XSEDE Campus Bridging Use Cases

November 3, 2016

Version 2.0

These use cases describe how campus information technology (IT) administrators and campus-based research projects: (a) treat XSEDE's resources as extended branches of their campus environments, and (b) offer their campus's services to others in the XSEDE community.

Use cases 1-7 and 9-11 were first documented during XSEDE's first five-year project period. (Use cases 9-11 were originally variations of use case 4.) Considerably more detail was provided in the original descriptions.[1]

[1] Stewart, Knepper, Foster, Morris, Bachmann, Grimshaw, Lifka. *XSEDE Campus Bridging Use Cases, version 1.5.* March 16, 2012. (http://hdl.handle.net/2142/43882)

CB-1: InCommon-based authentication to XSEDE resources

A **campus IT administrator** would like researchers at his/her campus to be able to login to XSEDE resources using InCommon-based authentication mechanisms. [1][2][3][4]

In most cases, the **campus IT administrator** would like to experience it as follows.

1. The administrator operates an InCommon IDP for his/her campus.
2. The administrator directs his/her campus researchers to login to XSEDE resources (e.g., XSEDE User Portal, service provider remote login servers) using their campus identity.
3. Any researcher at the administrator's campus can select their campus's InCommon IDP while logging in to an XSEDE resource, authenticate to the IDP, and then (if they are authorized) access the resource without further authentication.

It's always like the steps above, except when the user is logging in to a Science Gateway, in which case jobs are "translated" to ownership by a group account used for jobs launched on XSEDE resources by that Science Gateway (e.g. JimmyNeutrino or MRLEAD) that runs on XSEDE resources.

We'll accept any solution to this problem, as long as the following are true.

1. Once authentication is completed at the InCommon IDP, access to XSEDE resources doesn't take more than an additional 5 seconds.
2. Access via InCommon must be supported by Level 1 and 2 resources.
3. Access via InCommon must be available for XSEDE-operated services (XSEDE User Portal, XSEDE Resource Allocation Service, etc.).
4. The InCommon authentication mechanism must be available for use by Level 3 and campus IT administrators to enable access to their own resources via InCommon.

CB-2: Share the XSEDE "environment" with campus resources

A **campus IT administrator** needs to reuse XSEDE documentation and tools in his/her campus environment to reduce local costs and to smooth the path to XSEDE for campus researchers. We assume XSEDE has built a consensus around its own "common environment" expectations.

In most cases, the **campus IT administrator** would like to experience it as follows.

1. The administrator finds an editable documentation template on XSEDE's website that he/she can reuse to document his/her campus systems.
2. The administrator finds training materials on XSEDE's website that he/she can use with minimal alterations for campus researchers.
3. The administrator finds software packages on XSEDE's website that he/she can use on campus resources to provide the open source elements of a basic XSEDE system.

We'll accept any solution to this problem, as long as the following are true.

1. Documentation and training materials provided by XSEDE must be released with a license that allows reuse and modification, such as the CC BY 3.0 license. [5]
2. Documentation and training materials should be provided in editable, commonly used formats.

3. Software provided by XSEDE must be available for use under a free-use license.
4. The XSEDE "common environment" includes the following items, which we assume are already standard across Level 1 and Level 2 XSEDE SP resources: directory hierarchy, locations of standard software 'kits' and optional locally-installed or user-contributed software.

CB-3: Remote desktop services for researchers

A **researcher** on a college or university campus needs to open and maintain a remote desktop session on an XSEDE resource for a period of time potentially measured in several days.

In most cases, the **researcher** would like to experience it as follows.

1. The researcher's campus IT administrator installs any required client software on a campus computer.
2. The researcher uses the campus computer and initiates a remote desktop session.
3. The session stays active as long as the researcher specifies. (This could be multiple days.)
4. The researcher may temporarily disconnect from the desktop session and reconnect to it later. The state of the remote desktop should be retained across sessions.

We'll accept any solution to this problem, as long as the client software required on the campus computer is easy to obtain (free or low-cost) and easy to install (pre-installed on most systems or very simple installation). Ideally, the client software would be available for use on personal systems owned by the researcher as well as on campus systems.

CB-4: Access campus research data from XSEDE resources

An **XSEDE-allocated researcher** needs access to data stored on a campus resource when using an XSEDE Level 1 or 2 resource for analysis and/or visualization. [1][6][7][8][9]

In most cases, the **researcher** would like to experience it as follows.

