Grid meets Cloud: Using hosted service for reliable data movement in distributed computing

Raj Kettimuthu

Argonne National Laboratory and

The University of Chicago

the globus alliance
www.globus.org

Outline

- Grid
- Cloud
- Grid vs Cloud
- Globus.org
- GridFTP
- Globus.org – What does it bring for GridFTP

What is a Grid?

- Resource sharing
  - Computers, storage, sensors, networks, …
  - Sharing always conditional: issues of trust, policy, negotiation, …
- Coordinated problem solving
  - Beyond client-server: distributed data analysis, computation, collaboration, …
- Dynamic, multi-institutional virtual organizations
  - Community overlays on classic org structures
  - Large or small, static or dynamic

the globus alliance
www.globus.org

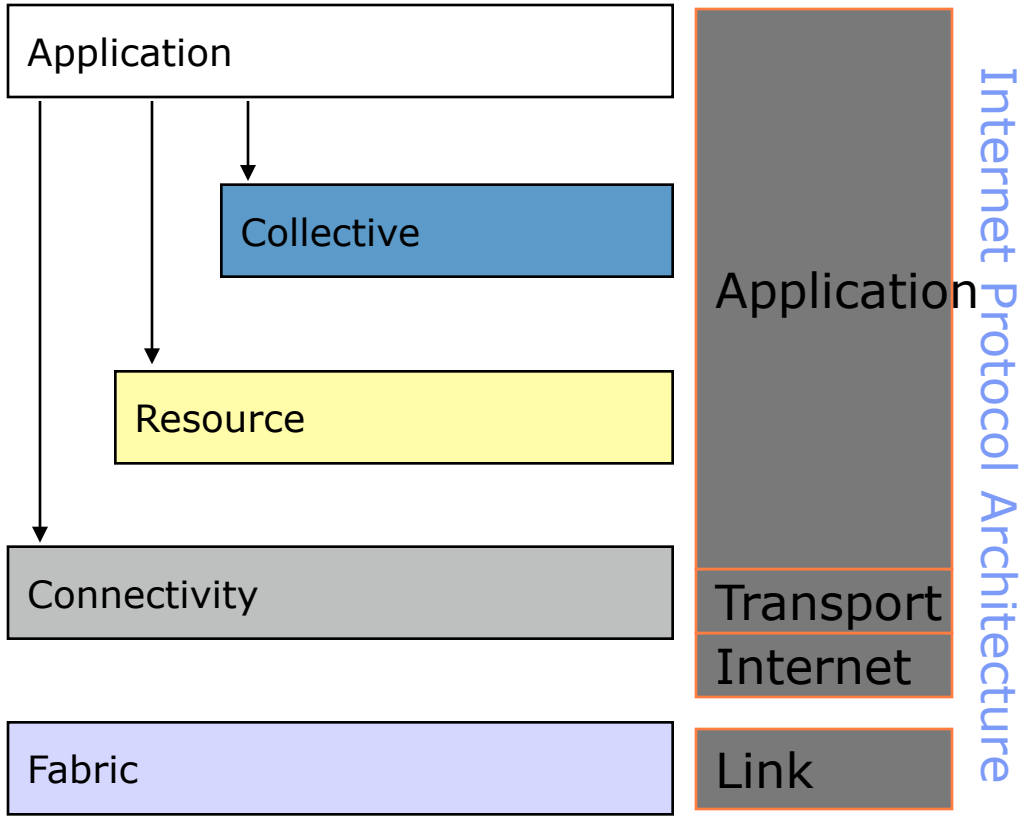
The Grid

"Coordinating multiple resources": ubiquitous infrastructure services, app-specific distributed services

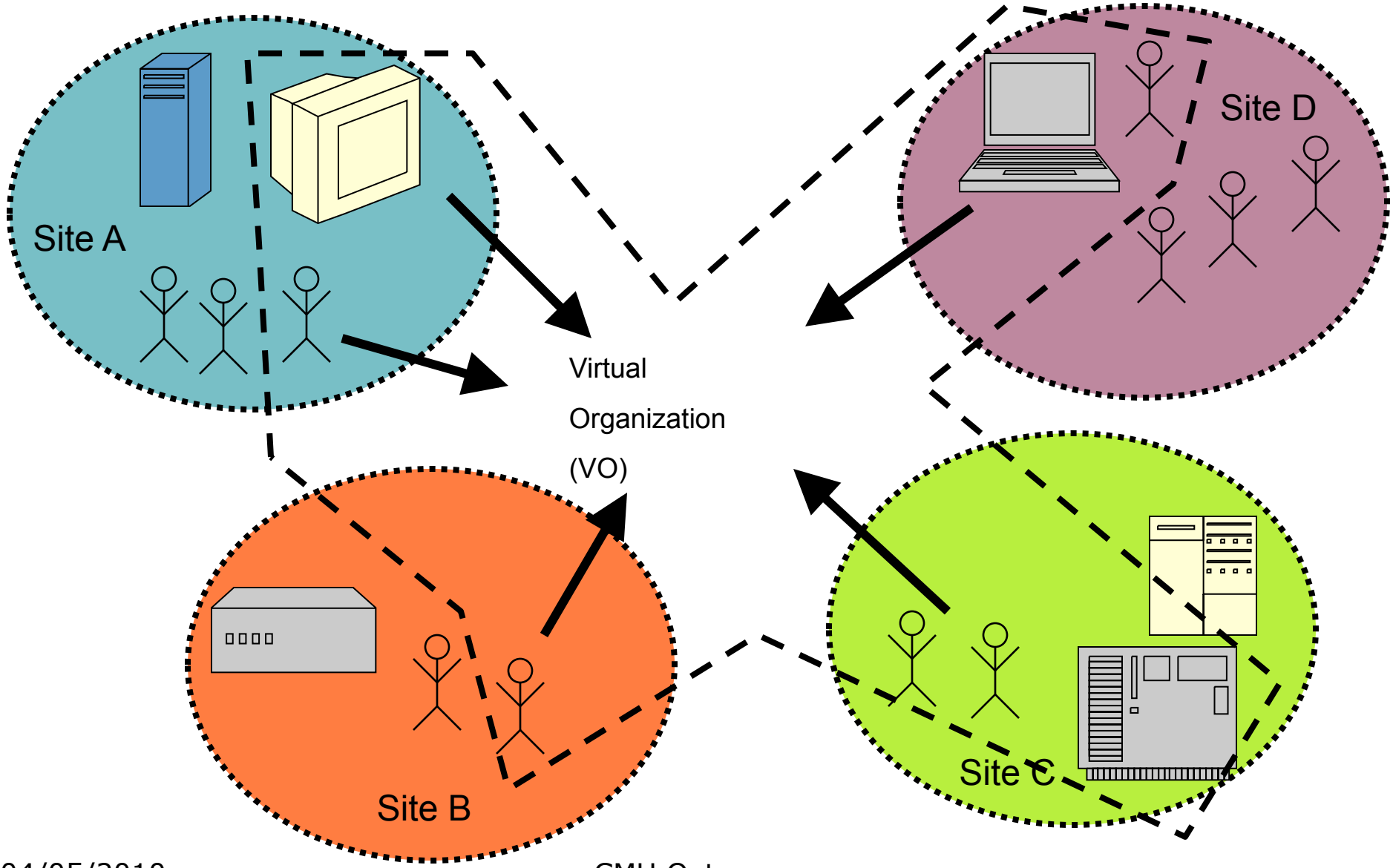
"Sharing single resources": negotiating access, controlling use

