Enhancing the Earth System Grid Security Infrastructure through Single Sign-On and Autoprovisioning

F. Siebenlist
Argonne National Laboratory
Argonne, IL, USA
franks@mcs.anl.gov

R. Ananthakrishnan
Argonne National Laboratory
Argonne, IL, USA

D. E. Bernholdt
Oak Ridge National Laboratory
Oak Ridge, TN, USA

L. Cinquini
National Center for Atmospheric Research, Boulder, CO, USA

I. T. Foster
Argonne National Laboratory
Argonne, IL, USA

D. E. Middleton
National Center for Atmospheric Research, Boulder, CO, USA

N. Miller
Argonne National Laboratory
Argonne, IL, USA

D. N. Williams
Lawrence Livermore National Laboratory, Livermore, CA, USA

ABSTRACT
In this paper, we discuss the recent ESG's development and implementation efforts concerning its authentication infrastructure. ESG's requirements are to make the user's logon-experience as easy as possible, and to facilitate the integration of the security services and the Grid components for both the developers and system administrators. To meet that goal, we leverage existing primary authentication mechanisms, deploy a "light-weight" but secure OpenID WebSSO, deploy a "light-weight" X.509-PKI, and use autoprovisioning to ease the burden of security configuration management. We're close to finalizing the associated development and deployment.

Categories and Subject Descriptors
K.6.5 [**Management of Computing and Information Systems**]: Security and Protection --- *Authentication*

General Terms
Management, Security, Human Factors.

Keywords
Authentication, Authorization, OpenID, PKI, SLCS, SAML, ESG.

1. INTRODUCTION
Climate scientists face a wide variety of practical problems, but there exists an overarching need to efficiently access and manipulate climate model data. Increasingly, for example, researchers must assemble and analyze large datasets that are archived in different formats on disparate platforms, and must extract portions of datasets to compute statistical or diagnostic metrics in place. The need for a common virtual environment in which to access both climate model datasets and analysis tools is therefore keenly felt. The software infrastructure to support such an environment therefore not only must provide ready access to climate data but also must facilitate the use of visualization software, diagnostic algorithms, and related resources. To this end, the Earth System Grid [3] Center for Enabling Technologies (ESG-CET) was established in 2006 by the Scientific Discovery through Advanced Computing program of the U.S. Department of Energy through the Office of Advanced Scientific Computing Research and the Office Biological and Environmental Research within the Office of Science. ESG-CET is working to advance climate science by developing computational resources for accessing and managing model data that are physically located in distributed multiplatform archives[1].

1.1 Earth System Grid Scale-up.
In coming years, the ESG-CET will scale up existing capabilities to meet the needs of the following scientific projects:

- The North American Regional Climate Change Assessment Program (NARCCAP) will disseminate high-resolution regional climate model data through ESG portals located at both the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at Lawrence Livermore National Laboratory (LLNL) and the National Center for Atmospheric Research (NCAR).
- The Computational Climate End Station (CCES) at the DOE leadership computing facility at Oak Ridge National Laboratory (ORNL) will advance climate science through both an aggressive model development activity and an extensive suite of climate simulations.
- Phase 5 of the Coupled Model Intercomparison Project (CMIP5) will support the challenging climate data needs of

the planned Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC).

These projects, and especially CMIP5, will drive future development of ESG technologies in order to connect a large number of users with geographically distributed climate model archives and to provide them with advanced data analysis tools. Together with its institutional collaborators, ESG-CET will extend its present capabilities in order to supply additional types of climate model data and metadata, to provide more powerful server-side access and analysis services, to enhance interoperability among commonly used climate analysis tools, and to enable end-to-end simulation and analysis workflow.

