What OGF Can Do for Enterprises

(A view from the CIO office)

Charlie Catlett
Founding Chairman, GGF (1999-2004)
Member OGF Board of Directors
Chief Information Officer, Argonne National Laboratory
Senior Fellow, Computation Institute (UChicago/ANL)

OGF Approaching Ten Years

Over 4,400 participants

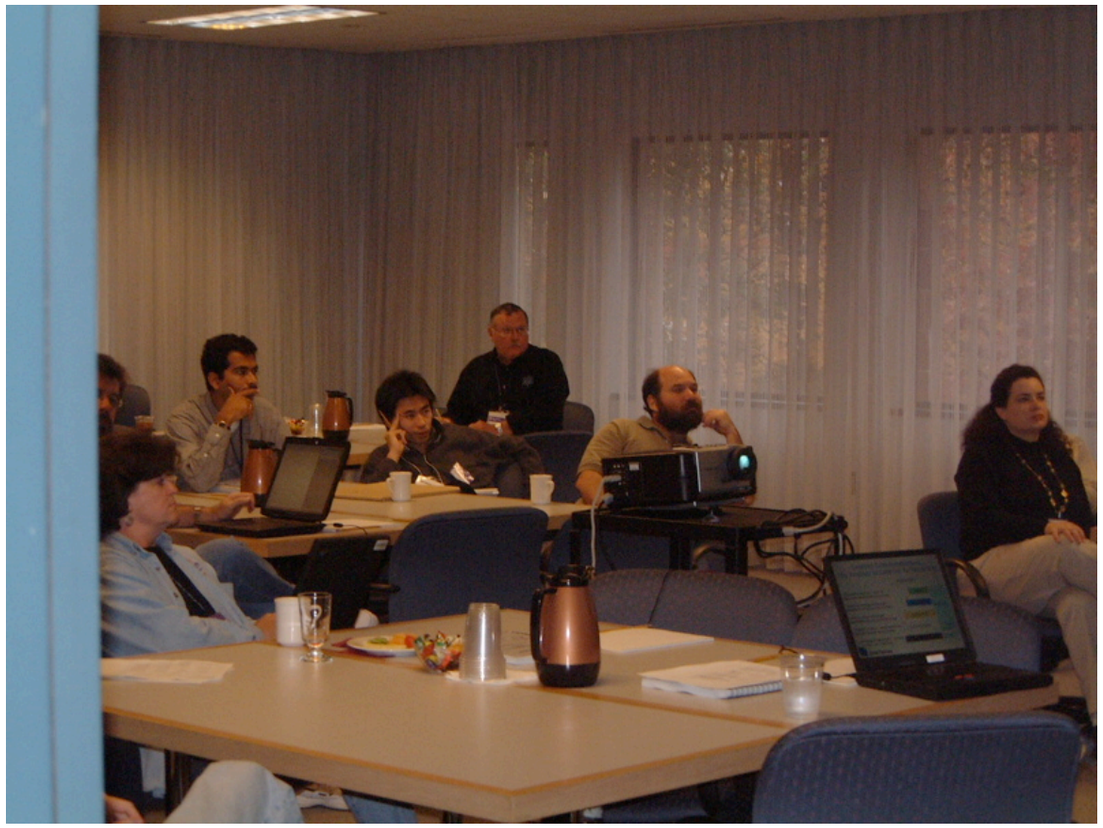
Over 2,800 one-timers (tourists...)

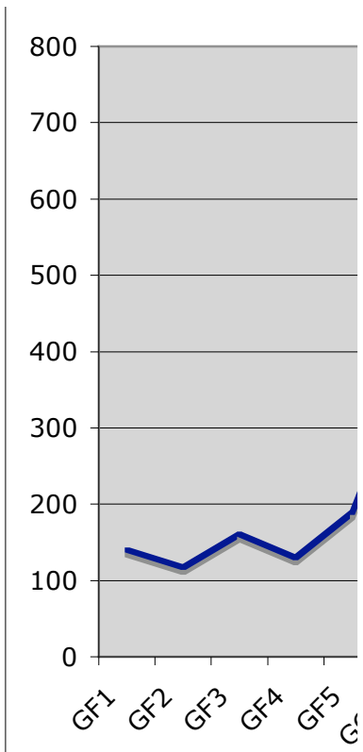
That leaves 1600 repeat participants

Charlie Catlett
(catlett@anl.gov)

1.5 Years

Boston: Grid Forum 5 (October 2000)





Worked our way up to about 200 people by the 5th meeting.