"Talking to things": communication (Internet protocols) & security

"Controlling things locally": Access to, & control of, resources



Virtual Organization



Virtual Organization

- Enable disparate groups of organizations and/or individuals to share resources in a controlled fashion
  - Members may collaborate to achieve a shared goal.
- Grid architecture defines basic mechanisms by which VO users & resources negotiate, establish & manage sharing relationships.
  - The key is provide seamless access to the resources for the users
  - Providing uniform security mechanism to access all the resources in a VO is critical

Globus Toolkit

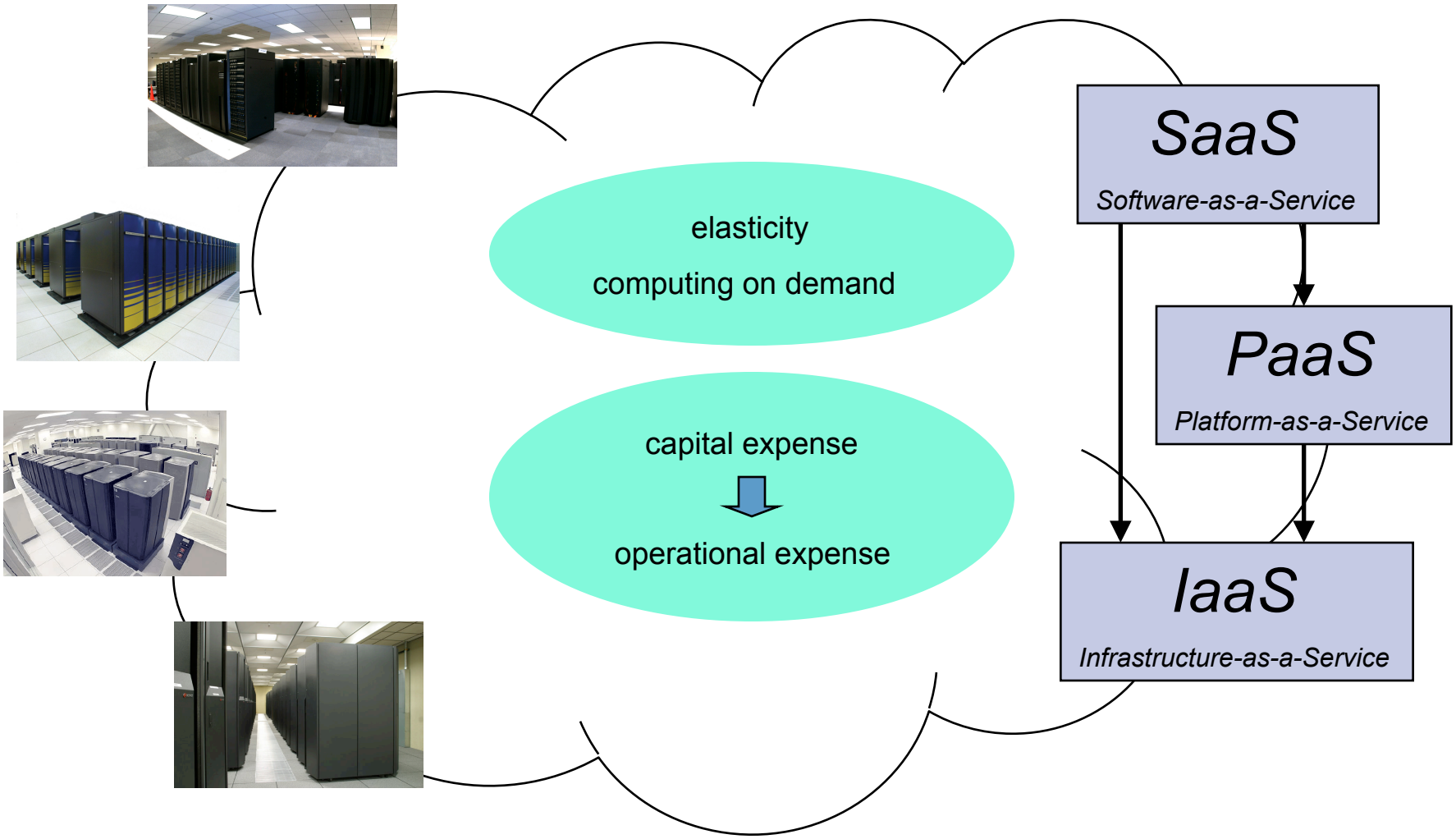
Assortment of Components for Grid Builders

- Focus on Connectivity and Resource layers
  - GRAM, GSI-OpenSSH: Run Programs
  - GridFTP: Access file systems
  - OGSA-DAI, caGrid: Access databases
  - GSI, MyProxy, GAARDS: Security
  - XIO, Java Core, C Core: Communication
- A few simple collective layer components
  - RLS: Replica tracking
  - RFT: Reliable file transfer

Many production grids

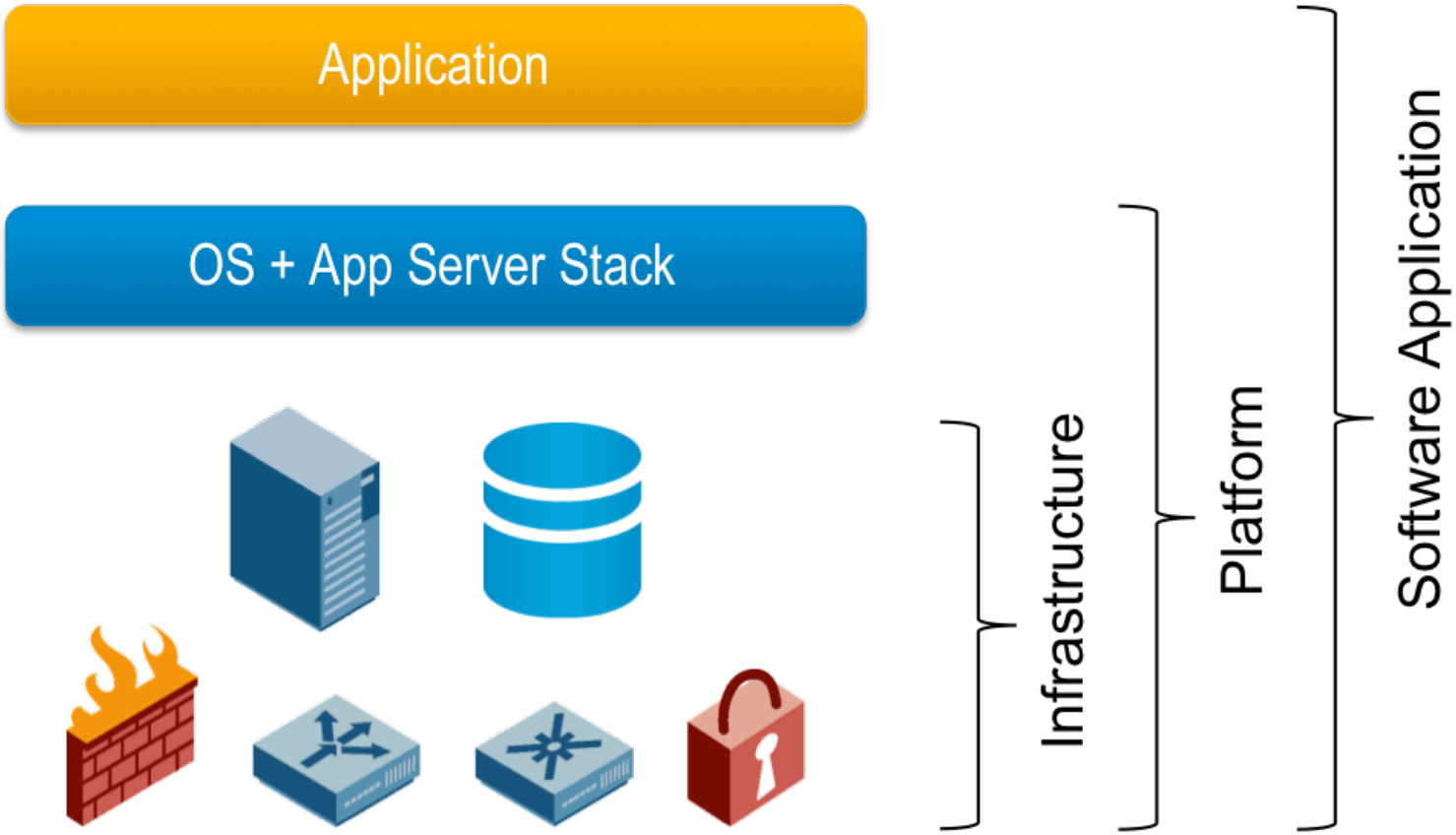
- BESTGrid
- BIRN
- caGrid
- ChinaGrid
- dGrid
- EGEE
- Garuda Grid
- LHC Computing Grid
- LIGO
- Open Science Grid
- TeraGrid
- ThaiGrid
- TIGRE
- UC Grid
- …

