



# Globus Online: File Transfer Made Easy!

**Matteo Lanati**  
[matteo.lanati@lrz.de](mailto:matteo.lanati@lrz.de)

**Initiative for Globus in Europe**  
Leibniz Supercomputing Centre



- Introduction and acknowledgments
- Motivation
- Demo session
- Future developments
- Wrap up



## TOP500 List - June 2012 (1-100)

$R_{\max}$  and  $R_{\text{peak}}$  values are in TFlops. For more details about other fields, check the [TOP500 description](#).

Power data in KW for entire system

[next](#)

Rank	Site	Computer/Year Vendor	Cores	$R_{\max}$	$R_{\text{peak}}$	Power
1	DOE/NNSA/LLNL United States	<b>Sequoia</b> - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom / 2011 IBM	1572864	16324.75	20132.66	7890.0
2	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect / 2011 Fujitsu	705024	10510.00	11280.38	12659.9
3	DOE/SC/Argonne National Laboratory United States	<b>Mira</b> - BlueGene/Q, Power BQC 16C 1.60GHz, Custom / 2012 IBM	786432	8162.38	10066.33	3945.0
4	Leibniz Rechenzentrum Germany	<b>SuperMUC</b> - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR / 2012 IBM	147456	2897.00	3185.05	3422.7



- Initiative for Globus in Europe
- Achievements:
  - GT packages for Linux (DEB, RPM)
  - Integration of Globus related software for the European Globus community: OGSA-DAI, GridSAM, GridWay, GSI-SSHTerm, ...
  - European Globus Community Forum:
    - Shape a community of European Globus users
    - Leverage future Globus developments

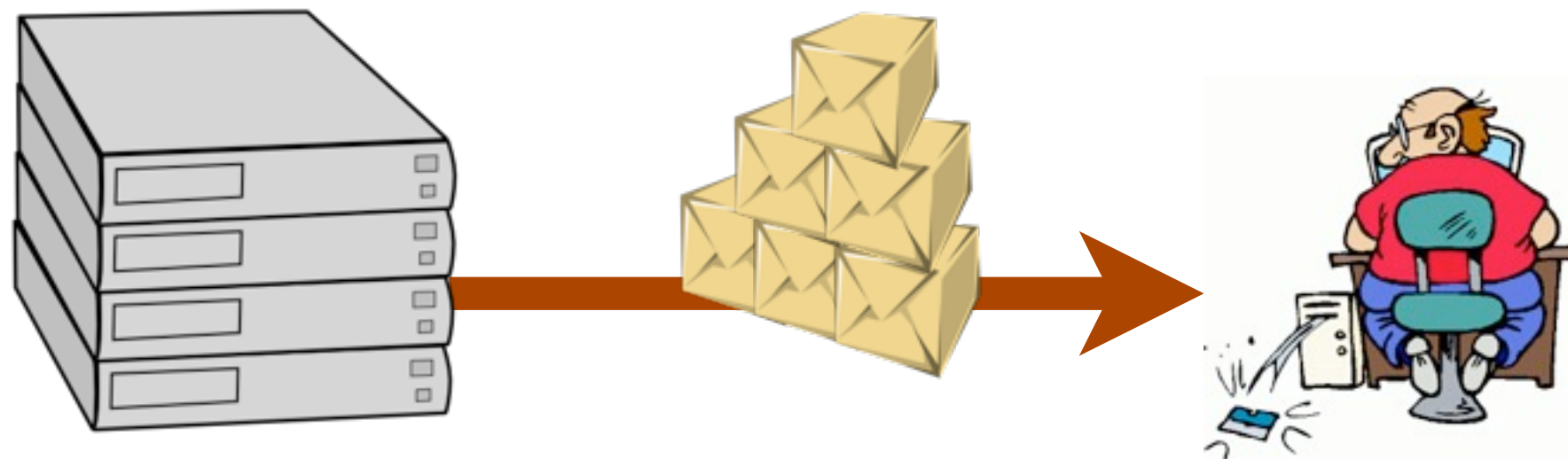


- Achievements (cont.):
  - Porting of the LoadLeveler adaptor from GT4 to GT5
  - Training events: GridKa Summer School 2010/2011, PRACE, Joint European DCI Summer School
  - Cooperation with other projects



