globus
online

Reliable File Transfer.
No IT Required.

June 27, 2011

globus
online

Work by:

Bryce Allen, John Bresnahan, Lisa Childers, Ian Foster, Gopi Kandaswamy, Raj Kettimuthu, Jack Kordas, Mike Link, Stuart Martin, Karl Pickett, Steven Tuecke, and many others.

The Challenge: Moving Big Data

- **What should be trivial…**

- **… can be painfully tedious and time-consuming**

Scientists aren't IT Experts!

- **Nor should they be!**
  - It can take days or weeks to move data using traditional methods
  - It can require numerous tweaks and retrys
  - What's lost:
    - Hours of more productive work
    - Projects that go undone
    - Potentially expanded scope gone when data movement barriers are too great
- **Your time gained = New innovations or discoveries**

L. Childers, L. Liming, and I. Foster, "Perspectives on Distributed Computing: Thirty People, Four User Types, and the Distributed Computing User Experience," Argonne National Laboratory Technical Report ANL/MCS/CI-31, September 2008.

An Example: Genomics

Will Computers Crash Genomics?

Pennisi, E., Science 2011 **331**:6018 pp. 666

CREDIT: ALVARO ARTEAGA/ALVAREJO.COM

"The various members of the genome informatics ecosystem are now facing a potential tsunami of genome data that will swamp our storage systems and crush our compute clusters."

Stein, L.D., *Genome Biology* 2010 **11**:5 pp. 207

Globus Products

Globus Toolkit

Build the Grid

Components for building custom grid solutions

globustoolkit.org

Globus Online

Use the Grid

Reliable file transfer
Software-as-a-Service

globusonline.org

What is Globus Online?

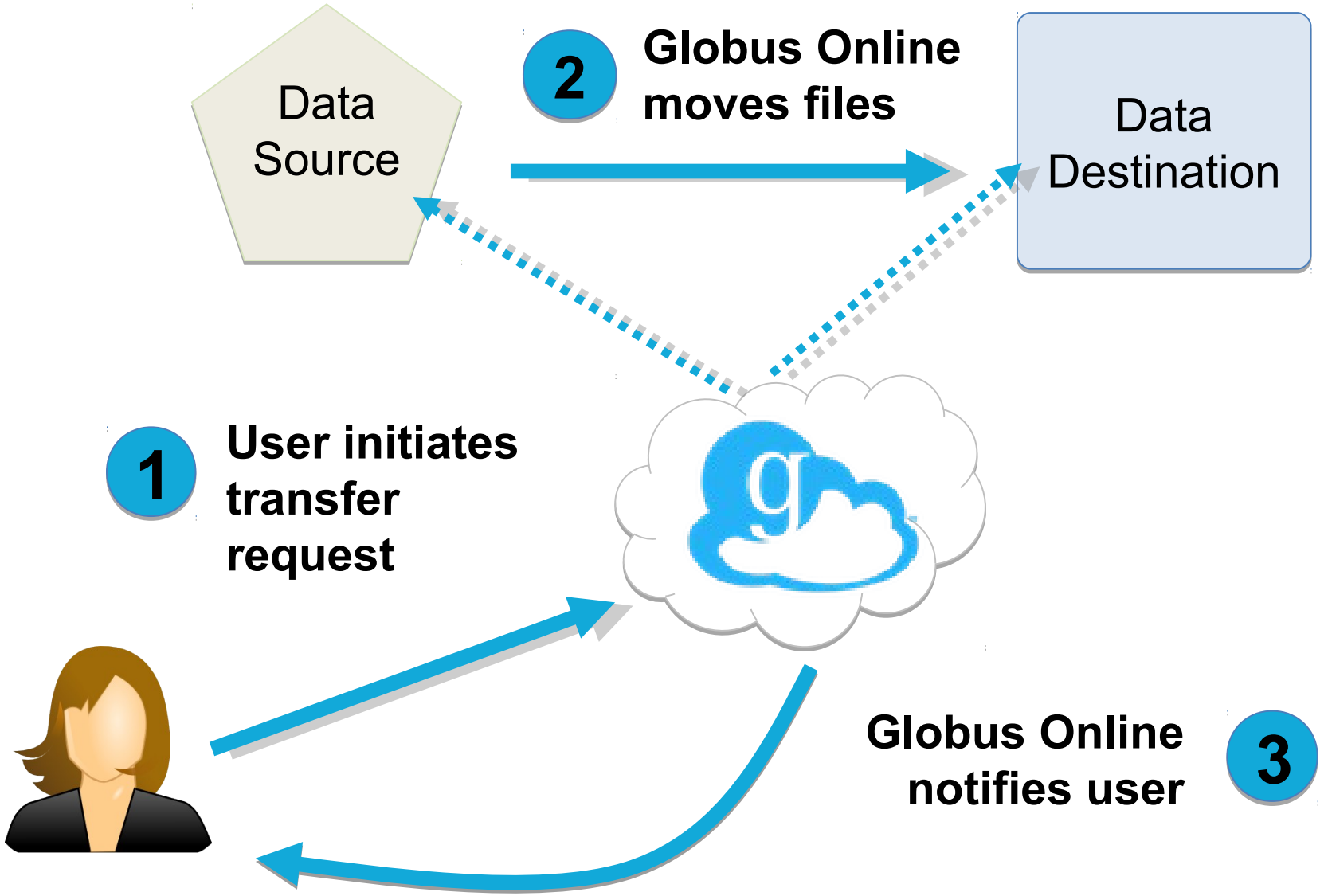
- **Reliable file transfer.**
  - Easy “fire and forget” file transfers
  - Automatic fault recovery
  - High performance
  - Across multiple security domains
- **No IT required.**
  - No client software installation
  - New features automatically available
  - Consolidated support and troubleshooting
  - Works with existing GridFTP servers
  - Globus Connect solves “last mile problem”