1. The researcher logs in to an XSEDE resource to use its analysis/visualization capabilities.
2. The researcher accesses his/her data on a campus resource from the XSEDE resource.
3. The researcher analyzes and/or visualizes data on the XSEDE resource.
4. The researcher writes/updates/deletes data back to the campus resource.

We will accept any solution to this problem so long as the following are true.

1. In the event of a transient failure (e.g., network glitch, client or server fault, expired credentials), the system should be able to restart transfers and notify the user once the transfer has completed successfully.
2. A graphical user interface should be available to initiate and manage file transfers.
3. It shouldn't take longer than one working day for a researcher with proper permissions to install the necessary software on their campus system.
4. The mechanism should allow access to flat files, file archives, and database files.

CB-5: Workflow automation combining XSEDE and campus resources

An **XSEDE-allocated researcher** needs to perform an analysis via a distributed workflow in which tasks automatically execute on and access data from various XSEDE and campus resources. [10] We assume that XSEDE has selected, documented, tested, deployed, and configured at least one distributed workflow service.

In most cases, the **researcher** would like to experience it as follows.

1. The researcher opens an interactive session with the workflow system, specifies the workflow to be executed, and initiates the workflow.
2. The workflow system carries out the tasks specified by the workflow.
3. The workflow system notifies the researcher when the workflow has completed.

It'll always be like that, except in the following cases.

1. The researcher may prefer or need to use a batch (non-interactive) interface to the workflow system, in which case the user must submit the workflow specification and delegated user credentials (to be used by the workflow system when submitting tasks or accessing data on other systems) to the system via a batch interface. The user can check the status of the workflow (queued, executing, completed). The system will notify the researcher when the workflow has completed.
2. The workflow to be executed is submitted on the researcher's behalf by a science gateway, in which case the workflow system must employ a delegation mechanism that allows the researcher to delegate credentials to the gateway for use on XSEDE and campus resources. The gateway will notify the researcher when the workflow has completed.

We'll accept any solution to this problem so long as the following are true.

1. The workflow system must be supported by XSEDE Level 1 resources.
2. The workflow system must be available for Level 2 and Level 3 resource SPs to install and configure on their resources.
3. The workflow system must be available for campus IT administrators to install and configure on their campus systems.
4. The workflow system must support tasks that execute on Level 1 resources or campus resources, consistent with the researcher's authorization(s).
5. The workflow system must support tasks that access data on Level 1 resources or campus resources--via a stage in/out model or via a CRUD (create, read, update, delete) model--consistent with the researcher's authorization(s).
6. After the workflow is initiated, the workflow system should not require further user interaction to complete the workflow. Transient issues (e.g., system downtime, network glitches, expired credentials, system misconfiguration) should be handled automatically by the workflow system or by support personnel.

CB-6: Sharing computational facilities among campuses

A **researcher** needs to share a computational facility on his/her campus with other researchers--possibly at other institutions--using elements of the XSEDE system to mediate the sharing and establish a familiar environment. [1][6][7][8] We assume that use cases CB-1 and CB-2 are supported by XSEDE.

In most cases, the **researcher** would like to experience it as follows.

1. The researcher builds a cluster, Condor flock, cloud, or other sort of computational resource using campus computing resources.
2. The researcher installs and configures the InCommon authentication mechanism from CB-1 on his/her resource.
3. The researcher uses the documentation, training, and/or software packages from CB-2 to make his/her resource similar to other XSEDE resources.
4. The researcher manages his/her own accounting, "value exchange," policy compliance, and security responses.
5. The researcher defines user groups and sets access control policies for his/her resource based on these groups.

It'll always be like this except when the researcher elects to become an XSEDE Level 3 service provider (SP).[2] In this case, XSEDE would be able to manage exchange rates among the campus-contributed resources, Level 1 resources, and Level 2 resources, and would be able to return service units to campus contributors in accordance with a negotiated Service Level Agreement. In this case, XSEDE will also provide security notifications in the event of a security breach related to accounts or services that use campus-based authentication mechanisms, and the researcher will be expected to participate in XSEDE's integrated help desk function.

We'll accept any solution to this problem as long as the InCommon authentication mechanism from CB-1 does not require users of the researcher's computational facility to have XSEDE allocations or participate in the allocation process.