A big issue was whether or not it made sense to talk about a Grid architecture

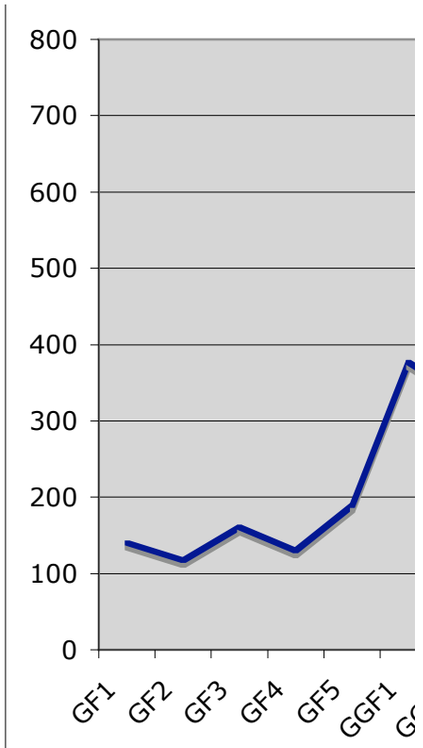
Charlie Catlett
(catlett@anl.gov)

2.5 Years

Amsterdam: GGF-1 (March 2001)

Shall we go global?

YES!



Some were worried that moving to Europe would stall momentum. *They were wrong.*

GGF-1 (joint with European Datagrid) had some 360 participants and (sadly) we had to turn 60 more away due to facility limitations.

Charlie Catlett
(catlett@anl.gov)

Perspectives from 2001 (Social)

Social Issues [circa 2001]

- Our present concerns
  - Building mindshare [evangelism]
  - Middleware technology [standards and interoperability]
  - Ecumenism and inclusively [process and governance]
  - Building application exemplars [accelerating adoption]
- Near future concerns
  - Academic vs commercial drivers for sustainability
  - Moving the Grid out of the Laboratories and Universities
  - Open source vs proprietary Grid technologies

Rick Stevens, March 2001 Argonne National Laboratory + University of Chicago

(GGF-1, Amsterdam)

Charlie Catlett
(catlett@anl.gov)

Perspectives from 2001 (Technical)

Middleware Concerns [circa 2001]

- Production
  - Core internet protocols
  - Core security infrastructure and single signon tools
  - Directory services
  - Virtual private networking
- On the Horizon
  - Robust co-scheduling and co-reservation services
  - Parallel and secure third party transfers
  - Data caching and replica management
  - Policy based resource management and optimization

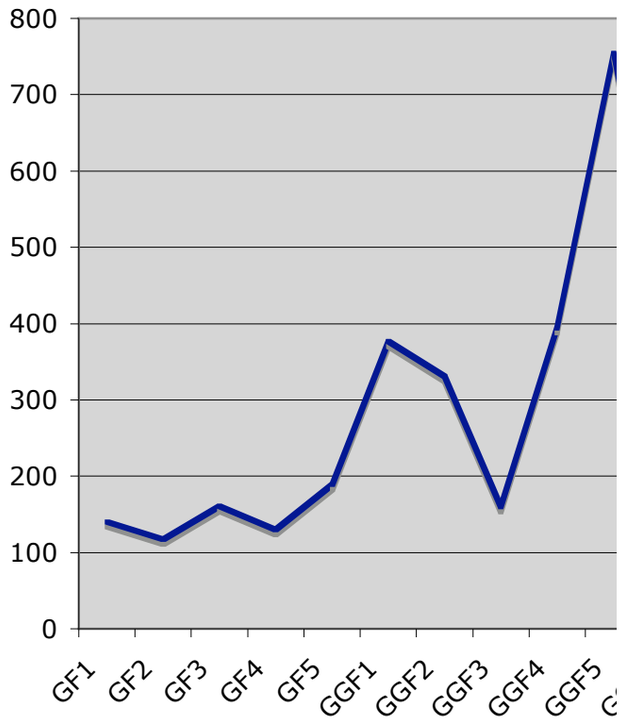
Rick Stevens, March 2001 Argonne National Laboratory + University of Chicago

