Grid in what sense?

People Interactions

Identify available resources and request allocation
User Documentation, User Portal, POPS
Learn how to use resources and their status
User Documentation, Knowledge Base
User Portal
Conferences and other events
Ask for help
E-mail
Telephone

(User) Software Interactions

Standard service interfaces (login, move data, run jobs, WAN file-systems, manage data, etc.)
Coordinated software (Grid clients, development tools, science workflow, etc.)
Coordinated Unix interaction (standard variables, SoftEnv)
Provide an Information Services Infrastructure:

- Applying grid concepts to information publishing
- Have RP/GIG/partner operated information services
- Centrally indexing and/or aggregated for discovery
- Primarily focused on public information
- Primarily accessible thru software interfaces
- Using standards based interfaces
- Reliable, scalable, and fast
- Publishing TeraGrid information
- And partner/community information
TG Information Services IS [NOT]

<table>
<thead>
<tr>
<th>IS NOT</th>
<th>IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A central database (Data Warehouse)</td>
<td>A central index/aggregation (Google)</td>
</tr>
<tr>
<td>A new user interface</td>
<td>A way user interfaces access information</td>
</tr>
<tr>
<td>A single implementation/tool</td>
<td>Includes several tools</td>
</tr>
<tr>
<td>A single software interface</td>
<td>Accessed using several useful interfaces</td>
</tr>
<tr>
<td>A specific set of data</td>
<td>Phased growing set of data</td>
</tr>
<tr>
<td>Changed data ownership</td>
<td>Ownership maintained as appropriate</td>
</tr>
<tr>
<td>Way to manage scientific information</td>
<td>Way to manage Grid meta-data</td>
</tr>
<tr>
<td>A data management system (database)</td>
<td>An information publishing system</td>
</tr>
</tbody>
</table>

A coordinated way to index and publish public [Tera]Grid information using software interfaces.
High-Level Components

TeraGrid Wide Services

Resource Provider Services

Clients

WS/REST HTTP GET

Apache 2.0

Tomcat
WebMDS

WS MDS4

Cache

WS MDS4

WS/SOAP

Clients

Clients

TeraGrid Repositories

Partners

TeraGrid Repositories

Partners
TeraGrid Wide vs RP Services

RP Information Services

Content:
- RP owned and maintained information
- Data can originate in local systems

Infrastructure:
- 2 scheduling MDS services: authenticated and public (merging)
- 1 general purpose MDS service

TeraGrid Wide Information Services

Content:
- Aggregate/index RP Information Services content
- Additional central information (TGCDB, GIG operated services, …)

Infrastructure:
- Several redundant servers
- Information caching (persistence)
- Several MDS4 services (WS/SOAP)
- WebMDS/Tomcat, Apache 2.0, … (WS/REST)
- Content published in: HTML, XHTML/XML, XML, Atom, RSS, …
WS/* (Tomcat 5.0, Apache 2.0)

Benefits
- Very common web services platform
- Supports several web service interfaces (including simple)
- Supports multiple styles like REST, Web 2.0
- Can be highly scalable

Content
- Many formats: HTML, XHTML/XML, XML, Atom, RSS, ...

WS/SOAP (Globus 4.0.5/VDT 1.7.1 MDS4)

Benefits
- Indexing, Trigger
- Registration, Publish, Subscribe
- Security/Authorization
- Robust WSRF interface

Content
- XML

WebMDS

Benefits
- XPath support
- XSLT transforms
High-Availability Design

TeraGrid Dynamic DNS

Dynamically direct clients to one or more servers
Set by Information Services administrators
Changes propagate globally fast (TTL = 15 minutes)

(info.teragrid.org)

(info.dyn.teragrid.org)

Dynamically Changes

Doesn’t Change

RP/partner services

TG wide servers

Clients

(Patrick Dorn & NCSA NetEng)
Information Services Users

User Documentation (Michael, Diana)

User Portal (Maytal & team)

Inca (Kate & team)

Database?

info.teragrid.org

Gateways

Peer Grids

User Applications
Completed Milestones

Infrastructure information services are in production
- RP Information Services
- TeraGrid Wide Information Services

Scheduling information published to User Portal (since Spring)
- Scheduler load
- Queue contents

CTSS 4 capabilities published in publicly (Since August)
- Accessed by
  - User Documentation
  - Inca system
- Content
  - Which capability kits are available on each resource
  - What software is available in each kit on each resource
  - What services are available from each kit on each resource
## Queue Contents in User Portal

### TeraGrid User Portal

- **Home**
- **My TeraGrid**
- **Resources**
- **Documentation**
- **Training**
- **Consulting**
- **Allocations**

### Systems Monitor

- **Science Gateways**
- **Data Collections**
- **HPC Queue Prediction [Beta]**
- **Remote Visualization [Beta]**
- **User Responsibilities**

#### TeraGrid Systems Monitor

- **Job summary for login-abe.ncsa.teragrid.org:**
  - **47 Running Jobs**
  - **68 Queued Jobs**
  - **50 Other Jobs**

