

2018-02-20

Data Sharing Workflow for Large Datasets (>1TB) with Globus

Borda, Susan

<https://hdl.handle.net/2027.42/142389>

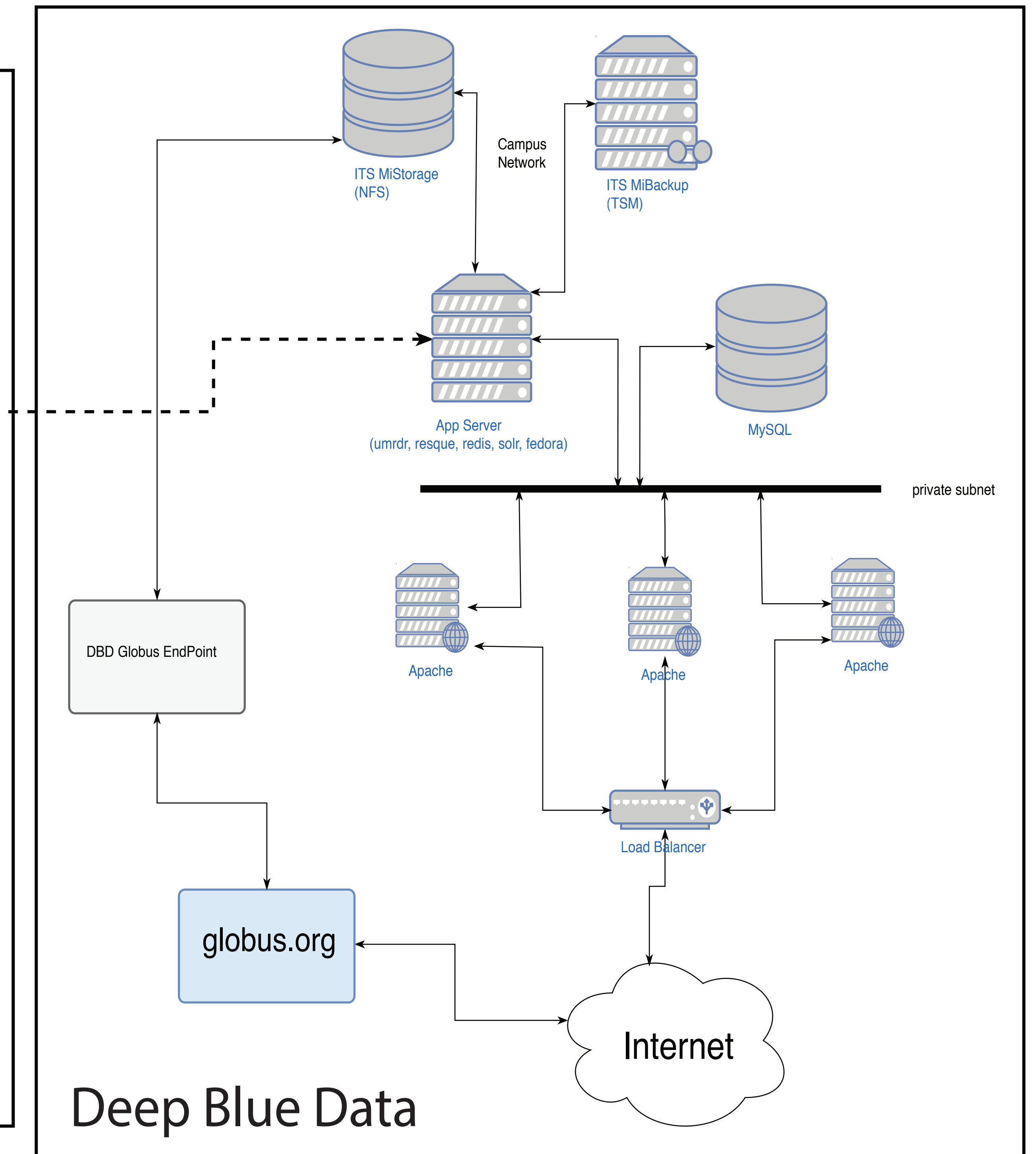
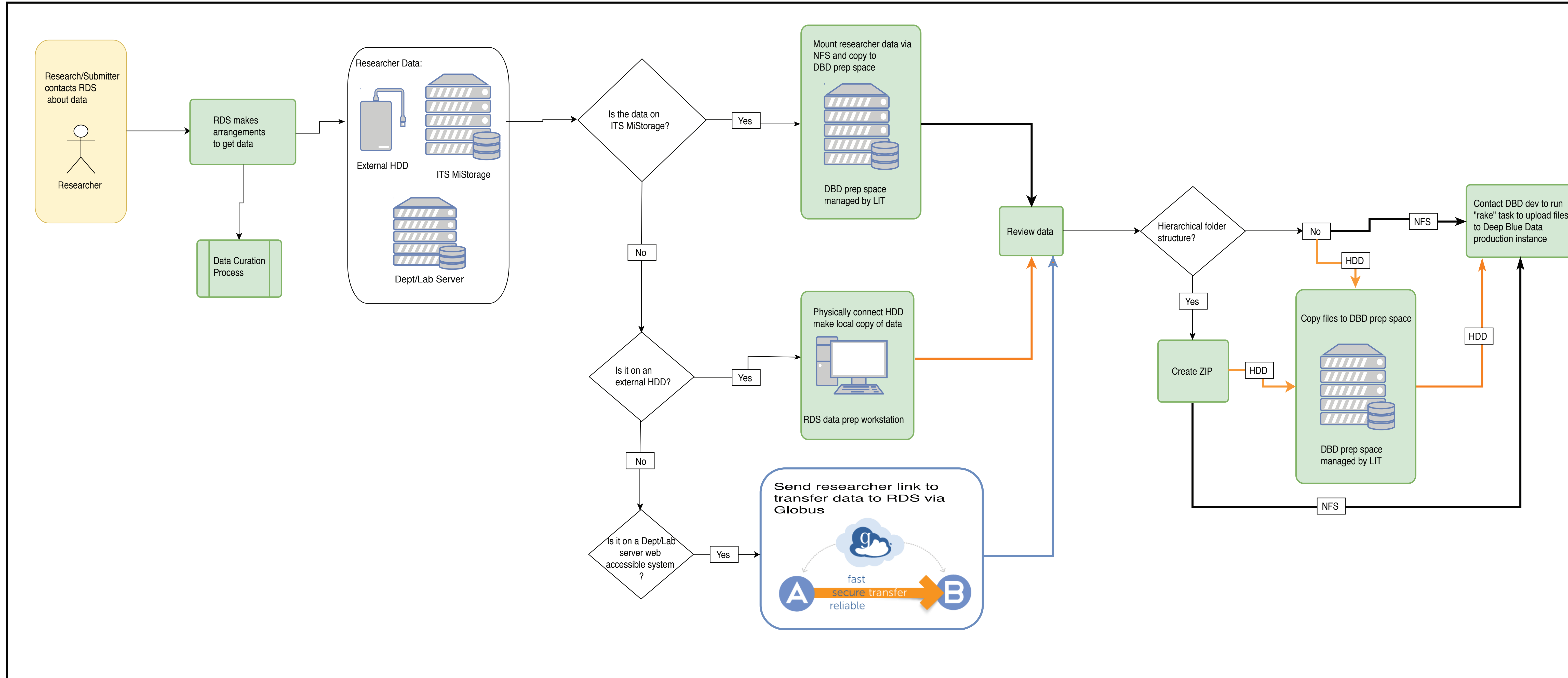
<http://creativecommons.org/publicdomain/zero/1.0/>

Downloaded from Deep Blue, University of Michigan's institutional repository

Data Sharing Workflow for Large Datasets (>1TB) with Globus

Susan Borda - Data Workflows Specialist University of Michigan - sborda@umich.edu

