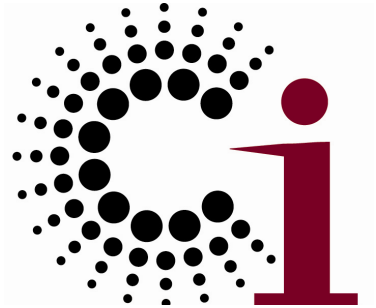
Scaling eScience Impact

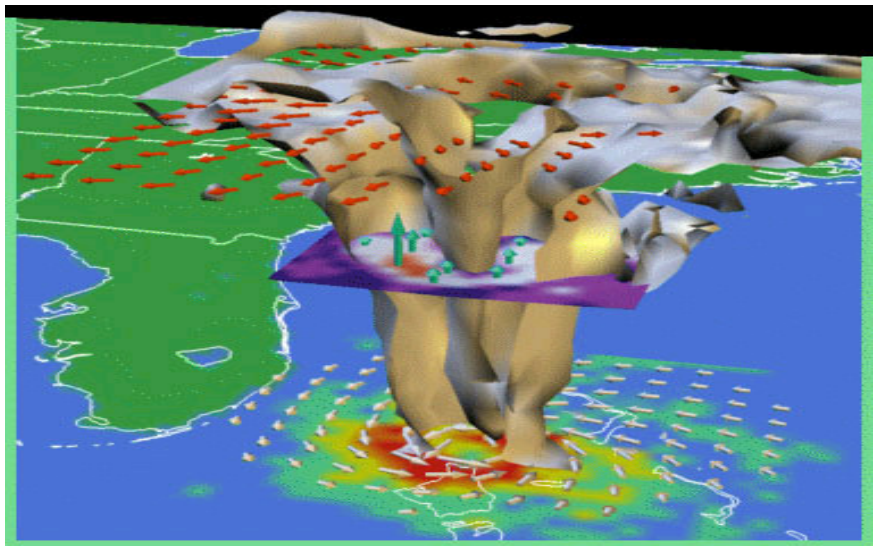
Ian Foster



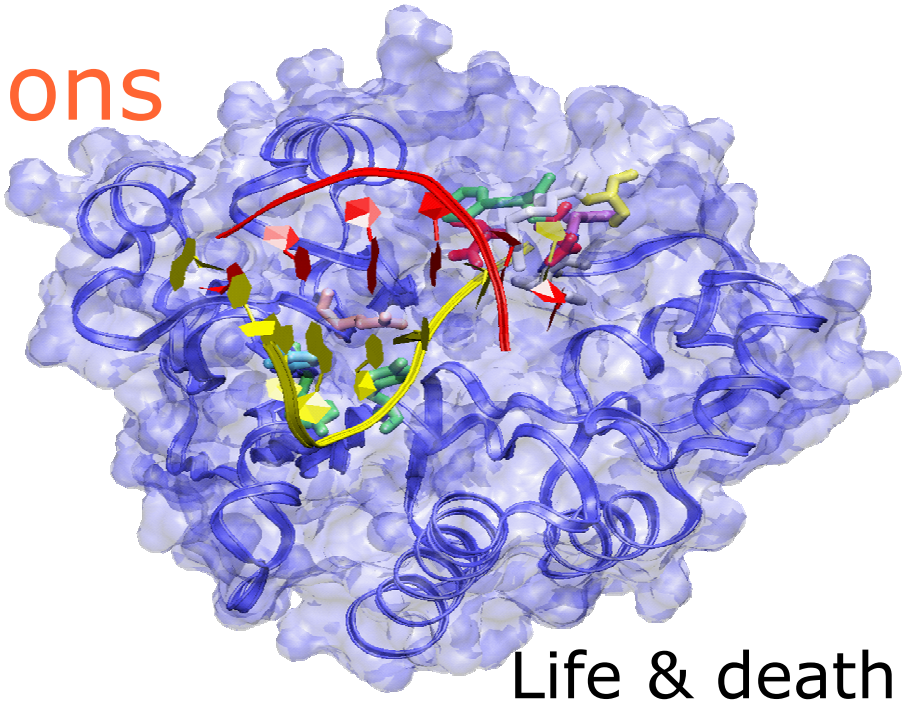
Computation Institute

Argonne National Lab & University of Chicago

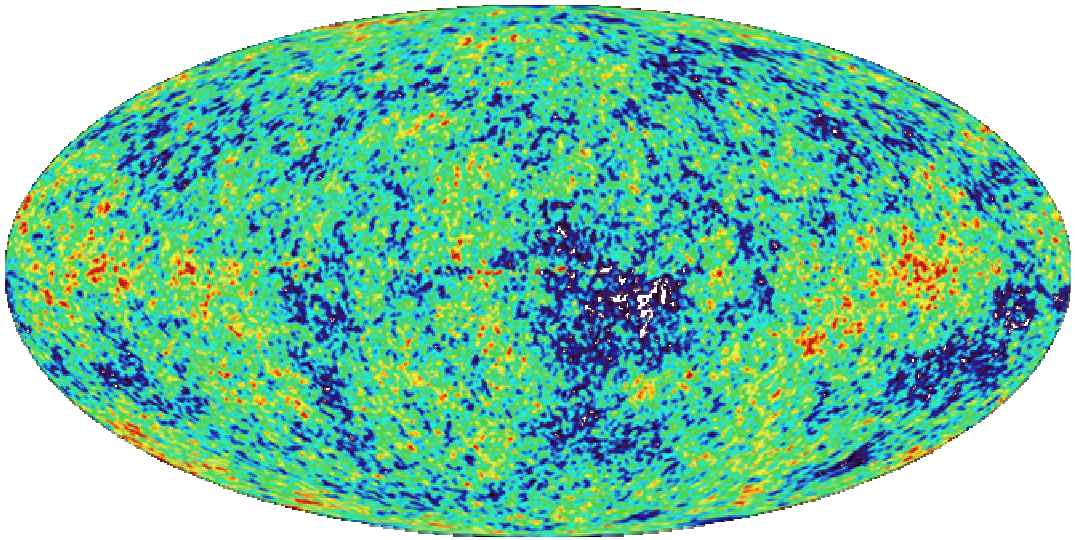
The Big Questions



Future of the planet



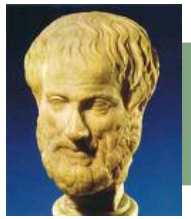
Life & death



Nature of the universe

Consciousness

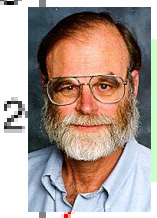
How Do We Answer Them?



Empirical

Theory

Simulation



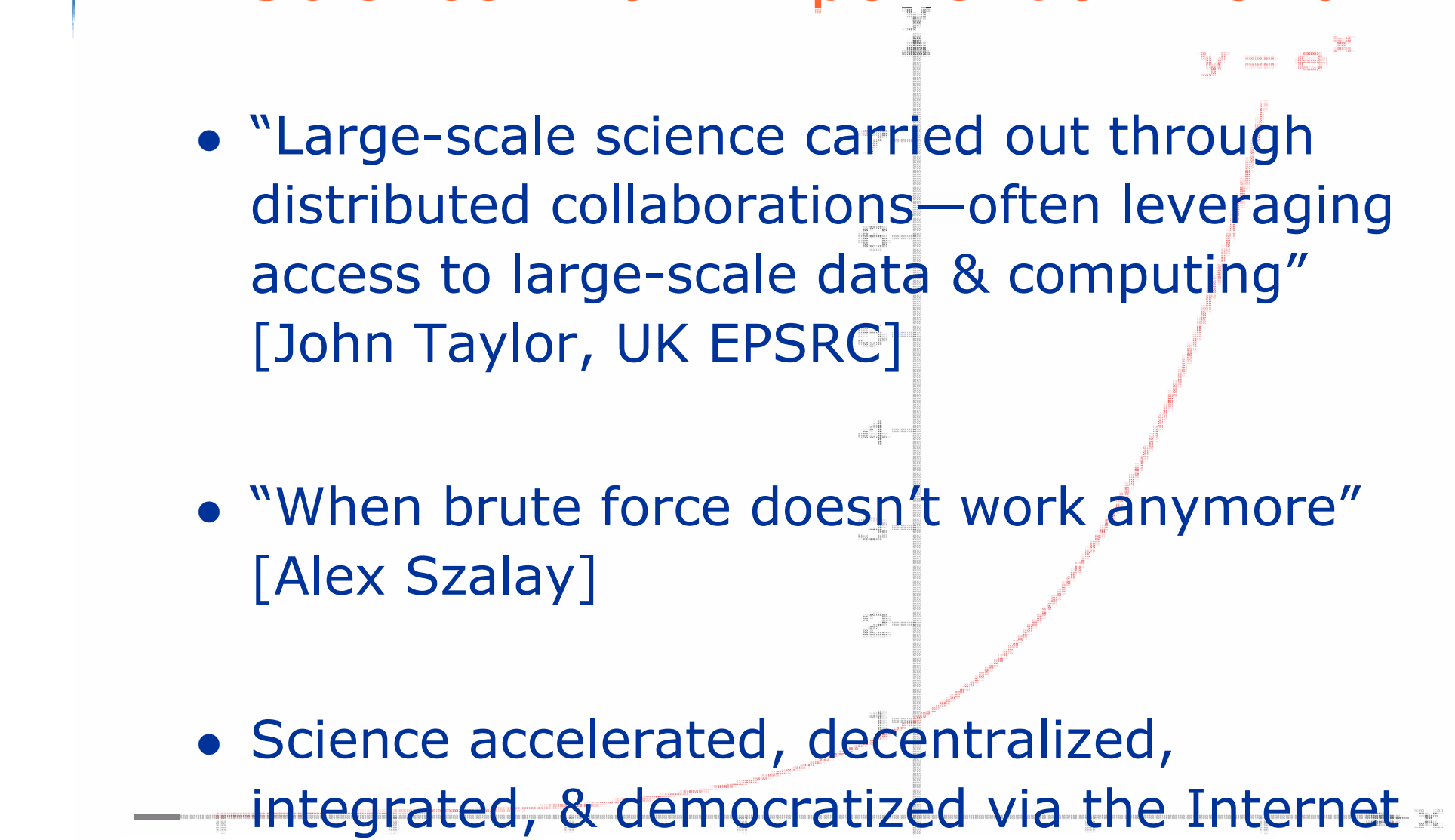
Data

The Same is True of Smaller Questions

- Designing new chemical catalysts
- Selling advertising
- Creating entertainment
- Finding parking

the globus alliance
www.globus.org

eScience: Science in an Exponential World



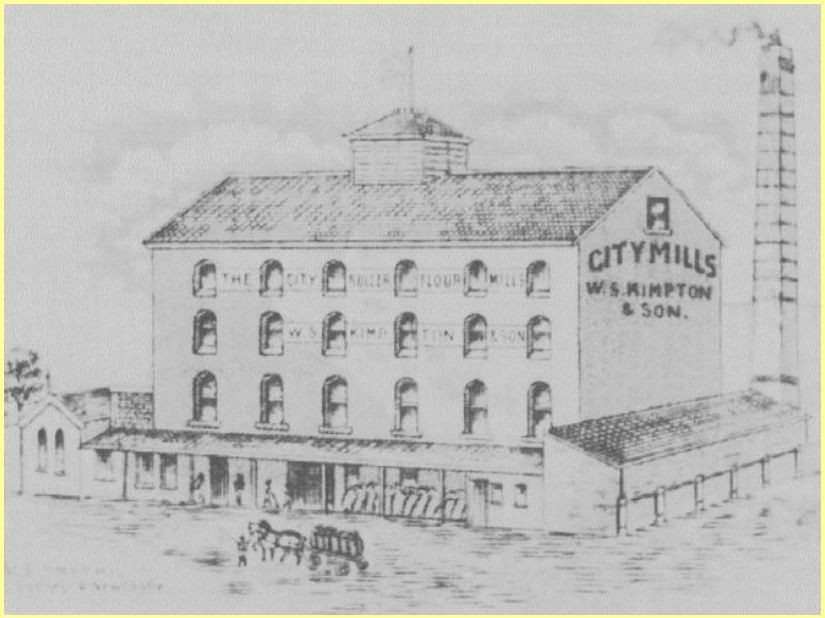
- "Large-scale science carried out through distributed collaborations—often leveraging access to large-scale data & computing" [John Taylor, UK EPSRC]
- "When brute force doesn't work anymore" [Alex Szalay]
- Science accelerated, decentralized, integrated, & democratized via the Internet

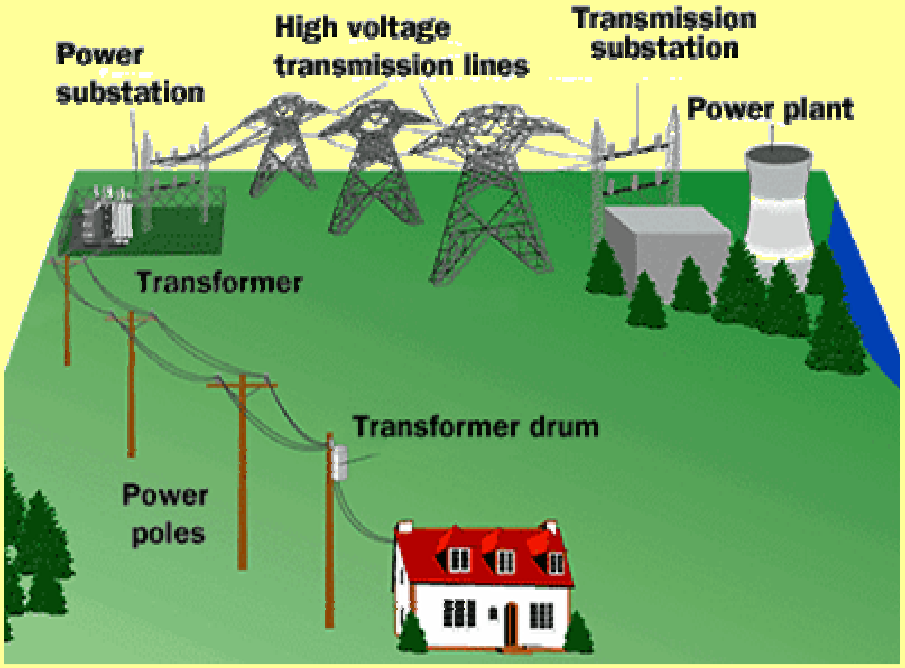
the globus alliance
www.globus.org

Grid: An Enabler of eScience

The dubious electrical power grid analogy

Must we buy (or travel to) a power source?