Charlie Catlett
(catlett@anl.gov)

(GGF-1, Amsterdam)

3.5 Years

Edinburgh: GGF-5
(July 2002)



GGF-5 (joint with HPDC-11) was at the peak of "Grid Hype." Reporters were asking if this was going to be the next "Web" phenomenon.

GGF-5 saw a surge in industry participation.

Charlie Catlett
(catlett@anl.gov)

8.5 Years

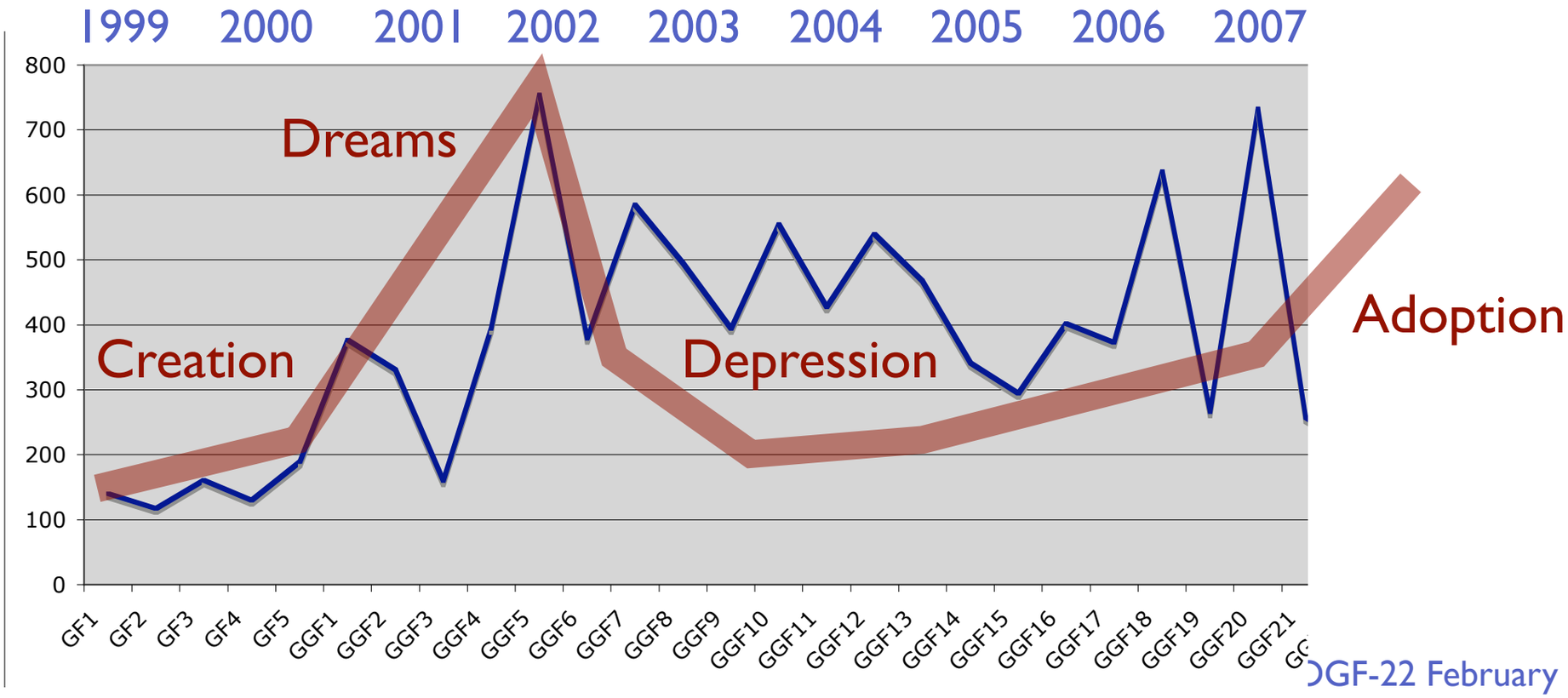
Total participant numbers suggest the standard "trough of despair" that follows the Hype S-Curve