- What should be trivial
  - Move data
- What can go wrong
  - Configuration
  - Firewall/NAT issues
  - Transfer failures
  - Cover the “last mile”

- Easy
- Resilient
- Error proof
- Comprehensive
- Covering the “last mile”





## About GridFTP and Globus Online

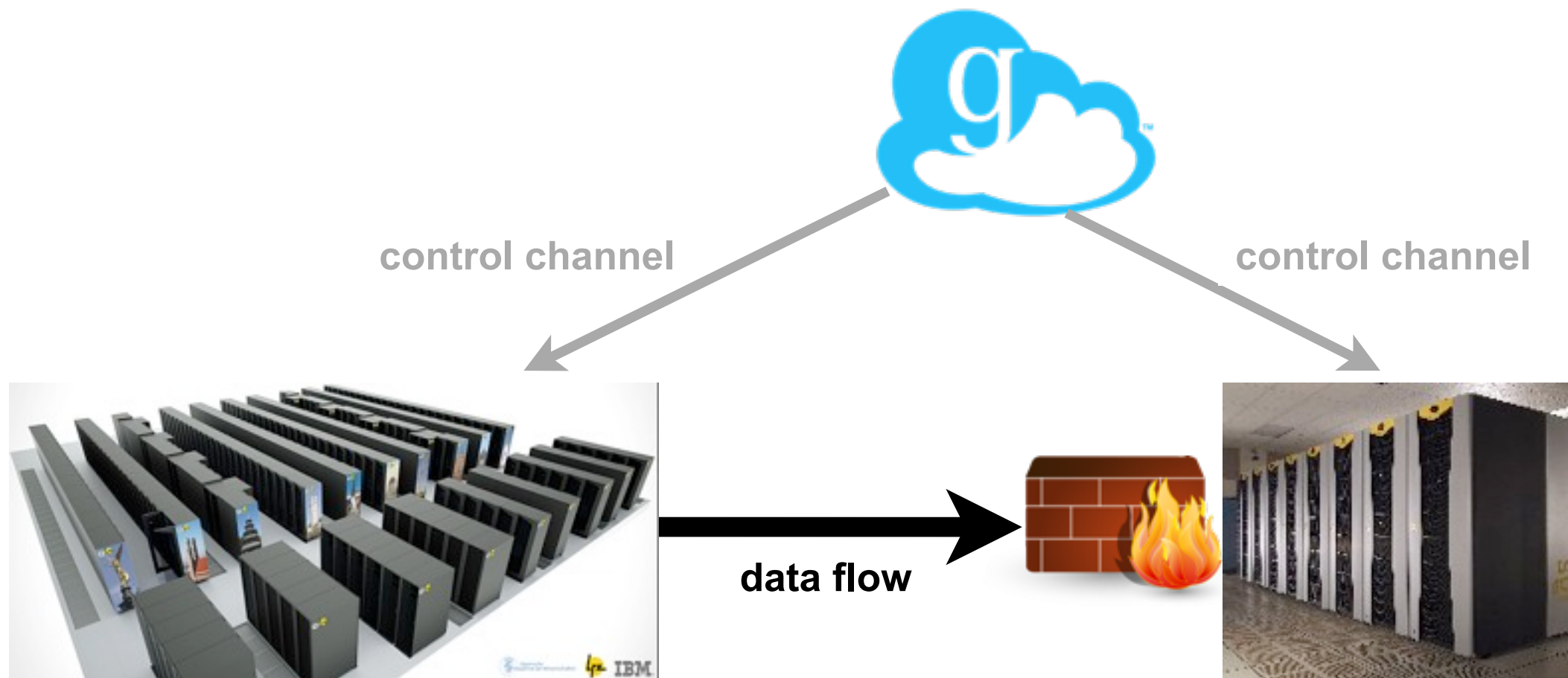
- GridFTP is one of the best protocol for efficient transfer of huge files
- Globus Online: “Reliable File Transfer. No IT required”, multiplatform cloud-based service (Nov. 2010)