Workflow for Large Datasets



Globus File Download Process

1 Files (65.3 GB in 24 files)

File	Filename	Date Uploaded	File Size	Access	Actions
	rcm_era_atm_present.nc	2017-08-25	4.80 GB	Open Access	View Details
	rcm_era_clm_present.nc	2017-08-25	4.31 GB	Open Access	View Details
	rcm_era_rad_present.nc	2017-08-25	3.64 GB	Open Access	View Details
	rcm_era_srf_present.nc	2017-08-25	1.03 GB	Open Access	View Details
	rcm_era_sts_present.nc	2017-08-25	471.06 MB	Open Access	View Details
	rcm_era_atm_future.nc	2017-08-25	4.56 GB	Open Access	View Details
	rcm_era_clm_future.nc	2017-08-25	4.92 GB	Open Access	View Details
	rcm_era_rad_future.nc	2017-08-25	4.09 GB	Open Access	View Details
	rcm_era_srf_future.nc	2017-08-25	4.42 GB	Open Access	View Details
	rcm_era_sts_future.nc	2017-08-25	3.46 GB	Open Access	View Details
	rcm_gfdl_atm_present.nc	2017-08-25	3.73 GB	Open Access	View Details
	rcm_gfdl_atm_future.nc	2017-08-25	978.40 MB	Open Access	View Details
	rcm_gfdl_clm_present.nc	2017-08-25	1.05 GB	Open Access	View Details
	rcm_gfdl_clm_future.nc	2017-08-25	447.24 MB	Open Access	View Details
	rcm_gfdl_rad_present.nc	2017-08-25	4.56 GB	Open Access	View Details
	rcm_gfdl_rad_future.nc	2017-08-25	4.92 GB	Open Access	View Details
	rcm_gfdl_srf_present.nc	2017-08-25	4.09 GB	Open Access	View Details
	rcm_gfdl_srf_future.nc	2017-08-25	4.42 GB	Open Access	View Details
	rcm_hadgem_atm_present.nc	2017-08-25	3.46 GB	Open Access	View Details
	rcm_hadgem_atm_future.nc	2017-08-25	3.73 GB	Open Access	View Details
	rcm_hadgem_clm_present.nc	2017-08-25	4.56 GB	Open Access	View Details
	rcm_hadgem_clm_future.nc	2017-08-25	4.92 GB	Open Access	View Details
	rcm_hadgem_rad_present.nc	2017-08-25	4.09 GB	Open Access	View Details
	rcm_hadgem_rad_future.nc	2017-08-25	4.42 GB	Open Access	View Details
	rcm_hadgem_srf_present.nc	2017-08-25	3.46 GB	Open Access	View Details
	rcm_hadgem_srf_future.nc	2017-08-25	3.73 GB	Open Access	View Details

Total work file size of 65.3 GB is too large to download directly, please contact deepblue@umich.edu.

[Download All for Globus](#)

2 Click submit to start downloading Regional Climate Model simulations to Globus. If you would like to be sent an email when the files are available fill in your email address:

Email:

Email again:

[Download All for Globus and Email Me When Complete](#) [Cancel X](#)

3 sborda@umich.edu 12/22/17

to me

Globus download is now available.
 Work: Regional Climate Model simulations
 At: http://training.deepblue.lib.umich.edu/data/concern/generic_works/75903f27-1331-406e-9f95-a4871d02e525
 By: Steiner, Allison; Bryan, Alex
 Deposited by: alsteine@umich.edu
 Globus link: https://www.globus.org/app/transfer?origin_id=99d8c648-a9ff-11e7-aedd-2200a92523b&origin_path=%2Fdownload%2FDeepBlueData_75903f27-1331-406e-9f95-a4871d02e525%2F

