

XSEDE use cases affected by the Globus Toolkit end-of-support notice: Service Providers

On May 26, 2017 the Globus team at the University of Chicago announced, “Support for open source Globus Toolkit will end as of January 2018; The Globus cloud service and Globus Connect are unaffected.”¹

XSEDE staff and research users currently use Globus Toolkit software components to accomplish roughly 40% of the activities described in XSEDE’s use case registry.² This document reviews specific use cases that XSEDE’s *service providers (SPs)* support that currently use Globus Toolkit software components and summarizes how the end-of-support announcement affects these programs. In addition to detailing specific changes XSEDE will need to make, we hope this will also serve as a basis for identifying additional use cases or activities that might be affected.

Interactive computing	1
Data services	2
Automated computing	3

Interactive computing

The most common way researchers interact with XSEDE SP resources is via a remote login shell. The researcher uses a terminal client on his/her local computer, connects to an XSEDE login server, and uses a Linux command-line interface. XSEDE currently offers two mechanisms to support this scenario: (1) the XSEDE Single Sign-on Hub (SSO Hub), which allows researchers to connect to the SSO Hub and authenticate using their XSEDE username and password and then access a login shell on the specific XSEDE resource they need to use; and (2) GSI-OpenSSH servers on specific XSEDE resources that allow researchers to connect directly to the resource they need to use and authenticate via an X.509 certificate. The former allows researchers to use standard SSH clients, and the latter requires both an X.509-enabled SSH client (currently obtained from the Globus Toolkit), and a tool for obtaining an X.509 certificate (MyProxy, also from the Globus Toolkit). The SSO Hub

¹ https://github.com/globus/globus-toolkit/blob/globus_6_branch/support-changes.md

² <https://software.xsede.org/xcsr/xsede-use-cases>

supports XSEDE two-factor authentication, and individual SPs may configure their GSI-OpenSSH servers to support two-factor authentication.

Use case **CAN-04**³ describes XSEDE's *enabling function* for interactive login. The specific tasks researchers perform once logged into a system are described in a large set of additional use cases as follows. Use cases **DA-02**, **-03**, **-04**, and **-05**⁴ describe the various methods used to analyze data on HPC and HTC systems, including the steps that are often required to prepare data for analysis. Use cases **HPC-01**, **HPC-02**⁵, and **HTC-01**⁶ describe how researchers submit HPC and HTC computations for processing via the command line. Use cases **VIZ-01**, **-02**, **-03**, **-04**, and **-05**⁷ describe the most common methods of data visualization performed on HPC systems.

To continue supporting XSEDE's interactive computing use cases without ongoing support for the Globus Toolkit, SPs need replacements for Globus Toolkit software in the following specific areas.

- We need a replacement for the **GSI-OpenSSH server** that allows access to SP login services via the XSEDE SSO Hub.
- Depending on what the replacement for the GSI-OpenSSH server is, we may need changes to the **AMIE reference implementation** that shares XSEDE X.509 subject names with SP login server administrators.
- We need to replace the **MyProxy and GSI-OpenSSH-clients** that are currently installed on SP systems to enable remote access to XSEDE login services.
- We need to replace the **Globus Toolkit clients** that are currently installed on SP systems to access XSEDE services (e.g., GridFTP data transfer) with any clients that are appropriate for newer XSEDE services (e.g., Globus Python CLI if necessary).

Data services

Access to data--and management of it after its initial production and use--is vital to research computing. While data movement from system-to-system has always been important, it has become increasingly so as HPC and HTC systems are rapidly replaced and as the research community has become more collaborative and more aware of its dual needs for transparency within the community and retrospective data archiving.

Use case **CAN-02**⁸ describes XSEDE's enabling function for managed data transfer: the ability to easily and reliably move research data from one location to another. Use case **CAN-03**⁹ describes XSEDE's enabling function for remote data access: the ability to access--and possibly change--data that's physically stored on one system from another system during processing. This remote data

³ <http://hdl.handle.net/2142/50328>

⁴ <http://hdl.handle.net/2142/45702>

⁵ <http://hdl.handle.net/2142/45703>

⁶ <http://hdl.handle.net/2142/47552>

⁷ <http://hdl.handle.net/2142/50017>

⁸ <http://hdl.handle.net/2142/50329>

⁹ <http://hdl.handle.net/2142/50327>

access function does not currently use Globus Toolkit software. The data transfer function, on the other hand, currently makes heavy use of software from the Globus Toolkit. GSI-OpenSSH servers that support X.509 authentication are used for scp and sftp-style transfers when transfers involve a small number of small files. For larger data transfers (many files or huge data volumes), XSEDE is currently transitioning from Globus Toolkit's GridFTP server software and client tools (e.g., globus-url-copy) to Globus's cloud services and Globus Connect software.