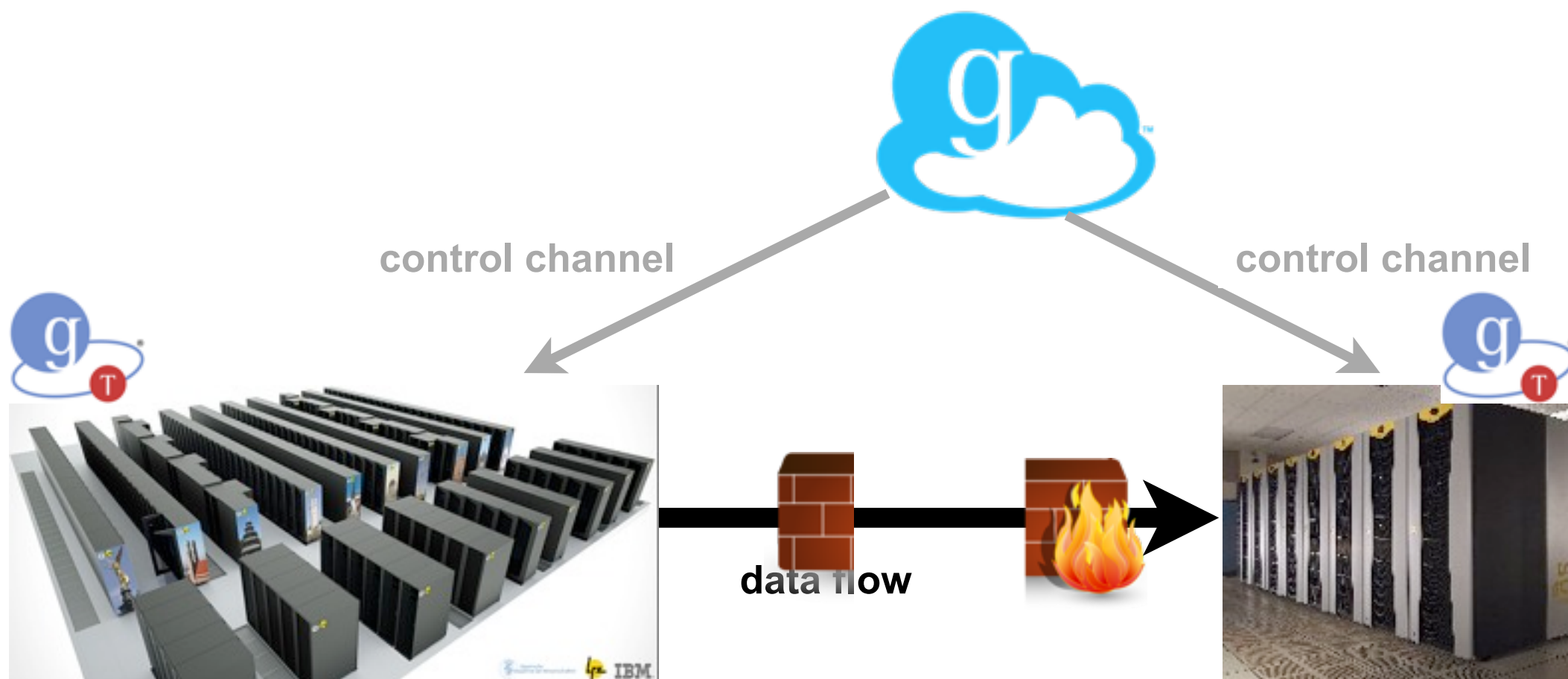
Move a file

- from **SuperMUC**
- to **Linux Cluster**



Move a file

- from **SuperMUC**
- to **Linux Cluster**





## Prerequisites



Grid Certificate or Short Lived Credential  
(<https://slcs.pca.dfn.de/gridshib-ca/>)



Make the proxy available online (upload to MyProxy) ([http://www.grid.lrz.de/en/mware/globus/client/gsissh\\_term.html](http://www.grid.lrz.de/en/mware/globus/client/gsissh_term.html))

