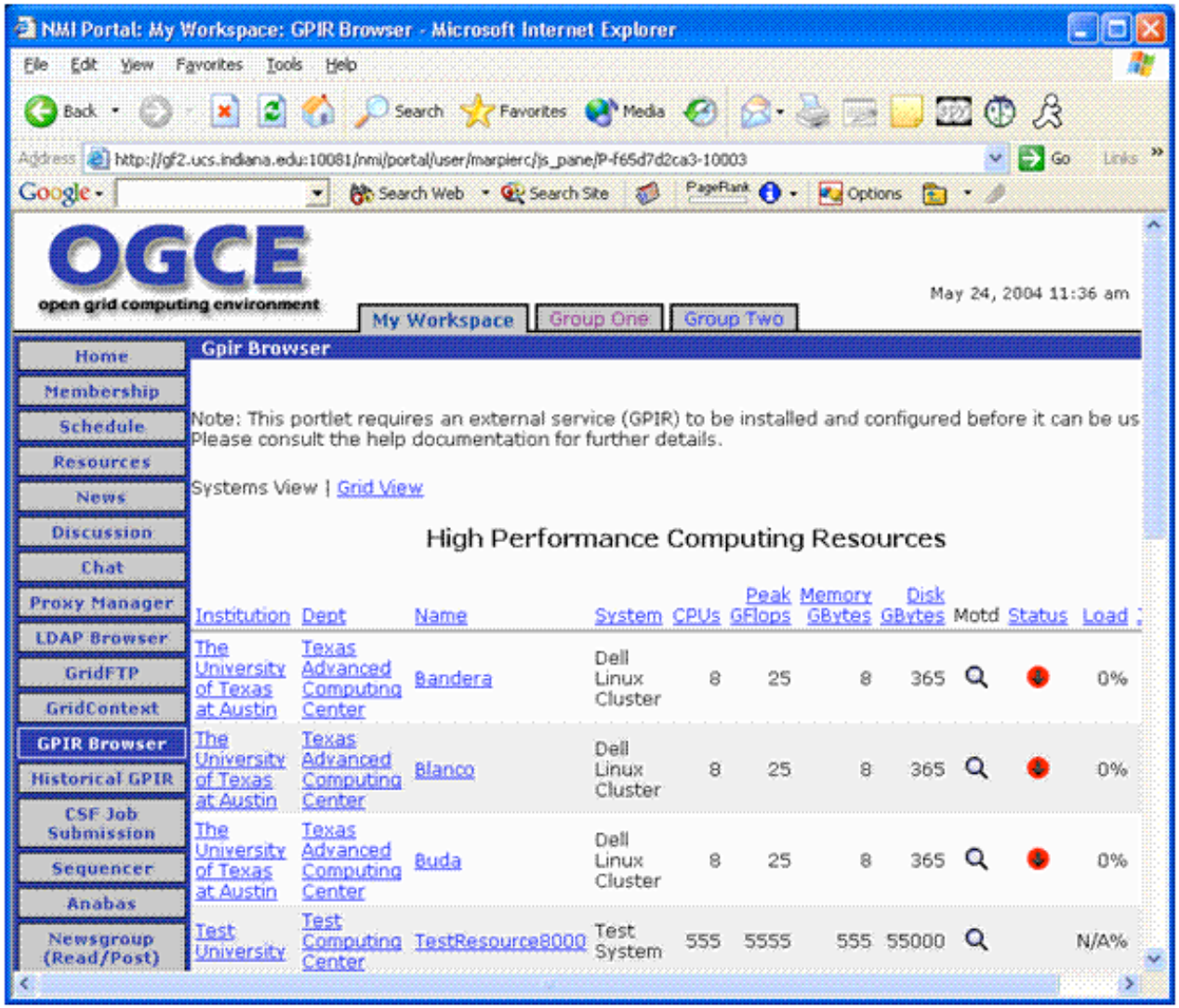
Web Portals - *Example*

Web Portals - *Technology*

- OGCE: **Open Grid Computing Environment**
  - Java-based Grid Portal toolkit framework
  - Supports job and credential management (MyProxy)
  - Supports GridFTP and Condor-G
  - Extensible
  - The NSF supported grid portal project
  - Requirements: Java, Tomcat, and Jetspeed
- Why a framework?
  - Frameworks allow for rapid prototyping
  - One portal will not meet all needs, thus..
  - Multiple portals for multiple projects