Cloud Computing



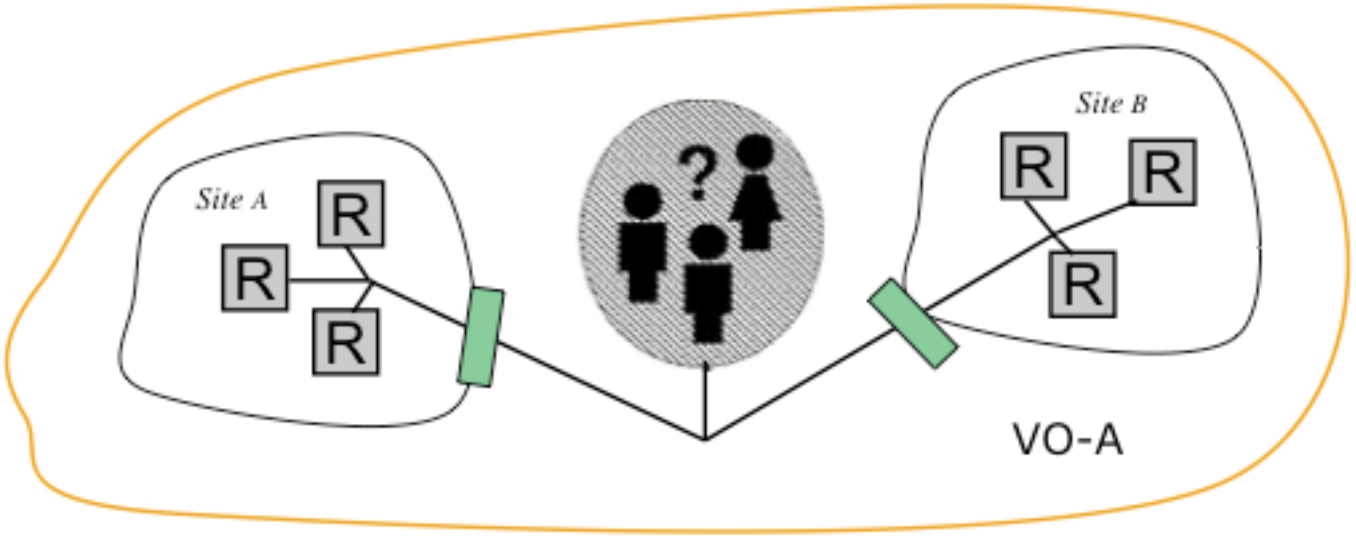
the globus alliance
www.globus.org

Cloud Computing



Grid Computing

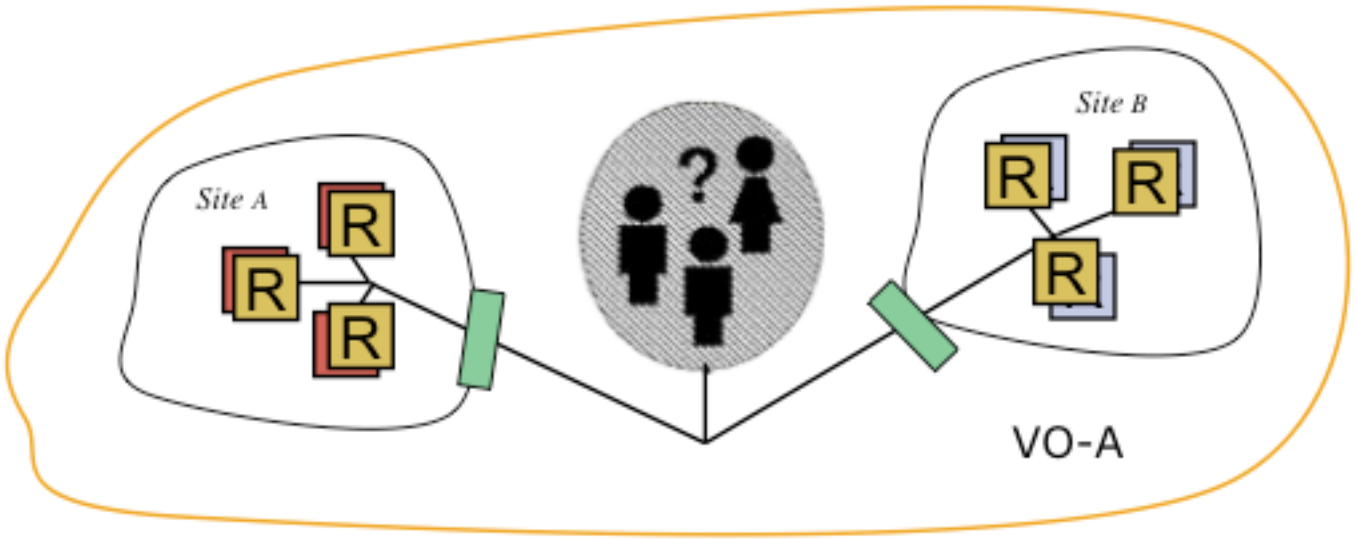
- *Assumption: control over the manner in which resources are used stays with the site*



- Site-specific environment and mode of access
- Site-driven prioritization

Cloud Computing

- *Change of assumption: control over the resource is turned over to the user*



- Enabling factors: virtualization and isolation
- Challenges our notion of a site