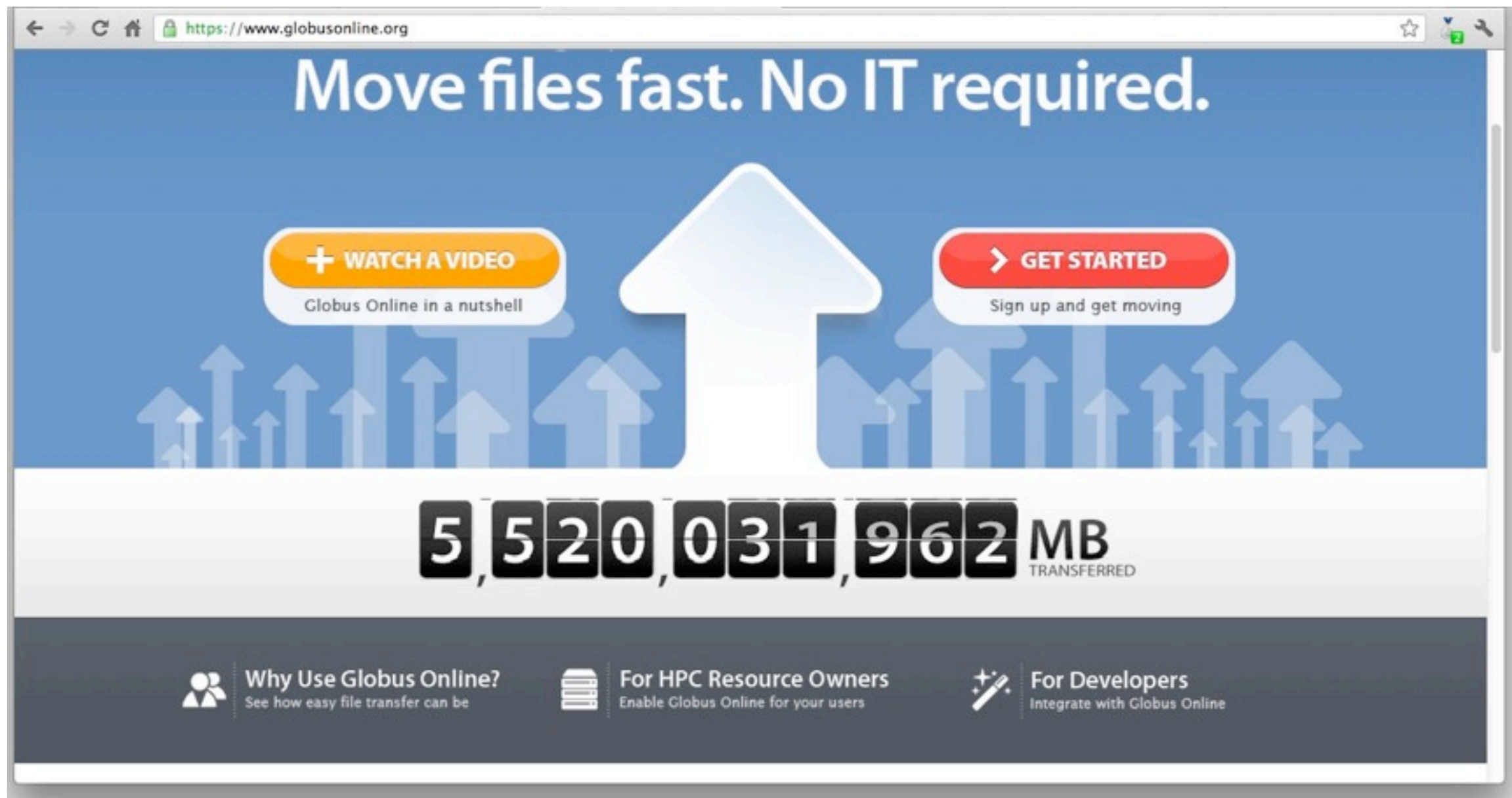


A Globus Online account



# Facts about Globus Online

From data hungry users to ...







... international awards ...

www.rdmag.com/Awards/RD-100-Awards/2012/06/R-D-100-2012-Winners-Overview/

**R&D Magazine**  
**2012**  
**R&D 100 Winners**

50th ANNIVERSARY

Press release: R&D Editors Announce 2012 R&D 100 Awards

The 2012 R&D 100 Award Winners are listed below in alphabetical order by the name of the primary developer company.

