



**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...

*"I need my data **over there** – at my \_\_\_\_\_" ( supercomputing center, campus server, etc.)*



Data Source



Data Destination

- ... can be painfully tedious and time-consuming

*"GAAAH!  
%&@#&"*



Data Source



Data Destination



# 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 retries
  - 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.”*

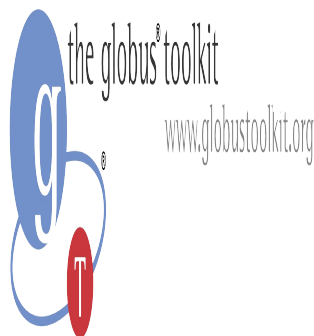
Stein, L.D., *Genome Biology* 2010 **11**:5 pp. 207



# Globus Products

## Globus Toolkit

Build the Grid



Components for building  
custom grid solutions

[globustoolkit.org](http://globustoolkit.org)

## Globus Online

Use the Grid



Reliable file transfer  
Software-as-a-Service

[globusonline.org](http://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



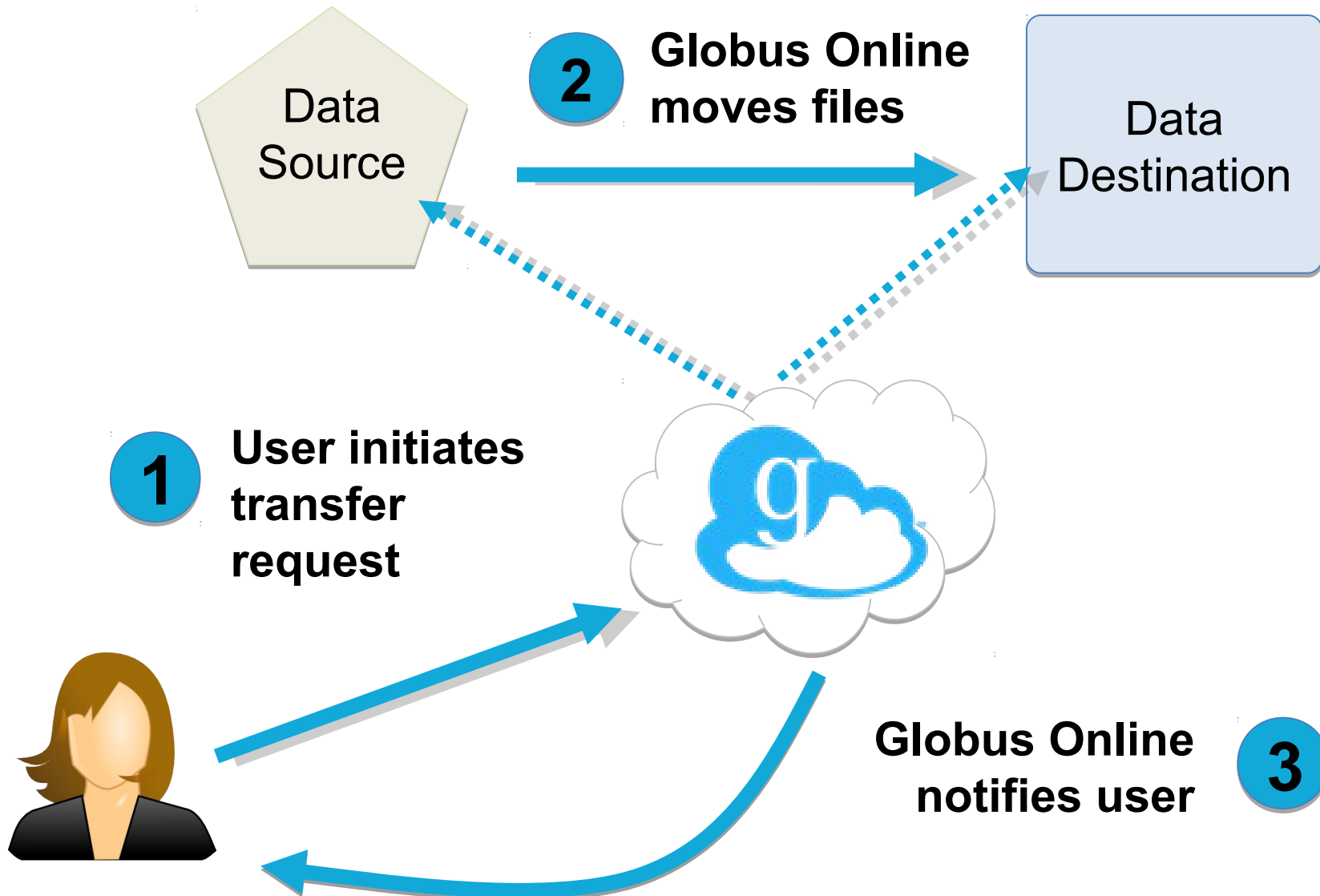
*“It’s just not a big deal to move big data anymore.”*  
– Initiative for Biomedical Informatics