"Cool new stuff"

"Sky's the limit; Eliminate old ways"

"New is crap; Back to basics"

"Hey this stuff actually works!"



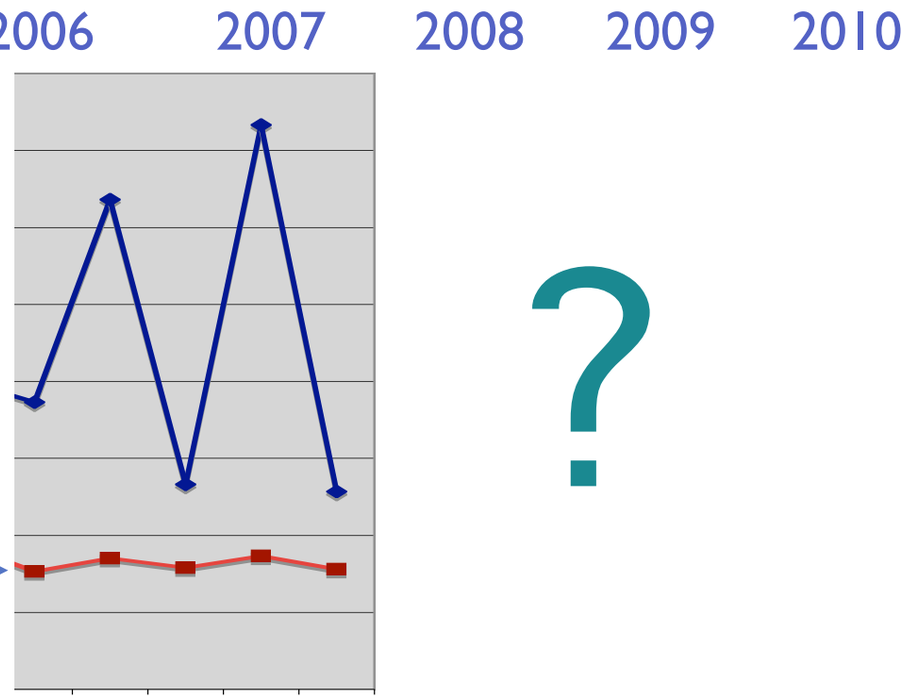
Charlie Catlett
(catlett@anl.gov)

How Do We Achieve Adoption?

Question: What do we want who to adopt?

Answer: It depends on **who we are** and what **we** want to accomplish.

