Update on the Globus Transition

MORGRIDGE
INSTITUTE FOR RESEARCH
CORE COMPUTATION

iris hep

FEARLESS SCIENCE

Reminder: Where are we coming from?

In 2017, the Globus organization announced the end-of-support for the Globus Toolkit.

- Globus Toolkit provides the reference implementation of GridFTP and the Grid Security Infrastructure (GSI).
  - GridFTP -> bulk data movement; GSI -> authorization infrastructure.
  - Also was the only implementation of these technologies used within OSG!
- This triggered an immediate support issue (who will patch this software we use?) and a larger technology soul-searching (is this the right software for our community?).

From May 2017 – Globus Toolkit is being retired in January 2018

MAY 26, 2017 | VAS VASILIADIS

We are announcing that, starting in January 2018, the Globus team at the University of Chicago will no longer support the open source Globus Toolkit, except for its use with the Globus cloud service by Globus subscribers. By the end of 2018, all endpoints connected to the Globus cloud service using the open source Globus Toolkit GridFTP server must migrate to Globus Connect. At the end of 2018, we will discontinue all maintenance (including security patches) and distribution of the open source Globus Toolkit. Endpoints using Globus Connect Server or Globus Connect Personal will be unaffected, as long as they continue to perform routine software updates.

The OSG Transition Plan

Immediately OSG started to work on a transition plan:

- We started a fork of the Globus Toolkit, christened the "Grid Communtiy Toolkit".
- This allows us to help manage the upstream release process, including to EPEL and Debian.

And included work on technologies that can replace GridFTP and GSI:

- For GSI, we have been focusing on capability tokens – specifically, technologies derived from JSON Web Tokens (JWTs) and the SciTokens profile.
- For GridFTP, we have focused on developing the third-party-copy functionality in the WebDAV protocol into "HTTP-TPC".
  - Supporting the implementation in the XRootD software suite from SLAC.
- Several lesser-used components (MyProxy, GSISSH) are being replaced by an instance-by-instance basis.

This has integrated work across a wide variety of projects and contributors: OSG-N5Y, IRIS-HEP (OSG-LHC), SciTokens, FNAL SCD – and a variety of European-based projects.

OSG is providing international leadership and pushing the worldwide community forward.

Current Status

The Grid Community Toolkit is in a reasonable place:

- Releases are made, nothing is broken.
- The OSG Software team has sufficient internal knowledge for minor fixes; there's a modest international community contributing as well.
- We have subsequently also adopted the abandoned UberFTP under the same umbrella.

The transition-from-Globus aspect is starting to boom:

- Client credentials ("grid proxies") are being replaced by token-based systems.
- HTTP-TPC is starting to be tested at scale (in testing) and in production (small scale).
- We have a support contact CILogon to handle authentication needs for OSG services that currently authenticate with X509 (e.g., OASIS for software installs).

Recently we have been coordinating closely with FNAL – they have been ramping up activity in this area.

Highlights from the transition – authorization.

With the WLCG AAI Working Group, we helped host a "WLCG JWT Hackathon"

- **Goal**: Utilize the WLCG JWT token profile to do capability-based TPC.
- Integrated technologies like `IAM` (token issuer from INFN), `oidc-agent` (OAuth2 client from Indigo-DC), `scitokens-cpp` (SciTokens validation library), FTS, and Rucio.
  - Storage technologies included XRootD, StoRM, EOS, DPM, dCache.
- We were able to get all these pieces working together except Rucio by end of the week.

We have additionally been able to get the capability tokens working with HTCondor-CE.

WLCG JWT Hackathon

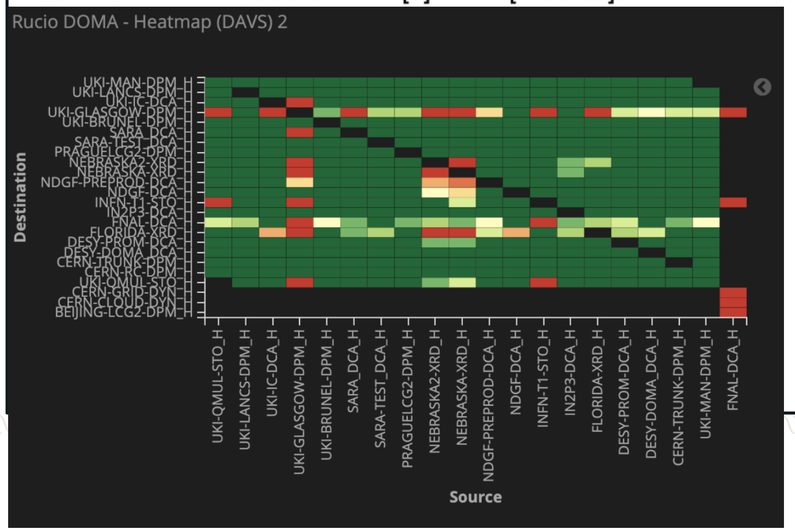
See GDB Update.