Primary Developer	2012 R&D 100 Winning Technology (Category)	Co-developers/Contributors
Adelphi Technology Inc.	High Output Neutron Generator (Beam Instruments)	Lawrence Berkeley National Laboratory
AdValue Photonics	2 Micron High Power Mode-Locked Fiber Laser (Lasers and Photonics)	
Agilent Technologies Australia (M) Pty Ltd	Agilent 4100 Microwave Plasma-Atomic Emission Spectrometer (MP-AES) (Analytical Instruments)	
Argonne National Laboratory	Globus Online (Software)	The University of Chicago
Argonne National Laboratory	High-Energy Concentration-Gradient Cathode Material for Plug-in Hybrids and All-Electric Vehicles (Energy Technologies)	Hanyang University ECOPRO Co. Ltd.
Argonne National Laboratory	Large Area Microchannel Plates (Imaging Technologies)	Incom, Inc. Berkeley Space Sciences Laboratory
Argonne National Laboratory	Ultra-fast and Large-scale Boriding (Process Science)	Istanbul Technical University (I.T.U.) Bodycote Thermal Processing
Arkema Inc.	Plexiglas Rnew Biopolymer Alloys (Materials Sciences)	Arkema France – G. R. L.
ArrayPower	ArrayPower Sequenced Inverter (Electrical Devices)	

Google to tangle in court  
8 hours ago

Key Energy lowers outlook  
10 hours ago

Polymet hires environmental expert as new CEO  
10 hours ago

Advertisement

SUBSCRIPTION CENTER

Magazine Newsletters



... top edge supercomputers ...

www.hpcwire.com/hpcwire/2012-01-25/blue\_waters\_supercomputer\_to\_use\_grid-based\_file\_service.html

**datanami**  
NO DATA • NO ANALYTICS • NO INSIGHTS

**Digital Manufacturing report**

**HPC In the Cloud**

**People to Watch 2012**

**View by Topic**

- Applications
- Developer Tools
- Interconnects
- Middleware
- Networks
- Processors
- Storage
- Systems
- Visualization


**View by Industry**

January 25, 2012

## Blue Waters Supercomputer to Use Grid-Based File Service

Michael Feldman

Who said grid computing and big iron supercomputers don't mix? On Monday it was announced that the National Center for Supercomputing Applications (NCSA) has selected Globus Online as its file transfer solution for the center's upcoming Blue Waters supercomputer.



Globus used GridFTP, a de facto standard for high-performance data movement across large-scale distributed systems. Globus has wrapped a cloud-based interface around it so that users can easily manage file transfers to and fro without having to become experts at Linux and file management commands.

NCSA is not alone in signing up Globus for its supercomputing needs. Other HPC centers including NERSC, Brookhaven National Lab, and facilities at the University of Colorado and University of Chicago have also adopted the grid-based file transfer service.

An article this week in HPC in the Cloud, digs deeper into the whys and wherefores of the selection of Globus for Blue Waters. In the report, Michelle Butler, NCSA's technical program manager, outlines the responsibilities that the center has handed over to Globus:

*Butler explains that Globus Online will be used for transferring data into and out of the Blue Waters machine. The SaaS-based tool will also be tasked with*

**Off the Wire** | **Most Read** | **Blogs**


**February 01, 2012**

- ▶ New Coalition Organizes First Nationwide Meeting for Modeling and Simulation Stakeholders
- ▶ Rogue Wave Delivers Early Version of TotalView Debugger for Blue Gene/Q Supercomputer