#### 47 Running Jobs on login-abe.ncsa.teragrid.org

<table>
<thead>
<tr>
<th>Status</th>
<th>Job Id</th>
<th>Name</th>
<th>Owner</th>
<th>Queue</th>
<th>Submission Time</th>
<th>Processors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running</td>
<td>35721.abem5.ncsa.uiu</td>
<td>d_1-40</td>
<td>petefred</td>
<td>normal</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Running</td>
<td>35730.abem5.ncsa.uiu</td>
<td>l_1-39</td>
<td>petefred</td>
<td>normal</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Running</td>
<td>36071.abem5.ncsa.uiu</td>
<td>rturb</td>
<td>pakshing</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36518.abem5.ncsa.uiu</td>
<td>A2Q_8-8</td>
<td>seabra</td>
<td>normal</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Running</td>
<td>36521.abem5.ncsa.uiu</td>
<td>A2Q_8-12</td>
<td>seabra</td>
<td>normal</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Running</td>
<td>36523.abem5.ncsa.uiu</td>
<td>A2Q_12-12</td>
<td>seabra</td>
<td>normal</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Running</td>
<td>36563.abem5.ncsa.uiu</td>
<td>AirNoFSTur</td>
<td>auzun</td>
<td>normal</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Running</td>
<td>36584.abem5.ncsa.uiu</td>
<td>DJ14</td>
<td>jhsin</td>
<td>normal</td>
<td></td>
<td>128</td>
</tr>
<tr>
<td>Running</td>
<td>36627.abem5.ncsa.uiu</td>
<td>rturb</td>
<td>pakshing</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36647.abem5.ncsa.uiu</td>
<td>P3</td>
<td>amyshih</td>
<td>normal</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>Running</td>
<td>36690.abem5.ncsa.uiu</td>
<td>rturb</td>
<td>pakshing</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36718.abem5.ncsa.uiu</td>
<td>cr-323-02</td>
<td>kjjin</td>
<td>normal</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Running</td>
<td>36828.abem5.ncsa.uiu</td>
<td>m0r8_RC0s</td>
<td>moo</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36833.abem5.ncsa.uiu</td>
<td>Estr_Prod</td>
<td>seabra</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36842.abem5.ncsa.uiu</td>
<td>Script.abe</td>
<td>dcolins</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36900.abem5.ncsa.uiu</td>
<td>L10_N64_4I</td>
<td>gbryan</td>
<td>normal</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Running</td>
<td>36903.abem5.ncsa.uiu</td>
<td>frthali</td>
<td>zetienne</td>
<td>normal</td>
<td></td>
<td>150</td>
</tr>
</tbody>
</table>
# Resources and Available Kits

## View CTSS Kits Available on TeraGrid

*Home > User Info > Software > CTSS > Results*

<table>
<thead>
<tr>
<th>Site: Resource</th>
<th>TeraGrid Core Integration Kit</th>
<th>Remote Login Capability Kit</th>
<th>TeraGrid Remote Compute Kit</th>
<th>Application Development and Runtime Support Capability Kit</th>
<th>TeraGrid Data Movement Kit</th>
<th>Parallel Application Capability Kit</th>
<th>Science Workflow Support Kit</th>
<th>TeraGrid Data Management Capability Kit</th>
<th>Data Visualization Support Kit</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCAR: Frost</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not Installed</td>
<td>Not Installed</td>
<td>Not Installed</td>
</tr>
<tr>
<td>PSC: Big Ben</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>currently available</td>
<td>currently available</td>
<td>Not Installed</td>
</tr>
<tr>
<td>PSC: Rachel</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not installed</td>
<td>currently available</td>
<td>Not Installed</td>
</tr>
<tr>
<td>Purdue: Condor</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not Installed</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not Installed</td>
</tr>
<tr>
<td>Purdue: Lear</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not installed</td>
<td>currently available</td>
<td>Not Installed</td>
</tr>
<tr>
<td>TACC: Lonestar</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not Installed</td>
</tr>
<tr>
<td>TACC: Maverick</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>currently available</td>
<td>Not installed</td>
<td>currently available</td>
</tr>
<tr>
<td>NCSA: Cobalt</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not Installed</td>
</tr>
<tr>
<td>NCSA: Abe</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not Installed</td>
</tr>
<tr>
<td>NCSA: IA-58 Cluster</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>currently available</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not installed</td>
<td>Not Installed</td>
</tr>
</tbody>
</table>
Where are the GridFTP services?

