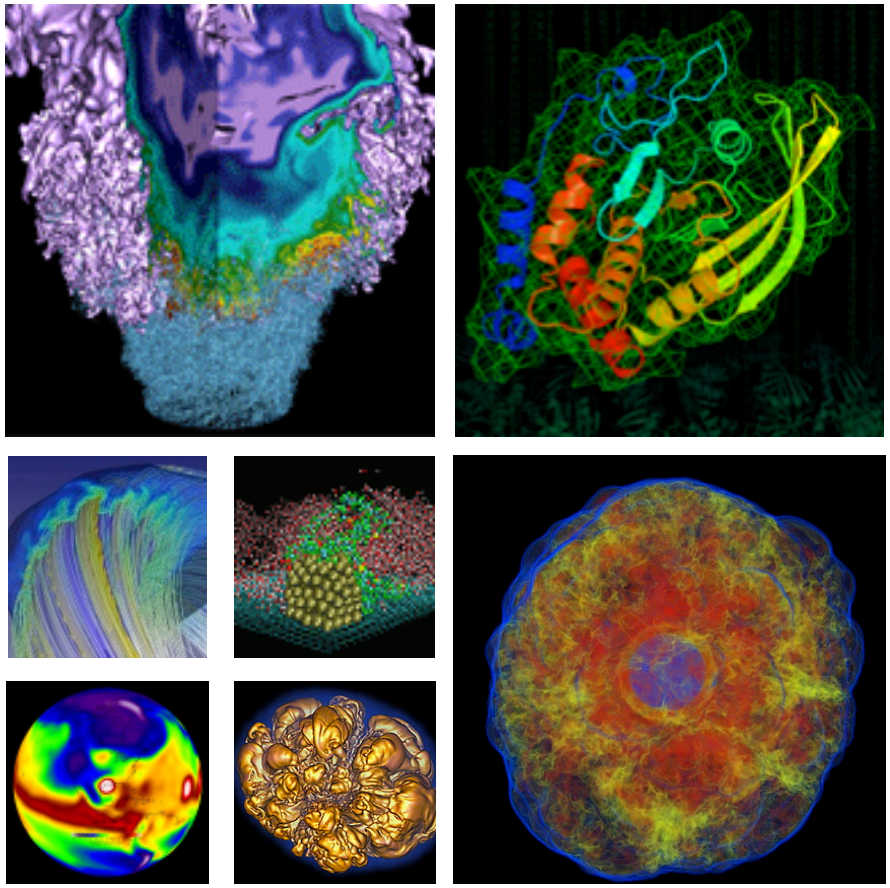
GlobusOnline/HPSS Demo



Nick Balthaser

LBNL/NERSC Storage Systems Group

HUF 2013
November 5, 2013

Agenda

- **Background**
  - GlobusOnline, gridFTP, and HPSS DSI
  - gridFTP/HPSS at NERSC
- **Motivation for Implementation**
- **Live Demo**
  - Web-enabled transfer
- **Pitfalls, Wish List, and To-Dos**
- **Further Info**

