

# HPC Data Management

---

[Douglas.Benjamin@duke.edu](mailto:Douglas.Benjamin@duke.edu)

[Mario.Lassnig@cern.ch](mailto:Mario.Lassnig@cern.ch)

# Current situation

---

- ↪ GlobusOnline implementation in Harvester
  - ↪ Files in - Preparator - [go\\_preparator.py](#)
  - ↪ Files out - Stager - [go\\_stager.py](#)
  
- ↪ Eventually, move to Rucio-managed workflow
  - ↪ Exploit all the machinery we already have
  - ↪ Transparently move data between Gridsites and HPC sites using the same rule method
  - ↪ Take care of deletion using central Rucio deletion machinery
  - ↪ (+added bonus: non-ATLAS experiments benefit from having GlobusOnline in Rucio)

# Issues seen using Globus w/ harvester

---

- ↪ Globus limits are fairly low
  - ↪ 3 active transfers
  - ↪ 100 queued transfers
- ↪ Globus transfers are point to point
- ↪ Doug B. gets email for all transfers (good or bad) - email DOS
- ↪ There is some latency between submitting transfer and being able to see the transfer in the Globus monitoring
- ↪ While Globus will record the transfer and save information about for 31 days, there is no way to know from Globus what files are associated with a particular transfer
  - ↪ Will need Rucio to keep track of the association between Globus transfer id and the individual files

# Other operational issues with US HPC's and Globus

---

- ↪ Need to have these transfers run under a different Globus account instead Doug Benjamin's - So this "service account" can spam itself
  - ↪ Require creating new Globus long lived credentials associate with an account
- ↪ Two of the 3 US DOE HPC sites (ALCF and OLCF) require two factor authentication to activate the Globus endpoints
  - ↪ Manual operation requiring external device
  - ↪ Cannot be made automatic as of now.
- ↪ NERSC (other US DOE site) is moving to two factor authentication
- ↪ Globus Endpoint authentication lasts 7 days
- ↪ Still to be tested - Is it possible for person A to activate the endpoints on behalf of the "service account"?

# Proposed Rucio-managed workflow



## ↪ Transfer to HPC

- ↪ 2 rules necessary: one to **HPC**, one to **Dual Stack Site**
- ↪ Transfer with *FTS* from **Grid Site** to **Dual Stack Site** (if not already on **Dual Stack Site**)
- ↪ Transfer with *GlobusOnline* from **Dual Stack Site** to **HPC** (force **Dual Stack Site** as only available source)
  - Expire rule at **Dual Stack Site** and delete with Rucio Reaper

## ↪ Transfer from HPC

- ↪ 2 rules necessary: one to **Dual Stack Site**, one to **Grid Site**
- ↪ Transfer with *GlobusOnline* from **HPC** to **Dual Stack Site**
  - Expire rule at **HPC** and delete with *GlobusOnline*
- ↪ Transfer with *FTS* from **Dual Stack Site** to **Grid Site** (if needed)
  - Expire rule at **Dual Stack Site** and delete with *Rucio Reaper*

# Proposed Rucio-managed workflow



## ↪ Deletion

- ↪ Has to be done asynchronously using GlobusOnline on **HPC**
- ↪ Cannot use direct reaper method, only for **Dual Stack Site** and **Grid Site**

## ↪ Conveyor is prepared for asynchronous operations

- ↪ Right now, we only use a the REQUEST\_TYPE = T (transfer)
- ↪ Extend to use REQUEST\_TYPE = D (deletion) for asynchronous deletion

# Changes needed for AGIS

---

For the existing RSE's the Agis changes are minimal - using BNL as an example

↪ on the Agis Storage element page -

[https://atlas-agis.cern.ch/agis/storage\\_element/edit/](https://atlas-agis.cern.ch/agis/storage_element/edit/)

<GOCDB/OIM Site:> BNL-ATLAS    <AGIS SE name:> - BNL-ATLAS-Globus-dcdoor15

(where dcdoor15 is the Display name BNL Globus Endpoint)

<Endpoint:> - go://d8981638-6af7-11e6-83d0-22000b97daec

(the string after go:// is the UUID for the endpoint - ie Globus endpoint ID)

↪ <SE type:> - add to pull down menu Globus

↪ For a new Globus Endpoint that is not at a site in GOCDB/OIM like alcf#dtn\_theta, what is value for <GOCDB/OIM Site:>

# AGIS changes (2)

---

On the create new DDM protocol page (again using BNL-OSG2-DATADISK as an example)

<https://atlas-agis.cern.ch/agis/ddmprotocol/create/>

- ↪ No additional changes are needed.
  - ↪ <ddmendpoint:> - BNL-OSG2\_DATADISK
  - ↪ <SE:> BNL-ATLAS-Globus-dcdoor15 (already defined previously)
  - ↪ <Endpoint:> /pnfs/usatlas.bnl.gov/BNLT0D1/rucio/

# DDM point of view for AGIS changes

---

- ↪ We need to know if a DDMEndpoint is FTS-only, GO-only, or FTS/GO Dualstack.
- ↪ "go" should become a new protocol (like root, https, davs, ...)
  - ↪ it is a stand-in for whatever data transfer mechanism the Globus Online servers do in the background (unlike FTS, where we know e.g. they are using whatever specified protocol).
- ↪ The Globus Online ID is unique per SE, not per DDMEndpoint.
- ↪ We need DDMEndpoints also for the Globus Online only endpoints, otherwise cannot put rules on them.
- ↪ If there is a separate authentication mechanism (e.g., the myproxy for BNL), it would be good to have this additional information (e.g., type, hostname, etc..) in the SE as well, so we can automate the authentication.

# Proposed timeline

---

- ↪ Changes needed in AGIS
  - ↪ Do we have an agreed upon plan - who will implement it and how long will it take?
- ↪ Changes needed in PanDA/JEDI
  - ↪ Tadashi what needs to be done?
- ↪ Changes needed in Harvester
  - ↪ Do we use the rucio CLI or just the rucio python libraries ?
- ↪ Changes needed in Rucio
  - ↪ Implementation of the GlobusOnline *transfertool* component in Rucio December 2017
  - ↪ Adapting the *conveyor* to recognise Dual Stack and HPC/GO-only RSEs February 2018
  - ↪ Implementation of new *deletiontool* component April 2018
  - ↪ Adapting the conveyor to use new asynchronous *deletiontool* June 2018
- ↪ Operation
  - ↪ Configuration Rucio October 2017
  - ↪ First Rucio managed transfers March 2018
  - ↪ Full operation April 2018