“I moved 400 GB of files and didn’t even have to think about it”
—Lawrence Berkeley National Lab

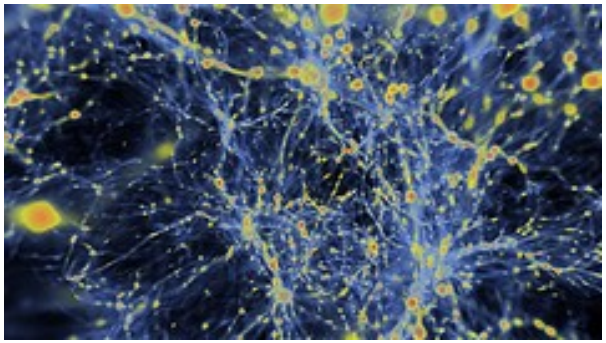
“It’s just not a big deal to move big data anymore.”
—Initiative for Biomedical Informatics

“Fantastic! I have started using Globus Connect to transfer data, and it only took me 5 minutes to set up. Thank you!”
—NERSC user

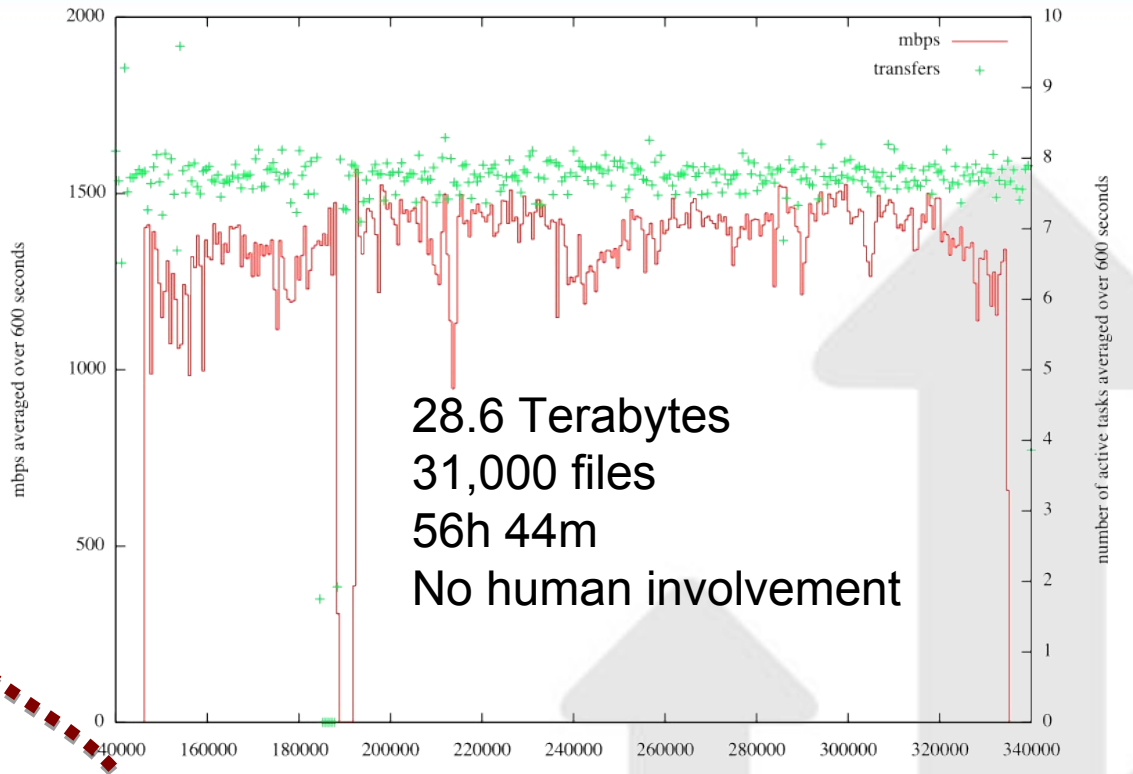
How It Works


Data Source
2
Globus Online moves files
Data Destination
1
User initiates transfer request
Globus Online notifies user
3