Grid = federation

Cloud = hosting

Many Custom Collective Layers

- Most Grid deployment have custom, domain-specific Collective layer
  - Built on common Resource layer components
  - E.g. Data transfer and mirroring, workflows, …
- Challenges:
  - Expensive to develop
  - Expensive to operating and supporting
  - Useful for narrow community
- How do we make these capabilities available to more users?

the globus alliance
www.globus.org

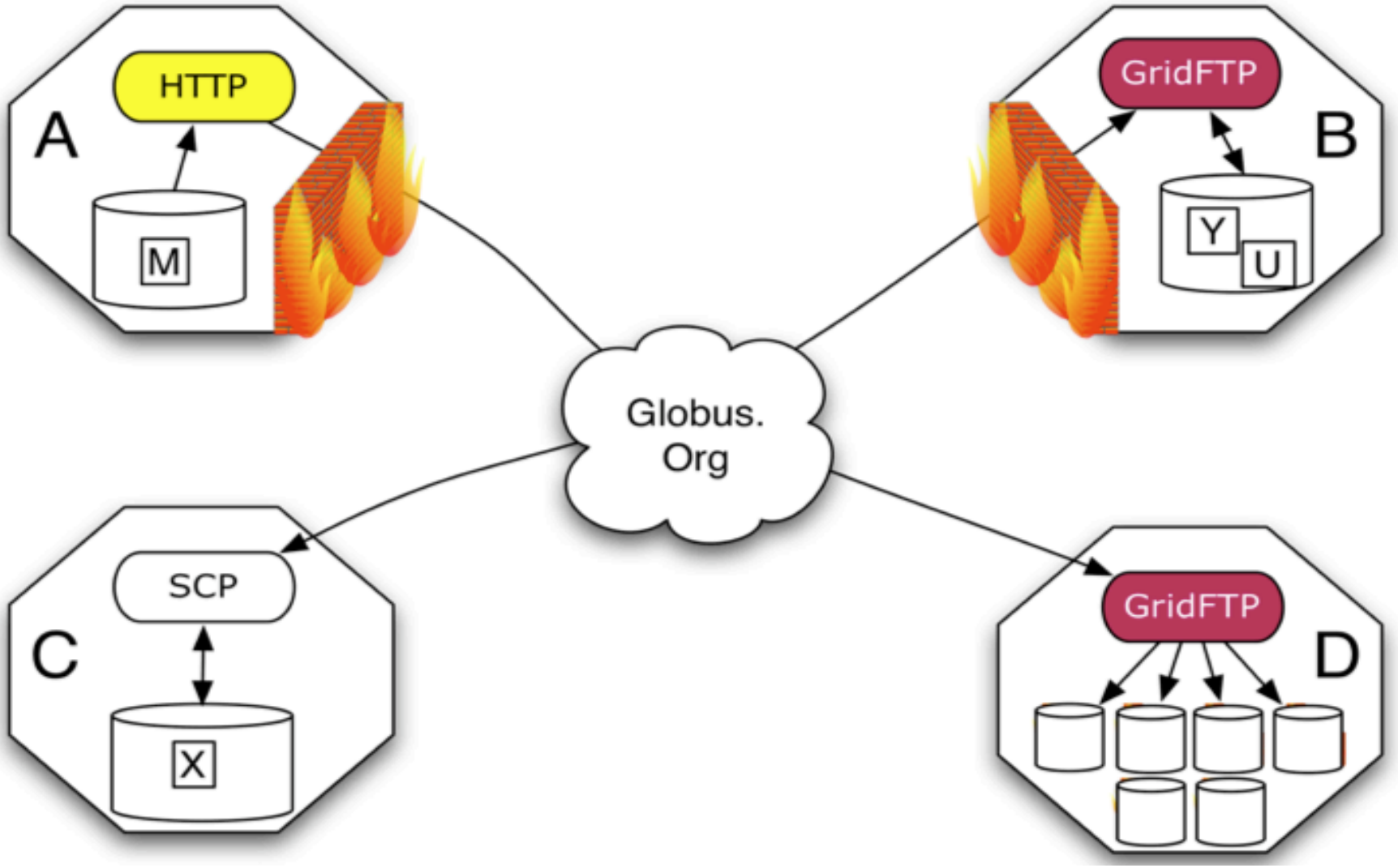
Globus.org Service
Collective Layer Grid Services via Cloud SaaS

- End-to-end collective layer functionality targeted toward end users
  - Generalize lessons from custom Grids
  - Focus on ease of use, federation
- Hosted and supported by Globus team
- Initial focus on file transfer
  - Near term: Add sync, mirroring, caching
  - Long term: Add job execution, workflows, VO management