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

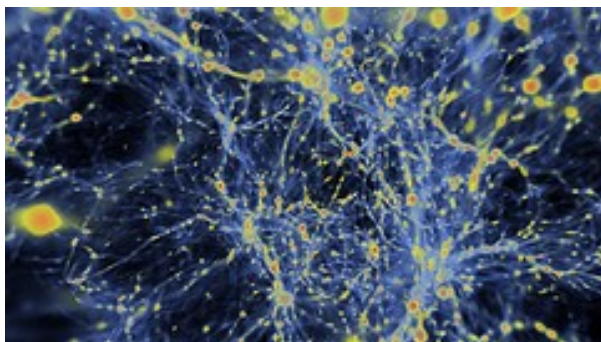


# How It Works

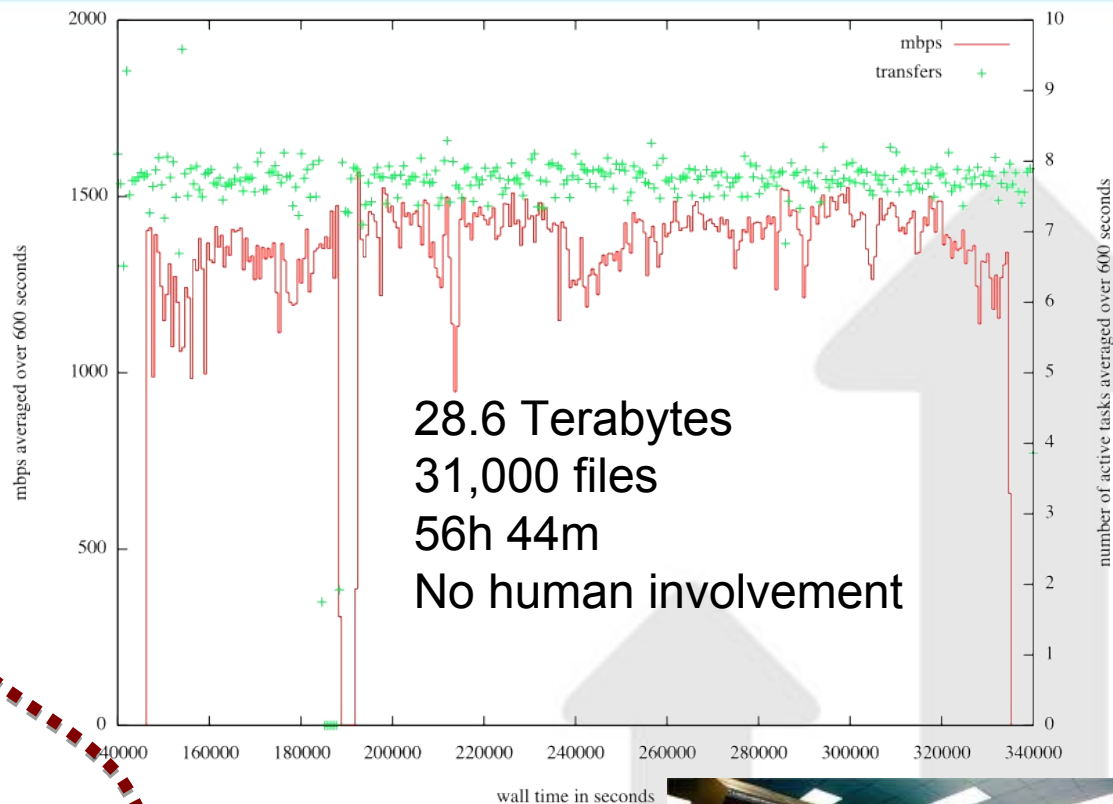




# Case Study: Astrophysics

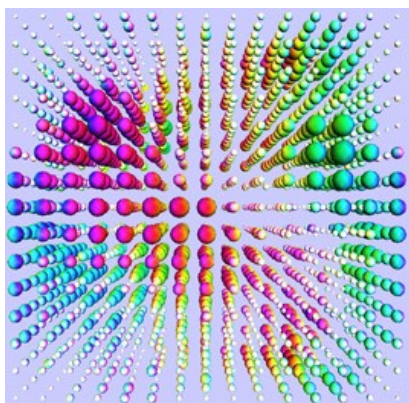


Astrophysics simulation data  
generated in Tennessee,  
moved to Illinois for visualization  
(Enzo, UCSD; Futures Lab, Argonne)