Case Study: Astrophysics

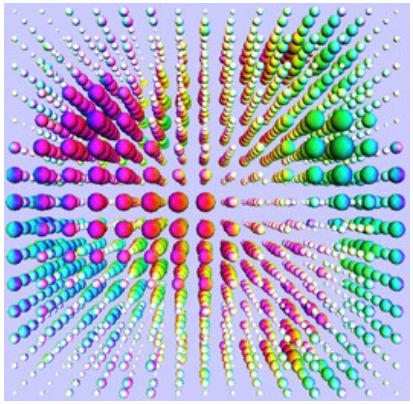


Astrophysics simulation data
generated in Tennessee,
moved to Illinois for visualization
(Enzo, UCSD; Futures Lab, Argonne)



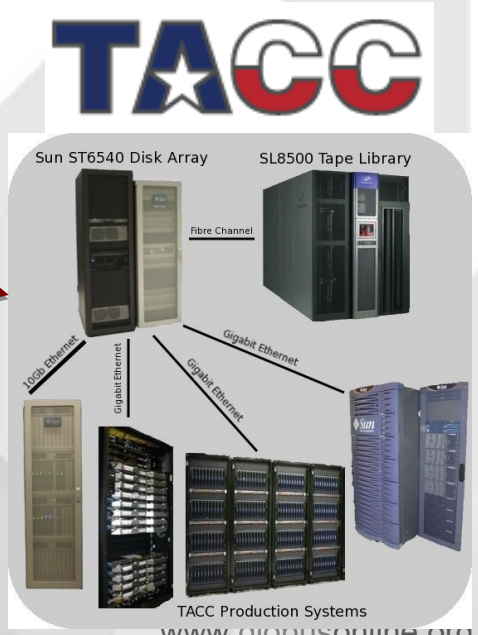
28.6 Terabytes
31,000 files
56h 44m
No human involvement

Case Study: Lattice QCD



- **Fast**: Reduced transfer times
- **Easy**: Fire-and-forget transfers
  - Automated retry
  - No file pre-staging
  - No complex infrastructure
  - Convenient CLI or GUI interfaces

Indiana University researcher moved ~6 TB from Oak Ridge to TACC in 2 days



"Globus Online frees up my time to do more creative work than typing scp commands or devising scripts to initiate and monitor progress to move many files."

"I moved 100 7.3 GB files tonight in about 1.5 hours. I am very impressed. I also like the new commands and help system."