NMI Portal: My Workspace: Proxy Manager - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://gf2.ucs.indiana.edu:10081/nmi/portal/user/marpierc/js_pane/P-f0545bae05-1001w
Go Links
Google Search Web Search Site PageRank Options
OGCE
open grid computing environment
May 24, 2004 11:33 am
My Workspace
Group One
Group Two
Home
Membership
Schedule
Resources
News
Discussion
Chat
Proxy Manager
LDAP Browser
GridFTP
GridContext
GPIR Browser
Historical GPIR
CSF Job Submission
Sequencer
Anabas
Newsgroup (Read/Post)
xportlets : ProxyManager
Help
The following GSI proxy credentials are loaded into your account:
(default proxy) /DC=org/DC=doegrids/OU=People/CN=Marlon Pierce 64229
View
Remove
Click the button below to add another GSI proxy credential to your account:
Get New Proxy
XPortlets: Gram Job Launcher
Help
Launch a Gram Job
Executable name *
xportlets : GridPinc
Service List:
chunk.extreme.indian
hunk.extreme.indiana
rainier.extreme.indian
Remove
Service:
rainier.extreme.indian
Directions:
First, load your proxy
ProxyManager portle
choose one or multip
new service URL in th
--> button to add th
service to the service
item(s) back to the b
check them all by clic
choose some of them
in the right box begir


NMI Portal: My Workspace: GridFTP - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://gf2.ucs.indiana.edu:10081/nmi/portal/user/marpierc/page/default.psml/js_peid/doList?action=xportlets.GridFTPC
Go Links
Google Search Web Search Site PageRank Options
OGCE
open grid computing environment
May 24, 2004 11:35 am
My Workspace
Group One
Group Two
Home
Membership
Schedule
Resources
News
Discussion
Chat
Proxy Manager
LDAP Browser
GridFTP
GridContext
GPIR Browser
Historical GPIR
CSF Job Submission
Sequencer
Anabas
Newsgroup (Read/Post)
xportlets : GridFTP Client
Help
View current transfers
Grid FTP Host 1: linbox1.extreme.indiana.edu:2811 /u/mpierce
[Home]
Name
Size
Time
../
ant
11:38 Feb 22
ASC_New_Cert.cer
855
2003 Aug 13
B7.gif
32279
12:36 Mar 16
bin
2003 Nov 15
Biocomplexity
18427103
11:05 Feb 25
chain
2003 Nov 3
changeIt
111
21:51 Mar 16
cog-1.1
2003 Jul 14
cog-1.1-bin.tar.gz
2739777
20:20 May 20
cog.properties
70
2003 Nov 15
Condor-CA.tar.gz
3594
09:36 May 18


NMI Portal: My Workspace: GPIR Browser - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://gf2.ucs.indiana.edu:10081/nmi/portal/user/marpierc/js_pane/P-f65d7d2ca3-10003
Go Links
Google Search Web Search Site PageRank Options
OGCE
open grid computing environment
May 24, 2004 11:36 am
My Workspace
Group One
Group Two
Home
Membership
Schedule
Resources
News
Discussion
Chat
Proxy Manager
LDAP Browser
GridFTP
GridContext
GPIR Browser
Historical GPIR
CSF Job Submission
Sequencer
Anabas
Newsgroup (Read/Post)
Gpir Browser
Note: This portlet requires an external service (GPIR) to be installed and configured before it can be us
Please consult the help documentation for further details.
Systems View | Grid View
High Performance Computing Resources
Institution Dept Name System CPUs Peak GFlops Memory GBytes Disk GBytes Motd Status Load
The University of Texas at Austin Texas Advanced Computing Center Bandera Dell Linux Cluster 8 25 8 365 0%
The University of Texas at Austin Texas Advanced Computing Center Blanco Dell Linux Cluster 8 25 8 365 0%
The University of Texas at Austin Texas Advanced Computing Center Buda Dell Linux Cluster 8 25 8 365 0%
Test University Test Computing Center TestResource8000 Test System 555 5555 555 55000 N/A%