**January 31, 2012**

- ▶ Solarflare Announces New 10GbE Offerings for

**10GbE Speed at Your Fingertips!**




www.mellanox.com/ethernet

**Feature Articles**

**Nimbus Brings Half Petabyte Flash Storage to Market**

Storage maker Nimbus Data Systems has launched

**Addison Snell**

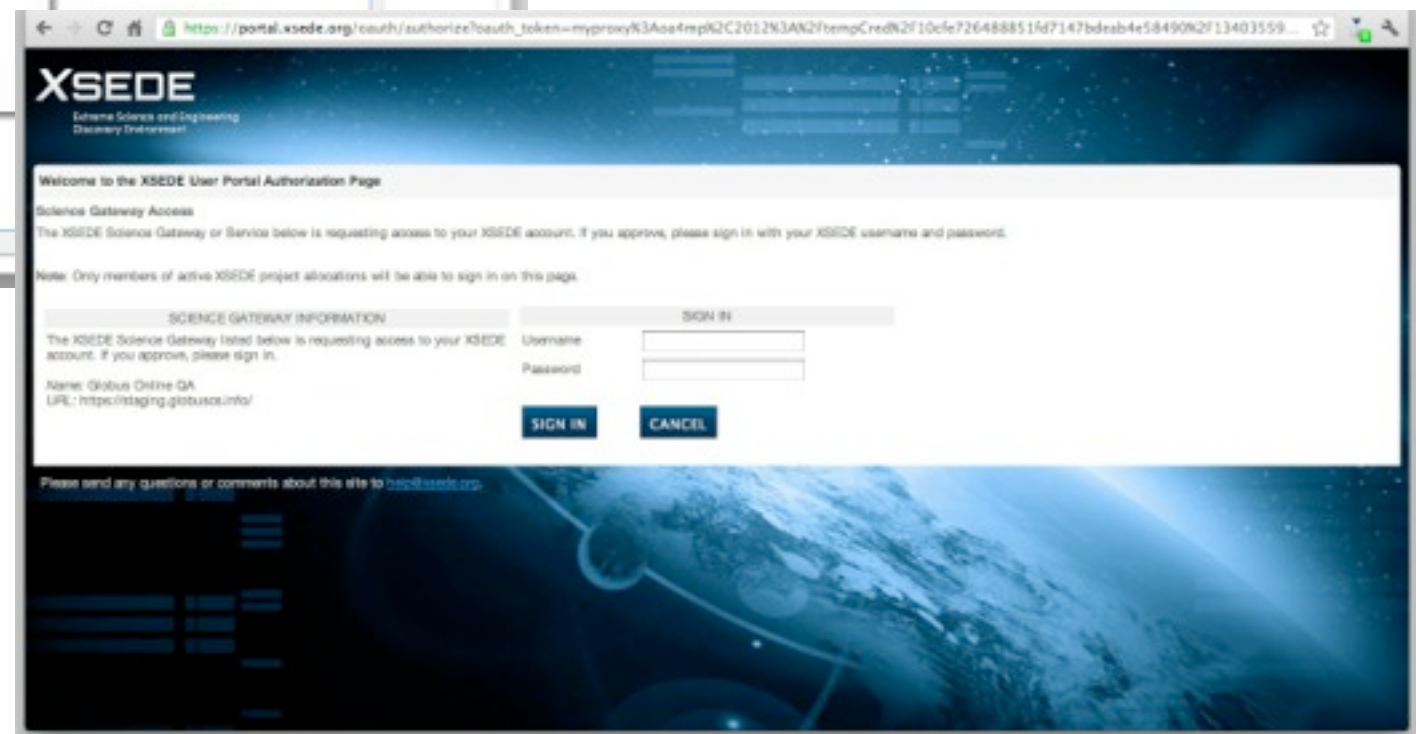
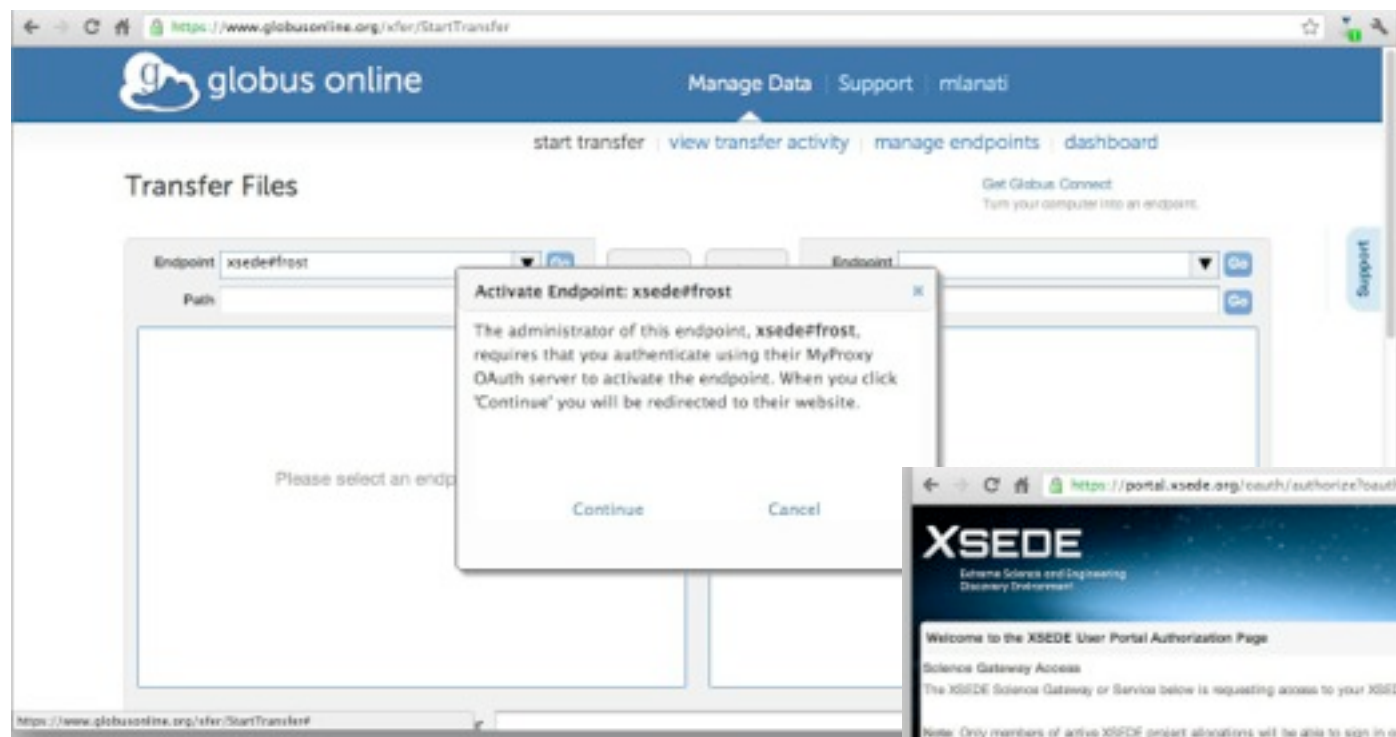