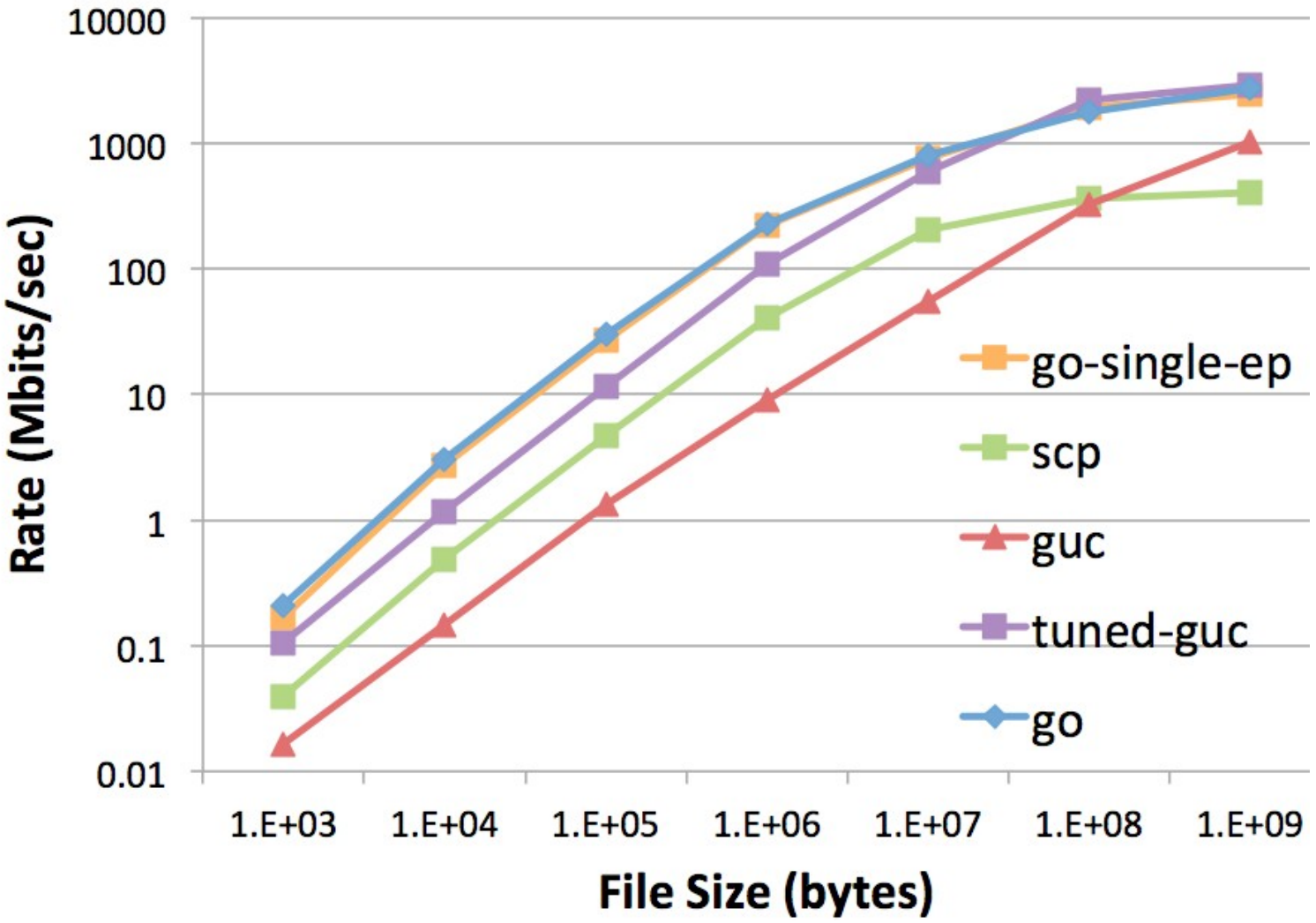
Case Studies

- **MIMD Lattice Computation (MILC) Research Collaboration**
  https://www.globusonline.org/casestudies/milc/
  - Moved 100 7.3-GB files in 88 minutes
  - Improved transfer rate from 20 Mb/sec to 1100 Mb/sec
  - Same transfer would have taken over 3 days with scp
- **University of Colorado**
  https://www.globusonline.org/casestudies/ucolorado/
  - *"I'm averaging 40 Mb/s and can even reach 400 Mb/s – that's insanely fast."*
- **National Energy Research Scientific Computing Center (NERSC)**
  https://www.globusonline.org/casestudies/nersc/
  - Globus Online is the preferred method of data movement and archiving for thousands of researchers who use NERSC resources on a daily basis.
  - *"Globus Online actually makes web-based data syncing an easy, nearly trivial process, so you don't have to be an IT or middleware expert to move your files."*