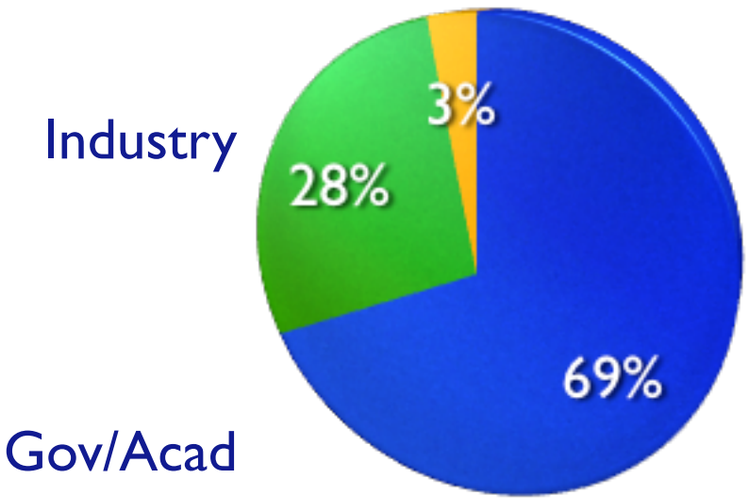
160-170 committed people



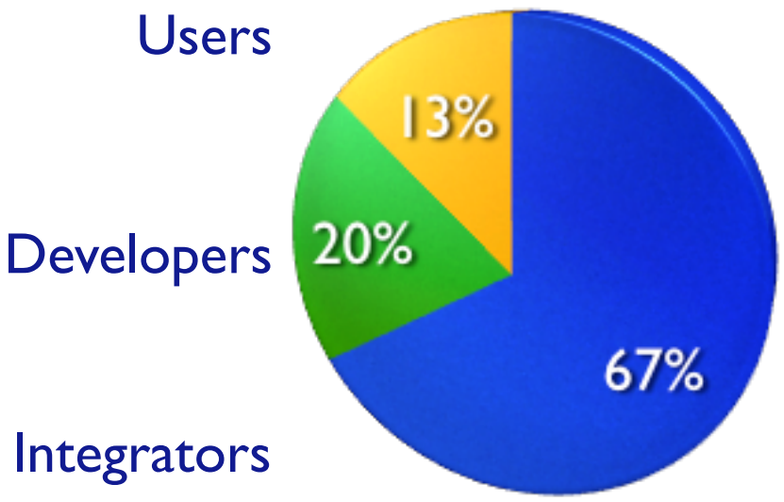
Charlie Catlett
(catlett@anl.gov)

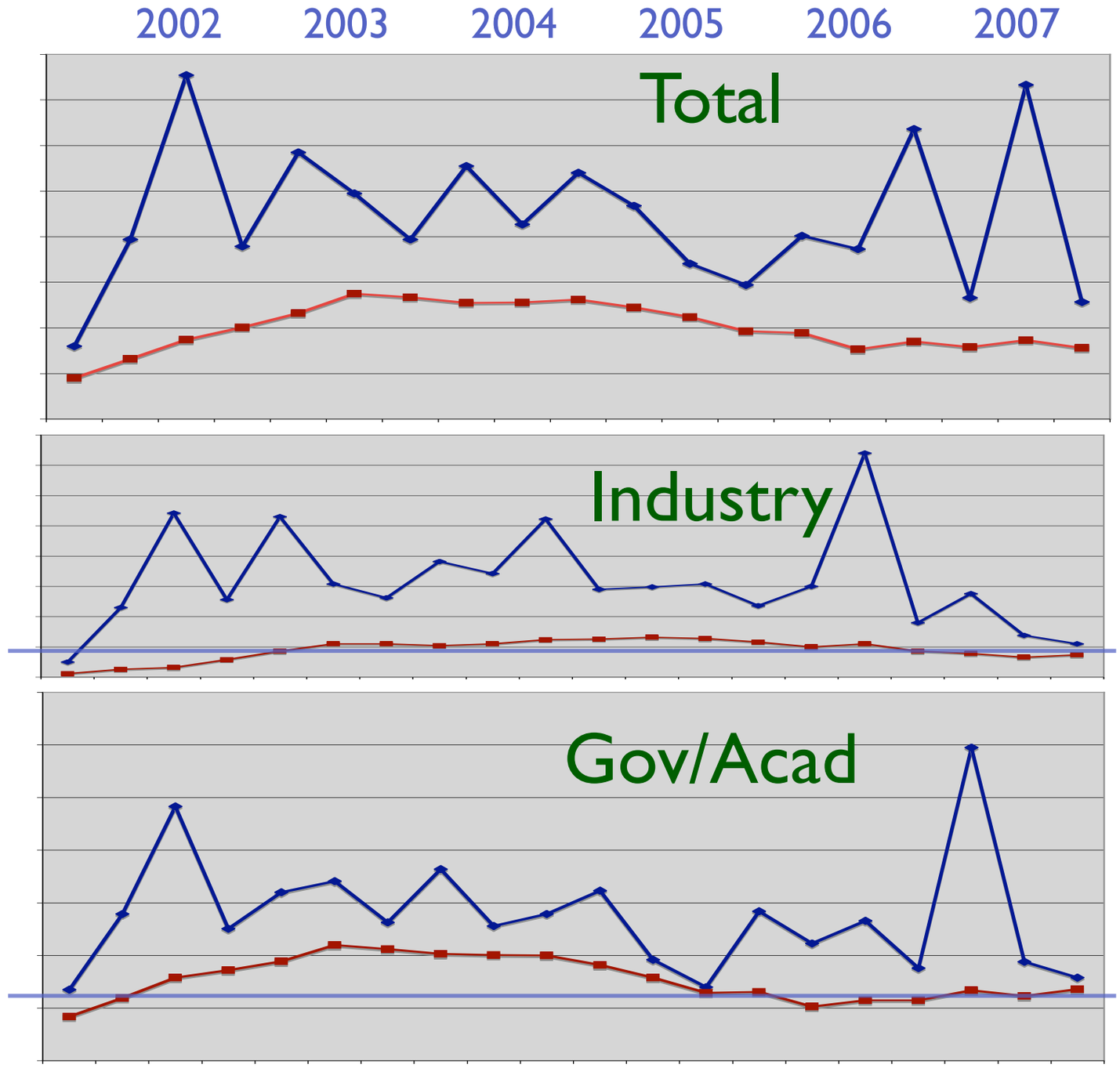
So Who Are We?

