

Leveraging Globus to support access and delivery of scientific data

Thomas Cram
Doug Schuster
Steven Worley
NCAR, Boulder, CO



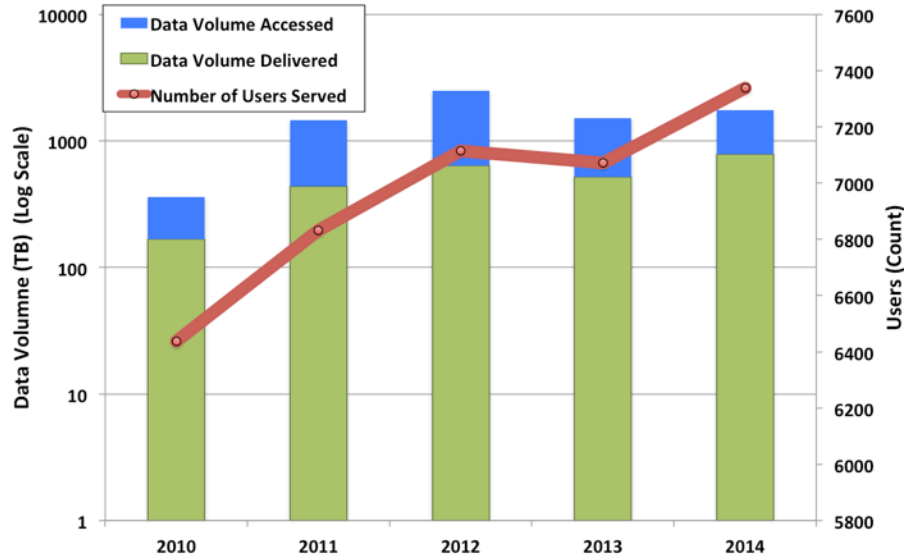
What is the RDA?



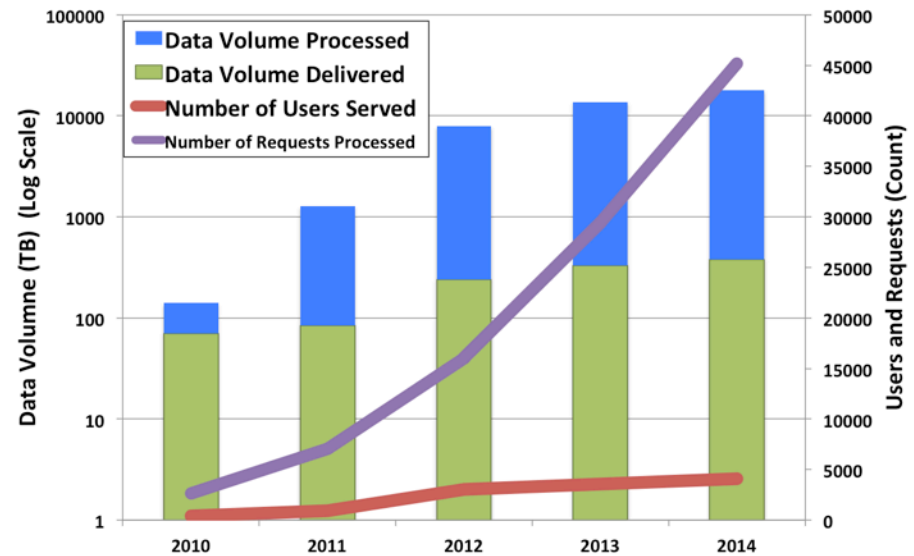
- 600+ datasets for climate and weather research
- Free and open access
- Worldwide usage
- Various data access pathways
 - **HTTP (wget, cURL, etc.)**
 - OPeNDAP, WCS, WMS
 - Web services (CLI, API)
 - Direct analysis on HPC systems (NCAR users)

RDA Web Usage

Yearly RDA User Access from Web Interface
Direct Archive File Downloads



Yearly Customized RDA User Access from Web Interface
Automated Archive Subsetting and Format Conversion Requests



FY 2015:

- 17+ PB virtual processing
- 11,000+ unique users
- 1.6 PB served

**That's a lot of data.
How can we improve download
efficiency and reliability?**



Advantages

- *Reliable, secure, high-performance* file transfer
- “Fire and forget”
- Automatic fault recovery
- Powerful GUI and APIs



www.globus.org



Globus @ RDA

- Globus Connect GridFTP server at NCAR
 - 5 Gb/s internal, 2 Gb/s external (ideal conditions)
- RDA managed endpoints
 - **rda#datashare** and **rda#data_request**
 - Data shared from subdirectories under endpoint source path
 - Access managed programmatically via Globus CLI
 - Data shares deleted after 5 days for custom orders

RDA example user interface



NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999
ds083.2

For assistance, contact [Grace Peng](#) (303-497-1218).