Globus.org hosted services: Data replication as an example

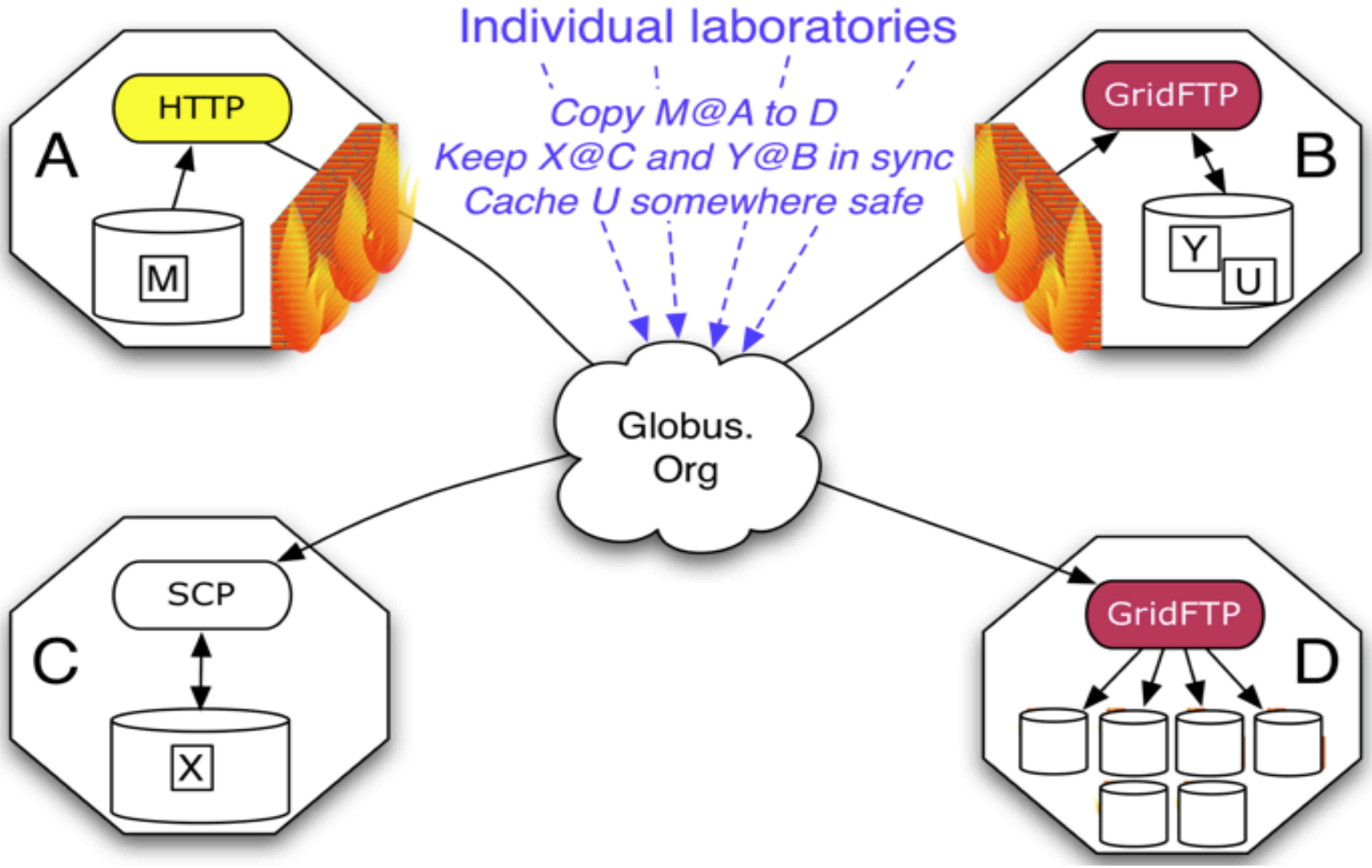
the globus alliance
www.globus.org

Globus.org



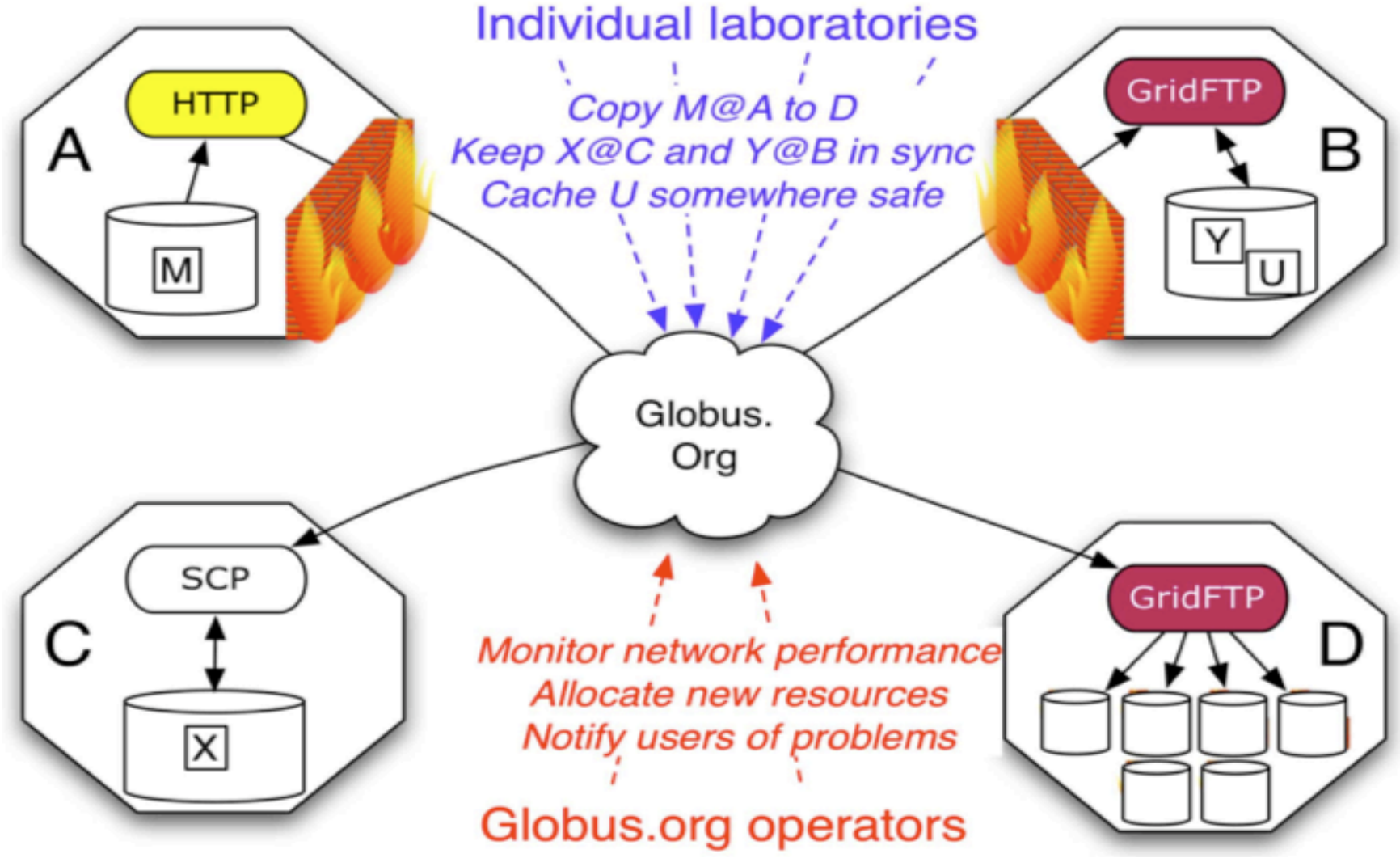
the globus alliance
www.globus.org

Globus.org



the globus alliance
www.globus.org

Globus.org



the globus alliance
www.globus.org

GridFTP

- High-performance, reliable data transfer protocol optimized for high-bandwidth wide-area networks
- Based on FTP protocol - defines extensions for high-performance operation and security
- Standardized through Open Grid Forum (OGF)
- GridFTP is the OGF recommended data movement protocol

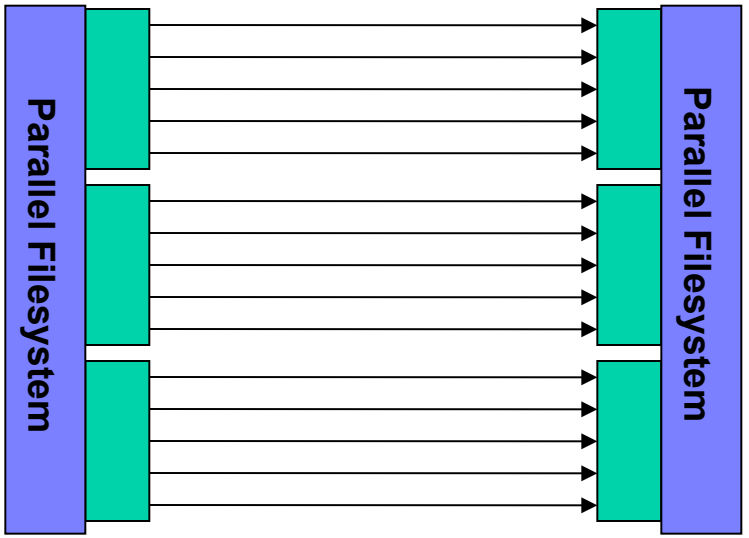
the globus alliance
www.globus.org

Globus GridFTP

- Performance
  - Parallel TCP streams, optimal TCP buffer
  - Non TCP protocol such as UDT
- Cluster-to-cluster data movement
- Multiple security options
  - Anonymous, password, SSH, GSI
- Support for reliable and restartable transfers