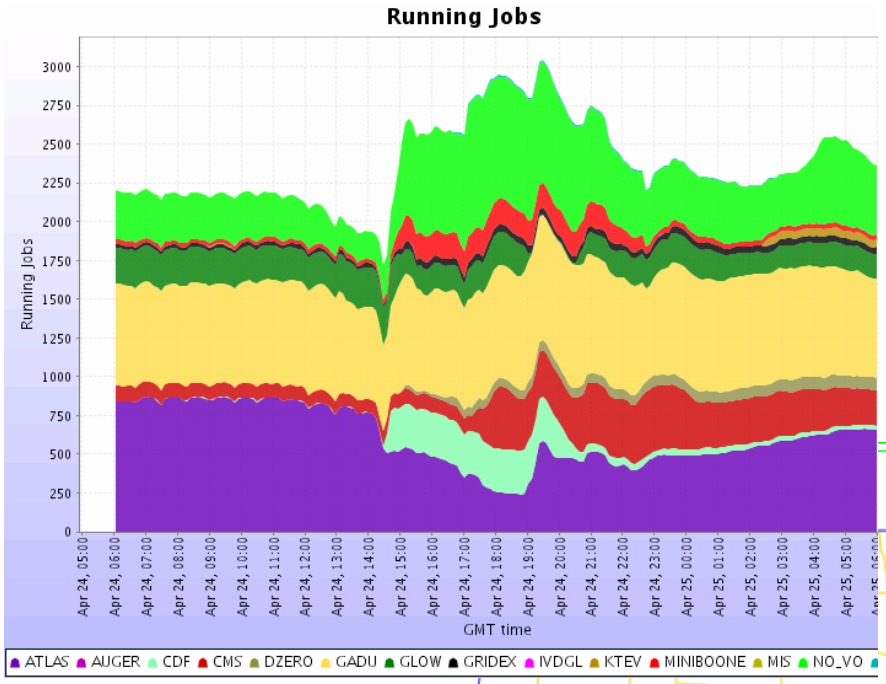
Or can we ship power to where we want to work?

Enable on-demand access to, and integration of, diverse resources & services, regardless of location

the globus alliance
www.globus.org

1st Generation Grids

Focus on aggregation of many resources for massively (data-)parallel applications



EGEE

Open Science Grid

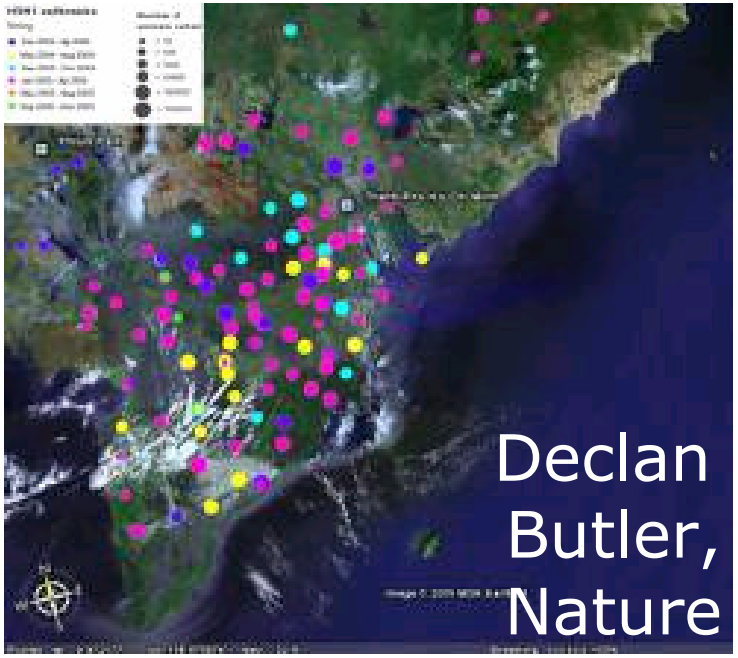
the globus alliance
www.globus.org

Second-Generation Grids

- Empower many more users by enabling on-demand access to **services**
- Science gateways (TeraGrid)
- Service oriented science
- Or, "Science 2.0"

"Service-Oriented Science", *Science*, 2005

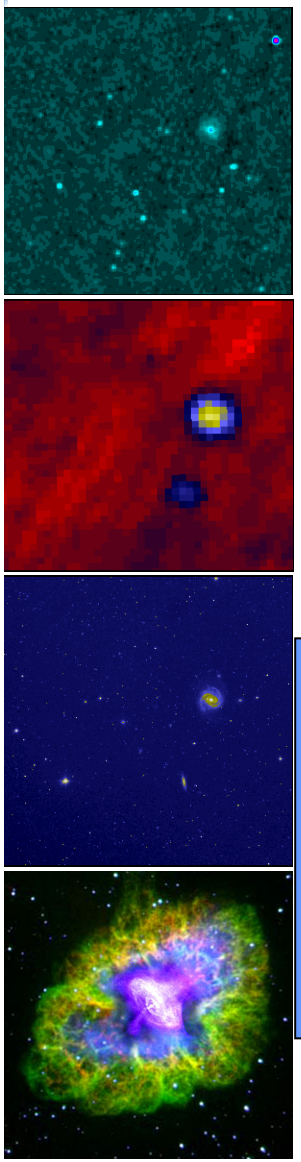
"Web 2.0"



- Software as services
  - Data- & computation-rich network services
- Services as platforms
  - Easy composition of services to create new capabilities ("mashups")—that themselves may be made accessible as new services
- Enabled by massive infrastructure buildout
  - Google projected to spend $1.5B on computers, networks, and real estate in 2006
  - Many others are spending substantially
- Paid for by advertising

the globus alliance
www.globus.org

Science 2.0: E.g., Virtual Observatories



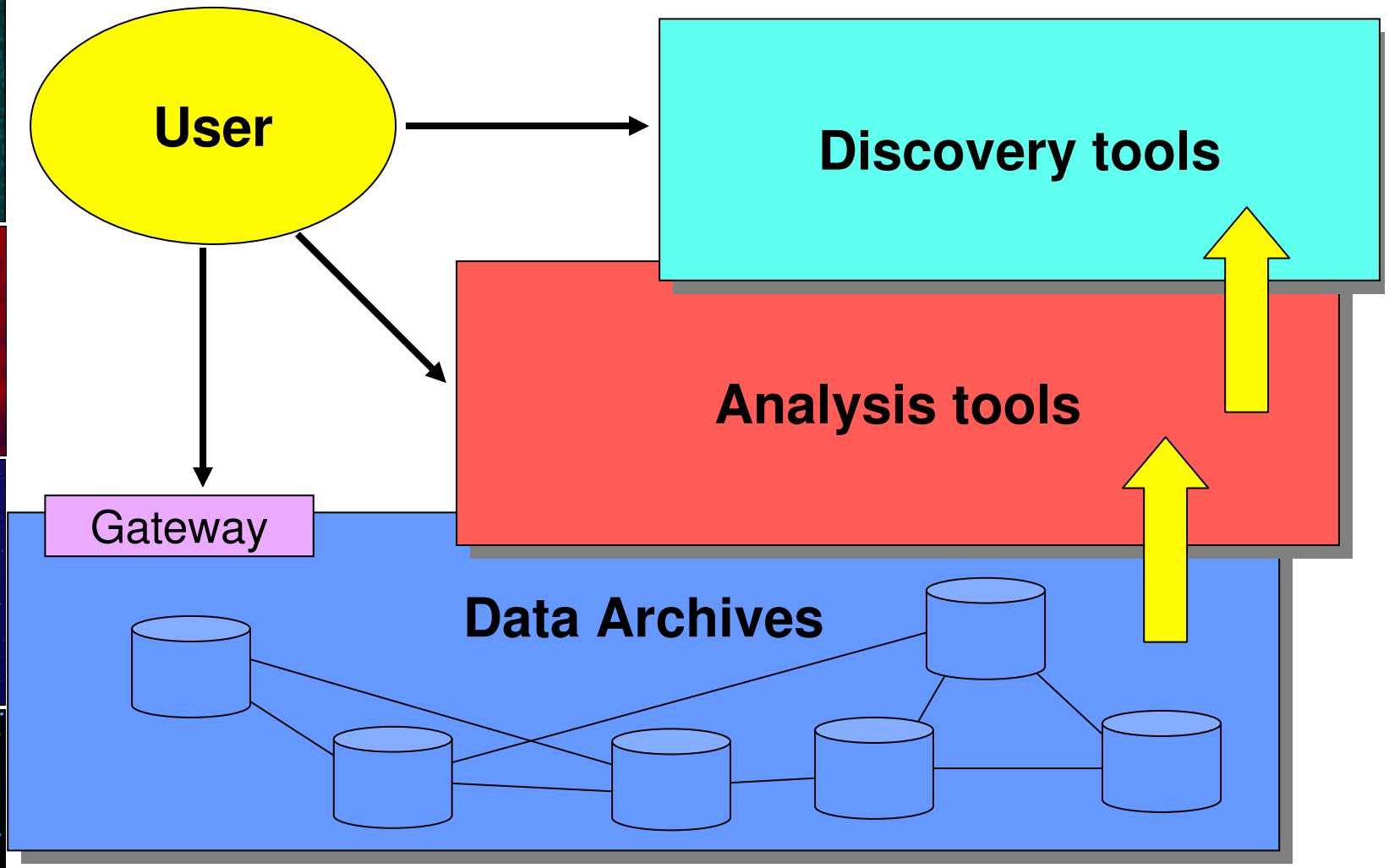
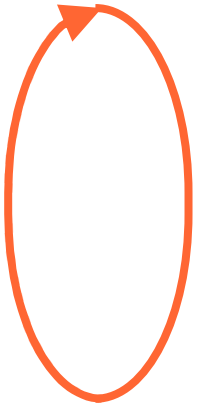


Figure: S. G. Djorgovski

the globus alliance
www.globus.org

Service-Oriented Science



People **create** services (data or functions) …
which I **discover** (& decide whether to use) …
& **compose** to create a new function …
& then **publish** as a new service.

→ *I find "someone else" to **host** services, so I don't have to become an expert in operating services & computers!*

!!

→ *I hope that this "someone else" can **manage** security, reliability, scalability, …*

TeraGrid™ EMPOWERING DISCOVERY

amazon web services™

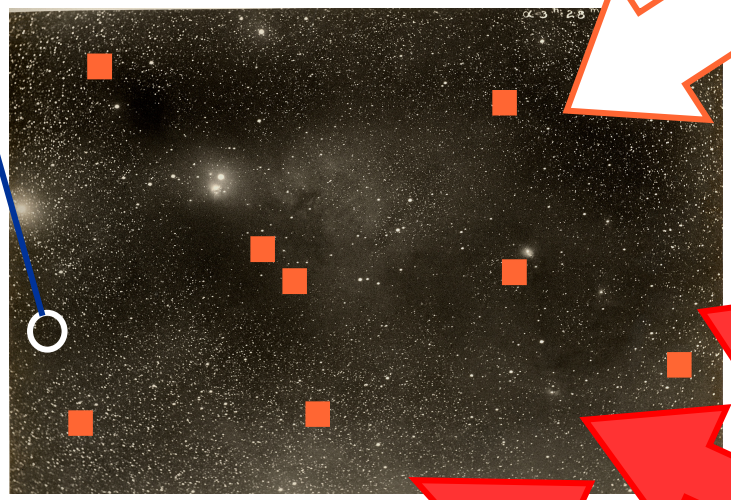
D-GRID

"Service-Oriented Science", *Science*, 2005

The Importance of "Hosting" and "Management"