Description

Data Access

Documentation

Software

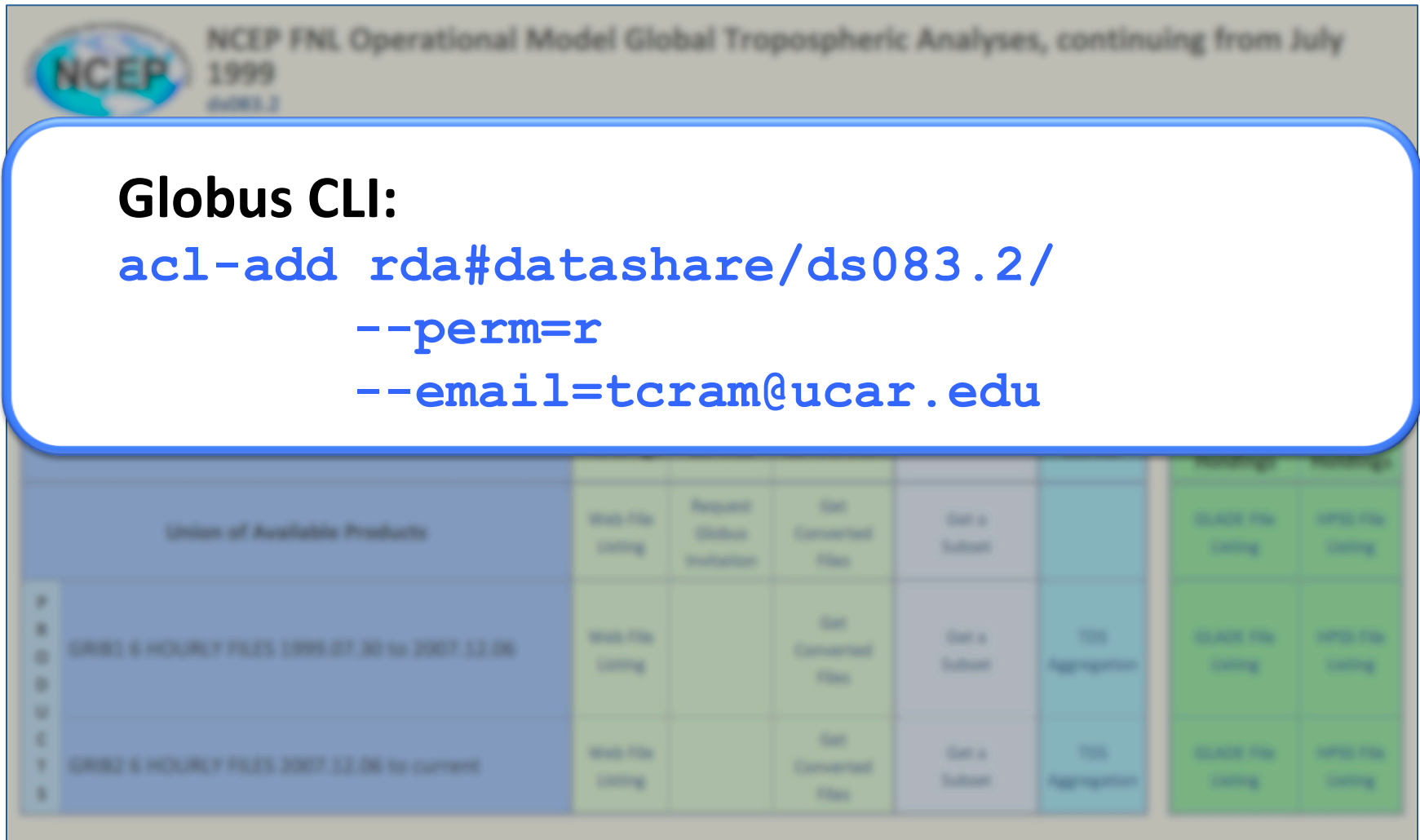
Mouse over the table headings for detailed descriptions

Data Description		Data File Downloads			Customizable Data Requests	Other Access Methods	NCAR-Only Access	
		Web Server Holdings	Globus Transfer Service	Data Format Conversion	Subsetting	THREDDS Data Server	Central File System (GLADE) Holdings	Tape Archive (HPSS) Holdings
Union of Available Products		Web File Listing	Request Globus Invitation	Get Converted Files	Get a Subset		GLADE File Listing	HPSS File Listing
PRODUCTS	GRIB1 6 HOURLY FILES 1999.07.30 to 2007.12.06	Web File Listing		Get Converted Files	Get a Subset	TDS Aggregation	GLADE File Listing	HPSS File Listing
	GRIB2 6 HOURLY FILES 2007.12.06 to current	Web File Listing		Get Converted Files	Get a Subset	TDS Aggregation	GLADE File Listing	HPSS File Listing