NMI Portal: My Workspace: Condor - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://gf2.ucs.indiana.edu:10081/nmi/portal/user/marpierc/js_pane/P-f65d7d2ca3-10018
Go
Google Search Web Search Site PageRank Options
OGCE
open grid computing environment
May 24, 2004 11:40 am
My Workspace
Group One
Group Two
Home
Membership
Schedule
Resources
News
Discussion
Chat
Proxy Manager
LDAP Browser
GridFTP
GridContext
GPIR Browser
Historical GPIR
CSF Job Submission
Sequencer
Anabas
Newsgroup (Read/Post)
XPortlets: Condor Portlet
Condor Portlet
Condor
High Throughput Computing
Submit Condor Jobs
Submit Condor DAGMan Jobs
Monitor Submitted Jobs

Tools summary – *Globus Toolkit*

- Globus 2.x great for “traditional supercomputing”
  - Pioneered many protocols, still widely used
- Globus 3.x introduced service model
- Globus 4.x well-backed in industry
  - WSRF, Grid Services

Tools summary – *Data Movement*

- GridFTP provides FTP compatibility with GSI authentication, third-party transfers, high performance
- SRB/MCAT provide *resource brokerage* services as well as *metadata cataloging*
- RLS is an efficient data *replication* service

Tools summary – *Condor, MPICH-G2*

- Condor
  - High performance distributed computing via cycle scavenging
  - Availability metrics
  - Multi-platform checkpointing
  - Lots of additional neat tools: DAGMan, Condor-G
- MPICH-G2
  - Harness multiple MPI resources as one
  - High latency, tough configuration, single programming model

Tools summary - *Windows*

- Using Windows in grid computing:
  - Portals are key
  - Condor can harness computational power (cluster-ize)
  - Globus grid services are now available for Windows, but no “old-style” services

Outline

- Introduction and Motivation
- Examples
- Architecture, Components, Tools
- **Lessons Learned and The Future**
- Questions?

Political Challenges

- Unsolved problems are politically-based
  - Federation in Certificate Authorization, e.g.
- Purpose-built grids are successful
  - “Field of Dreams” does not work here, yet
- Users are set in their ways
  - Grid technology is a big change requiring a paradigm shift!
- Virtual Organization partner relationships
  - Security response policies, compromises, software stacks
  - Who is responsible for what? Many “bosses”

Political Challenges *cont.*

- Security is hard:
  - GSI helps this as a single base, but:
    - Common authentication
    - Accounting
    - CA policy
  - Cross-site trust is hard! No solutions yet.
  - Why political?
    - “Grid” security is just security + X.509
    - Rough spots are in the inter-realm areas

Technical Challenges

- Software stacks:
  - Common base of interoperable software
  - Synchronized upgrades?
- Support:
  - Is problem local or remote? How does user tell?
- Network issues:
  - No one group to control commodity network between
- Standards
  - The Global Grid Forum
  - Lots of proposed standards, loosely collected
  - No O'Reilly / Dummies books here!

Words of Advice

- Policy first!
  - Make sure everyone is in agreement
  - “Fire plan” (playbook) especially important
  - Agree on software policies as well as versions
- Define goals to meet
  - “Grid Computing” is task oriented, not a general provision

So, What Is The Grid, Then?

- “The Grid” is a collection of applications, protocols, policies, and procedures which allow collaboration to occur in an efficient manner
- It’s everywhere!
  - Modern systems have global aspirations
  - Design with extensibility in mind

And Why Do I Care?

- Collaboration is the name of the game
- Use unused resources efficiently
  - Tragedy of the Commons
- Resource sharing is new key to funding
- Most importantly..
  - It's a state of mind! Get into the habit of building resources with "The Grid" from the beginning.

Summary

- What is Grid?
  - *Common protocols allowing large problems to be solved in a distributed multi-resource multi-user environment.*
- Why Grid?
  - Exponential increase in computing needs
  - Collaboration extremely important
- Modern Grids
  - Industry/academia, large/small, focused/utility

Summary *cont.*

- Grid architecture
  - Security
    - X.509 certificates
    - GSI
  - Tools
    - Globus 2.x, 3.x, 4.x
    - Web/Grid Services
    - Condor, MPI / MPICH-G2
    - Web portals
    - Data movement

Summary *cont.*

- Challenges
  - Political challenges
  - Technical challenges
- Advice
  - Start early
  - Policy
  - Well-defined goals

Questions?

Ken MacInnis

ken.macinnis@gmail.com