Background

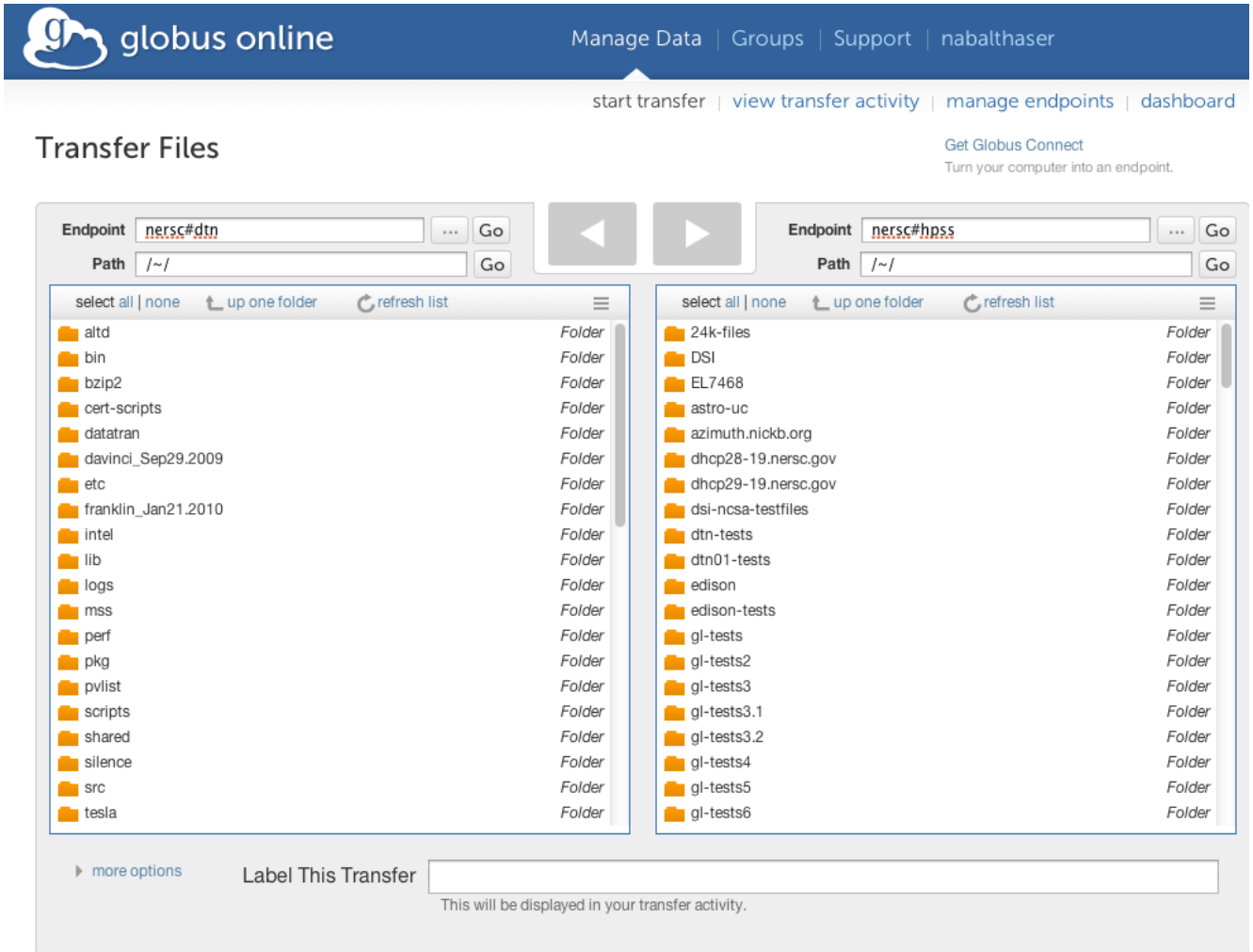
- **GlobusOnline is a user-friendly interface for managing gridFTP data transfers**
  - Both web and CLI transfer management interfaces are supported
  - gridFTP/HPSS communication is implemented via Globus Toolkit Data Storage Interface (DSI) support
    - DSI is a shared library interface for gridFTP communication with non-file system storage backends
    - HPSS DSI uses client API to communicate with HPSS
- **gridFTP/HPSS interface has been in production at NERSC since HPSS 5.x**
  - HPSS 5.x pre-DSI gridFTP authentication code for PFTP developed in-house
  - HPSS 6.2 DSI developed at ANL
    - enabled at NERSC in 2009
    - **nersc#hpss** GO endpoint enabled 04/2012 but disabled after HPSS 7.3 upgrade 03/2013
  - HPSS 7.3.3 DSI developed at NCSA
    - **nersc#hpss** GO endpoint re-enabled 09/2013 after extensive re-testing

Motivation

- **User demand**
  - GO provides two features frequently requested by NERSC archive users:
    - GUI/drag-drop interface for HPSS
    - Directory sync feature (like rsync)
- **Leverages existing gridFTP infrastructure**
  - Uses NERSC CA, LDAP-integrated myproxy auth service, existing data transfer (DTN) nodes