The specific scenarios in which these two basic functions are used in conjunction with SP resources are described in a series of use cases as follows. Use case **DM-01**¹⁰ describes a community of researchers sharing a common data repository for collaborative work. Use cases **DM-02** and **DM-03** explain how researchers manage the data needed for--and created by--distributed simulation and analysis activities and large-scale sensor networks. Use case **DM-04** describes the need to migrate data when one HPC or HTC system is replaced by another. Use case **SGW-02** describes how researchers who build and use science gateways (aka web applications) need to build data movement into their applications and how those functions interact with SP resources.

To continue supporting XSEDE's data transfer use cases without ongoing support for the Globus Toolkit, SPs need replacements for Globus Toolkit software in the following specific areas.

- XSEDE service providers are replacing their **GridFTP server** software from the Globus Toolkit with equivalent Globus Connect Server software. Some customizations will need to be carried over from the older to the newer distributions.
- XSEDE user documentation that describes using **GridFTP services and client tools** directly should be updated to describe using Globus cloud services instead, and researchers who currently use **globus-url-copy** should transition to using the Globus cloud CLI.
- We need a replacement for the **GSI-OpenSSH server** that allows SCP-style access to XSEDE systems.

Automated computing

The section above on interactive computing described use cases where researchers use XSEDE's computation resources directly. As research computing becomes more sophisticated and problem complexity grows, it becomes more common for computing tasks to be automated: controlled by software agents rather than by researchers. There are two common scenarios in which this happens. First, a single researcher or research team might develop a high-level plan for a set of tasks and use a *workflow engine* to manage their execution. Second, a research community with its own computing system might form a partnership with XSEDE or a specific XSEDE SP and arrange to have some portion of its computing tasks "outsourced" to one or more XSEDE resources. In the latter scenario, there may or may not be a single "plan" for the tasks, but the mechanism by which tasks are executed on XSEDE resources on behalf of the other system is automated.

¹⁰ <http://hdl.handle.net/2142/48909>

Use case **CAN-01**¹¹ describes XSEDE's enabling function to allow computing tasks to be submitted and managed by a software agent. There are three common mechanisms currently in use within XSEDE to accomplish this function.

- The first mechanism is the one described in XSEDE's architectural documentation, and it involves use of UNICORE services running on SP resources to provide a remote job management interface. Within this mechanism, authentication is performed via X.509 certificates, which are obtained via XSEDE's MyProxy service, obtained via the Globus Toolkit.
- The second mechanism is a historical legacy still in use by some applications: the GRAM service from the Globus Toolkit. Use of GRAM has not been encouraged by XSEDE for many years, but it is still available on a handful of SP resources.
- The third mechanism is the most commonly used in practice, and it works via the GSI-OpenSSH services running on XSEDE SP interactive login servers. In essence, the remote software agent simulates a human using the command-line interface directly, using an ssh connection (authenticated using an X.509 certificate) as the communication interface. X.509 certificates are obtained via XSEDE's MyProxy service, as in the UNICORE case above.

Specific use cases for the first research scenario above (a single team using a workflow engine) include: **HPC-03**¹² (high-performance computing using a workflow engine), **HTC-02, -03, and -04**¹³ (high-throughput computing of various types), and **SW-01**¹⁴ (a generic workflow use case). Specific use cases for the second research scenario above (outsourcing a portion of computing tasks to one or more SP resources) include: **FI-01**¹⁵ (a generic "outsourcing" use case), **SGW-03**¹⁶ (using SP computing services to support a science gateway), and **VIZ-06**¹⁷ (using SP computing services to support a scientific visualization gateway).

To continue supporting XSEDE's automated computing use cases without ongoing support for the Globus Toolkit, SPs need replacements for Globus Toolkit software in the following specific areas.

- We need a replacement for the **GSI-OpenSSH server and clients** that allows remote submission and management of tasks on XSEDE compute resources.
- Depending on what the replacement for the GSI-OpenSSH server is, we may need changes to the **AMIE reference implementation** that shares XSEDE X.509 subject names with campus login server administrators.
- Anyone currently using **GRAM** for remote task execution on XSEDE resources should transition to one of the other available mechanisms.

¹¹ <http://hdl.handle.net/2142/50326>

¹² <http://hdl.handle.net/2142/45703>

¹³ <http://hdl.handle.net/2142/47552>

¹⁴ <http://hdl.handle.net/2142/50016>

¹⁵ <http://hdl.handle.net/2142/43878>

¹⁶ <http://hdl.handle.net/2142/43883>

¹⁷ <http://hdl.handle.net/2142/50017>

- XSEDE SPs who are using **UNICORE** for remote task execution either need a replacement authentication mechanism that doesn't rely on the Globus Toolkit for X.509 support, or need to transition to another remote task execution mechanism.