Highlights from the transition – HTTP-TPC for OSG-LHC

DOMA-TPC smoke test, started 2020-02-14T12:00+0100, took 60:26.

SOUND ENDPOINTS

SCORE	ENDPOINT	SOFTWARE		WORK-AROUNDS
20	AGLT2	dCache		[in 00:56]
20	BEIJING	DPM		[in 03:08]
20	BRUSSELS	dCache		[in 00:22]
20	CALTECH2	xrootd-R/HDFS	[L]	[in 02:14]
20	CALTECH	xrootd-D/HDFS		[in 01:48]
20	CERN-DYN-S3	DynaFed/S3		[in 00:44]
20	CERN-TRUNK	DPM		[in 01:16]
20	DESY-DOMA	dCache		[in 00:25]
20	DESY-PROM	dCache		[in 00:24]
20	FNAL	dCache		[in 00:51]
20	IN2P3	dCache		[in 00:22]
20	IN2P3-TEST	dCache		[in 00:28]
20	INFN-T1	StoRM		[in 00:49]
20	KIT	dCache	[2]	[in 00:30]



Our Xcache packaging exports over HTTP.

- SciTokens support is in progress.

IRIS-HEP's OSG-LHC has a goal of migrating 30% of the traffic at one U.S. LHC site over to non-GridFTP yet this spring.

- Many U.S. LHC sites are participating in the relevant WLCG working group – only a few missing (notably the T1s).
- WLCG HTTP-TPC scale tests are about 5% of global traffic.
- This month, we're starting to move production traffic with these new protocols.

Current Status

Date	Milestone or Deliverable	Completed
Aug 2019	Beginning of OSG 3.5 release series (last release series depending on GCT)	✅
Aug 2019	Including HTCondor 8.9.2 in the 'upcoming' repository (first HTCondor version with SciTokens support).	✅
Oct 2019	OSG no longer carries OSG-specific patches for the GCT. All patches are upstreamed or retired.	✅
Jan 2020	**DELAYED, expected completion Mar 2020:** "GSI free" site demo. Show, at proof-of-concept / prototype level, all components without use of GCT.	❌
July 2020	All GCT-free components are in OSG-Upcoming.	
Jan 2021	OSG series 3.6, without GCT dependencies, is released.	
Jan 2022	End of support for OSG 3.5.	

Early milestones have gone well.

Current deliverable is to show a prototype "of everything"

- Single missing piece: GlideinWMS that can submit to CEs using SciTokens.
- Other GlideinWMS pieces are going well.

Migration Status Page

Ongoing Risks

Within OSG, things are looking good.

We are expecting the GlideinWMS release to occur in May (prototype in March).

- This likely pushes the rest of our deadlines by at least 2 months.
- The relevant deadline is to get OSG 3.6 out-the-door before Run3 starts: this is on track.
- **Mitigation**: More actively attending the GlideinWMS development meetings; coordinating with FNAL management and the FNAL Federated ID project.

We run the risk of "running ahead" of international sites:

- This would create a "split system": Tokens for US pilots, Proxies otherwise. An operational cost for our stakeholders!
- Key technology to watch: ARC-CE. Unclear what their long-term plans are!
- **Mitigation**: Frank is raising the issue with the WLCG Management Board.

We run the risk of running ahead of the LHC VOs:

- Their jobs will need to have tokens; they will need to push sites to upgrade to new transfer protocols.
- Key technology to watch: For CMS, WMAgent. For ATLAS, PanDA & Rucio.
- **Mitigation**: OSG Software is making endpoints available to these groups.

MORGRIDGE
INSTITUTE FOR RESEARCH
CORE COMPUTATION

morgridge.org

This material is based upon work supported by the National Science Foundation under Grant No. 1836650. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

FEARLESS SCIENCE