Tell me about this star

Tell me about these 20K stars



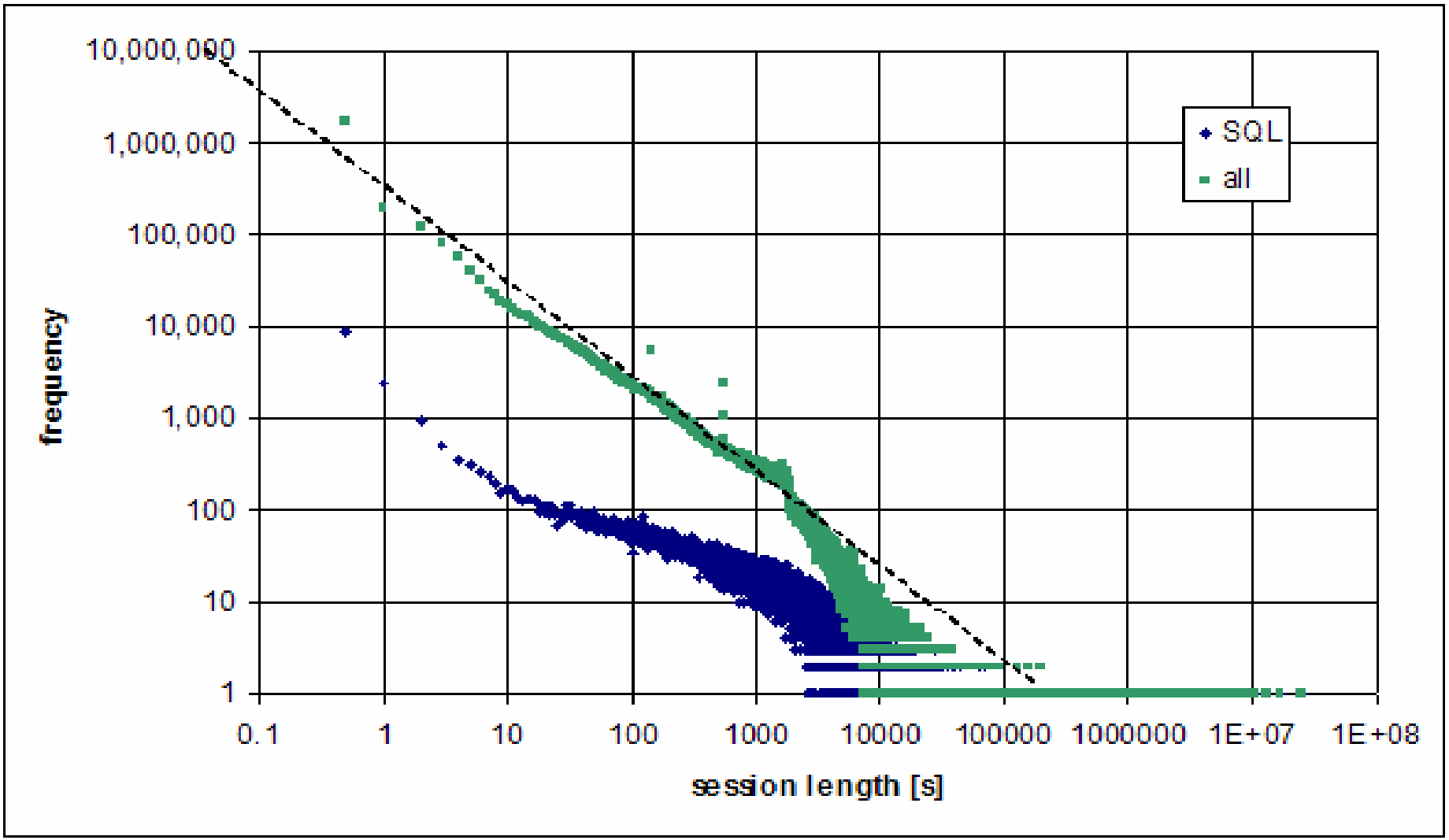
Support 1000s of users

E.g., Sloan Digital Sky Survey, ~10 TB; others much bigger

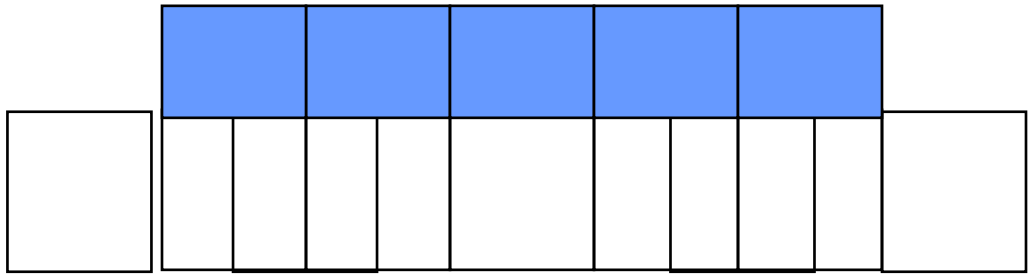
Load Comes from All Over …

A few hours of Globus.org access, early one morning …

Skyserver Sessions (Thanks to Alex Szalay)



The Two Dimensions of Service Oriented Science

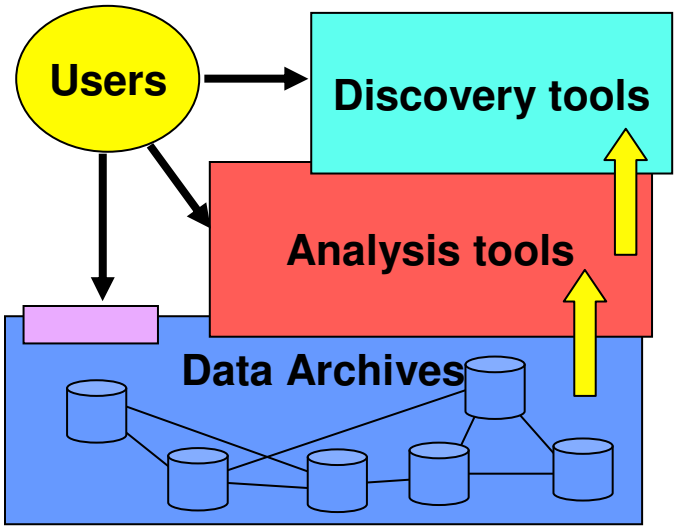


Function

Resource

- **Decompose** across network

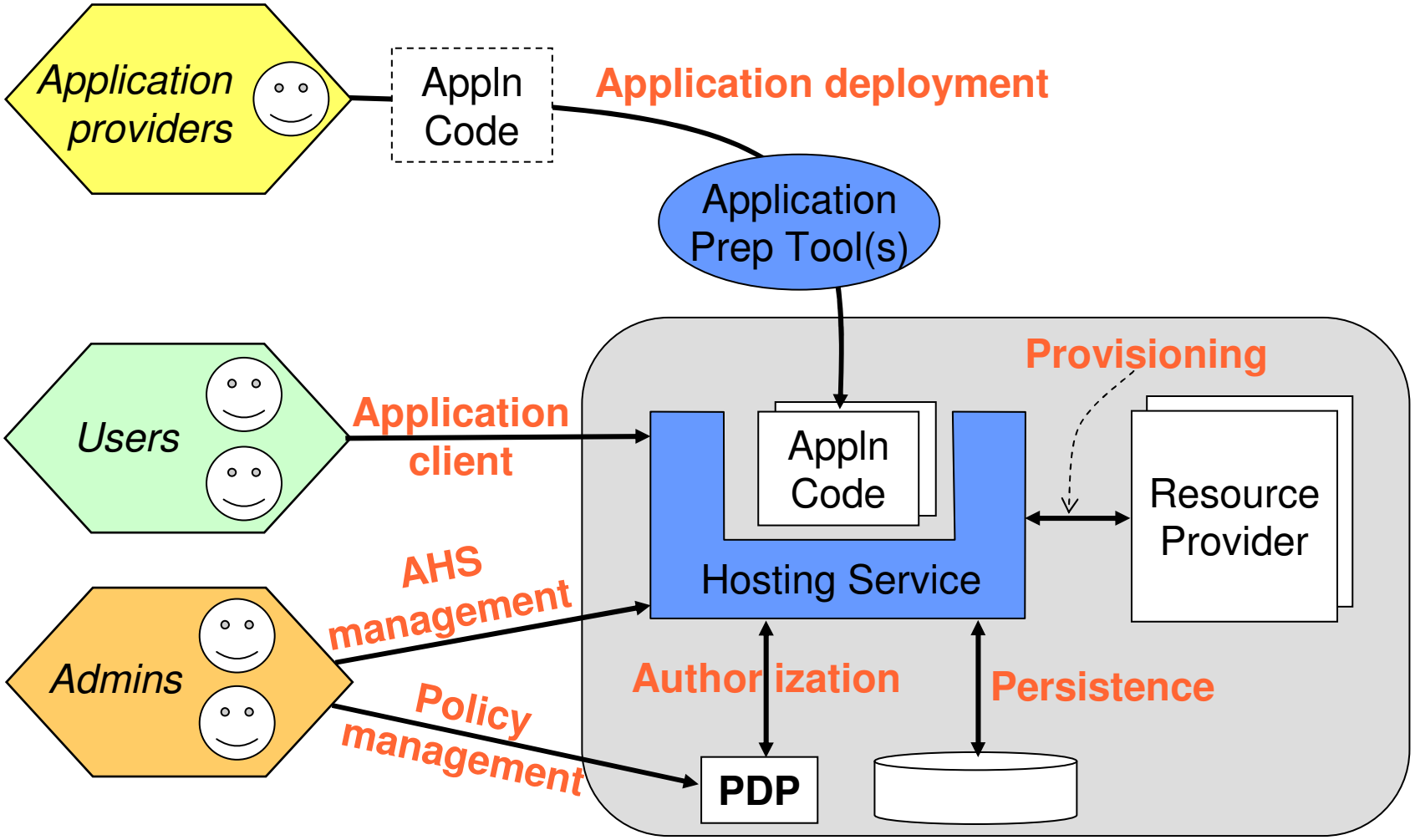
  Clients **integrate** dynamically
  - Select & compose services
  - Select "best of breed" providers
  - Publish result as new services



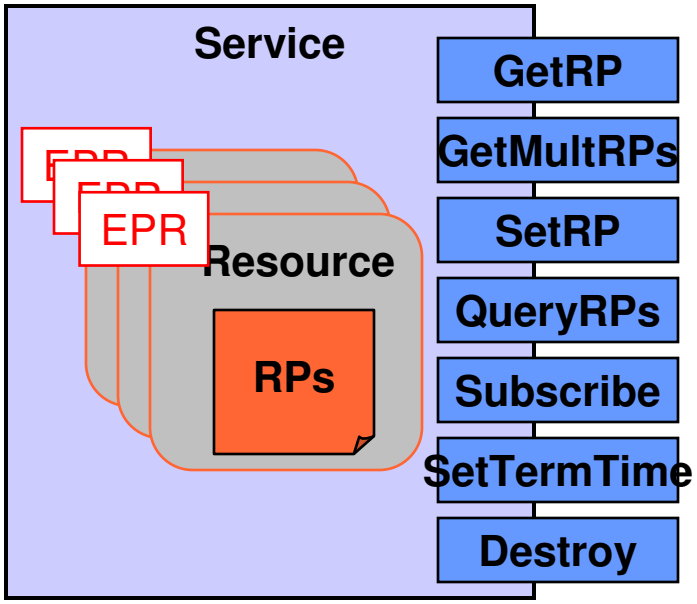
- Decouple **resource** & **service** providers

Fig: S. G. Djorgovski

Hosting & Management: Application Hosting Services



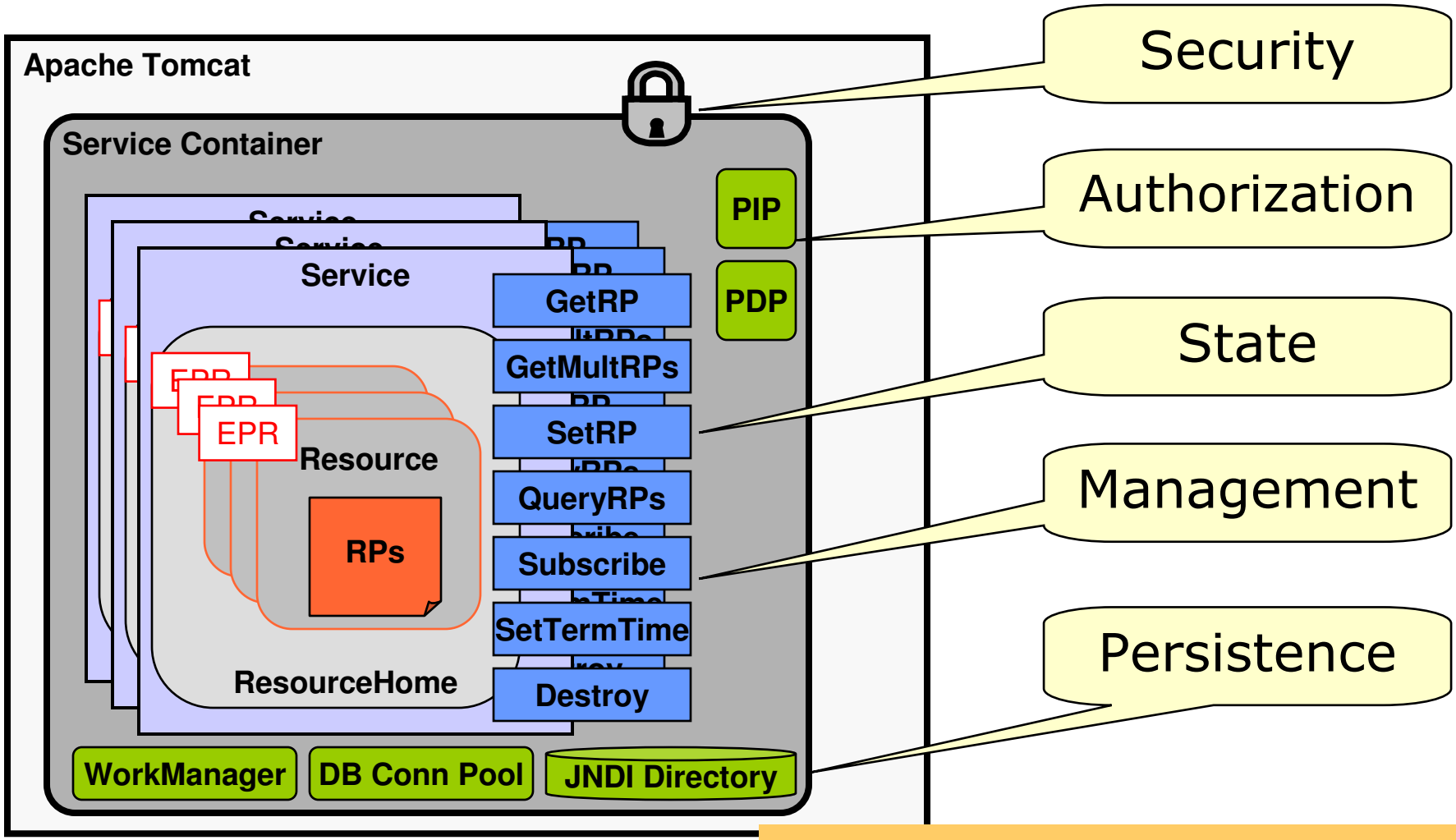
Web Services Resource Framework in a Nutshell