Cluster-to-Cluster transfers

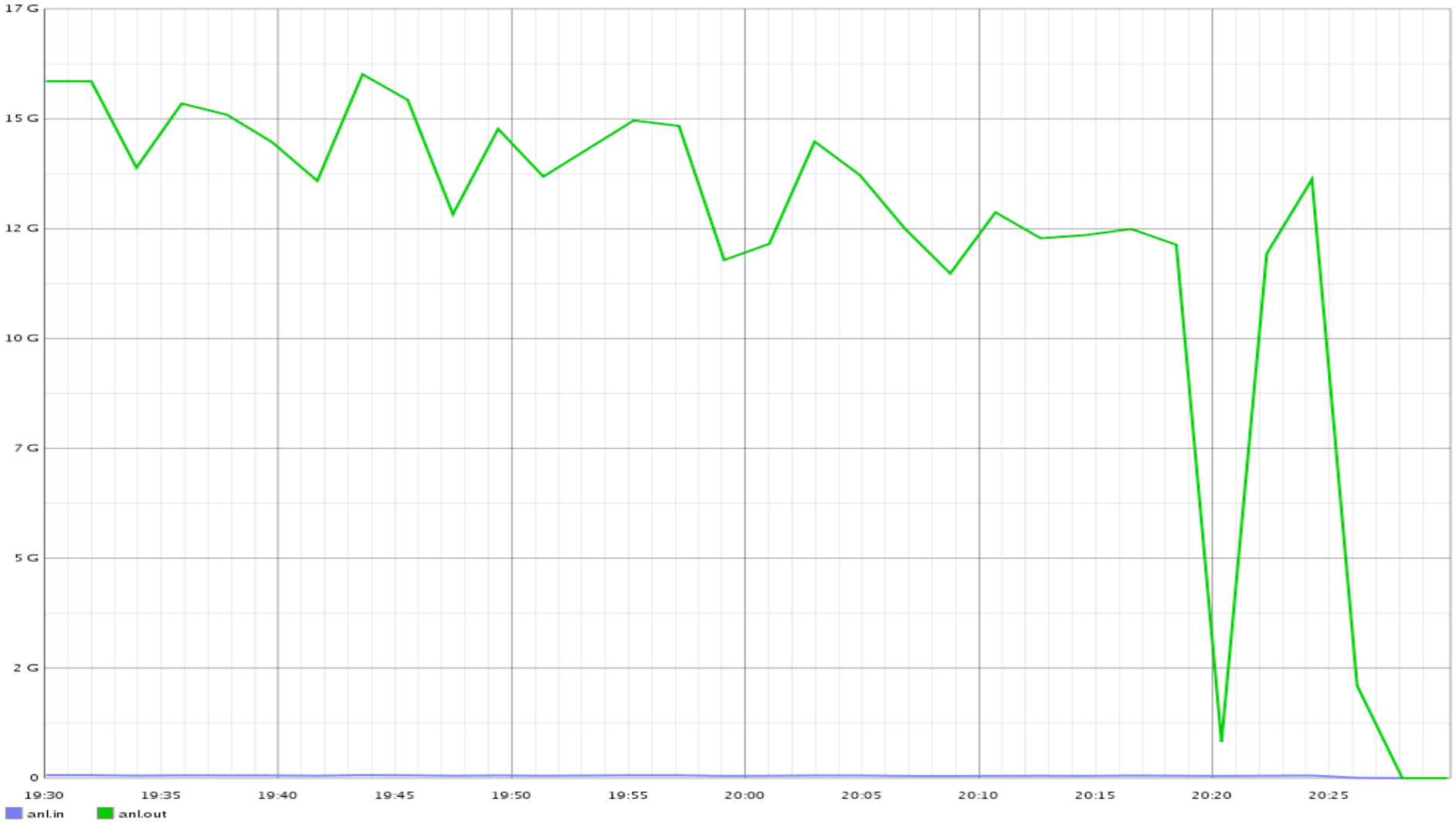
- Multiple nodes work together as a single logical GridFTP server
- Multiple nodes are used to transfer data into/ out of the cluster
  - Each node reads/writes only pieces they're responsible for
  - Head node coordinates transfers
- Multiple levels of parallelism
  - CPU, bus, NIC, disk etc.
  - Maximizes use of Gbit+ WANs



Striped Transfer
Fully utilizes bandwidth of Gb+ WAN using multiple nodes.