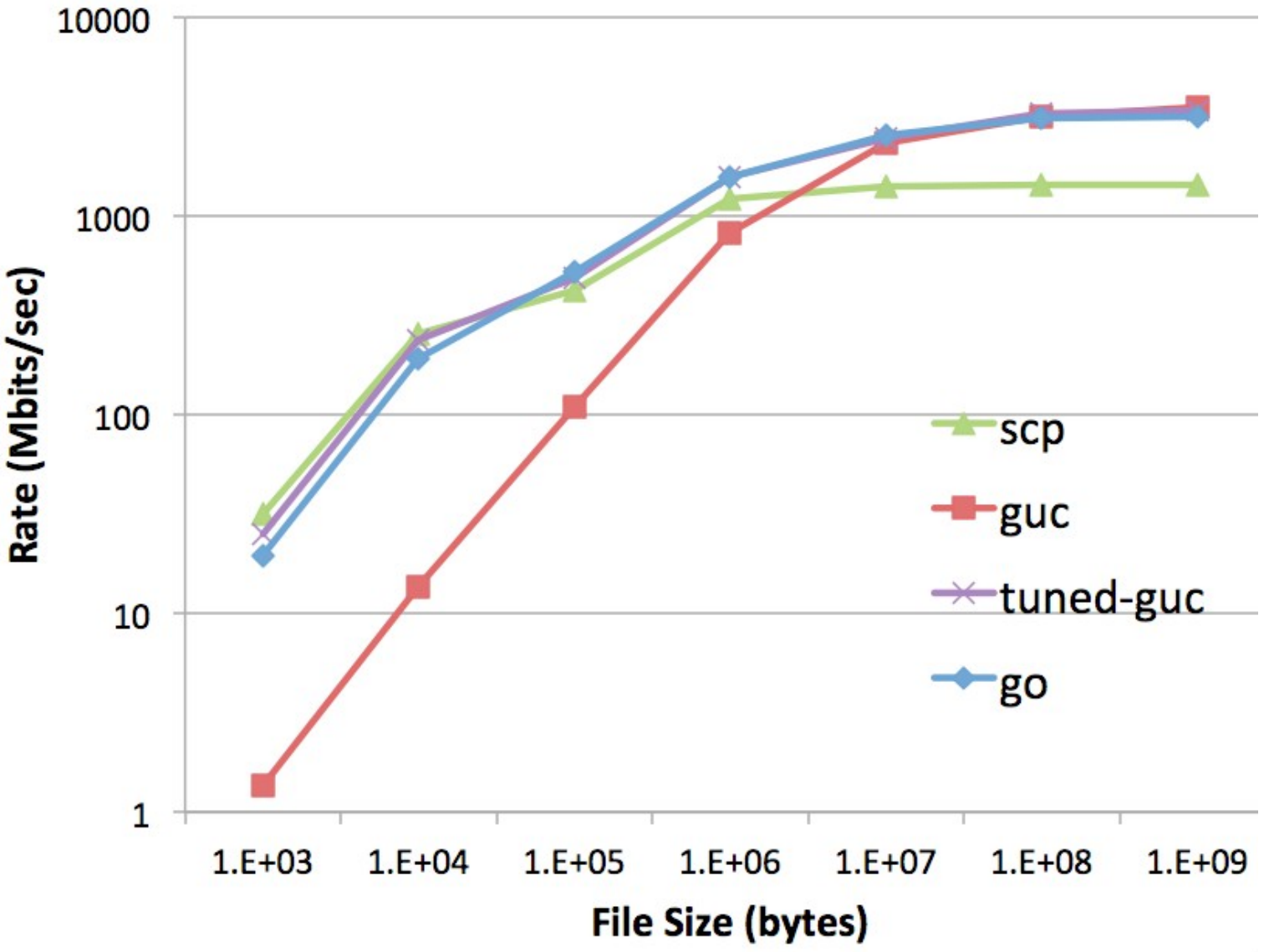
Who can benefit from Globus Online?

- **Scientists** (with no IT background) who need to move many files can use Web GUI
- **Programmers** who want to create automated workflows can use Command Line Interface (CLI)
- **System builders** who don't want to re-engineer file transfer solutions can use REST API