- Service
- State representation
  - Resource
  - Resource Property
- State identification
  - Endpoint Reference
- State Interfaces
  - GetRP, QueryRPs, GetMultipleRPs, SetRP
- Lifetime Interfaces
  - SetTerminationTime
  - ImmediateDestruction
- Notification Interfaces
  - Subscribe
  - Notify
- ServiceGroups

the globus alliance
www.globus.org

Globus Toolkit Web Services Container

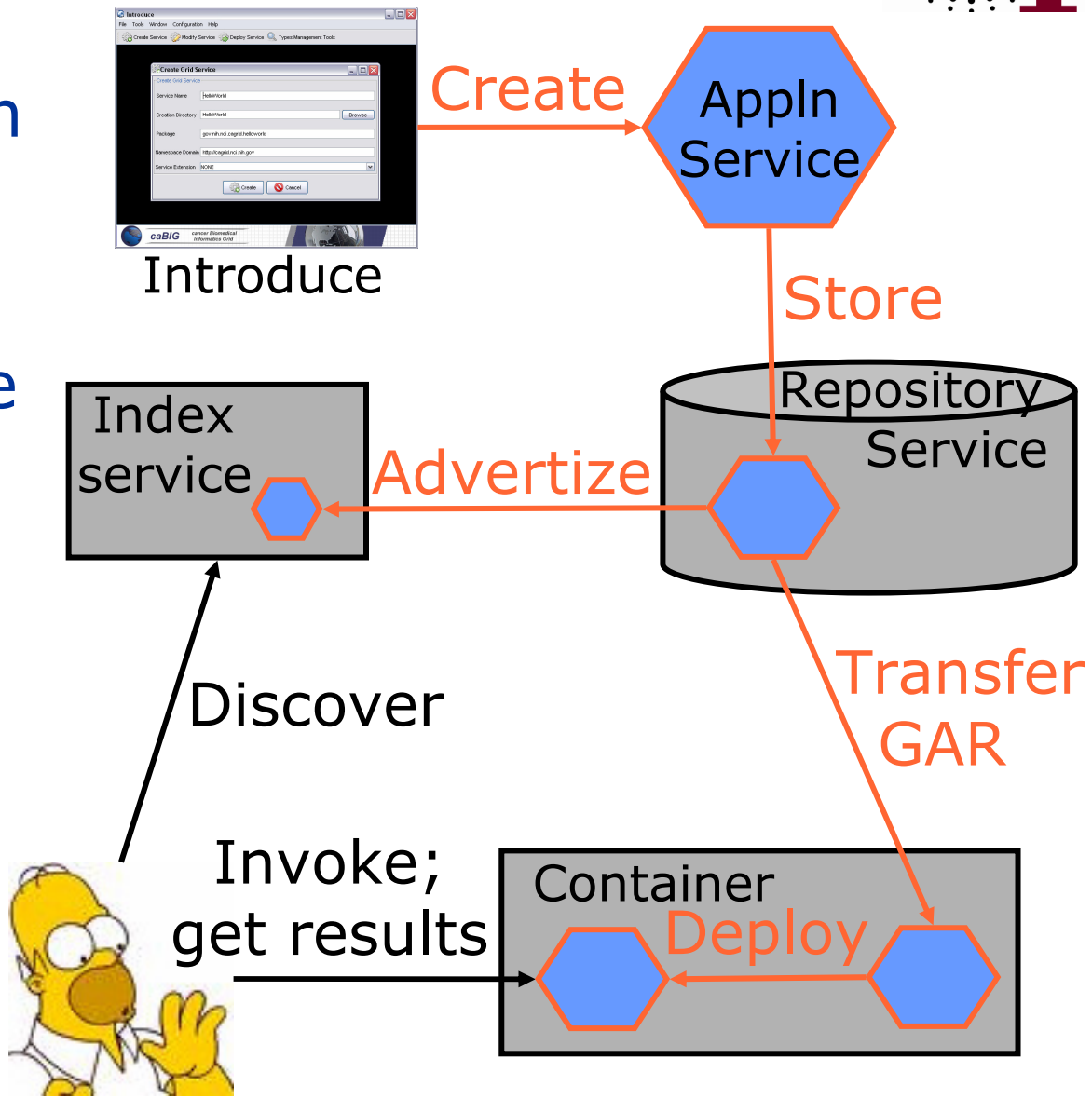


GT4 Web Services Container

Globus Toolkit Version 4: Software for Service-Oriented Systems, LNCS 3779, 2-13, 2005

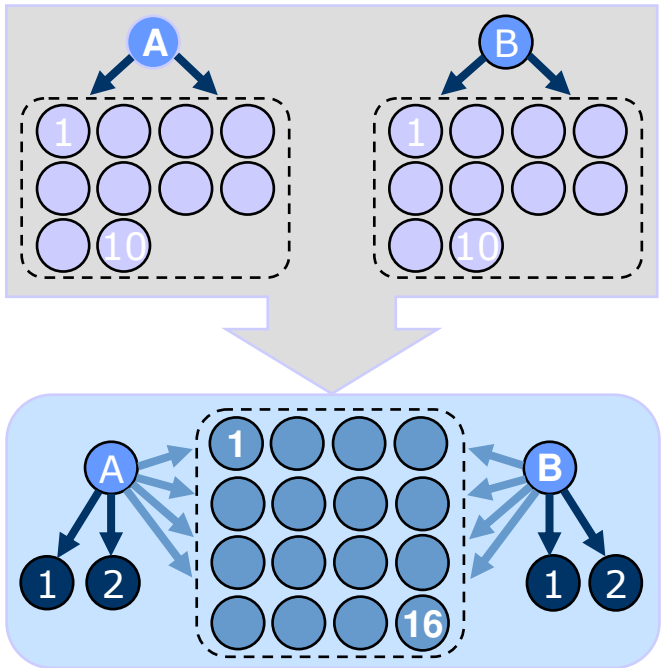
RAVE

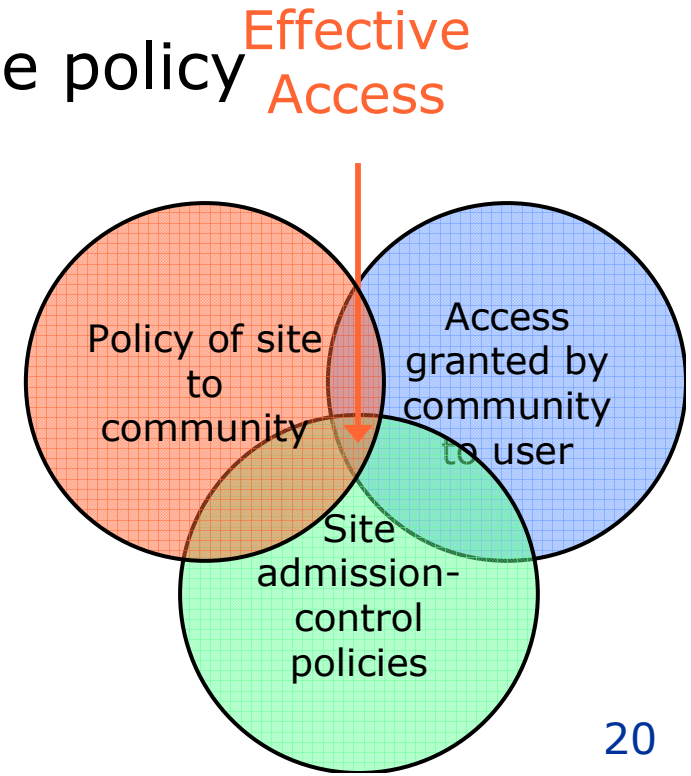
- **R**emote **A**pplication **V**irtualization **E**nvironment
- Builds on Introduce
  - Define service
  - Create skeleton
  - Discover types
  - Add operations
  - Configure security
- Wrap arbitrary executables

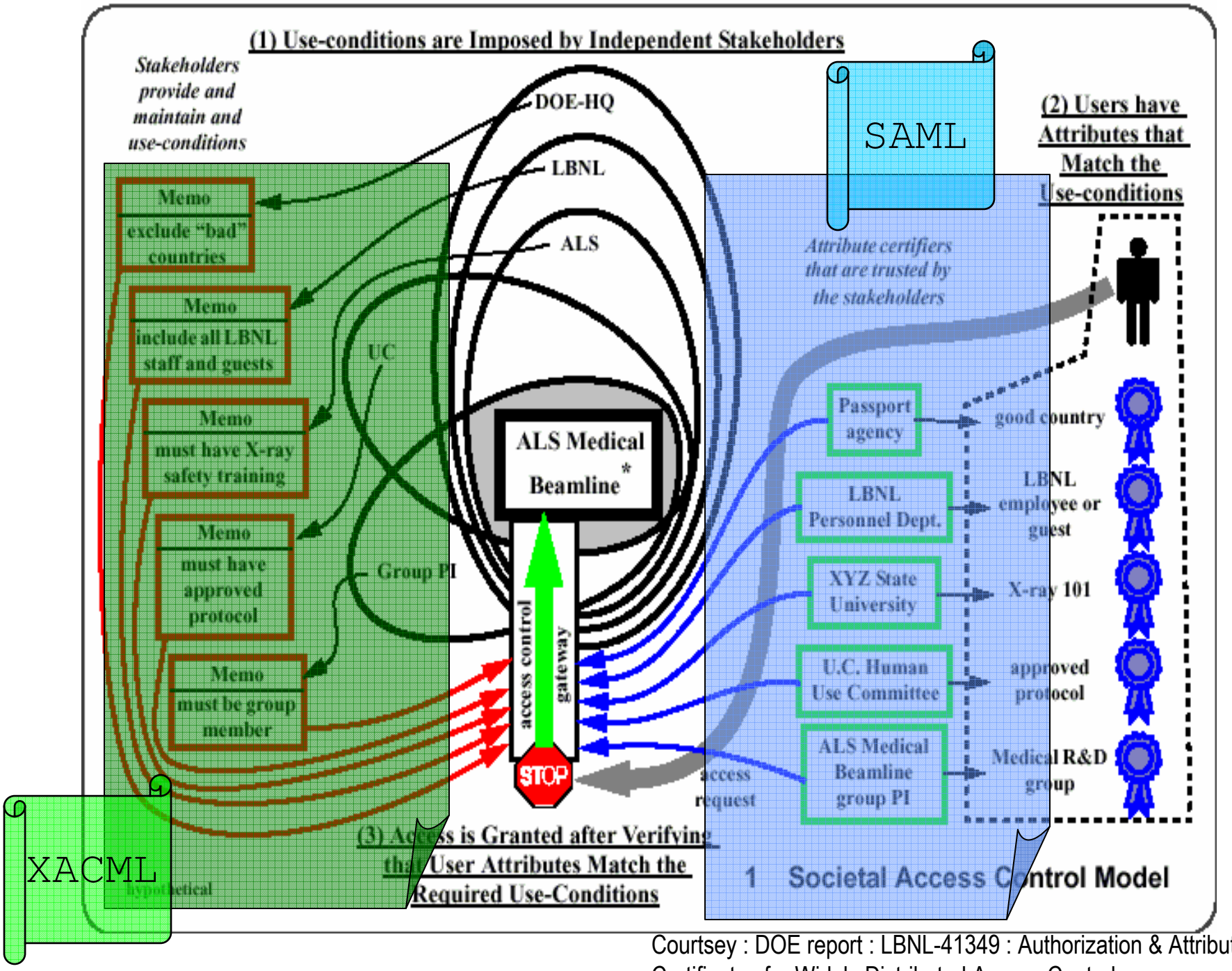


Defining Community: Membership and Laws

- Identify VO participants and roles
  - For people and services
- Specify and control actions of members
  - Empower members → delegation
  - Enforce restrictions → federate policy