Live Demo

- **GO web initiated transfer**



Live Demo

- **GO transfer status**

globus online | Manage Data | Groups | Support | nabalthaser

start transfer | view transfer activity | manage endpoints | dashboard

Transfer Activity

Cancel ◀◀ ◀ 1 of 1 ▶ ▶▶ View 25 Records

	Status	Label	Task Progress	Completion Time	Request Time
☐	✓	Task Id:4efd884e-2adc-11e3-99f8-12313d2005b7	1 / 1	10/01/2013 01:59 PM	10/01/2013 01:59 PM
☐	✓	Task Id:1ad4cfe8-2166-11e3-99f6-12313d2005b7	1 / 1	09/19/2013 05:05 PM	09/19/2013 01:00 PM
☐	!	Task Id:480d0b58-1fe3-11e3-99f6-12313d2005b7	3113 / 4929 / 8042	09/18/2013 10:19 AM	09/17/2013 02:51 PM
☐	✓	Task Id:ebc6c618-1fbf-11e3-99f6-12313d2005b7	8042 / 8042	09/17/2013 06:04 PM	09/17/2013 10:38 AM
☐	✓	Task Id:e881de1e-1efe-11e3-99f6-12313d2005b7	8042 / 8042	09/16/2013 11:26 PM	09/16/2013 11:36 AM
☐	✓	Task Id:e0259b16-1efe-11e3-99f6-12313d2005b7	8042 / 8042	09/17/2013 01:13 AM	09/16/2013 11:36 AM
☐	!	Task Id:ee628744-1efd-11e3-99f6-12313d2005b7	116 / 7926 / 8042	09/16/2013 11:33 AM	09/16/2013 11:29 AM
☐	!	Task Id:e240d4ac-1efd-11e3-99f6-12313d2005b7	12 / 8030 / 8042	09/16/2013 11:33 AM	09/16/2013 11:29 AM
☐	✓	Task Id:cc3bfa9a-1cc5-11e3-99f6-12313d2005b7	24003 / 24003	09/13/2013 05:50 PM	09/13/2013 03:42 PM
☐	✓	Task Id:c24c2866-1cc5-11e3-99f6-12313d2005b7	24003 / 24003	09/13/2013 05:00 PM	09/13/2013 03:42 PM
☐	✓	Task Id:bbf95772-1cc5-11e3-99f6-12313d2005b7	24003 / 24003	09/13/2013 06:59 PM	09/13/2013 03:42 PM
☐	✓	Task Id:a443bf82-1cc5-11e3-99f6-12313d2005b7	24003 / 24003	09/13/2013 06:13 PM	09/13/2013 03:41 PM
☐	!	Task Id:6c7a68a0-1caa-11e3-99f6-12313d2005b7	5155 / 2887 / 8042	09/13/2013 11:33 PM	09/13/2013 12:26 PM
☐	!	Task Id:5bd718ea-1caa-11e3-99f6-12313d2005b7	7179 / 863 / 8042	09/13/2013 11:33 PM	09/13/2013 12:26 PM
☐	!	Task Id:45df373e-1caa-11e3-99f6-12313d2005b7	6965 / 1077 / 8042	09/13/2013 11:33 PM	09/13/2013 12:25 PM
☐	✓	Task Id:f146a43e-1a63-11e3-9f31-22000a972bd6	17 / 17	09/10/2013 02:57 PM	09/10/2013 02:57 PM
☐	✓	Task Id:78eec7d4-1a4d-11e3-9f31-22000a972bd6	1 / 1	09/10/2013 12:19 PM	09/10/2013 12:16 PM
☐	✓	Task Id:2523c4ba-779b-11e2-b74b-12313906b091	1 / 1	02/15/2013 10:12 AM	02/15/2013 10:11 AM
☐	!	Task Id:67877d1a-7700-11e2-b74b-12313906b091	949 / 7052 / 8001	02/14/2013 03:46 PM	02/14/2013 03:44 PM

Cancel ◀◀ ◀ 1 of 1 ▶ ▶▶ View 25 Records

Live Demo

- **GO Event Log**

globus online | Manage Data | Groups | Support | nabalt

start transfer | view transfer activity | manag

Event Log (1ad4cfe8-2166-11e3-99f6-12313d2005b7)