# 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.”*

## TACC





# 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.”*

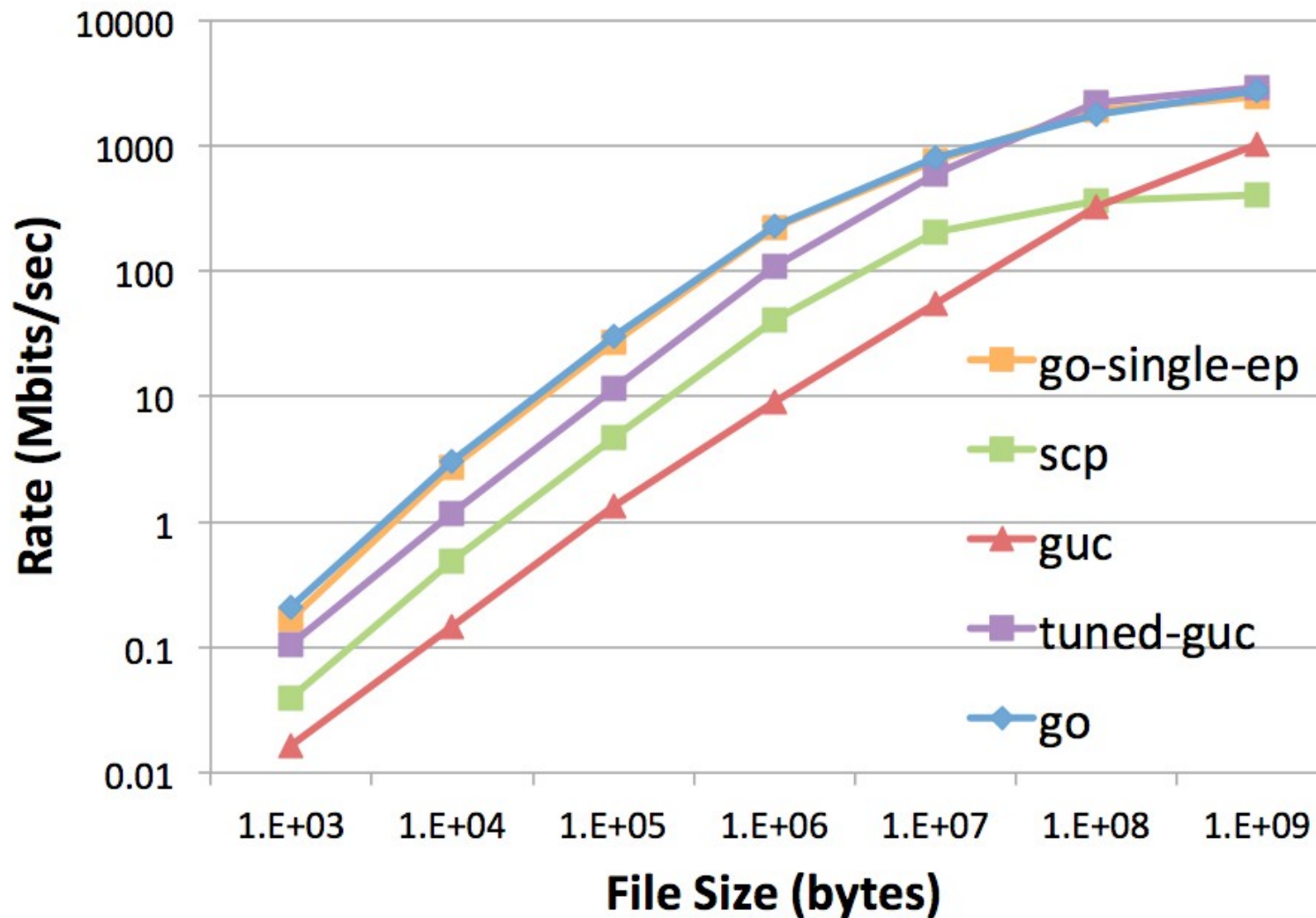


# Who can benefit from Globus Online?

- **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**

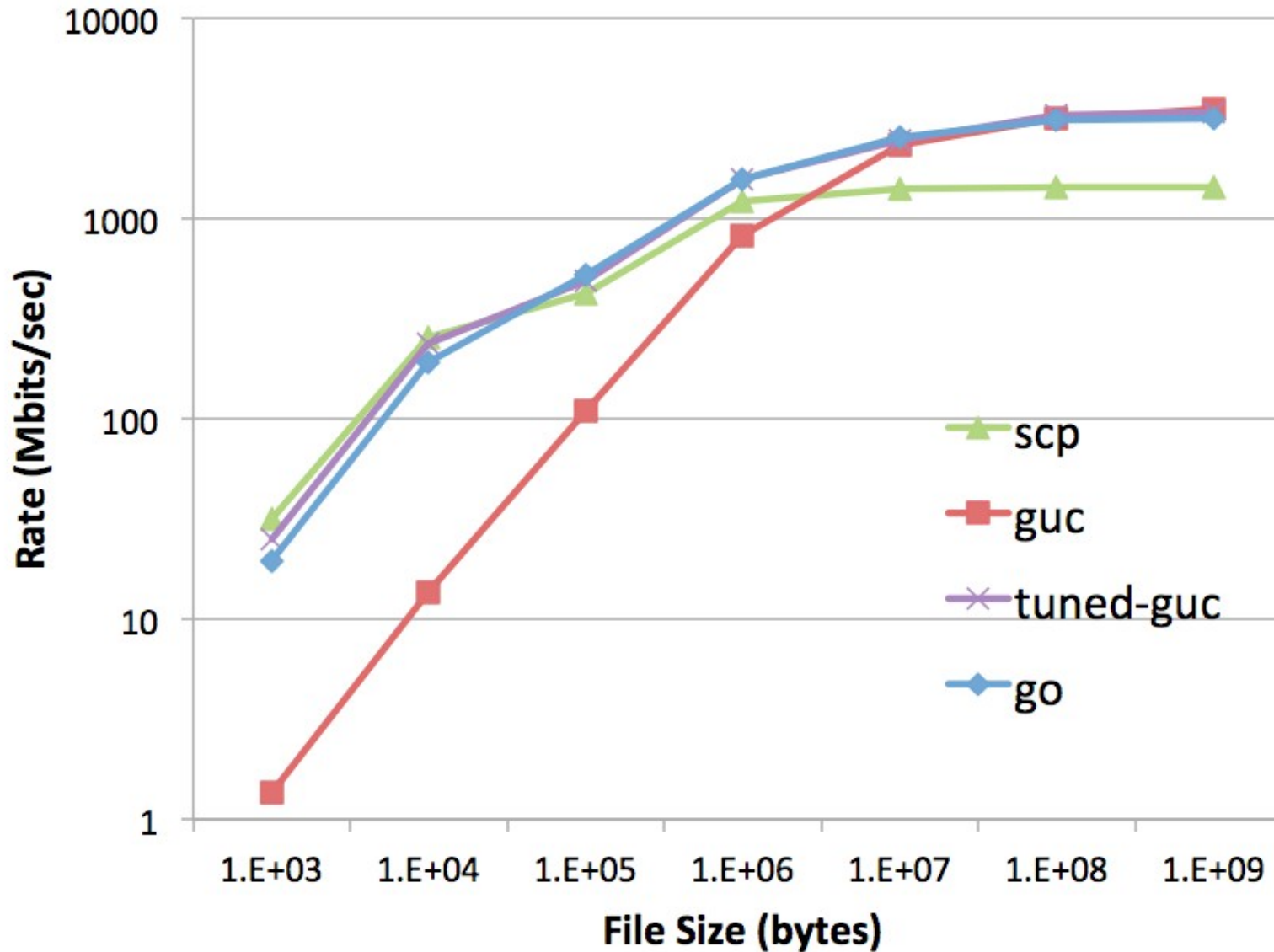


# Performance: Transfer between sites



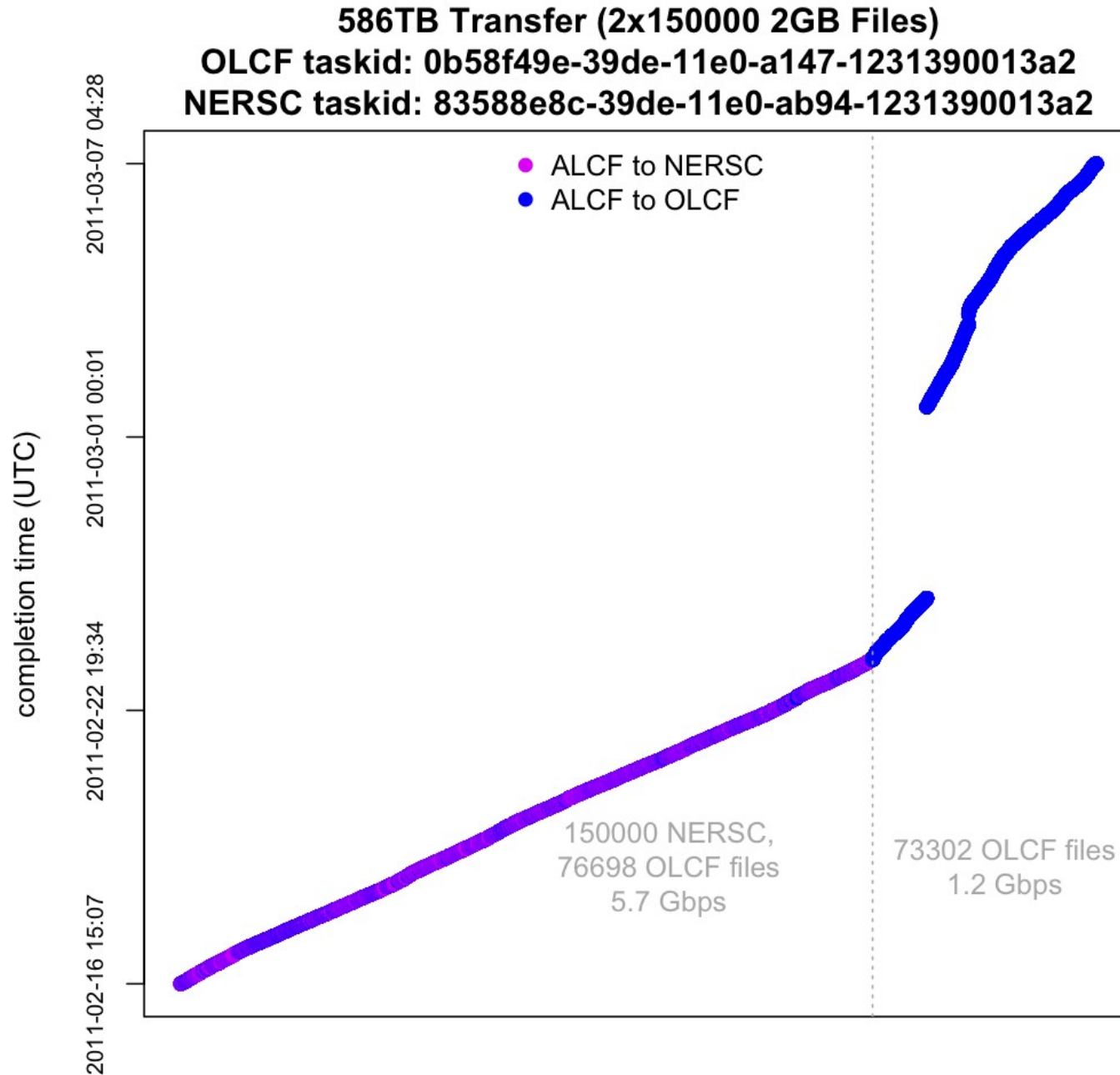


# Performance: Transfer inside a site





# Experiences: Transferring 586TB





# Getting Started (2 easy steps)

1. **Sign up:** Visit [www.globusonline.org](http://www.globusonline.org) to create an account