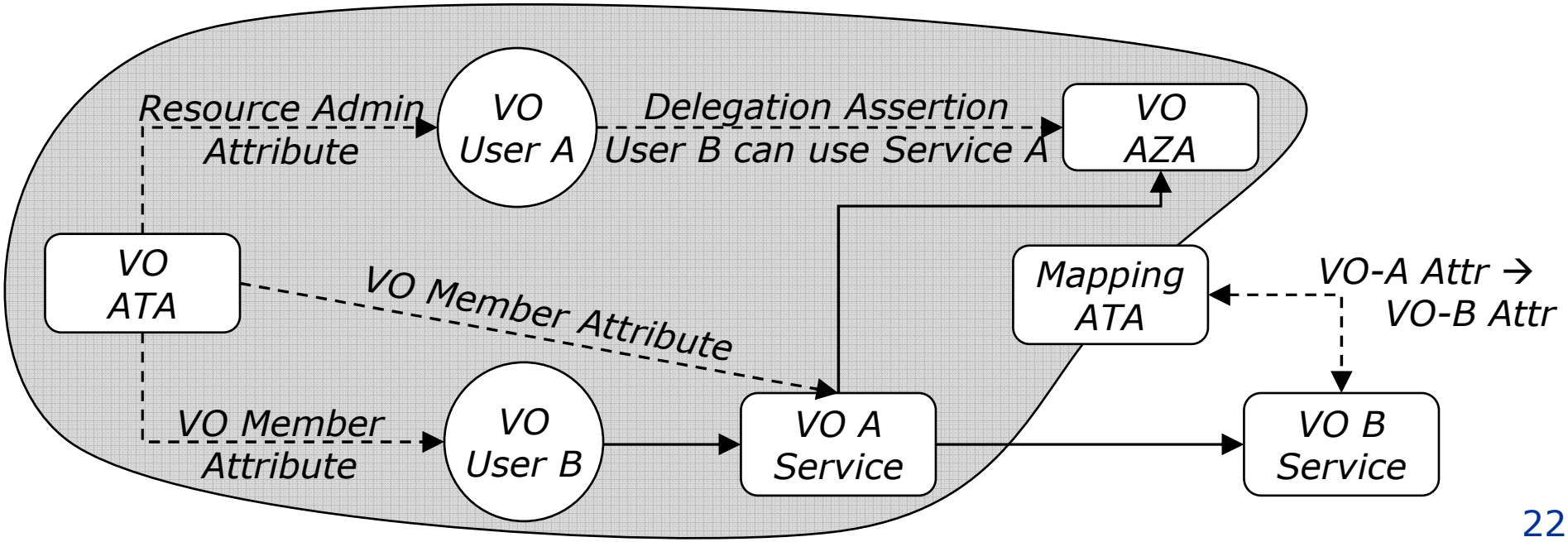




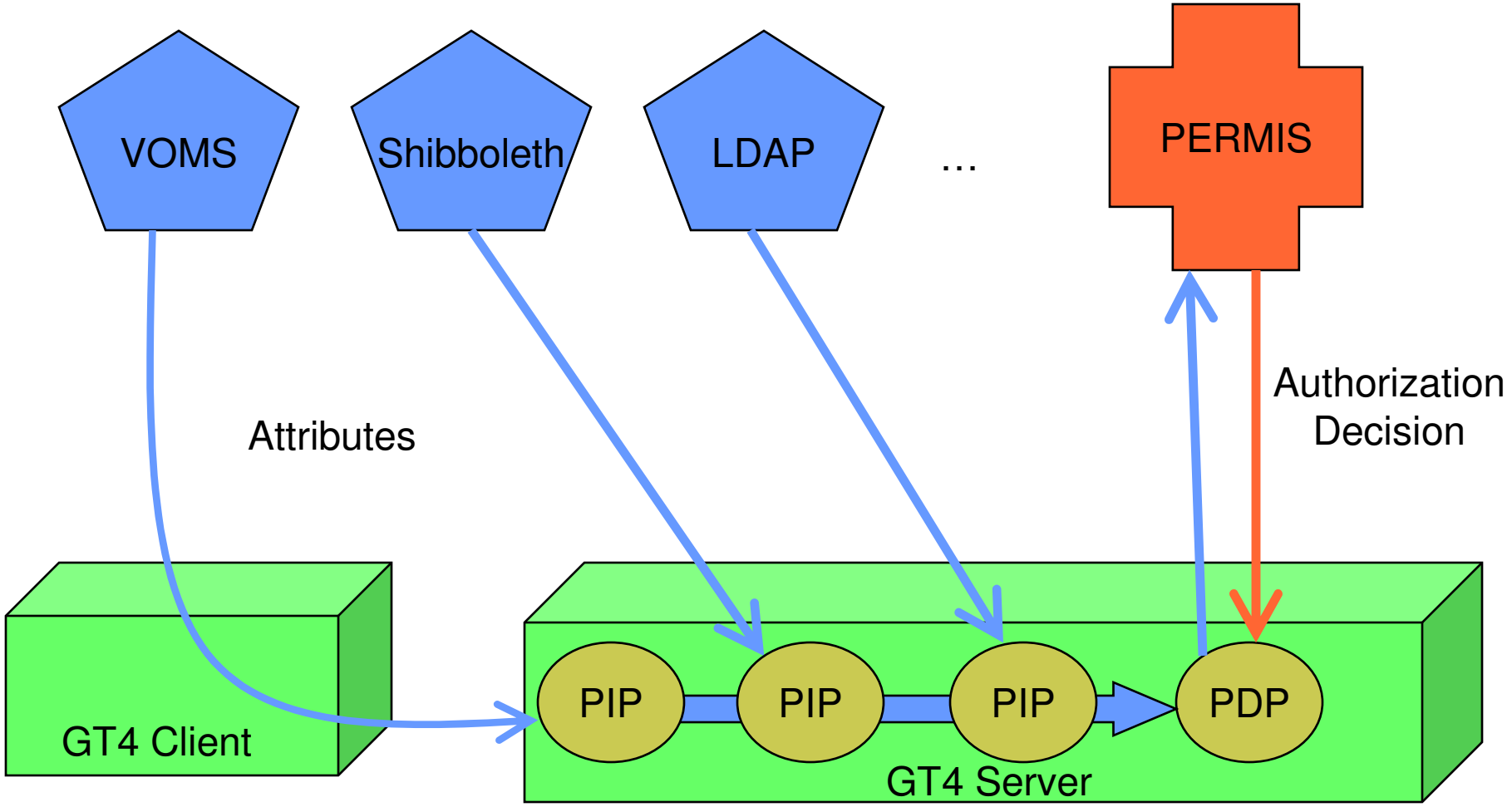
Courtesy : DOE report : LBNL-41349 : Authorization & Attribute Certificates for Widely Distributed Access Control

Security Services for VO Policy

- Attribute Authority (ATA)
  - Issue signed attribute assertions (incl. identity, delegation & mapping)
- Authorization Authority (AZA)
  - Decisions based on assertions & policy
- Use with message- or transport-level security

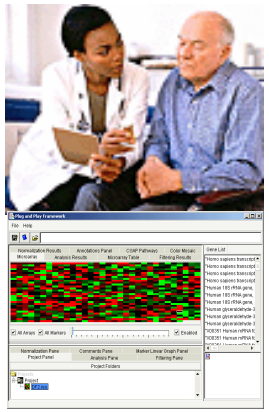


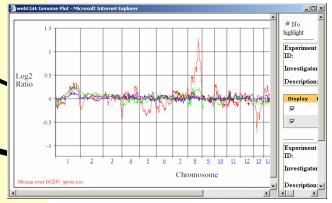
Globus Authorization Framework

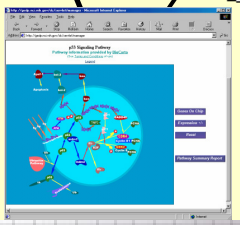


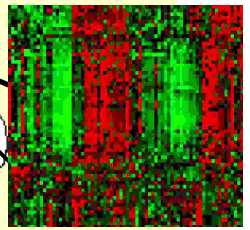
Service-Oriented Science & Cancer Biology

caBIG: sharing of infrastructure, applications, and data.







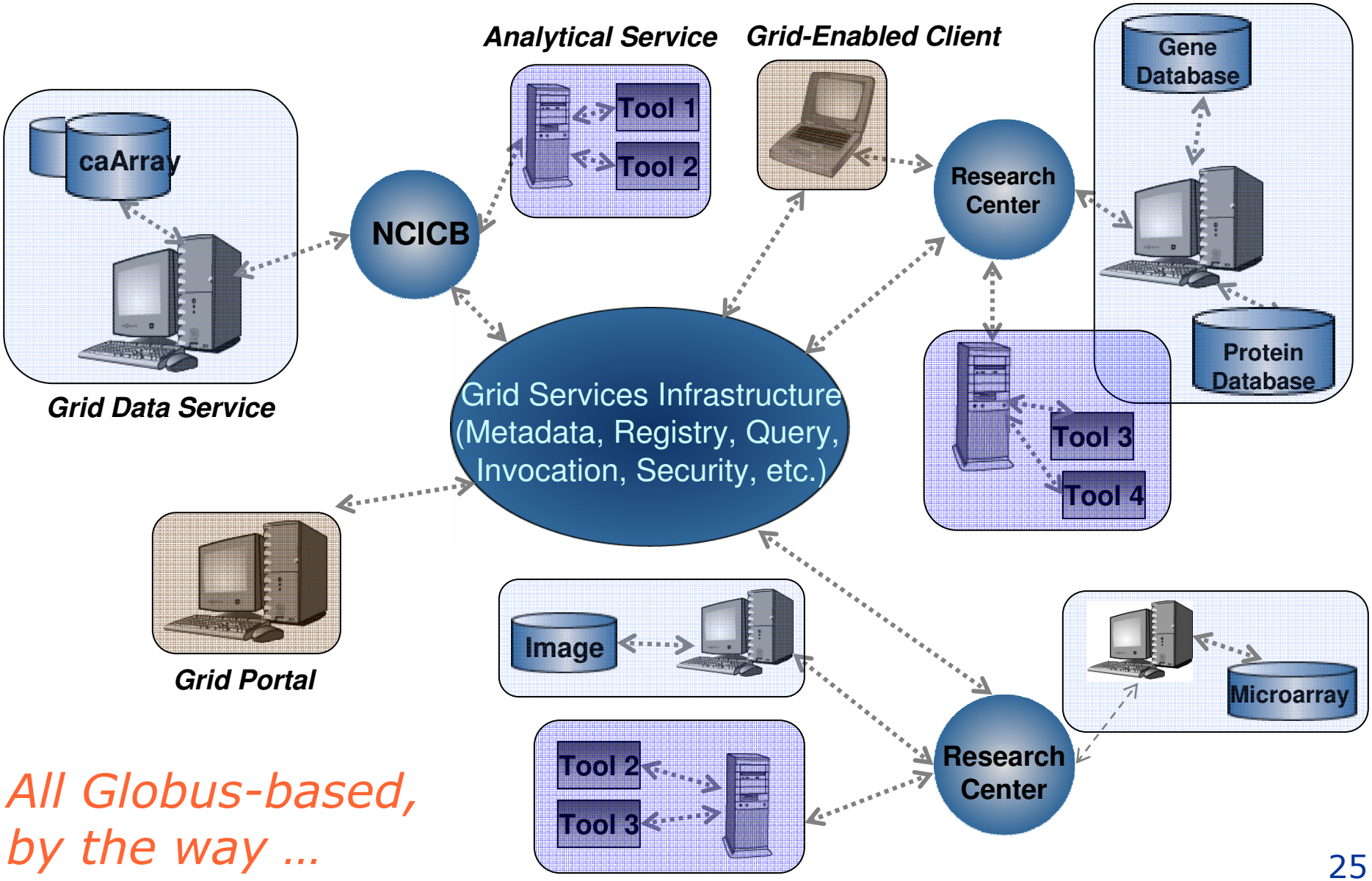


Data Integration!