CB-7: Support for commercial service providers

An **independent resource provider** wants to offer a privately operated computational facility for use by researchers at colleges and universities in a manner consistent with the expectations of users of XSEDE resources in exchange for a fee or other form of payment. [1] We assume that use cases CB-1 and CB-2 are supported by XSEDE.

In most cases, the **resource provider** wants to experience it as follows.

1. The resource provider builds a cluster, Condor flock, cloud, or other sort of computational resource using privately owned and operated computing resources.
2. The resource provider installs and configures the InCommon authentication mechanism from CB-1 on his/her resource.

[2] XD Service Providers Forum: Charter, Membership, and Governance (v10.1_120228). (https://www.xsede.org/documents/10157/281380/SPF_Definition_v10.1_120228.pdf)

3. The researcher uses the documentation, training, and/or software packages from CB-2 to make his/her resource similar to other XSEDE resources.
4. The resource provider manages his/her own accounting, payment mechanism, and security responses.
5. The resource provider advertises the mechanism by which to contract for time (credit card or Purchase Order).
6. A **researcher** sets up an account with the resource provider.
7. The researcher uses the private resource, in ways that are convenient because they leverage system similarity with XSEDE resources and user familiarity with same, and the company receives payment from the researcher.

We'll accept any solution to this problem as long as the InCommon authentication mechanism from CB-1 does not require users of the researcher's computational facility to have XSEDE allocations or participate in the allocation process and the resources from CB-2 are available under licenses that permit commercial use.

CB-8: Use XSEDE SSO with campus login servers

A **campus IT administrator** wants to allow XSEDE-registered researchers to login to campus login servers (remote command shell) using their XSEDE usernames/passwords.

In most cases, the **campus IT administrator** wants to experience it as follows.

1. First, the administrator registers his/her campus login service(s) with XSEDE. The result should be that the XSEDE login service becomes aware of the campus login service(s) and allows end users to select it/them as login targets.
2. Then, the administrator configures his/her login services to authorize specific XSEDE users to use the GSISSH service.
3. Then, the administrator notifies users that they can use the XSEDE login service to access the campus service(s).
4. Finally, an end user authorized by the campus administrator uses the XSEDE login service, selects a campus service for login, and is given a login shell (under his/her local ID) on the campus service.

It'll always be like that, except when the end user doesn't have an XSEDE identity, in which case the end user will need to register with XSEDE before Step 2 can be completed.

We'll take any solution, as long as...

1. The solution should not require the campus services to participate in XSEDE allocation processes.
2. The solution should not require the researcher to have an active or prior XSEDE allocation.
3. The administrator must retain control over who can/cannot login to his/her service.
4. The solution should use the standard XSEDE XUP registration mechanism for end users to sign up with XSEDE.
5. The solution should use the standard XSEDE login service interface in Step 4.

CB-9: Access an XSEDE-hosted research data collection from campus

An **XSEDE-allocated researcher** needs access to data in an XSEDE-hosted data collection when using a campus resource for analysis and/or visualization. [1][6][7][8][9]

In most cases, the **researcher** would like to experience it as follows.

1. The researcher searches or browses a directory of XSEDE-hosted data collections to find the one he/she will be using.
2. The researcher specifies the data to be moved from the XSEDE data collection to a campus resource.
3. The system moves the data and lets the researcher know when it's ready.

We'll accept any solution to this problem so long as the following are true.

1. In the event of a transient failure (e.g., network glitch, client or server fault, expired credentials), the system should be able to restart transfers and notify the user once the transfer has completed successfully.
2. A graphical user interface should be available to initiate and manage file transfers.
3. It shouldn't take longer than one working day for a researcher with proper permissions to install the necessary software on their campus system.
4. The mechanism should allow access to flat files, file archives, and database files.

CB-10: Synchronize research data between campus and XSEDE resources

An **XSEDE-allocated researcher** needs to keep data synchronized when copies are kept on a campus resource and XSEDE Level 1 or 2 resources. [1][6][7][8][9]

In most cases, the **researcher** would like to experience it as follows.

1. The researcher specifies a data set that s/he wishes to maintain, in a synchronized fashion, on a campus resource and one or more XSEDE resources.
2. When the researcher makes a change to one copy of the data, the other copies are automatically updated.

We'll accept any solution to this problem so long as the following are true.