RDA example user interface

Globus CLI:

```
acl-add rda#datashare/ds083.2/  
--perm=r  
--email=tcram@ucar.edu
```



Union of Available Products					
	Web File Listing	Request Global Information	Get Converted File	Get a Subset	
P					
R					
G					
B					
6					
H					
O					
U					
R					
E					
S					
1					
9					
9					
0					
7					
-					
0					
6					
to					
2					
0					
0					
7					
-					
1					
2					
-					
0					
6					
to					
c					
u					
r					
r					
e					
n					
t					

RDA user dashboard

View, re-send, and delete Globus shares

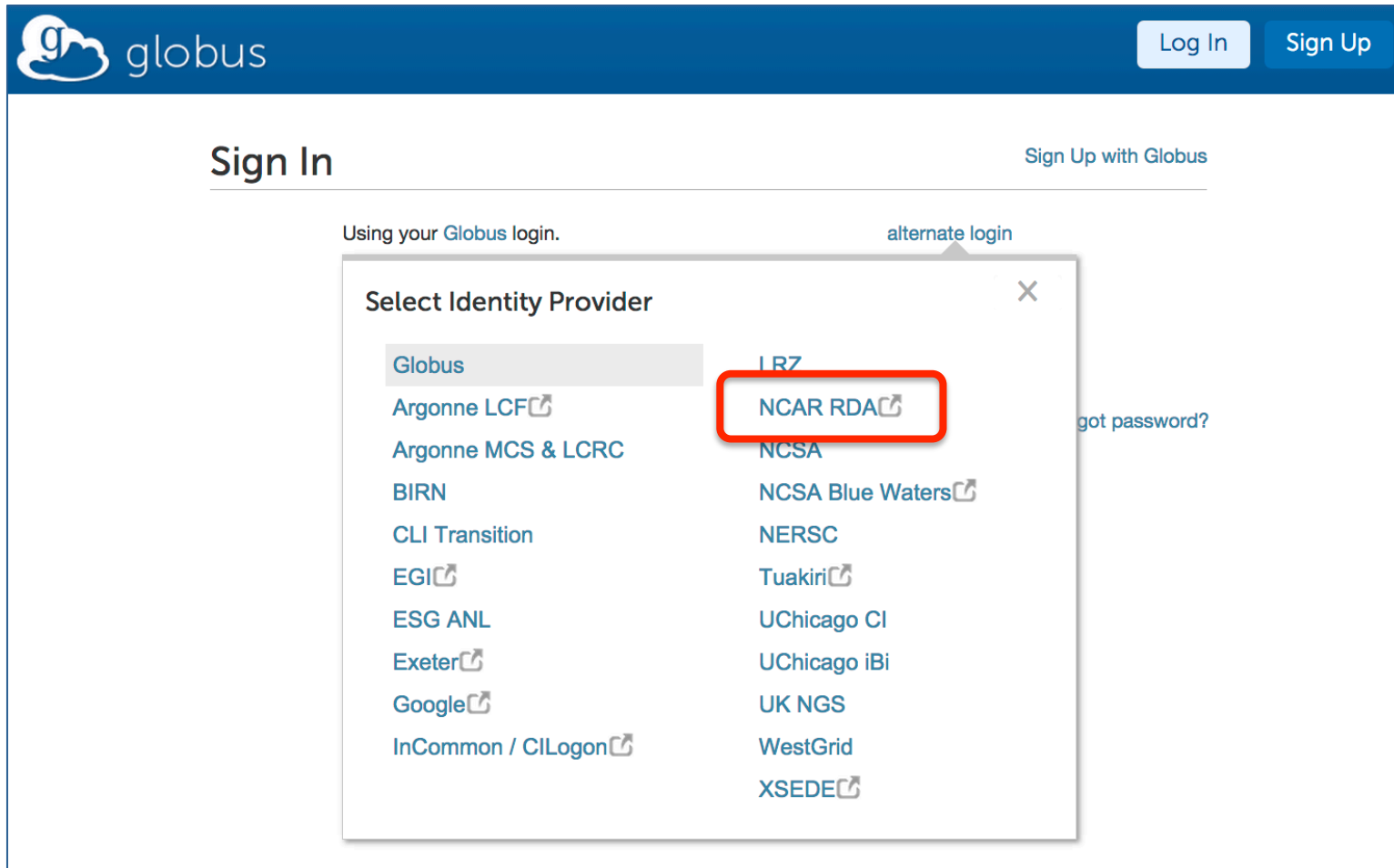
Globus Endpoints Shared With You

[Hide Shares](#)