caBIG cancer Biomedical Informatics Grid

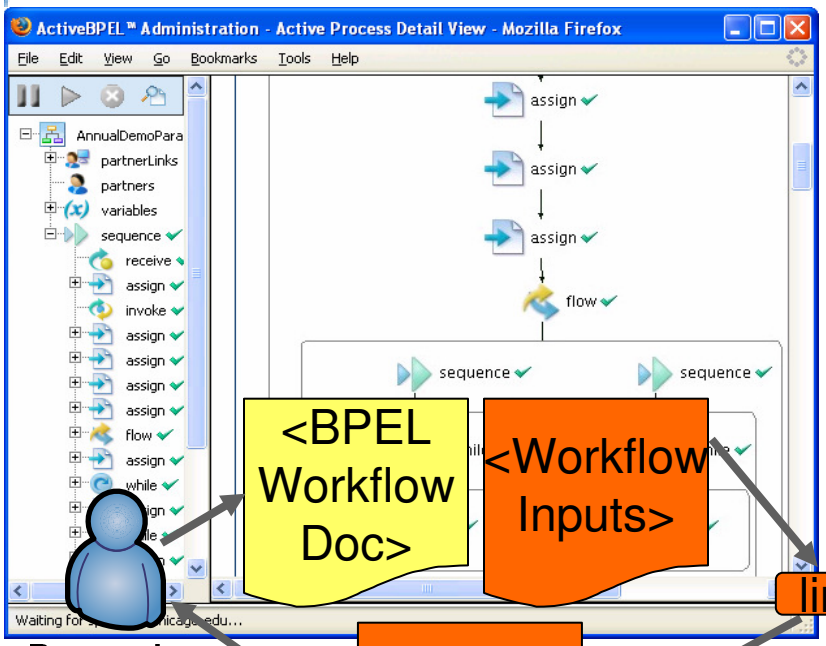
the globus alliance
www.globus.org

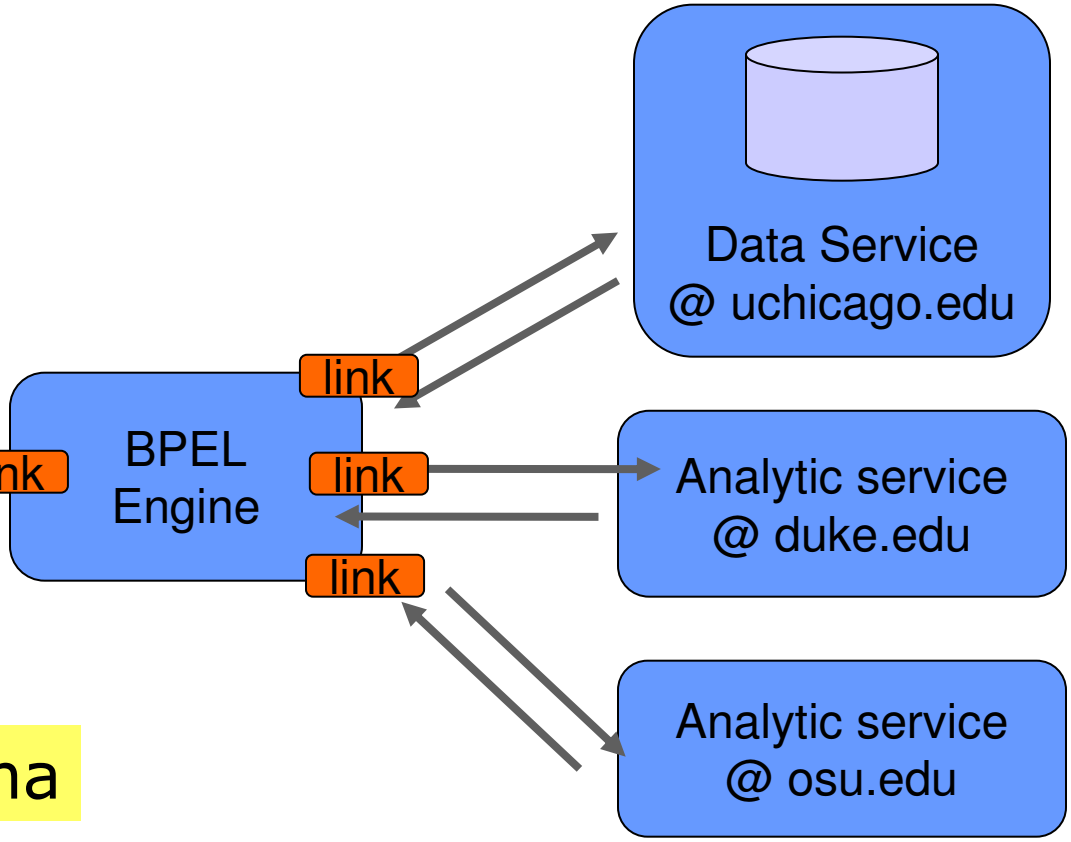
Cancer Bioinformatics Grid



All Globus-based, by the way …

Composing Services: E.g., BPEL Workflow System





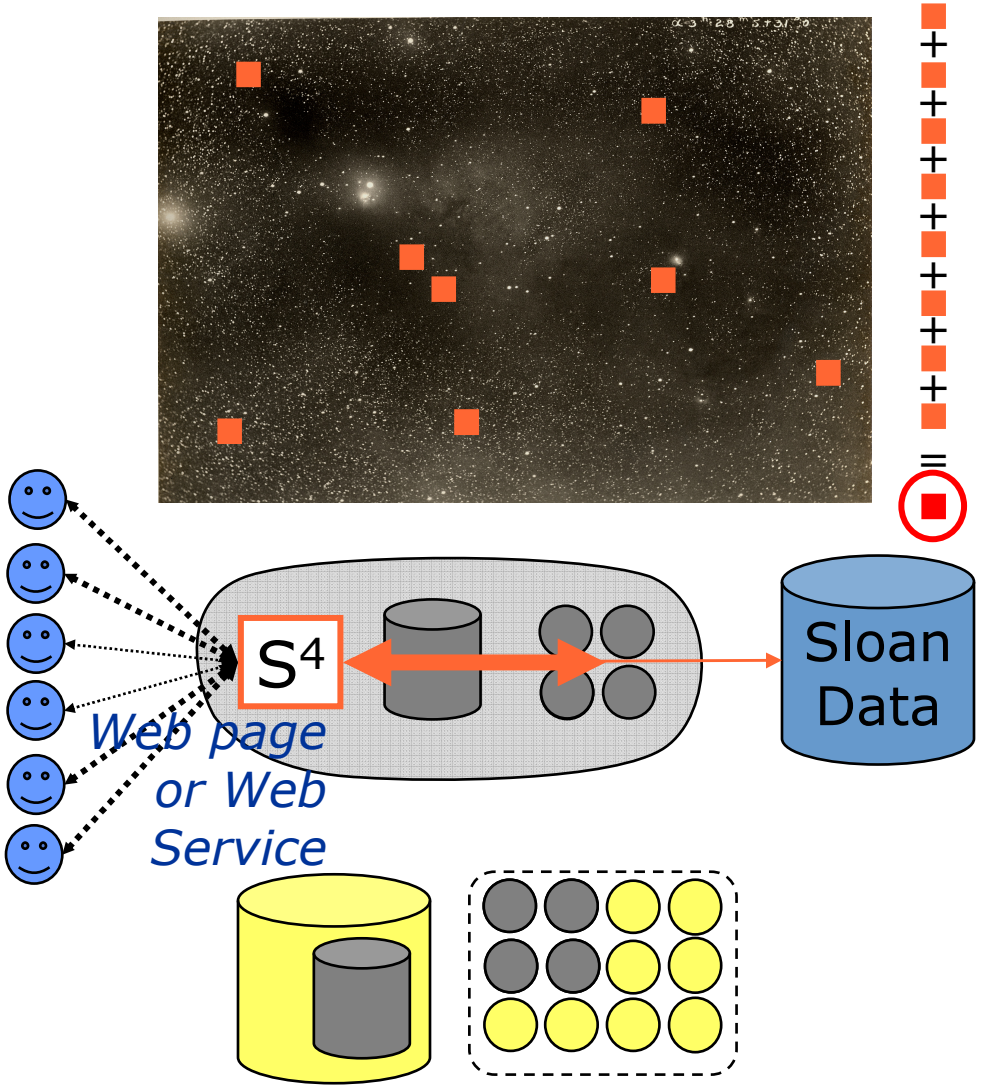
See also Kepler & Taverna

caBIG™ cancer Biomedical Informatics Grid™ an initiative of the National Cancer Institute

caBiG: https://cabig.nci.nih.gov/; BPEL work: Ravi Madduri et al.

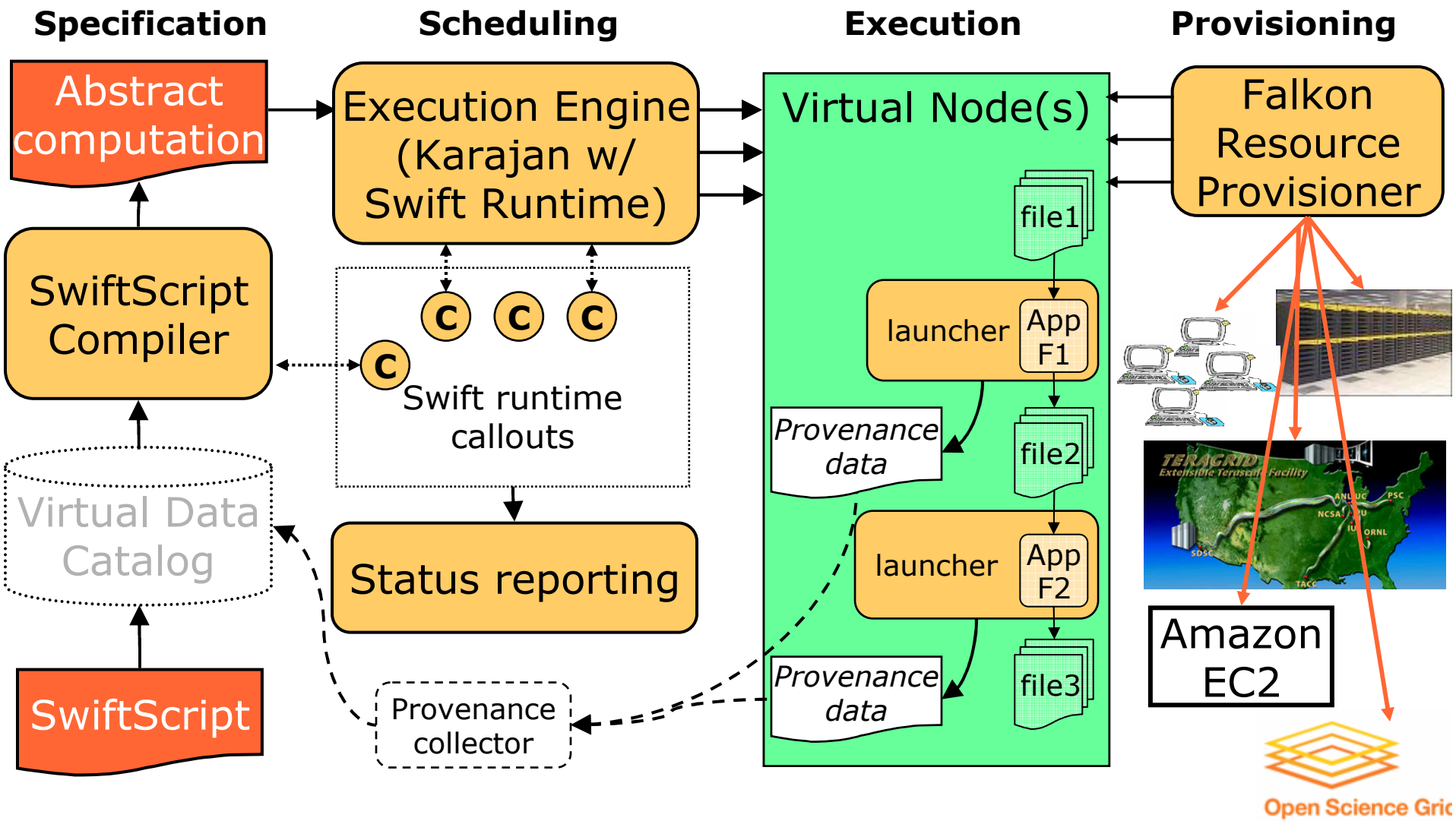
Provisioning: Astro Portal Stacking Service

- Purpose
  - On-demand "stacks" of random locations within ~10TB dataset
- Challenge
  - Rapid access to 10-10K "random" files
  - Time-varying load
- Solution
  - Dynamic acquisition of compute, storage



Joint work with Ioan Raicu & Alex Szalay

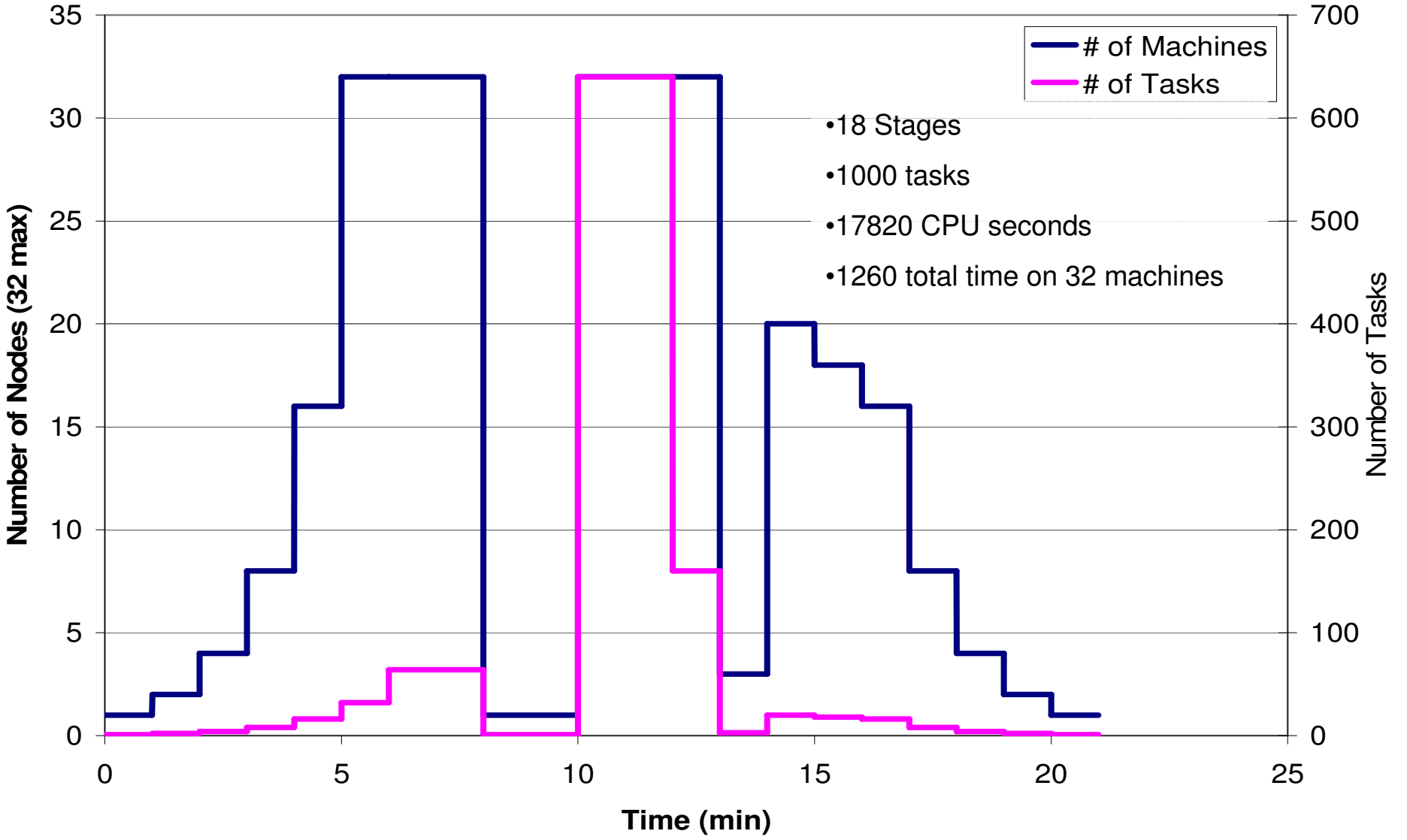
Dynamic Provisioning: Swift Architecture