1.2 National and International Collaborations.

Future ESG-CET activities will be framed by relationships with other institutions that share common data-management interests, organized as the Global Organization for Earth System Science Portal (GO-ESSP) consortium. GO-ESSP will develop a common software infrastructure for acquisition and analysis of climate model data. ESG-CET consortium members that will take leading roles as gateways and/or data nodes in the CMIP5 (IPCC AR5) testbed include PCMDI, NCAR, ORNL, and LANL. Other members that will play a vital role in the CMIP5 effort include the Geophysical Fluid Dynamics Laboratory (GFDL), the British Atmospheric Data Centre (BADC), the World Data Center for Climate (WDCC), and the University of Tokyo Center for Climate System Research. Because GO-ESSP extends beyond U.S.-based partnerships, it will need to develop software to accommodate components from the U.K. Natural Environment Research Council (NERC) DataGrid (NDG), the European Union (EU) MetaFor project, and the German C3-Grid initiative.

1.3 Future Usage.

Under the current ESG system, a user first accesses and queries a remote database by means of a Web browser and then retrieves desired data records via the ESG data portal, a Data Mover Light (DML) tool [2], or a Web "get" scripting-operation (wget/curl). After downloading these records to the local site, the user usually regrids, reduces, and further analyzes the data. This process often requires many data movements that can overtax network, storage, and computing resources. With the next-generation ESG architecture, the user instead will browse, search, and discover (i.e., determine the properties of) distributed data on remote sites. These may include nontraditional data products (e.g., biogeochemical and dynamical vegetation variables simulated by CMIP5-coupled climate–carbon cycle models). The user then will be able to regrid and analyze the desired data in place before downloading the data to the local site. This approach will place new data-management demands on ESG hosting sites but will allow scientific issues, rather than the organization and movement of data, to receive primary attention. In future ESG services, the existing Web portal capabilities will be augmented by applications to streamline data download, as well as provide powerful analysis and visualization capabilities. For example, it will be feasible to use popular and free climate analysis and visualization tools (e.g., CDAT, NCL, GrADS, and Ferrret) directly within the ESG system.

1.4 Functional Specification and Architecture Design.

Computer processing capabilities of the order of 10^{15} floating-point operations per second (petaflops) are expected to be the norm by 2010. We expect the climate modeling requirements to match the increase of available compute power. In order to meet these petascale computational needs, the future ESG architecture must allow for the networking of a large number of distributed sites with varying capabilities. Such "federation" implies that users will have to authenticate only once in order to gain access to data across multiple systems and institutions. In order to accomplish this objective, the future ESG architecture will be based on three tiers of data services.

Tier 1 services will operate across the entire ESG-CET federation. These include user registration and management, common metadata and notification services to communicate data changes, and global monitoring services to detect data problems. Because all ESG-CET sites will share a common database, a user will be able to find data of interest throughout the entire federation, independent of the site where a data search is initiated. However, access to specific datasets and related resources will still require approval by the data "owners."

Tier 2 data services will comprise multiple ESG gateways that manage limited access to specified data (e.g., the CMIP5 database). Such gateway- deployed services will include the user interface for searching and browsing metadata, for requesting data products (including analysis and visualization tools), and for orchestrating complex workflows. Because the relevant software will require considerable expertise to maintain, Tier 2 gateways will be monitored directly by ESG-CET engineers.

Tier 3 will include the actual data holdings and the services used to access these data, which will reside on ESG data nodes. Tier 3 typically will host the services needed to publish data to ESG and to execute data product requests made through an ESG gateway that may serve data requests to many associated data nodes: for example, more than 20 institutions are expected to operate ESG data nodes for the CMIP5 database. Because personnel with varying levels of expertise will operate ESG data nodes, the Tier 3 software will come with extensive documentation.

1.5 Security.

Maintaining the security of data and resources is crucial, but this should not place an undue burden on data users and administrators. A practical security protocol is to require only a single sign-on in order for a user's browser or client software to gain access to distributed data. Single sign-on will allow the security function of the ESG portal to be split among multiple servers while users authenticate only within their home domain.