Red curve shows individuals who have been to at least 3 of the previous 5 meetings. For the past two years this has been steady at 160-180.



But Industry vs. gov/acad isn't a useful way to think about this community. A better way to parse the community:





Charlie Catlett
(catlett@anl.gov)

Skipping the Tired Discussions

- Number of groups versus number of people
- Marketing
- Productivity and time-to-market
- Organization
- Locations and meeting formats
- Other silver bullets

Charlie Catlett
(catlett@anl.gov)

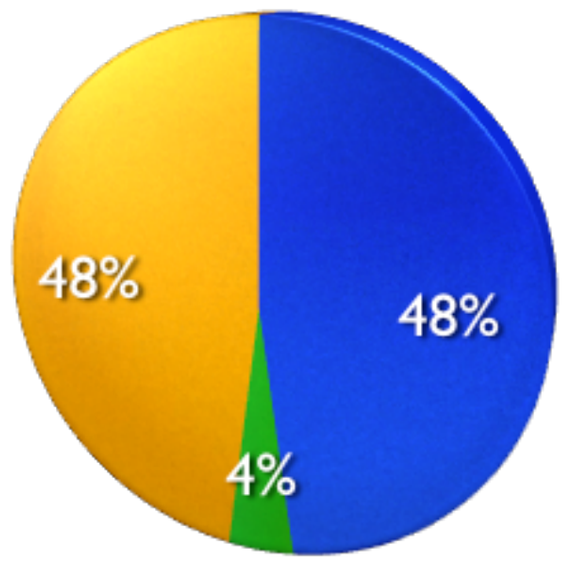
About HPC and Grids

- Continued Exponential Change in HPC
  - In TeraGrid (and I suspect DEISA and other HPC grids), the biggest challenge is harnessing multicore (aka the latest architecture of *standalone* supercomputer)
    - But there are multiple supercomputers in a Grid, and authentication and authorization are among the biggest challenges (ref GIN).
    - Much effort also to provide a consistent user experience.
- But HPC is a niche within a niche
  - Smaller experiments that need to stitch together databases and modest computing capabilities.
  - Science Gateways - "portals" supporting workflow - are addressed to this large community.
  - Emerging commercial and academic grid *service providers*.

Charlie Catlett
(catlett@anl.gov)

Laboratory CIO Challenges

- Science and Engineering infrastructure
  - Data is my community's biggest technical concern
- Facilitating the Enterprise
  - Business systems that don't waste my community's time
- Providing the right Environment
  - Sharing and collaborating services to help people to form and sustain teams.