1. When data needs to be moved between systems, in the event of a transient failure (e.g., network glitch, client or server fault, expired credentials), the system should be able to restart transfers and notify the user once the transfer has completed successfully.
2. A graphical user interface should be available to specify the data set to be synchronized.
3. It shouldn't take longer than one working day for a researcher with proper permissions to install the necessary software on their campus system.
4. The mechanism should allow synchronization of flat files, file archives, and database files.

CB-11: Archive research data on an XSEDE resource

An **XSEDE-allocated researcher** needs to create an archival copy of research data on an XSEDE Level 1 or 2 resource. [1][6][7][8][9] We assume that at least one data archive service is provided by one or more XSEDE SPs.

In most cases, the **researcher** would like to experience it as follows.

1. The researcher authenticates with an archive service or resource.
2. The researcher identifies the data set (stored on a campus system or an XSEDE L1 or L2 system) and provides appropriate metadata describing the data set.
3. The researcher specifies a license governing use of the data.
4. The system makes an archival copy of the data set on the archive resource and adds the metadata and license information so that the data appears in appropriate search results and is accessible per the license.

We'll accept any solution to this problem so long as the following are true.

1. In the event of a transient failure (e.g., network glitch, client or server fault, expired credentials), the system should be able to restart transfers and notify the user once the transfer has completed successfully.
2. A graphical user interface should be available to initiate and manage file transfers.
3. It shouldn't take longer than one working day for a researcher with proper permissions to install the necessary software on their campus system.
4. The mechanism should support flat files, file archives, and database files.

References

[1] NSF Advisory Committee for Cyberinfrastructure Task Force on Campus Bridging. Final Report. March 2011. http://www.nsf.gov/od/oci/taskforces/TaskForceReport_CampusBridging.pdf. Available as print-on-demand book from: https://www.createspace.com/3597300

[2] Barnett, W., V. Welch, A. Walsh and C.A. Stewart. A Roadmap for Using NSF Cyberinfrastructure with InCommon. 2011. http://hdl.handle.net/2022/13024 or http://www.incommon.org/nsfroadmap.html. Available as print-on-demand book from https://www.createspace.com/3630011

[3] Barnett, W., V. Welch, A. Walsh and C.A. Stewart. A Roadmap for Using NSF Cyberinfrastructure with InCommon: Abbreviated Version. 2011. http://hdl.handle.net/2022/13025 or http://www.incommon.org/nsfroadmap.html

[4] Jim Basney, Terry Fleury, and Von Welch, "Federated Login to TeraGrid," 9th Symposium on Identity and Trust on the Internet (IDtrust 2010), Gaithersburg, MD, April 2010. http://dx.doi.org/10.1145/1750389.1750391

[5] http://creativecommons.org/licenses/by/3.0/ Attribution 3.0 Unported license (CC BY 3.0)

[6] Almes, G.T.; Jent, D.; Stewart, C.A. 2011. Campus Bridging: Data and Networking Issues Workshop Report. http://hdl.handle.net/2022/13200. Available as print-on-demand book from: https://www.createspace.com/3592681

[7] Dreher, P., S.C. Ahalt, G. Almes, M. Mundrane, J. Pepin and C.A. Stewart, (eds.), 2011. Campus Bridging: Campus Leadership Engagement in Building a Coherent Campus Cyberinfrastructure Workshop Report. 2011. http://hdl.handle.net/2022/13194

[8] McGee, J.; V. Welch; G.T. Almes. 2011. Campus Bridging: Software & Software Service Issues Workshop Report. http://hdl.handle.net/2022/13070

[9] Open Data Commons. ODC Public Domain Dedication and License (PDDL). http://www.opendatacommons.org/licenses/pddl/1-0/

[10] Scott Michael, Stephen Simms, W. B. Breckenridge, III, Roger Smith, and Matthew Link. 2010. A compelling case for a centralized filesystem on the TeraGrid: enhancing an astrophysical workflow with the data capacitor WAN as a test case. In Proceedings of the 2010 TeraGrid Conference (TG '10). ACM, New York, NY, USA, , Article 13 , 7 pages. DOI=http://dx.doi.org/10.1145/1838574.1838587

History

	Version	Date	Changes	Author
Entire Document	1.4	3/09/2012	Formatted	Stewart, Knepper, Grimshaw, et al.
Entire Document	1.5	3/14/2012	Hopefully penultimate version before public release	Stewart
Entire document	2.0	11/3/2016	Use cases 1-7 rewritten using XSEDE-2 format; added use case 8; new use cases 9-11 extracted from use case 4	Liming