8 RDA data paths have been shared with you.

Endpoint:Path	Globus URL	Manage
rda#datashare:/ds083.2/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds083.2%2F	 
rda#datashare:/ds084.3/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds084.3%2F	 
rda#datashare:/ds131.2/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds131.2%2F	 
rda#datashare:/ds132.0/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds132.0%2F	 
rda#datashare:/ds316.1/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds316.1%2F	 
rda#datashare:/ds502.0/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds502.0%2F	 
rda#datashare:/ds502.1/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds502.1%2F	 
rda#datashare:/ds601.0/	https://www.globus.org/xfer/StartTransfer?origin=rda%23datashare%2Fds601.0%2F	 

RDA Alternate Identity login



The screenshot shows the Globus Sign In interface. At the top, the Globus logo is on the left, and 'Log In' and 'Sign Up' buttons are on the right. Below the header, the 'Sign In' section is active, with a link to 'Sign Up with Globus'. The main content area has two options: 'Using your Globus login.' and 'alternate login'. The 'alternate login' option is selected, which has triggered a 'Select Identity Provider' modal. This modal lists various identity providers in two columns. The 'NCAR RDA' provider is highlighted with a red rectangle. Other providers include Globus, Argonne LCF, Argonne MCS & LCRC, BIRN, CLI Transition, EGI, ESG ANL, Exeter, Google, InCommon / CILogon, IRZ, NCSA, NCSA Blue Waters, NERSC, Tuakiri, UChicago CI, UChicago iBi, UK NGS, WestGrid, and XSEDE.

globus

Log In Sign Up

Sign In Sign Up with Globus

Using your Globus login. alternate login

Select Identity Provider

Globus	IRZ
Argonne LCF	NCAR RDA
Argonne MCS & LCRC	NCSA
BIRN	NCSA Blue Waters
CLI Transition	NERSC
EGI	Tuakiri
ESG ANL	UChicago CI
Exeter	UChicago iBi
Google	UK NGS
InCommon / CILogon	WestGrid
	XSEDE

got password?

RDA Alternate Identity login



NCAR's Research Data Archive (RDA) MyProxy Client Authorization

Welcome to NCAR's RDA OAuth for MyProxy Client Authorization Page. The Client below is requesting access to your account. If you approve, please sign in with your RDA email/username and RDA password.

Client Information

Name: Globus
URL: <https://www.globus.org>

NCAR RDA Email/Username

tcram@ucar.edu *

NCAR RDA Password

..... *

Sign In

Cancel

Globus Management Console

Monitor activity and retrieve usage metrics

[Transfer Files](#) | [Activity](#) | [Manage Endpoints](#) | [Dashboard](#) | [Console](#)

Management Console Beta

[endpoint overview](#) [active tasks](#)

nda#datashare



Tasks matching:

nda#datashare

Completed Tasksx



☐ — select action —

Tasks: [active](#) [completed](#) [with faults](#) [custom filter](#)



from nda#datashare (ncar#datashare) to <private endpoint> initiated by tcram 30 minutes ago

SUBTASKS	SUCCEEDED	FAILED	BYTES TX	MBit/s	CUM. FAULTS	DURATION	COMPLETED
4	4	0	67.21 MB	170.43	0	00:00:04	2015-06-30T17:33:58+00:00



from nda#datashare (ncar#datashare) to <private endpoint> initiated by tcram an hour ago

SUBTASKS	SUCCEEDED	FAILED	BYTES TX	MBit/s	CUM. FAULTS	DURATION	COMPLETED
5	5	0	1.26 GB	529.24	0	00:00:20	2015-06-30T17:16:14+00:00



from nda#datashare (ncar#datashare) to <private endpoint> initiated by pthoya a day ago

SUBTASKS	SUCCEEDED	FAILED	BYTES TX	MBit/s	CUM. FAULTS	DURATION	COMPLETED
2	2	0	57.48 MB	7.72	1	00:01:03	2015-06-29T15:14:34+00:00



from nda#datashare (ncar#datashare) to <private endpoint> initiated by miroziak 2 days ago

Upcoming features

- Globus HTTPS downloads
- Globus Auth service
 - single sign-on with RDA identity
 - no Globus identity requirement
- Custom file list transfers

Contact us

[*rda.ucar.edu*](http://rda.ucar.edu)



@NCAR_RDA



NCAR CISL Research Data Archive



[*ncarrda.blogspot.com*](http://ncarrda.blogspot.com)