Beyond HPC - Example supercomputer center: 4800 of 5000 users account for 88TB out of over 2 PB. Fewer than 100 of these users require >500 GB of space for their data.

Charlie Catlett
(catlett@anl.gov)

My Community's Priorities

- Science and Engineering infrastructure
  - Sharable, reliable data and document storage
  - High-volume data movement
  - Secure, remote user access
- Facilitating the Enterprise
  - As easy as the web; transparent and accountable delivery
    - SOA and web services; evolving large, expensive legacy systems
- Providing the right Environment
  - Notification and (action/commitment) Tracking
  - Social Networking (finding collaborators, creating virtual teams)
- Underlying Challenges
  - Security - moving beyond castles and moats

Charlie Catlett
(catlett@anl.gov)

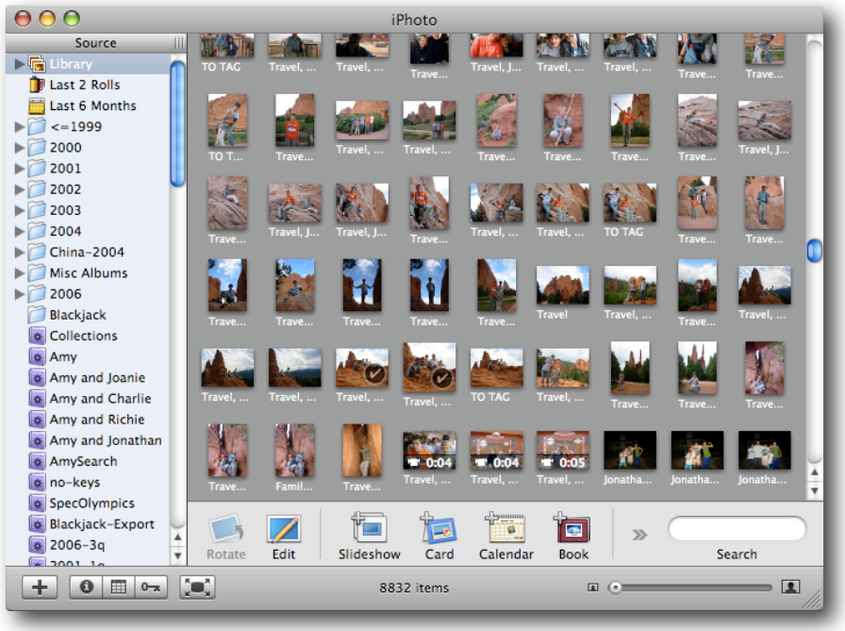
Our Strategy

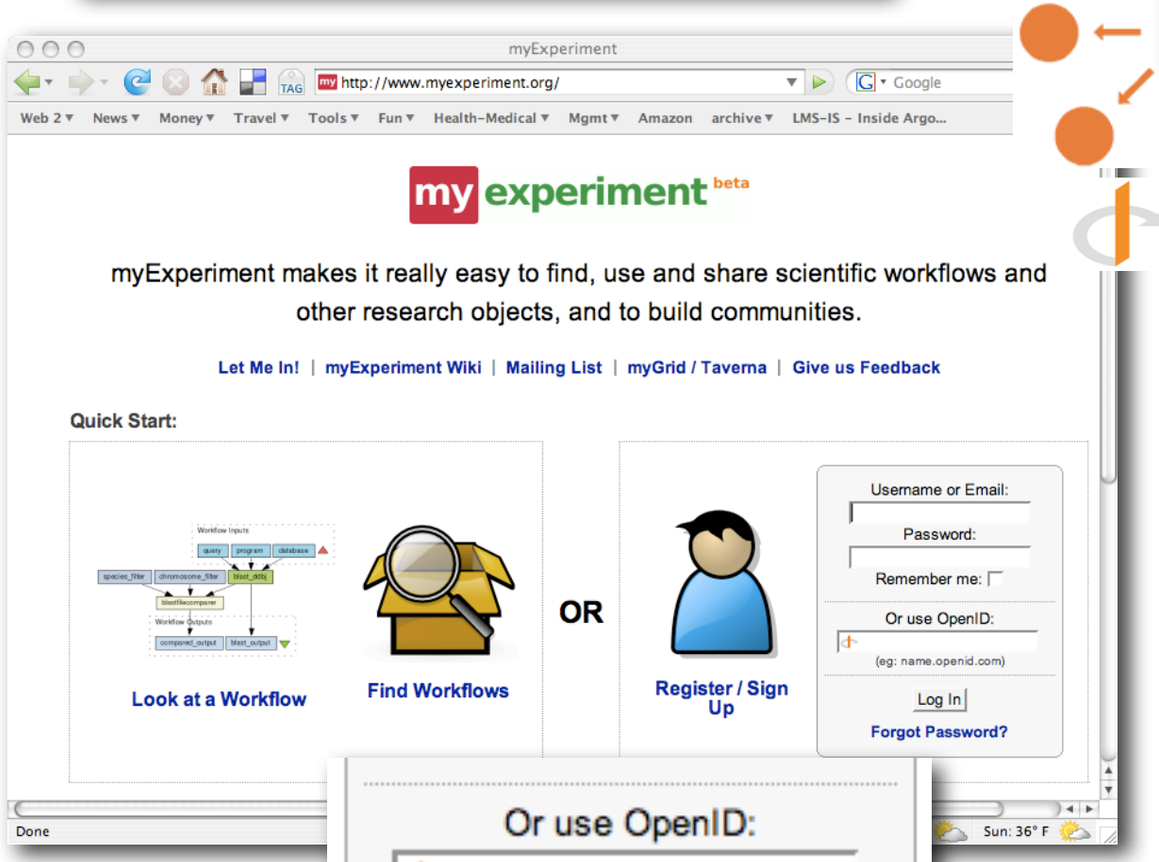
- Three Choices
  - Build, Buy, Integrate (SOA)
- Mostly Integrate!
  - Internet Industry is Rapidly Rolling out "Grid" development platforms with markets that dwarf HPC
    - Microsoft .NET (2.2 million developers)
    - Adobe AIR (aimed at the >1 million Flash developers)
    - AJAX (community size?)
    - Sun JavaFX, Google Gears, Mozilla Prism...
    - An entire industry is emerging atop commercial grid services
- In 2001 we did not have this breadth or quality of choices for building distributed systems.
  - Explosion of new services and capabilities
  - How can we ride this wave?

Harnessing Internet Innovation

I *accidentally* used two grid service providers!

(Amazon S3 and OpenID)