The screenshot shows the 'Create an account' page on the Globus Online website. The page has a blue header with the Globus Online logo and the text 'Create an account'. A link 'Already a member? Sign-in.' is in the top right. The form includes fields for 'Full Name' (with a hint 'First Last'), 'Email' (containing 'float@institute.org'), 'Username' (containing 'floater'), 'Password' (with a strength indicator 'Strong'), and 'Confirm Password'. Below these is the 'Terms of Service' section, which includes a checkbox for 'I agree to the Terms of use.' and 'Cancel' and 'Register' buttons. A red arrow points to the 'Sign Up' button in a blue box. A second blue box highlights the 'Sign Up' and 'Sign In' buttons. A third box shows a message: 'Please standby while your file transfer account is created.' with a progress indicator and the text 'Provisioning in Progress (This may take a few minutes.)'.

**globus online**

Already a member? [Sign-in.](#)

**Create an account**

Full Name

Email

Username

Password  **Strong**

Confirm Password

**Terms of Service** Last Updated: 10-Nov-2010

This is a legal agreement between you and Globus Consortium, Inc. By using or accessing this website or any website hosted by Globus Consortium, Inc. ("Sites"), or by accessing or using any part of the Globus Online file transfer service ("Service"), you agree to be bound by these Terms of Service and the Globus Online Privacy Policy, which are collectively referred to as the "Agreement". If you do not agree to these terms then you have no right to access or use the Service. The Service is provided by Globus Online from time to time in accordance with the Agreement.

☒ I agree to the Terms of use.