the globus alliance
www.globus.org

Performance

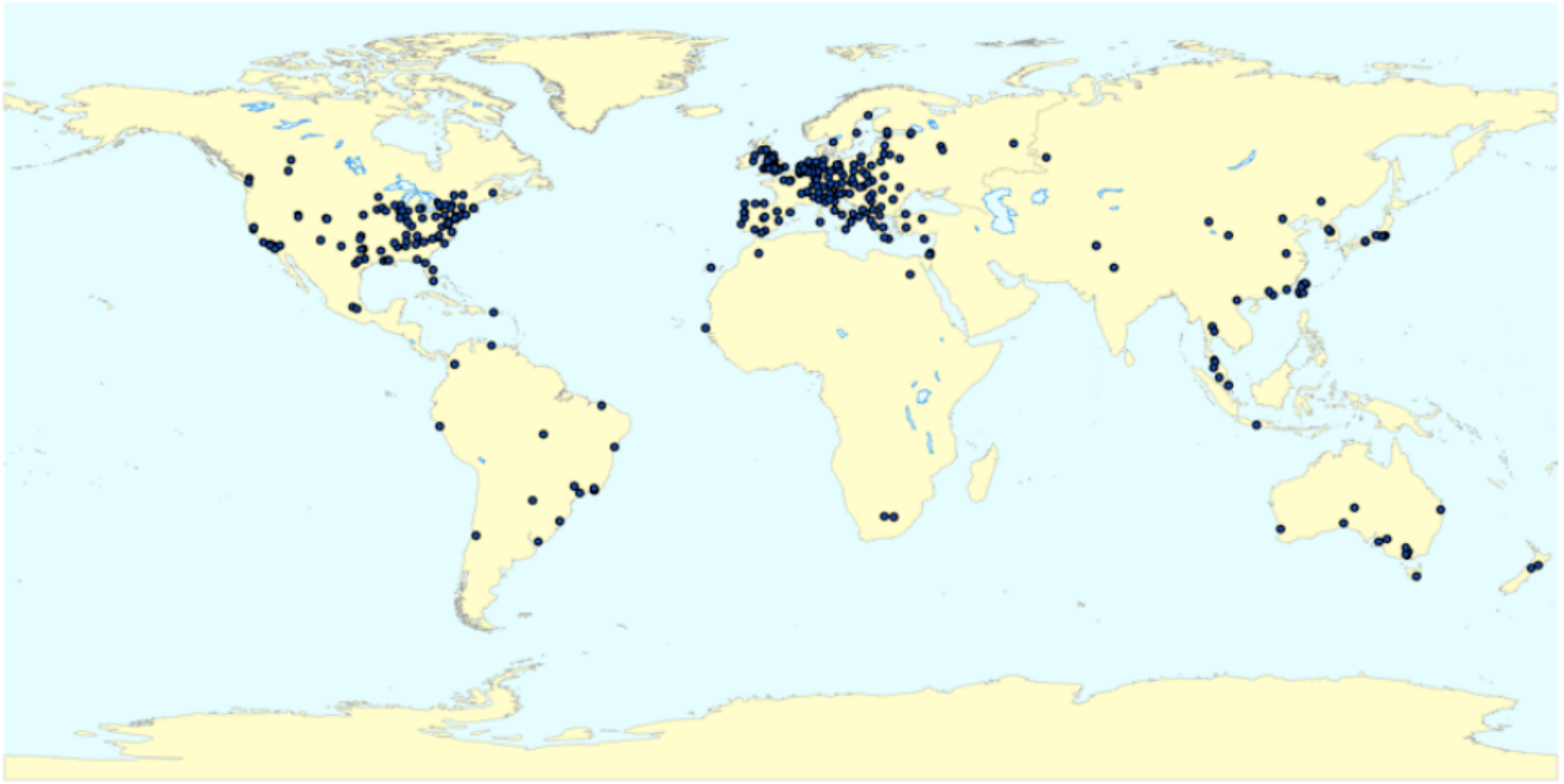


GridFTP in production

- Many Scientific communities rely on GridFTP
  - High Energy Physics - LHC computing Grid
  - Southern California Earthquake Center (SCEC), Earth Systems Grid (ESG), Relativistic Heavy Ion Collider (RHIC), European Space Agency, BBC use GridFTP for data movement
- GridFTP facilitates an average of more than 7 million data transfers every day

the globus alliance
www.globus.org

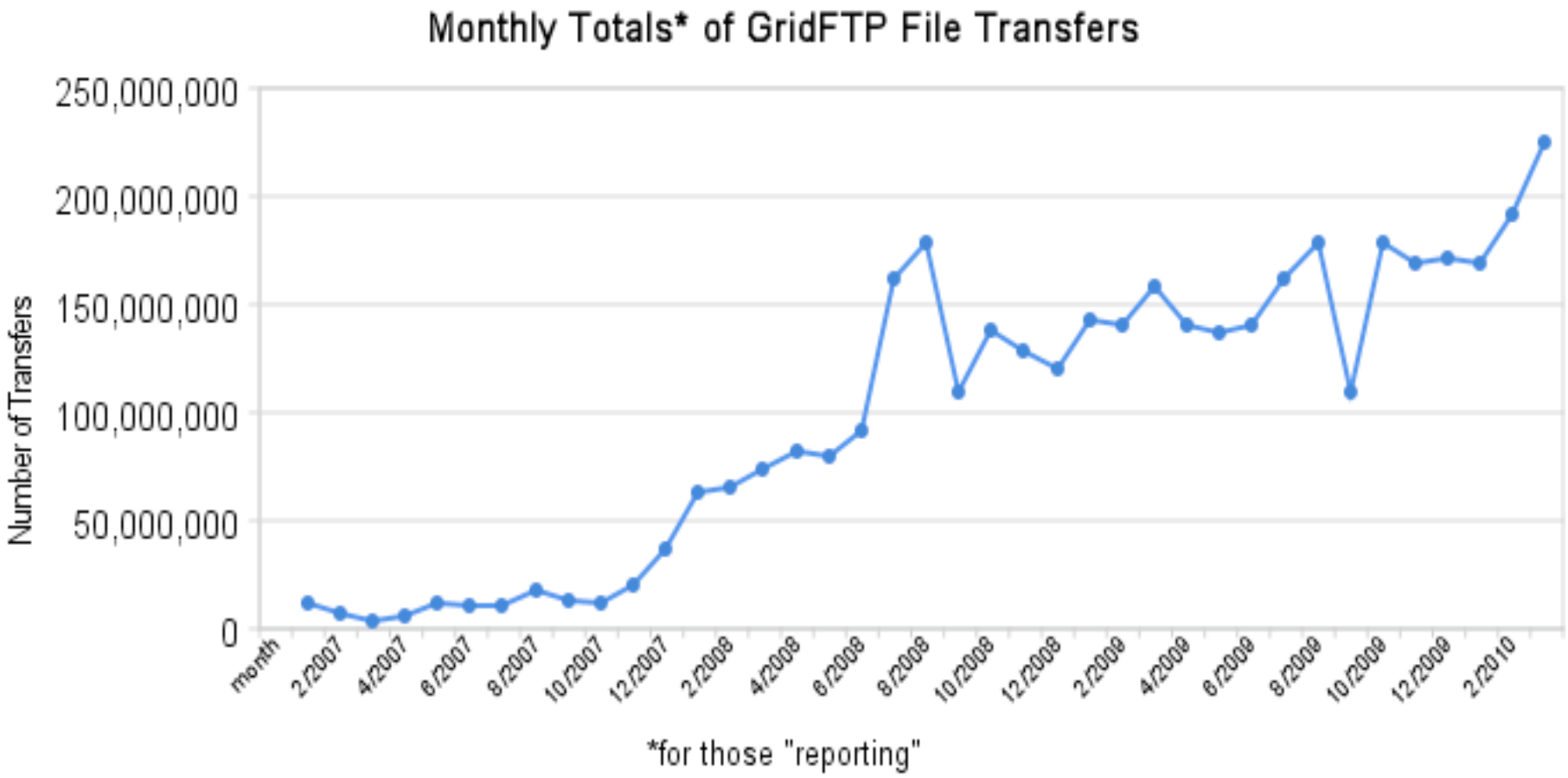
GridFTP Servers Around the World



Created by Tim Pinkawa (Northern Illinois University) using MaxMind's GeoIP technology (http://www.maxmind.com/app/ip-locate).

the globus alliance
www.globus.org

GridFTP Usage



Handling failures

- GridFTP server sends restart and performance markers periodically
  - Default every 5s - configurable
- Helpful if there is any failure
  - No need to transfer the entire file again
  - Use restart markers and transfer only the missing pieces
- GridFTP supports partial file transfers

Server failure

- Command-line client - globus-url-copy - support transfer retries
  - Use restart markers
- Recover from server and connection failures
- What if the client fails in the middle of a transfer?