<< View Transfer Activity

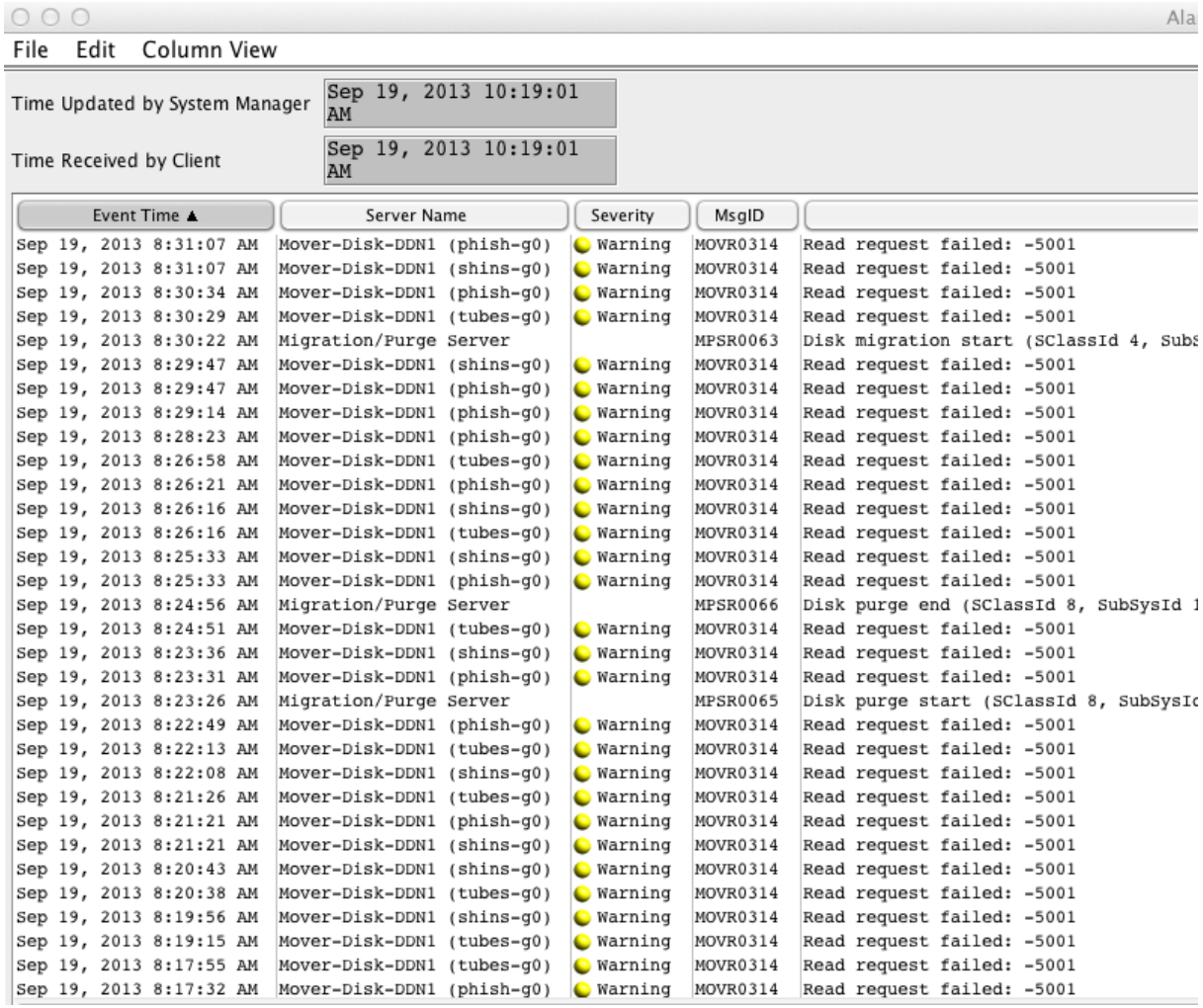
1 of 9

Date	Time	Code	Description	Details
09/19/2013	05:05:38 PM	SUCCEEDED	The operation succeeded	bytes=1073741824000 mbps=671.655
09/19/2013	05:05:29 PM	PROGRESS	Performance monitoring event	bytes=13583515648 mbps=1811.135
09/19/2013	05:04:29 PM	PROGRESS	Performance monitoring event	bytes=12526026752 mbps=1670.137
09/19/2013	05:03:29 PM	PROGRESS	Performance monitoring event	bytes=12663128064 mbps=1688.417
09/19/2013	05:02:29 PM	PROGRESS	Performance monitoring event	bytes=7547650048 mbps=1006.353
09/19/2013	05:01:29 PM	PROGRESS	Performance monitoring event	bytes=5641601024 mbps=752.213
09/19/2013	05:00:29 PM	PROGRESS	Performance monitoring event	bytes=5006950400 mbps=621.981
09/19/2013	04:59:25 PM	PROGRESS	Performance monitoring event	bytes=5037359104 mbps=666.097
09/19/2013	04:58:24 PM	PROGRESS	Performance monitoring event	bytes=3989569536 mbps=531.058
09/19/2013	04:57:24 PM	PROGRESS	Performance monitoring event	bytes=4679008256 mbps=623.868
09/19/2013	04:56:24 PM	PROGRESS	Performance monitoring event	bytes=4453826560 mbps=593.844
09/19/2013	04:55:24 PM	PROGRESS	Performance monitoring event	bytes=4957143040 mbps=660.952
09/19/2013	04:54:24 PM	PROGRESS	Performance monitoring event	bytes=5469372416 mbps=729.250
09/19/2013	04:53:24 PM	PROGRESS	Performance monitoring event	bytes=5642387456 mbps=752.318
09/19/2013	04:52:24 PM	PROGRESS	Performance monitoring event	bytes=5238947840 mbps=698.526
09/19/2013	04:51:24 PM	PROGRESS	Performance monitoring event	bytes=5377359872 mbps=716.981
09/19/2013	04:50:24 PM	PROGRESS	Performance monitoring event	bytes=5881200640 mbps=724.955