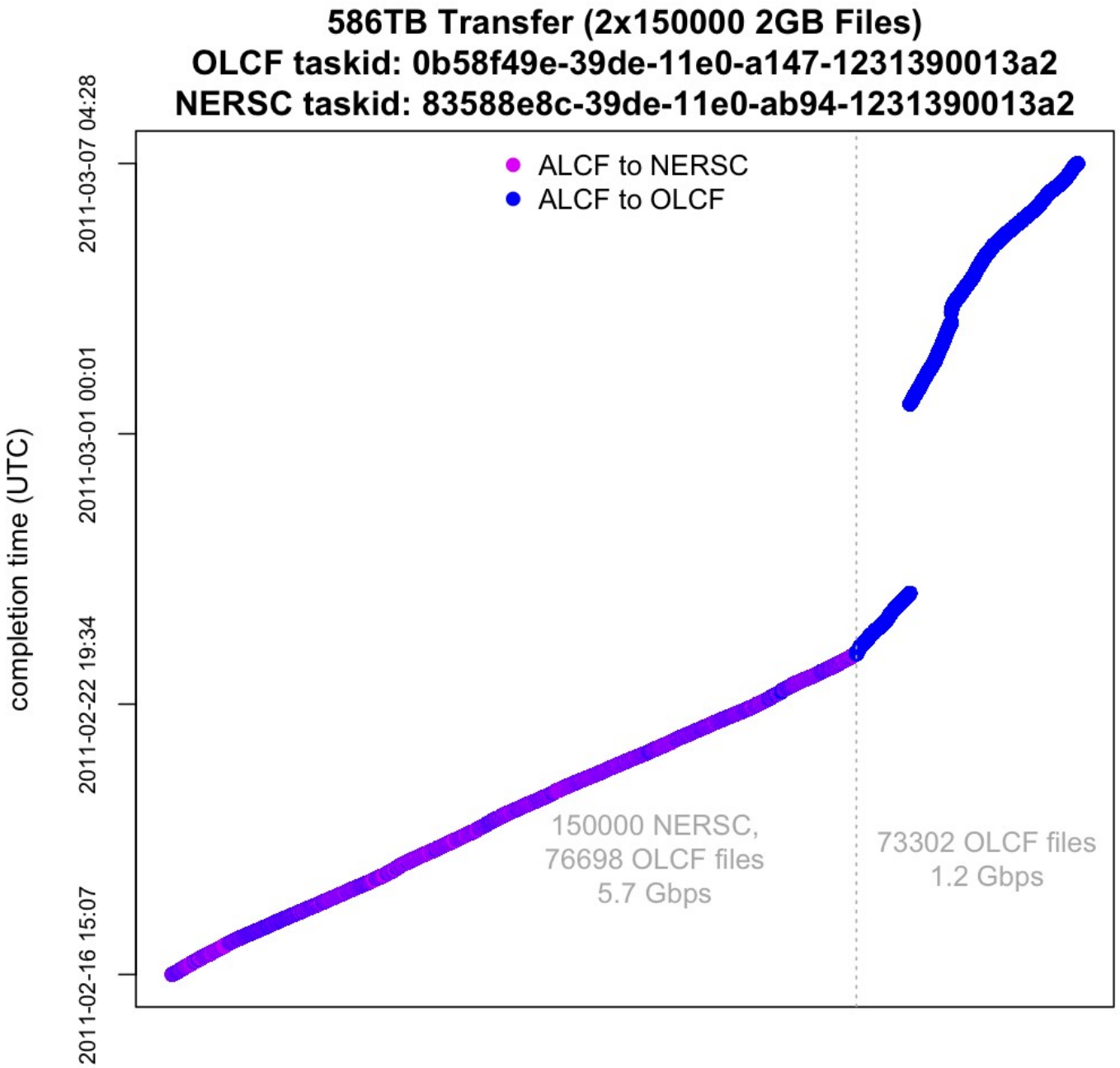
Performance: Transfer between sites



Performance: Transfer inside a site

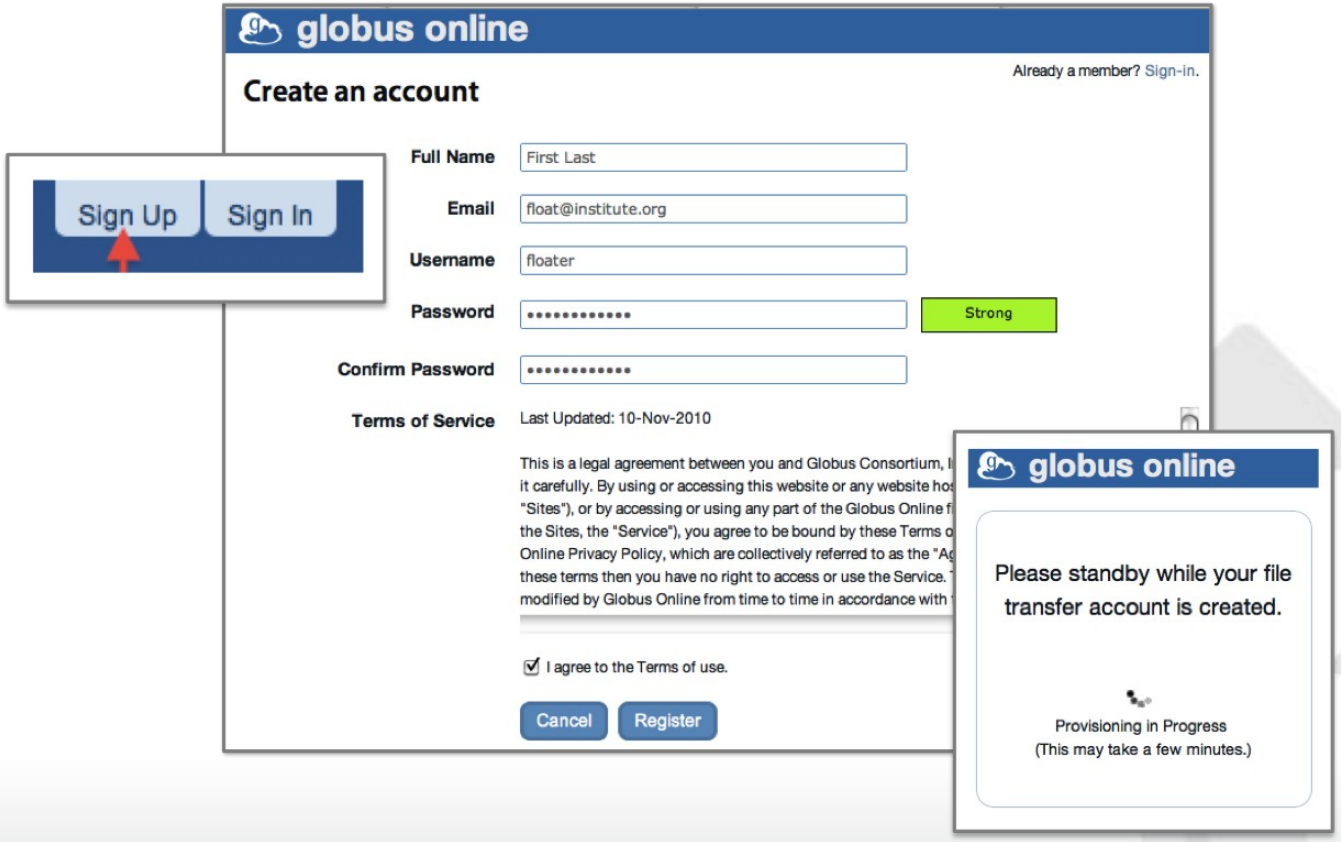


Experiences: Transferring 586TB



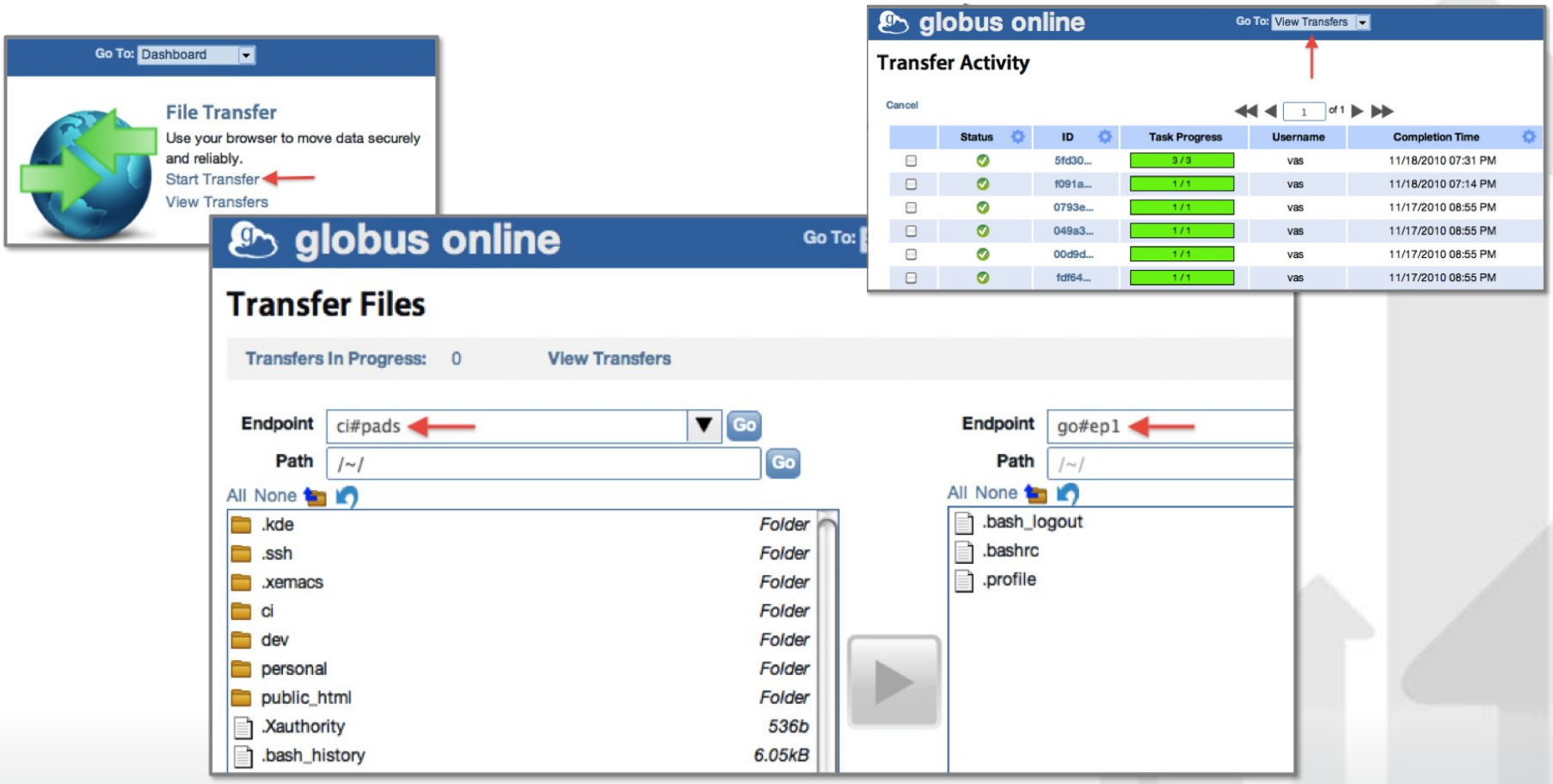
Getting Started (2 easy steps)

1. **Sign up:** Visit www.globusonline.org to create an account



Getting Started (2 easy steps)

1. **Start moving files:** Pick your data and where you want to move it, then click to transfer