Name: http://charliecatlett.smugmug.com/

Charlie Catlett
(catlett@anl.gov)

Some Thoughts

- The HPC "Trickle-down" Myth
  - Moore's Law means that HPC is a time-machine for computing
  - Internet innovation moves much *faster* than Moore's Law
    - and much faster than the academic/government
- Our choices today are much richer than in 2001
  - Build our own trust infrastructure or leverage commercial?
    - Users live day-to-day in the commercial Internet space, not our grids
  - Aim for broad community or only the HPC niche?
    - computing and job submission details versus general service description
    - directory of computer details versus general resource description
    - co-scheduling/co-reservation versus workflow orchestration
    - parallel data movement versus ease-of-sharing

What is Uniquely Valuable about OGF?

- Product-Neutral Experience and Perspective
  - In 2001 this would mean helping me work with things like Kerberos, X.509, Condor, Unicore, Globus, etc.
  - In 2008 the world is much bigger, and the choices richer
- As a CIO, this is potentially of tremendous value
  - A specification could be helpful to my integrators/developers
    - E.g. usage record or something multiple vendors agree to use (e.g. S3)
  - Analysis of options for various services and standards
    - Security considerations and performance evaluation of web storage offerings
    - Framework for management of virtual system images and evaluation of service offerings
  - Examples of successful integration of standards and services

Charlie Catlett
(catlett@anl.gov)