4 Transfer Files

Endpoint: DeepBlueData

Path: /download/DeepBlueData_75903f27-1331-406e-9f95-a4871d02e525

select all up one folder refresh list permissions

rcm_era_atm_present.nc	4.80 GB
rcm_era_clm_present.nc	4.31 GB
rcm_era_rad_present.nc	3.64 GB
rcm_era_srf_present.nc	1.03 GB
rcm_era_sts_present.nc	471.06 MB
rcm_era_atm_future.nc	4.56 GB
rcm_era_clm_future.nc	4.92 GB
rcm_era_rad_future.nc	4.09 GB
rcm_era_srf_future.nc	4.42 GB
rcm_era_sts_future.nc	3.46 GB
rcm_gfdl_atm_present.nc	3.73 GB
rcm_gfdl_atm_future.nc	978.40 MB
rcm_gfdl_clm_present.nc	1.05 GB
rcm_gfdl_clm_future.nc	447.24 MB
rcm_gfdl_rad_present.nc	4.56 GB
rcm_gfdl_rad_future.nc	4.92 GB
rcm_gfdl_srf_present.nc	4.09 GB
rcm_gfdl_srf_future.nc	4.42 GB
rcm_hadgem_atm_present.nc	3.46 GB
rcm_hadgem_atm_future.nc	3.73 GB
rcm_hadgem_clm_present.nc	4.56 GB
rcm_hadgem_clm_future.nc	4.92 GB
rcm_hadgem_rad_present.nc	4.09 GB
rcm_hadgem_rad_future.nc	4.42 GB
rcm_hadgem_srf_present.nc	3.46 GB
rcm_hadgem_srf_future.nc	3.73 GB

Label This Transfer

Transfer Settings

- sync - only transfer new or changed files
- delete files on destination that do not exist on source
- preserve source file modification times
- verify file integrity after transfer
- encrypt transfer

4a Activity

Back to Transfer Files Task List

DeepBlueData to sborda lib macbook transfer started 5 minutes ago

Overview Event Log

Task ID: f85e1da8-fb08-11e7-831b-0a208f818180

Owner: susan borda (sborda@umich.edu)

Source: DeepBlueData (owner: skomer@umich.edu)

Destination: sborda lib macbook (owner: sborda@umich.edu)

Condition: ACTIVE

Requested: 2018-01-16 05:02 pm

Deadline: 2018-01-17 05:02 pm

Transfer Settings

- verify file integrity after transfer
- transfer is not encrypted
- overwriting all files on destination

Files: 11
 Directories: 0
 Bytes Transferred: 12.47 GB
 Effective Speed: 37.79 MB/s

Pending: 8
 Succeeded: 3
 Cancelled: 0
 Expired: 0
 Failed: 0
 Retrying: 0
 Skipped: 0

5 Globus

Search

Name	Date Modified	Size	Kind
rcm_era_atm_present.nc	Yesterday at 5:06 PM	4.81 GB	Document
rcm_era_clm_present.nc	Yesterday at 5:06 PM	4.31 GB	Document
rcm_era_rad_present.nc	Yesterday at 5:09 PM	3.65 GB	Document
rcm_era_srf_present.nc	Yesterday at 5:07 PM	1.03 GB	Document
rcm_era_sts_present.nc	Yesterday at 5:07 PM	471.1 MB	Document
rcm_gfdl_atm_future.nc	Yesterday at 5:10 PM	4.56 GB	Document
rcm_gfdl_clm_present.nc	Yesterday at 5:13 PM	4.93 GB	Document
rcm_gfdl_atm_present.nc	Yesterday at 5:13 PM	4.1 GB	Document
rcm_gfdl_clm_future.nc	Yesterday at 5:16 PM	4.42 GB	Document
rcm_gfdl_rad_future.nc	Yesterday at 5:16 PM	3.46 GB	Document
rcm_gfdl_rad_present.nc	Yesterday at 5:18 PM	3.74 GB	Document

Globus Code on Github:

<https://github.com/mlibrary/umrdr/tree/master/app/jobs>

