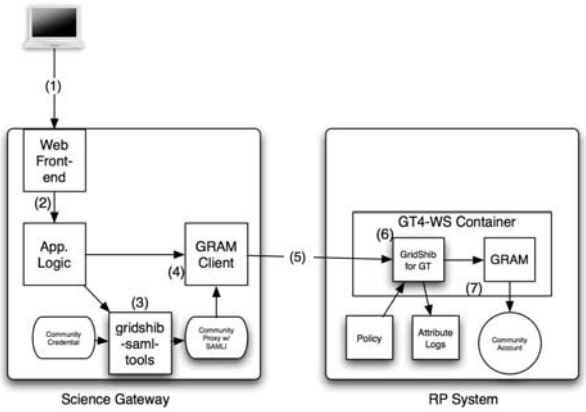
TeraGrid Science Gateways: Scaling TeraGrid Access

Aaron Shelmire[1], Jim Basney[2], Jim Marsteller[1], Von Welch[2], Tom Scavo[2], Terry Fleury[2], and Nancy Wilkins-Diehr[3]

[1]Pittsburgh Supercomputing Center, [2]National Center for Supercomputing Applications, and [3]San Diego Supercomputer Center

The TeraGrid Science Gateways program provides community-scale access to TeraGrid resources. Rather than requiring individual community members to apply for access to the TeraGrid, TeraGrid allows gateways to apply for resources on behalf of an entire community, enabling TeraGrid to scale to larger numbers of users than its current accounting mechanisms can handle. The gateway provides a community-tailored interface for members to register, submit jobs, and access data collections, transparently using TeraGrid resources. By opening up access to new communities of users who have not registered directly with TeraGrid, science gateways create new challenges for accounting and risk management. TeraGrid resource providers need the ability to gather user statistics as input to planning for future developments and for success metrics, both for use internally and for reporting to funding agencies. TeraGrid resource providers also need to minimize exposure to compromise and respond to security incidents. We present new accounting, authorization, and access control mechanisms that TeraGrid has developed to meet these needs.



TeraGrid will be using attribute assertions encoded in the Security Assertion Markup Language (SAML) to convey information about individual gateway users to TeraGrid resource providers to support accounting and authorization. Each gateway obtains a community credential issued by a TeraGrid certificate authority and maps that community credential to the gateway's community accounts on TeraGrid resources. When a community member accesses the gateway, the gateway creates a proxy credential for that member from its community credential and embeds a SAML assertion in the proxy credential that contains the member's username, IP address, email address, and a timestamp indicating when the member authenticated to the gateway. Then, when the community member accesses TeraGrid resources via the gateway, the gateway authenticates to the TeraGrid services using the proxy credential it created for that member. TeraGrid services validate the proxy credential, using the certificate distinguished name to determine the target community account and using the SAML assertion to obtain information about the specific community user behind the request. The TeraGrid services record this information for accounting purposes and also use it for authorization decisions. In particular, TeraGrid resource providers desire the ability to block individual community users in case of unwanted user behavior without needing to block the entire community, especially during off-hours when gateway administrators may not be immediately available to investigate. The software for inserting SAML assertions into proxy credentials and for logging and authorization on SAML attributes is provided by the GridShib project.

Some TeraGrid resource providers are also using the *community shell* to provide controlled system access to science gateway users. TeraGrid expects science gateways to facilitate appropriate use of resources by the community. Gateways must appropriately identify community users and provide capabilities targeted to the community via a suite of community-specific scientific applications and data sets. In contrast with individual TeraGrid users, who have direct access to run their own applications on TeraGrid systems, gateway users may run only those applications configured by the gateway developers and administrators. Together with mechanisms implemented by the gateway, TeraGrid resource providers can deploy the community shell to set policies as to what applications and files gateway users can access on TeraGrid systems, limiting possible targets for exploit. The community shell is integrated with job submission services (i.e., Globus GRAM) to enforce the RP policies for the community account.

Building on our experiences to-date, the process and implementation details surrounding science gateways and community accounts are evolving as more resource providers enable this model and more gateways are brought online. Attribute-based accounting and authorization provided by GridShib software, together with access control policies at the application and data access level enforced by the community shell, address new security challenges created by the science gateway model. Science gateways hold great promise for streamlining access

to TeraGrid computational resources by scientific users, allowing researchers to focus on science rather than programming and software details.

Bibliography

1. Kevin Price, "Restricted Community Accounts: Securing Science Gateways at the Account Level," TeraGrid 2006.
2. Von Welch, Ian Foster, Tom Scavo, Frank Siebenlist, Charlie Catlett, Jill Gemmill, and Dane Skow, "Scaling TeraGrid Access: A Testbed for Identity Management and Attribute-based Authorization," TeraGrid 2007.
3. Tom Scavo and Von Welch, "A Grid Authorization Model for Science Gateways," To appear in Concurrency and Computation: Practice and Experience, 2008.
4. Von Welch, Jim Barlow, Jim Basney, Doru Marcusiu, and Nancy Wilkins-Diehr, "A AAAA model to support science gateways with community accounts," Concurrency and Computation: Practice and Experience, Volume 19, Issue 6, March 2007.
5. GridShib Project, http://gridshib.globus.org/.
6. Community Shell Project, http://security.ncsa.uiuc.edu/research/commaccts/.