# Facts about Globus Online

... the nitty-gritty of science ...





# Facts about Globus Online

... and beyond







## Globus Online benefits

- Reliable
- Optimized
- Click-and-forget approach
- Easy monitoring

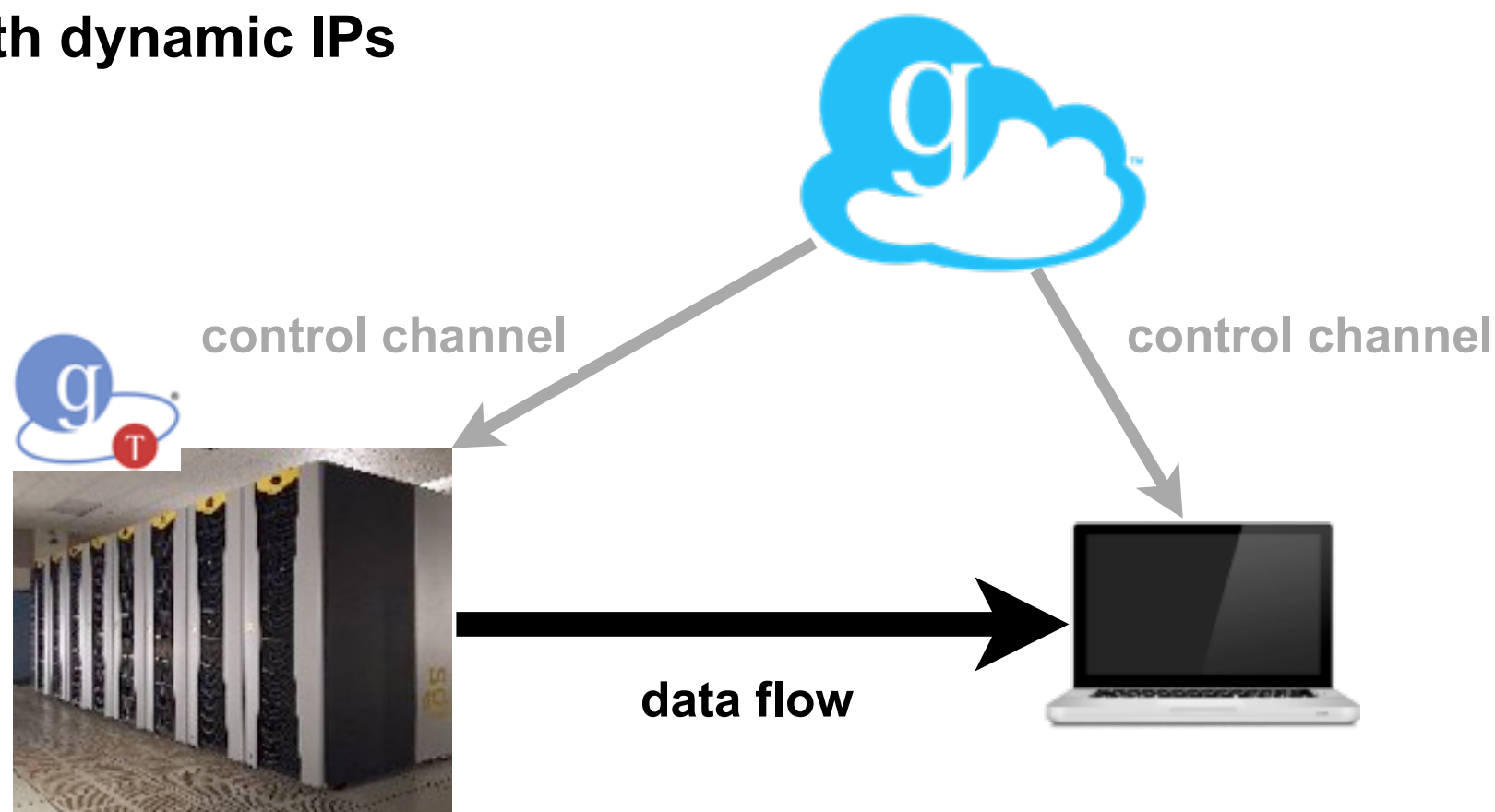


## Retrieve data

- from **an existing endpoint**
- to a laptop

Works behind NAT

Works with dynamic IPs





## For command line addicted

- Install GC on a Linux client
- Scripting interface:
  - upload your key from 'manage identities'
  - connect to [cli.globusonline.org](http://cli.globusonline.org)





- Globus Online is **Software-as-a-Service**
- Uses external services (GridFTP, MyProxy, emails, identity providers)
- Runs on Amazon instances, hosting:
  - DBs (Postgresql, Cassandra)
  - Webservers (Apache)
  - Code (Python 2.6)
  - Globus Toolkit (!)
  - Administration, monitoring and support tools



## Why the Cloud?

- It fits the agile development approach
- Heavy use of snapshots
- Live advice from experts
- Easy and dynamic scaling on demand
- Easy to tests new features
  - and revert back in case of problems



## Globus Online cares about security

- ✓ Exponential delay times for wrong passwords
- ✓ Disallow access to certain paths
- ✓ Use limited proxies
- ✓ Use OAuth
  - eduGAIN under investigation



## How to share resources

- Set up a GridFTP server installing GT 5.x
- Globus Connect Multi-User
  - Auth. mechanism for local users
  - GridFTP installation
  - Endpoint creation
- And ...



- Globus Online is part of a eco system for distributed cooperative environments
- Preview: Globus Storage
  - Wider range of access methods
  - Rent or share resources
  - Take snapshots
  - Default in the cloud





## Globus Online is

- Multiplatform
- User-friendly
- Ready for big data (Gigabytes, Terabytes, Petabytes)
- Automatic
- Cross domain
- Always up to date
- Becoming more and more popular



Make sure to join us at ...

- Join EGCF
  - [http://www.egcf.eu/site/?page\\_id=12](http://www.egcf.eu/site/?page_id=12)
- Attend GlobusEUROPE 2012
  - Sept. 17th 2012, Prague (EGI-TF)
  - <http://www.globuseurope.org/>



# Thank you for your attention