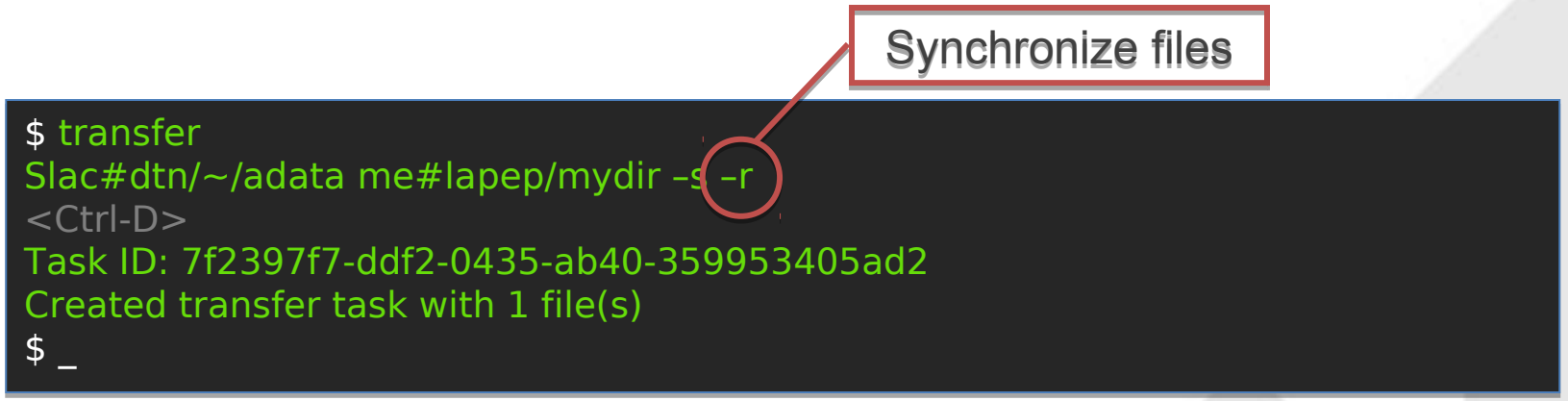


File Movement Options

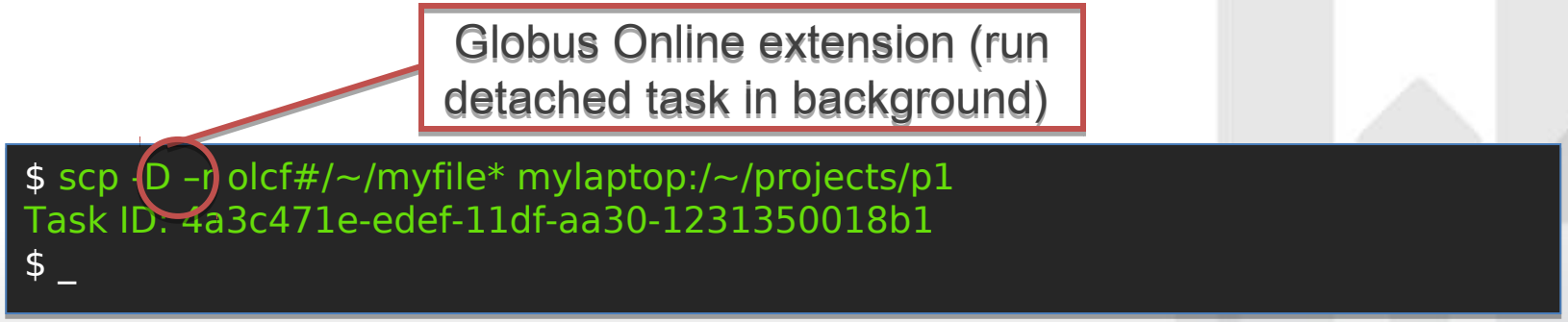
- **Web UI**: Using a web browser, files or entire directories from one endpoint can be selected and transferred to another endpoint in just a few clicks
- **CLI**: Provide extensive control of all aspects of the Globus Online service, with greater flexibility for creating and controlling file transfer requests.
- **Transfer REST API**: The Transfer REST API provides a "RESTful" interface to Globus Online for developers who need to deliver solutions to a community of users.

File Transfer with CLI (2 options)

- **transfer**: Offers increased flexibility in specifying a list of individual files or directories for transfer between endpoints



- **scp**: A subset of transfer's functionality packaged as the familiar scp, but with extensions for reliability and security



For More Information

- **Visit https://www.globusonline.org/signup to:**
  - Get a free account and start moving files
- **Visit www.globusonline.org for:**
  - Tutorials
  - FAQs
  - Pro Tips
  - Troubleshooting
- **Contact support@globusonline.org for:**
  - Help getting started
  - Help using the service