Yong Zhao, Mihael Hatigan, Ioan Raicu, Mike Wilde, Ben Clifford

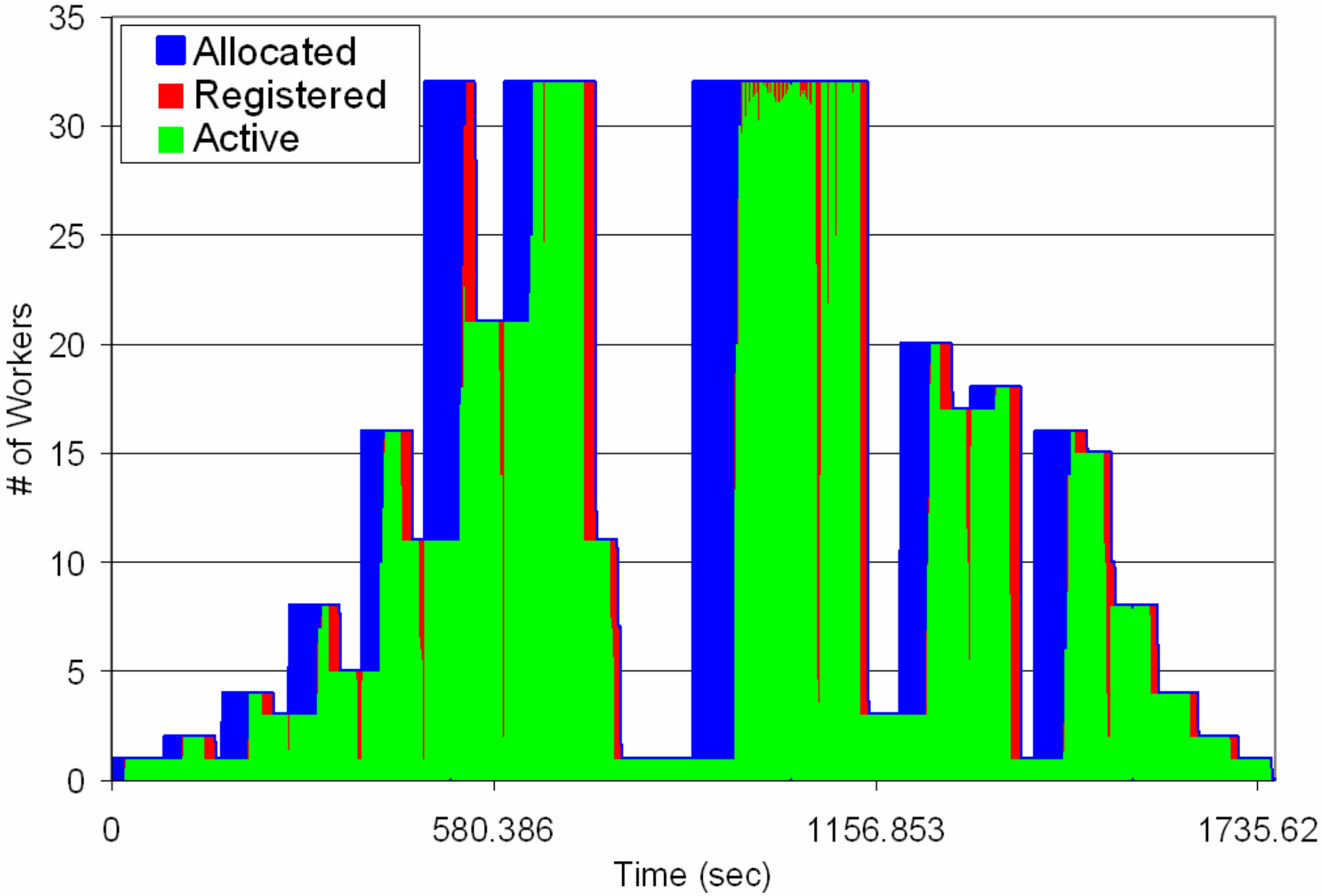
the globus alliance
www.globus.org

Synthetic Benchmark

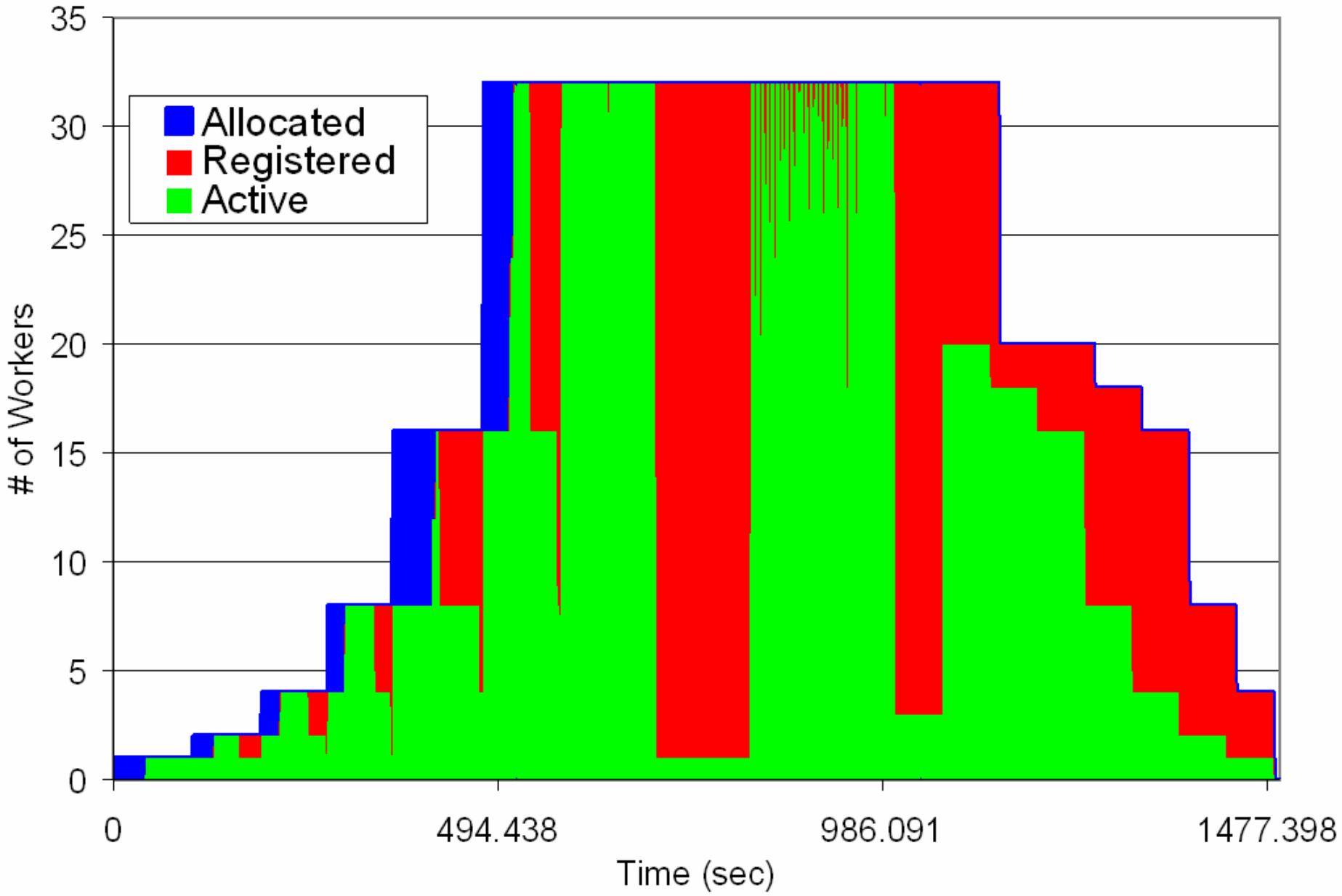


Ioan Raicu & Yong Zhao, U.Chicago

Release after 15 Seconds Idle

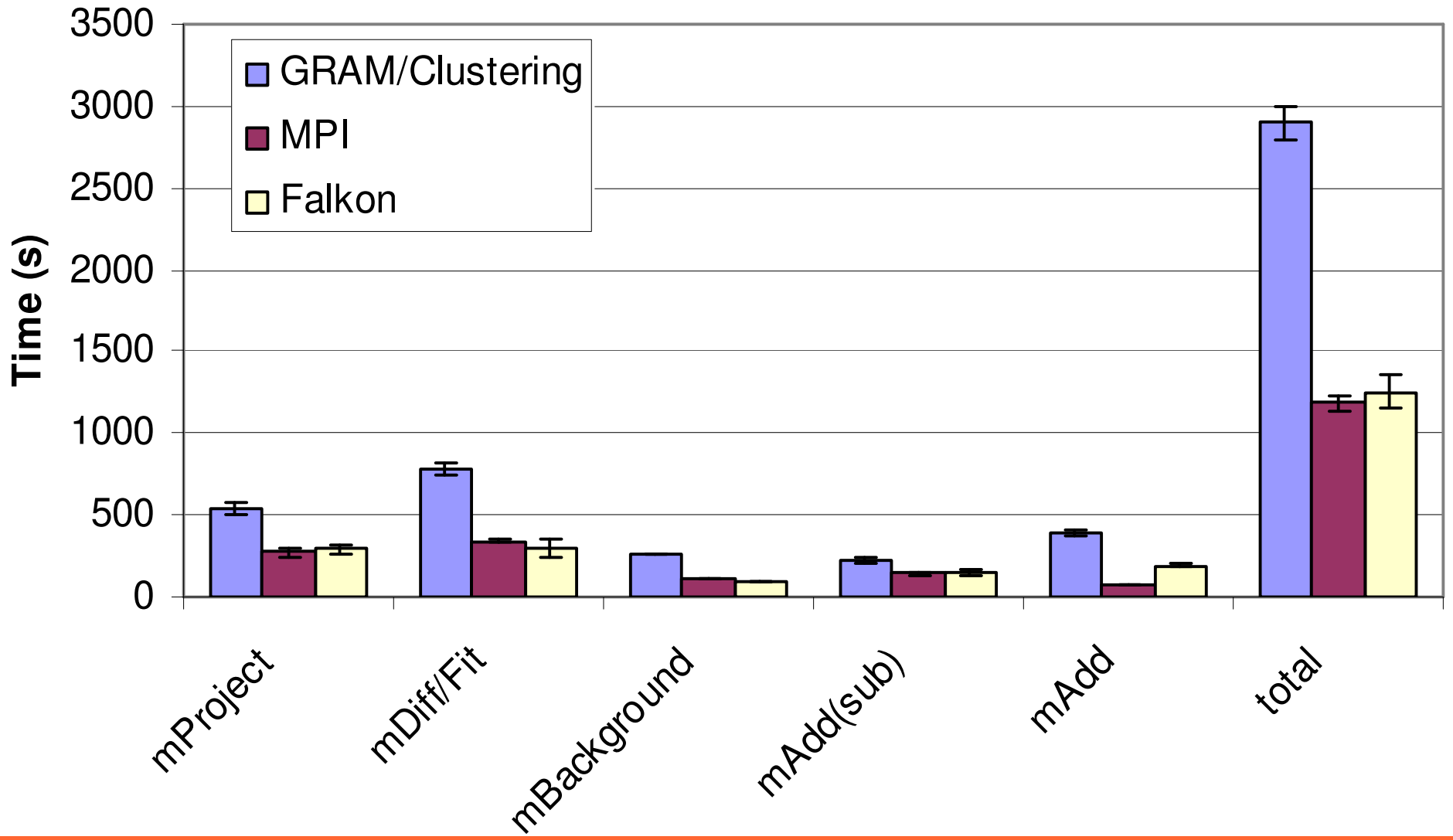


Release after 180 Seconds Idle

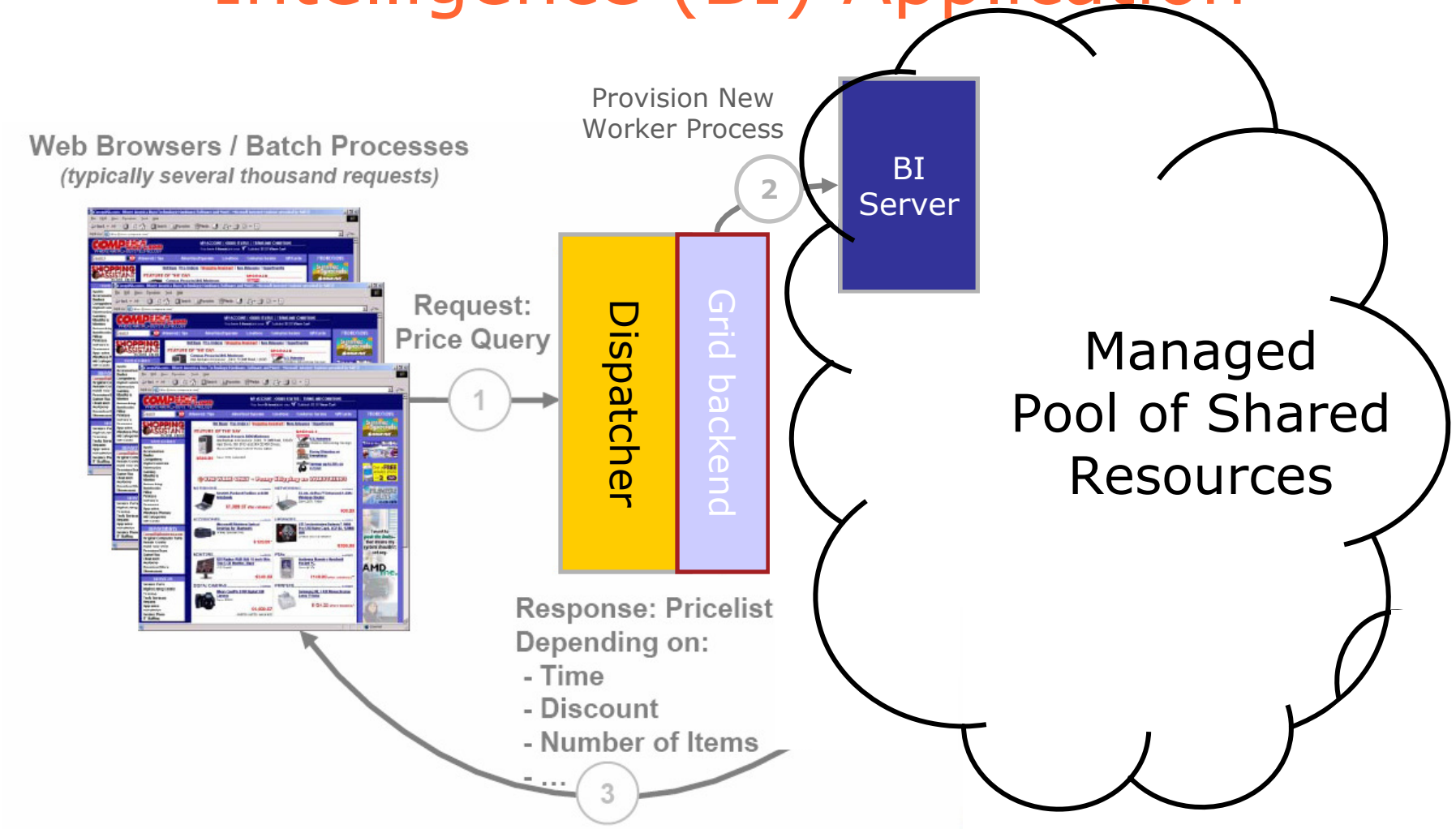


the globus alliance
www.globus.org

Montage



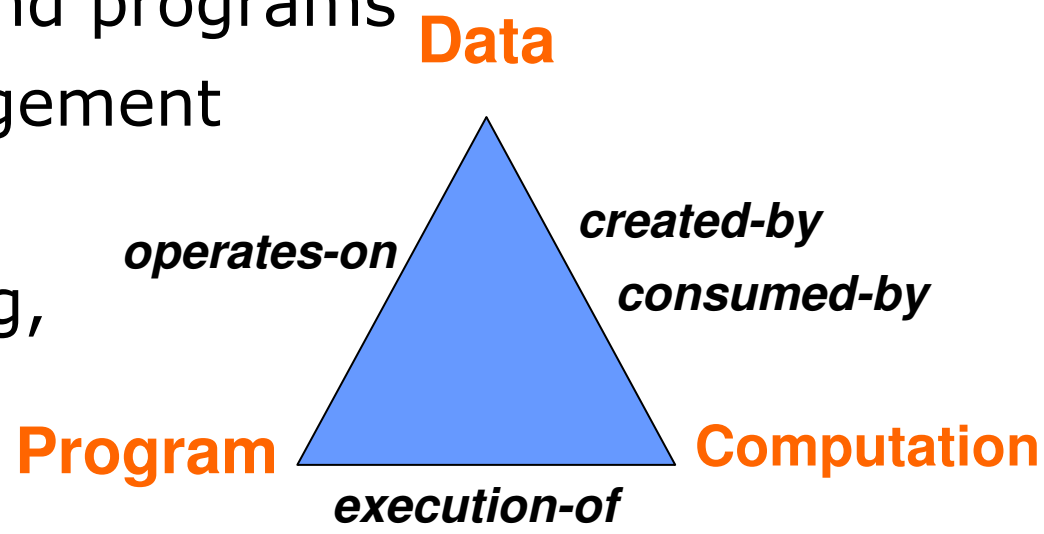
Grid-enabled Business Intelligence (BI) Application



BI server applications started and decommissioned by a Grid-enabled dispatcher

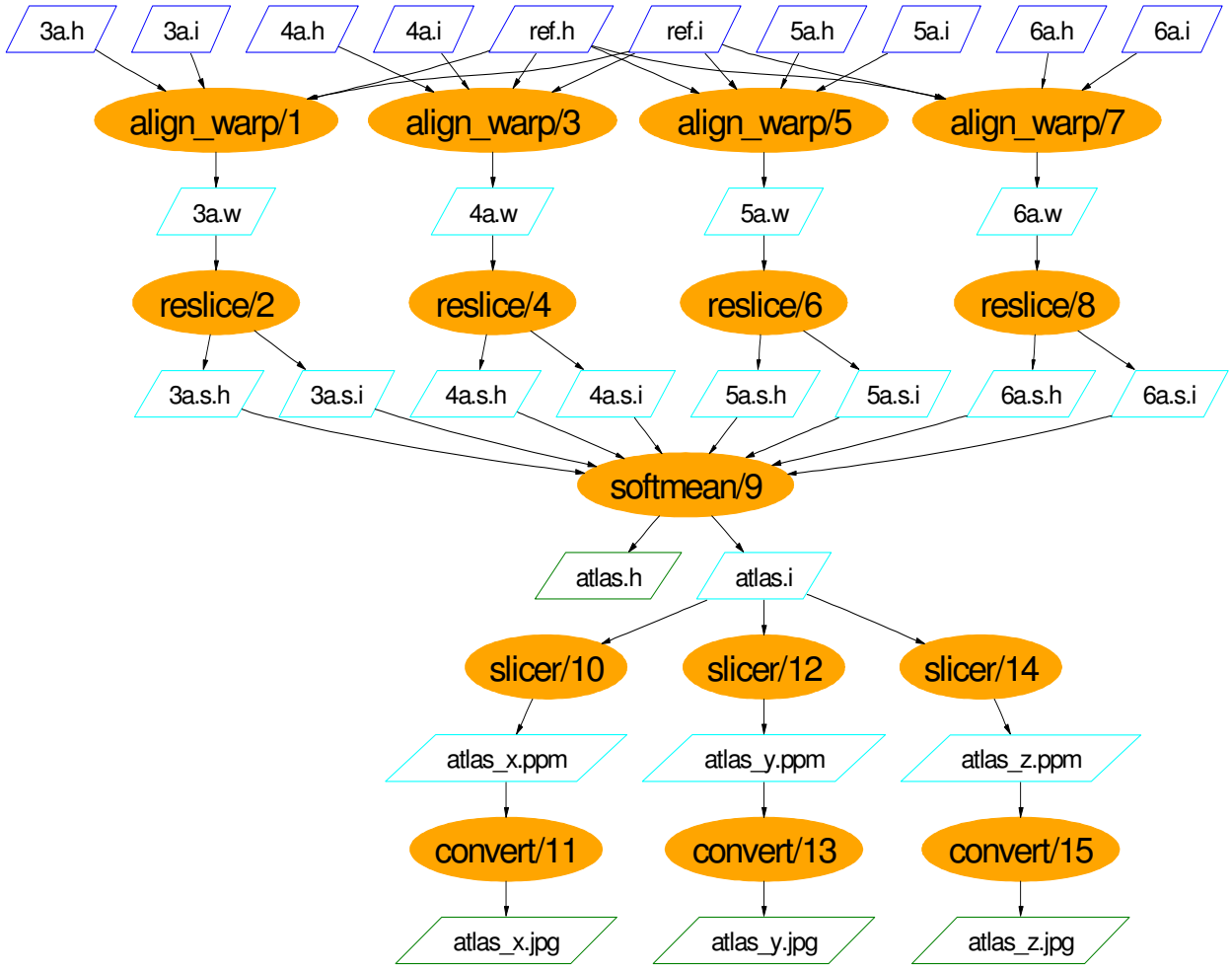
Computation as a First-Class Entity

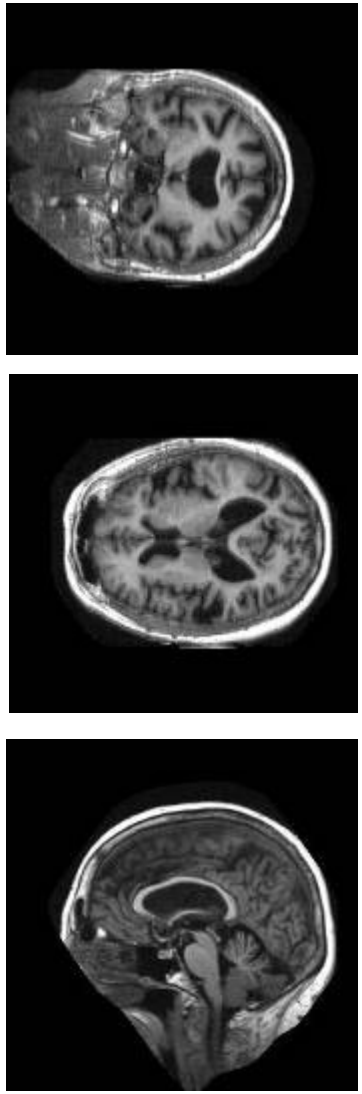
- Capture information about relationships among
  - Data (varying locations and representations)
  - Programs (& inputs, outputs, constraints)
  - Computations (& execution environments)
- Apply this information to:
  - Discovery of data and programs
  - Computation management
  - Provenance
  - Planning, scheduling, performance optimization



the globus alliance
www.globus.org

Example: fMRI Analysis





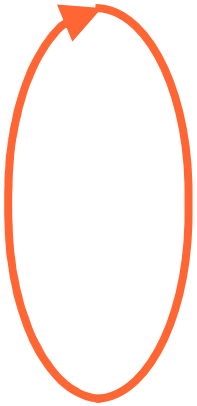
First Provenance Challenge, http://twiki.ipaw.info/ [CCPE06]

Query Examples

- Query by procedure signature
  - Show procedures that have inputs of type *subjectImage* and output types of *warp*
- Query by actual arguments
  - Show *align_warp* calls (including all arguments), with argument *model=rigid*
- Query by annotation
  - List anonymized subject images for young subjects:
    - Find datasets of type *subjectImage* , annotated with *privacy=anonymized* and *subjectType=young*
- Basic lineage graph queries
  - Find all datasets derived from dataset '5a'
- Graph pattern matching
  - Show me all output datasets of *softmean* calls that were aligned with *model=affine*

dev.globus — Community Driven Improvement of Globus Software, NSF OCI

Summary: Service-Oriented Science



People **create** services (data or functions) …

which I **discover** (& decide whether to use) …