[Cancel](#) [Register](#)

**globus online**

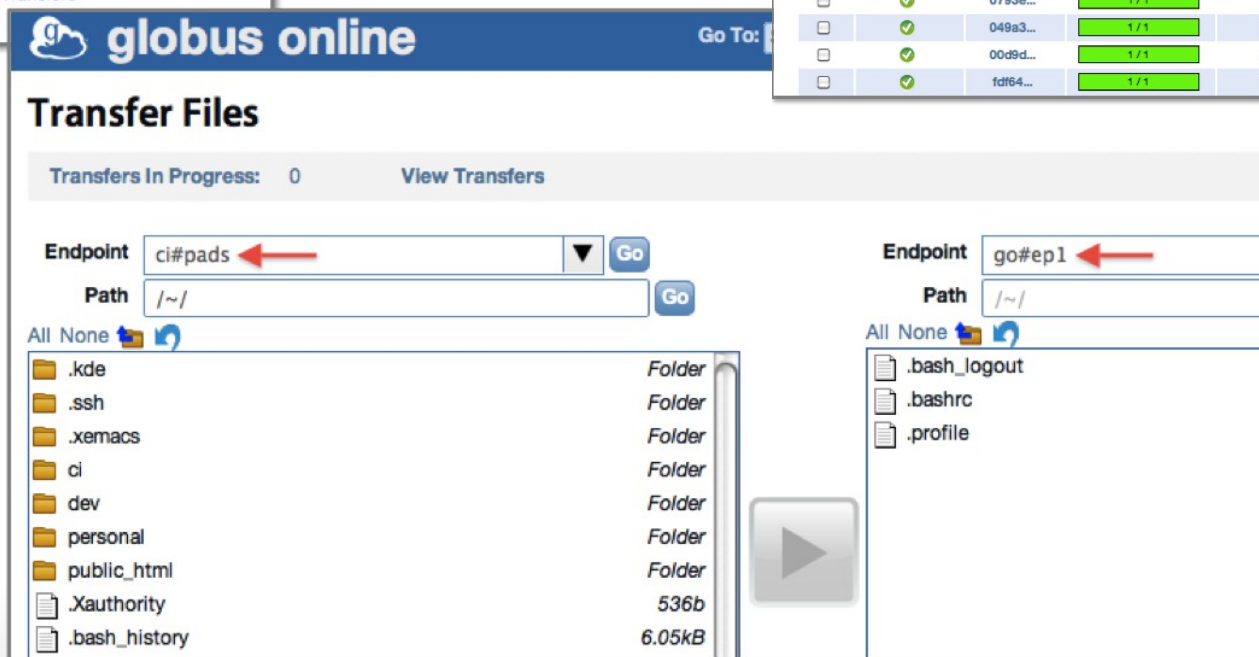
Please standby while your file transfer account is created.

Provisioning in Progress  
(This may take a few minutes.)



# Getting Started (2 easy steps)

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



	Status	ID	Task Progress	Username	Completion Time
<input type="checkbox"/>	✓	5fd30...	3 / 3	vas	11/18/2010 07:31 PM
<input type="checkbox"/>	✓	f091a...	1 / 1	vas	11/18/2010 07:14 PM
<input type="checkbox"/>	✓	0793e...	1 / 1	vas	11/17/2010 08:55 PM
<input type="checkbox"/>	✓	049a3...	1 / 1	vas	11/17/2010 08:55 PM
<input type="checkbox"/>	✓	00d9d...	1 / 1	vas	11/17/2010 08:55 PM
<input type="checkbox"/>	✓	fd64...	1 / 1	vas	11/17/2010 08:55 PM



# 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

Synchronize files

```
$ transfer  
Slac#dtn/~ /adata me#lapep/mydir -s -r  
<Ctrl-D>  
Task ID: 7f2397f7-ddf2-0435-ab40-359953405ad2  
Created transfer task with 1 file(s)  
$ _
```

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

Globus Online extension (run detached task in background)

```
$ scp -D -r olcf#/~ /myfile* mylaptop:~/projects/p1  
Task ID: 4a3c471e-edef-11df-aa30-1231350018b1  
$ _
```



# For More Information

- Visit <https://www.globusonline.org/signup> to:
  - Get a free account and start moving files
- Visit [www.globusonline.org](http://www.globusonline.org) for:
  - Tutorials
  - FAQs
  - Pro Tips
  - Troubleshooting
- Contact [support@globusonline.org](mailto:support@globusonline.org) for:
  - Help getting started
  - Help using the service