Pitfalls, Wish List, To-Dos

- **Pitfalls**
  - GO sync feature is not optimal for HPSS
    - Recursive directory syncs can store lots of small files
    - Recursive retrievals can request files spread across hundreds or even thousands of cartridges
      - No tape ordering facility for optimizing retrievals
  - GO retry logic can cause unexpected issues
    - Repeated errors during failed transfers (see next slide)
    - No visibility into client side transfer problems
  - Early "Resume" bug
    - Sparse files in HPSS had to be located and deleted, user notifications sent
  - Empty checksum records in HPSS
- **Wish List**
  - Better gridFTP session logging
  - Administrative controls in GO
    - job management (list/pause/cancel)
    - Session limits
  - Facility to query volume/aggregate metadata to enable ordered retrieval
  - Improvements to retry logic
- **To Do**
  - Enable GridFTP checksums

GO Retries



File Edit Column View

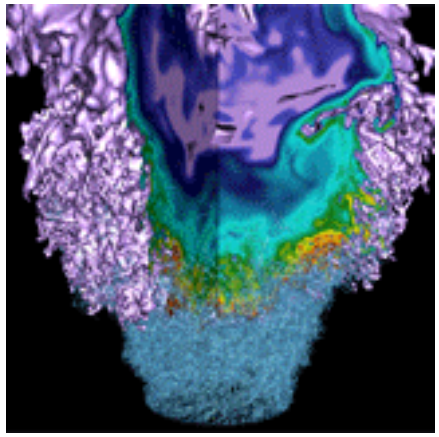
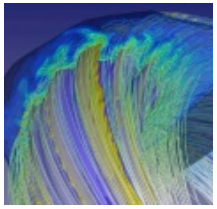
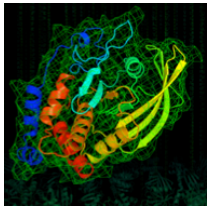
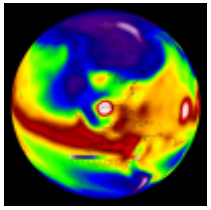
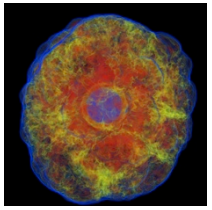
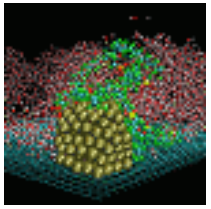
Time Updated by System Manager: Sep 19, 2013 10:19:01 AM

Time Received by Client: Sep 19, 2013 10:19:01 AM

Event Time ▲	Server Name	Severity	MsgID	
Sep 19, 2013 8:31:07 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:31:07 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:30:34 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:30:29 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:30:22 AM	Migration/Purge Server		MPSR0063	Disk migration start (SClassId 4, Subs
Sep 19, 2013 8:29:47 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:29:47 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:29:14 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:28:23 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:26:58 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:26:21 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:26:16 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:26:16 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:25:33 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:25:33 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:24:56 AM	Migration/Purge Server		MPSR0066	Disk purge end (SClassId 8, SubSysId 1
Sep 19, 2013 8:24:51 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:23:36 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:23:31 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:23:26 AM	Migration/Purge Server		MPSR0065	Disk purge start (SClassId 8, SubSysId
Sep 19, 2013 8:22:49 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:22:13 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:22:08 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:21:26 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:21:21 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:21:21 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:20:43 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:20:38 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:19:56 AM	Mover-Disk-DDN1 (shins-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:19:15 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:17:55 AM	Mover-Disk-DDN1 (tubes-g0)	Warning	MOVR0314	Read request failed: -5001
Sep 19, 2013 8:17:32 AM	Mover-Disk-DDN1 (phish-g0)	Warning	MOVR0314	Read request failed: -5001

Further Information

- **DSI Info: GridFTP Developer's Guide:** *http://www.globus.org/toolkit/docs/5.2/5.2.4/gridftp/developer/*
- **GlobusOnline Resource Provider's Guide:** *https://www.globusonline.org/media/publications/The_Resource_Providers_Guide_to_Globus_Online/*
- **NERSC Grid Software and Services:** *https://www.nersc.gov/users/software/grid/*
- **Grid implementation at LBNL/NERSC:**
  - Shreyas Cholia, Outreach, Software, & Programming: *SCholia@lbl.gov*
  - Nick Balthaser, Storage Systems: *nabalthaser@lbl.gov*

National Energy Research Scientific Computing Center

Section Title