& **compose** to create a new function …

& then **publish** as a new service.

→ *I find "someone else" to **host** services, so I don't have to become an expert in operating services & computers!*

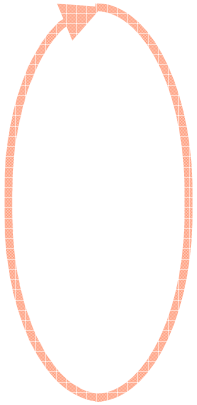
!!

→ *I hope that this "someone else" can **manage** security, reliability, scalability, …*

"Service-Oriented Science", *Science*, 2005

the globus alliance
www.globus.org

Summary: Service-Oriented Science



People **create** services (data or functions) …
which I **discover** (& decide whether to use) …
& **compose** to create a new function …
& then **publish** as a new service.

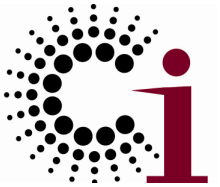
Profoundly revolutionary:

→ **Accelerates the pace of enquiry**

→ **Introduces a new notion of "result"**

→ **Requires new reward structures, training, infrastructure**

"Service-Oriented Science", *Science*, 2005



Science 1.0 → Science 2.0

Megabytes & gigabytes	→ Terabytes & petabytes
Tarballs	→ Services
Journals	→ Wikis
Individuals	→ Communities
Community codes	→ Science gateways
Supercomputer centers	→ Campus & national grids …
Makefile	→ Workflow
Computational science	→ Science as computation
Mostly physical sciences	→ All sciences (& humanities)
1000s of computationalists	→ Millions of scientists
Government funded	→ Government funded

the globus alliance
www.globus.org

Thanks!

- DOE Office of Science 

- NSF Office of Cyberinfrastructure 
- Colleagues at Argonne, U.Chicago, USC/ISI, and elsewhere
- Many members of the German DGrid community

Service-Oriented Science Challenges

- A need for new **technologies, skills, & roles**
  - Creating, publishing, hosting, discovering, composing, archiving, explaining … services
- A need for substantial **software development**
  - "30-80% of modern astronomy projects is software"—S. G. Djorgovski, Caltech
- A need for more & different **infrastructure**
  - Computers & networks to host services
- And certainly profound **research challenges**
  - In every part of the service & science lifecycle

For more information: **http://ianfoster.typepad.com**