the globus alliance
www.globus.org

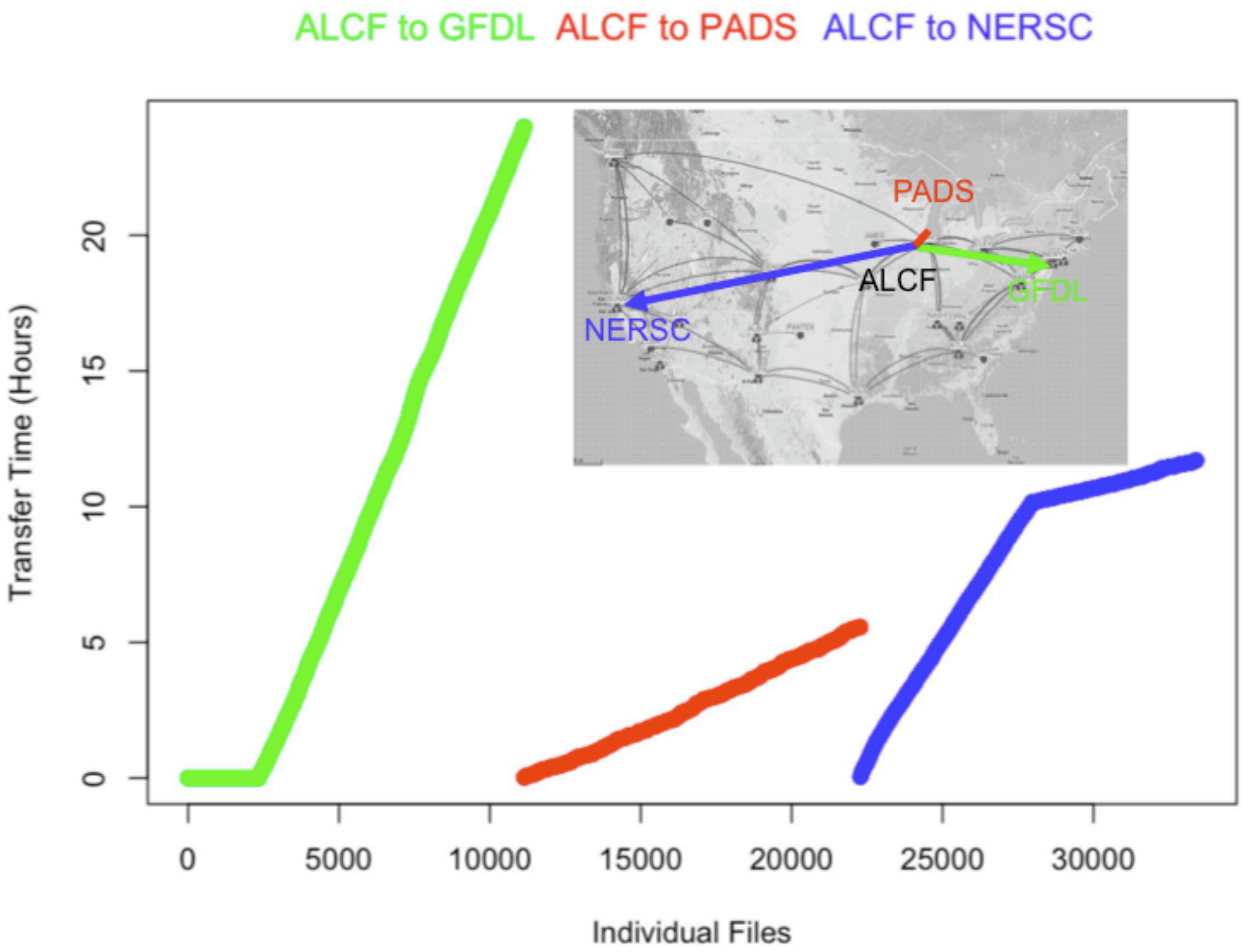
Globus.org
Value Additions for GridFTP

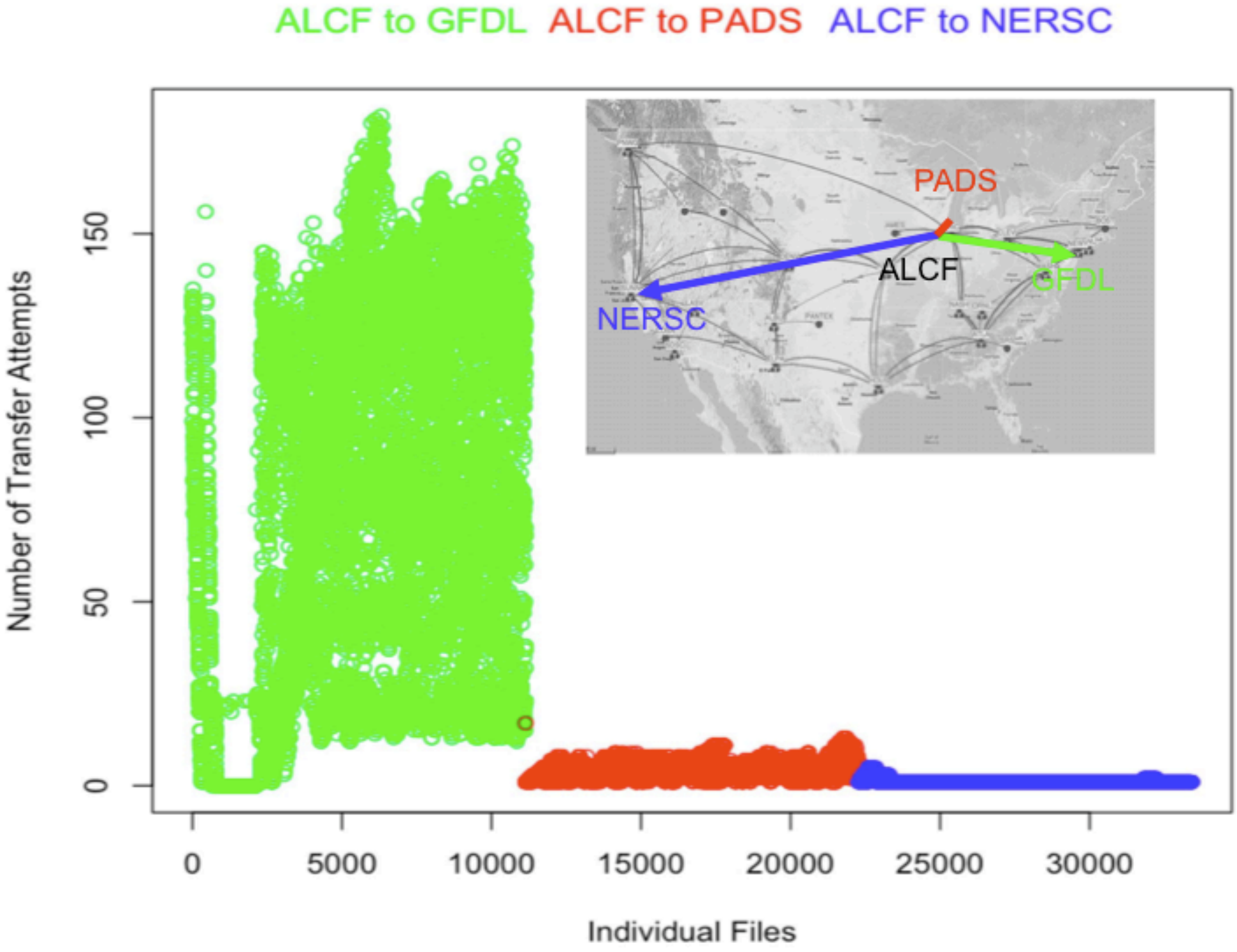
- Fire and forget
  - Less user interaction
  - Email notifications
  - No need to babysit transfers
- Failure handling
  - Automatic retries
- Familiar user interfaces
- Technology interactions requiring no special expertise
- No client software to install

the globus alliance
www.globus.org

Globus.org

- Enable users to focus on domain-specific work
  - Manage technology failures
  - Notifications of interesting events
  - Provide users with enough information to resolve problems
- Ease the infrastructure providers' support burden
  - Hosted and supported by Globus team


ALCF to GFDL
ALCF to PADS
ALCF to NERSC
Transfer Time (Hours)
0
5
10
15
20
0
5000
10000
15000
20000
25000
30000
Individual Files
PADS
ALCF
GFDL
NERSC


ALCF to GFDL
ALCF to PADS
ALCF to NERSC
Number of Transfer Attempts
0
50
100
150
0
5000
10000
15000
20000
25000
30000
Individual Files
PADS
ALCF
GFDL
NERSC

More Information at http://www.globus.org/service/ http://www.gridftp.org

Questions