<table>
<thead>
<tr>
<th>Service selection: globus-gridftp-server</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCAR Blue Gene/L (Frost)</td>
</tr>
<tr>
<td><strong>Type</strong>: gridftp</td>
</tr>
<tr>
<td><strong>Version</strong>: 4.0.3</td>
</tr>
<tr>
<td><strong>Endpoint</strong>: gsiftp://gridftp.frost.ncar.teragrid.org:2811/</td>
</tr>
<tr>
<td><strong>Name</strong>: globus-gridftp-server</td>
</tr>
<tr>
<td>PSC XT3 (BigBen)</td>
</tr>
<tr>
<td><strong>Type</strong>: gridftp</td>
</tr>
<tr>
<td><strong>Version</strong>: 4.0.3</td>
</tr>
<tr>
<td><strong>Endpoint</strong>: gsiftp://gridftp.bigben.psc.teragrid.org:2811/</td>
</tr>
<tr>
<td><strong>Name</strong>: globus-gridftp-server</td>
</tr>
<tr>
<td>PSC Alpha EV7 (Rachel)</td>
</tr>
<tr>
<td><strong>Type</strong>: gridftp</td>
</tr>
<tr>
<td><strong>Version</strong>: 4.0.3</td>
</tr>
<tr>
<td><strong>Endpoint</strong>: gsiftp://gridftp.psc.teragrid.org:2811/</td>
</tr>
<tr>
<td><strong>Name</strong>: globus-gridftp-server</td>
</tr>
<tr>
<td>Purdue Condor Pool</td>
</tr>
<tr>
<td><strong>Type</strong>: gridftp</td>
</tr>
<tr>
<td><strong>Version</strong>: 4.0.3</td>
</tr>
<tr>
<td><strong>Endpoint</strong>: gsiftp://tg-data.purdue.teragrid.org:2811/</td>
</tr>
<tr>
<td><strong>Name</strong>: globus-gridftp-server</td>
</tr>
<tr>
<td>Purdue EM64T Cluster (Lear)</td>
</tr>
<tr>
<td><strong>Type</strong>: gridftp</td>
</tr>
<tr>
<td><strong>Version</strong>: 4.0.3</td>
</tr>
<tr>
<td><strong>Endpoint</strong>: gsiftp://tg-data.purdue.teragrid.org:2811/</td>
</tr>
<tr>
<td><strong>Name</strong>: globus-gridftp-server</td>
</tr>
</tbody>
</table>
CTSS 4 Kit Capabilities

For each kit on each resource
- Current support level, and target support level
  - Development/prototype, Testing/pre-production, Production
- Support organization
- Inca status URL
- Multiple version of a kit with different support levels

For each kit software component on each resource
- Name, version, how to access it
- Multiple versions of a single component

For each kit service on each resource
- Name, type, version, and Endpoint (contact location)
- GSI OpenSSH, GridFTP, SRB servers, GRAM, MDS4
- Multiple services of the same type

The coordinated way TeraGrid publishes available CTSS capabilities.
Open Forward Process - Requirements

**Requirements Analysis**

**Design & Develop**

**Deploy & Support**

**What**
- Use cases (who needs it)?
- What type of information (attributes)?
- Who owns and provides the data (ownership, source)?
- Is any of it already available thru information services (gap analysis)?

**Who**
- Everyone: users, staff, partners

**Where**
- Discussion: tg-cat@teragrid.org
- Analysis results: wiki documents currently linked here:
Open Forward Process - Develop

What
Register a provider and/or publish data?
Design/adopt a schema?
develop publishing adaptors, services.
Choose and/or develop query interfaces.

Who
Several groups: information users, owner, consumers, SI facilitated

Where
Discussion: gig-info@teragrid.org
Open Forward Process - Deploy

What
Deploy adapters or publishing services.
Deploy query interfaces.
Update information services documentation.

Who
Several groups: information providers, RPs, SI

Where
Discussion: software-wg@teragrid.org, gateways, docgroup, ...
Current Activity Areas

Deploy & Support
- Upgrade/expand/merge scheduling information services
  - Part of WS GRAM 4.0.5 upgrade
  - Doesn’t address Scheduling WG requirements yet
- Implement 99.5% availability TeraGrid wide information services
- Improve information services documentation

Design & Develop
- Publish TGCDB resource and organization information
- Prototyping REST/Web 2.0 interfaces:
  - Publish HTML (staff/developer views), XHTML/XML, Atom, RSS, etc.
- Prototype universal command line client

Requirements Analysis
- HPC (non-CTSS) software
- Resource hardware specifications
- Scheduling information (BQP, Scheduling WG)
- Data movement information (Data WG)
[Not so] Farfetched Possibilities

Gateways
- User Services gateway database information
- Gateway published, capabilities, software, and services

Data collections
- User Services data collections database information
- Data collections access method, service Endpoint, paths

Community software areas
- Which resources have each CSA
- What software is available in each CSA, how to access it

Community accounts
- Which resources is each account active on

Outages
- Planned and unplanned outage information (gateways could use)

Resource Provider
- Policies

Peer grids/interoperability
- Resources, services available on peer grids (OSG, EGEE, ...)

.......
More Information

Discuss Information Services content, requirements, and design:
E-mail list tg-cat@teragrid.org

View current Information Services content
User Portal (scheduler load & queue contents):
https://portal.teragrid.org:443/gridsphere/gridsphere?cid=resources
User Documentation (CTSS 4 kits, software, services):
http://www.teragrid.org/userinfo/software/ctss.php
Staff and developer WebMDS views

Useful URLs:
Will link to other areas:
http://info.teragrid.org/
Information Services activities, plan, etc,