Cloud Based Services for Science

Steve Tuecke
Computation Institute
University of Chicago and Argonne National Laboratory
Cloud Layers

- Software-as-a-Service (SaaS)
- Platform-as-a-Service (PaaS)
- Infrastructure-as-a-Service (IaaS)
Big science has achieved big successes with advanced community services

Built on Globus Toolkit and Condor software

LIGO: 1 PB data in last science run, distributed worldwide

Substantial teams
Sustained effort
Leverage common technology
Application-specific solutions
Production focus

ESG: 1.2 PB climate data delivered to 23,000 users; 600+ pubs

OSG: 1.4M CPU-hours/day, >90 sites, >3000 users, >260 pubs in 2010
But small and medium science is suffering

- Data deluge
- Ad-hoc solutions
- Inadequate software, hardware & IT staff
Every night, they receive 100,000 files in Illinois

They transmit files to Texas for analysis ... then move results back to Illinois ... and make them available to users

Process must be reliable, routine, and efficient

The cyberinfrastructure team is not large!
Time-consuming Tasks in Research

- Run experiments
- Collect data
- Manage data
- Move data
- Acquire computers
- Analyze data
- Run simulations
- Compare experiment with simulation
- Search the literature
- Communicate with colleagues
- Publish papers
- Find, configure, install relevant software
- Find, access, analyze relevant data
- Order supplies
- Write proposals
- Write reports
- …
Accelerate discovery and innovation worldwide by providing **research IT as a service**

Leverage the **cloud** to

- provide millions of researchers with unprecedented access to powerful tools;
- enable a massive shortening of cycle times in time-consuming research processes; and
- reduce research IT costs dramatically via economies of scale
Why SaaS?

- **Deliver advanced functionality that:**
  - Requires no user software installation or operation
    - Minimal IT proficiency required
  - Can be cheaply and incrementally adopted
    - Usage-based subscription pricing; no big up-front costs
  - Consolidates troubleshooting and support
    - An expert group can proactively detect and correct problems
  - Utilizes an efficient software delivery lifecycle
    - Updates developed, tested and deployed quickly

- **Dominates commercial & consumer markets**
  - What about the research market?
Towards “research IT as a service”

Research Data Management-as-a-Service

Globus Transfer
Globus Storage
Globus Collaborate
Globus Compute

Globus Integrate

…SaaS

…PaaS

(1) Collect Data
(2) Move to Storage Store
(3) Ingest Processing
(4) Move to Community Store
(5) Publish in Registry
(6) Validate
(7) Backup
(8) Mirror
(9) Search, Browse, Analyze, Update, Annotate
Globus Transfer: For when you want to...

- **Transfer and synchronize files**
  - Easy “fire-and-forget” transfers
  - Automatic fault recovery
  - High performance
  - Across multiple security domains

- **Minimize IT costs**
  - Software as a Service (SaaS)
    - No client software installation
    - New features automatically available
  - Consolidated support & troubleshooting
  - Simple endpoint installation with Globus Connect and GridFTP

>4,500 registered users, 4PB / 400M files transferred
Globus Storage: For when you want to …

- **Place** your data where you want
- **Access** it from anywhere via different protocols
- **Update** it, **version** it, and take **snapshots**
- **Share** versions with who you want
- **Synchronize** among locations
Join with a few or many people to:

- Share documents
- Track tasks
- Send email
- Share data
- Do whatever

With:

- Common groups
- Delegated management
• **Run computational jobs**
  – On national facilities (e.g., OSG, XSEDE), campus clusters, lab and personal computers, cloud IaaS
  – Focus on high throughput jobs

• **Minimize IT costs**
  – SaaS, consolidated support & troubleshooting, etc.

• **Minimize end-user expertise required**
  – Simple GUIs for submission, management, policies

Based on **Condor High Throughput Computing**
• No one SaaS provider can deliver it all
• Must create ecosystem that:
  – Allows any SaaS provider to easily participate
  – Dramatically reduces the cost of creating and operating service within the ecosystem
  – Provides seamless user experience across services
  – Agnostic to / works across any cloud IaaS provider

• Ecosystem requires Platform as a Service
  – Target the unique needs of the research community
Globus Integrate: For when you want to…

- Integrate with the Globus research cloud ecosystem
- Write programs that access/manage:
  - user identities, profiles, groups, resources
  - data, compute and collaboration

... via REST APIs and command line programs
## Globus Transfer Premium Plans

<table>
<thead>
<tr>
<th></th>
<th><strong>Standard</strong></th>
<th><strong>Silver</strong>&lt;br&gt;<strong>Sign up now</strong></th>
<th><strong>Gold</strong>&lt;br&gt;<strong>Sign up now</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Support service level</strong></td>
<td>Monday-Friday 8am-5pm Central</td>
<td>Monday-Friday 8am-5pm Central 4-hour response</td>
<td>24hr x 7 days/week 2-hour response</td>
</tr>
<tr>
<td><strong>Globus Connect Multi-User (GCMU)</strong>&lt;br&gt;<strong>What's this?</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Concurrent transfers</strong>&lt;br&gt;<strong>What's this?</strong></td>
<td>2 per user</td>
<td>Unlimited to/from your endpoints</td>
<td>Unlimited to/from your endpoints</td>
</tr>
<tr>
<td><strong>Flight Control endpoint operations console</strong>&lt;br&gt;<strong>What's this?</strong></td>
<td>—</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Branded interface with your own DNS name</strong>&lt;br&gt;<strong>What's this?</strong></td>
<td>—</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Single Sign-On with default identity provider</strong>&lt;br&gt;<strong>What's this?</strong></td>
<td>—</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Pricing</strong></td>
<td>Free</td>
<td>$1,000/month or $10,000/year</td>
<td>$7,500/month or $75,000/year</td>
</tr>
</tbody>
</table>
• Visit [https://www.globusonline.org/signup](https://www.globusonline.org/signup) to:
  – Get a free account and start moving files

• Visit [www.globusonline.org](www.globusonline.org) for:
  – Tutorials, FAQs, Pro Tips, Troubleshooting
  – Papers, Case Studies

• Contact [support@globusonline.org](mailto:support@globusonline.org) for:
  – Help getting started & using the service

• Follow us at [@globusonline](https://twitter.com/globusonline) on Twitter
  and [Globus Online](https://www.facebook.com/GlobusOnline) on Facebook