For this paper, we will focus on requirement details and solutions for the authentication architecture that is being implemented by the ESG team. This paper is organized as follows. Section 2 discusses more of the details of ESG's requirements concerning authentication. Section 3 describes the single sign-on solutions that were chosen and are implemented. Section 4 briefly discusses the next technology choices concerning attribute and authorization facilities and services that are layered on top of the authentication infrastructure. Section 5 presents a summary of the paper.

1.6 ESG and external compute resources.

The climate models require vast amount of compute resources and ESG plans to make it as easy as possible for its users to deploy external compute resources, like TeraGrid (TG). Note that the PKI credentials issued by ESG's Certification Authorities (CAs) are already compatible with those used by TG.

2. AUTHENTICATION REQUIREMENTS

We discuss three aspects of authentication: single sign-on, public key, and security configuration.

2.1 Web Single Sign-On.

ESG's infrastructure consists of multiple Web portals and Web application servers that are hosted by the member organizations of the collaboratory. The main portals will manage their own user base, will require their users to login "locally," and will expect those authentication credentials to be honored throughout the ESG-virtual organization. Each main portal should also have the option to integrate their authentication mechanism with the portal-organization's existing Identity Management System, such that much of the organization's user-account management can be leveraged.

2.2 X.509 Public Key Authentication.

Besides the Web browser clients, there is also the requirement for the use of specialized clients for data-movement applications (GridFTP[5], OPeNDAP[14], DML, or the Live Access Server (LAS)[11] and for job-submission in compute-grid facilities, like TeraGrid[23]. In most of those cases, the application's security infrastructure is based on the Globus Grid Security Infrastructure, GSI[4][8], and requires X.509-public-key authentication credentials.

2.3 Security Configuration.

All client and servers, whether Web or Grid applications, require the configuration of security-related parameters, such as the trusted Certification Authorities, revocation lists, trusted identity providers, and attribute and authorization authorities. This information is not static and will have to be updated for every revoked identity, and for any change in ESG's membership as far as trust-roots, like identity providers and certification authorities are concerned. The timely and correct update of this security configuration information is crucial for the correct functioning and integrity of the whole ESG operation, Incorrect security configuration of for example, revoked credentials or CAs could lead to security compromises, while users of a new collaborating organization will only be able to access the ESG resource once their organization's CA is part of the security configuration of all ESG's resource providers.

3. AUTHENTICATION INFRASTRUCTURE

The basic authentication infrastructure of ESG is depicted in Figure 1. It shows how both OpenID and X.509 credentials are derived from a pluggable, primary authentication mechanism, such as username/password. This section presents details of the OpenID and X.509 components and the associated security configuration management.

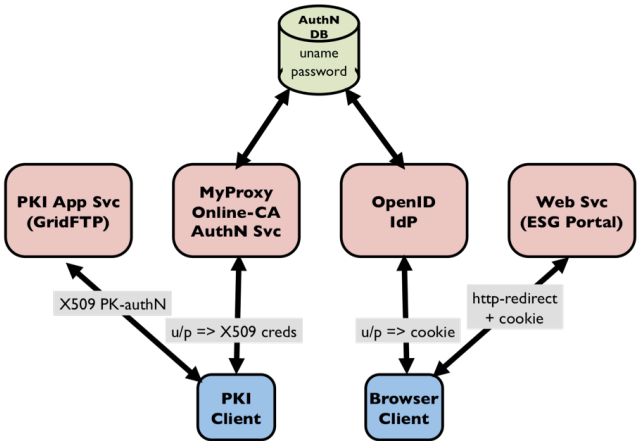

Figure 1. ESG's Authentication Architecture

3.1 OpenID WebSSO.

The Web Single Sign-On (WebSSO) solution chosen by ESG is OpenID [15][16]. It provides for a mechanism and trust infrastructure to transparently redirect unauthenticated web clients from a service provider (SP) to the identity provider (IdP) of their home-organizations to login. The client is then transparently redirected back to the SP's

application server, which will honor the IdP's proof of authentication that is presented.

The OpenID technology is similar to, but arguably less heavy-weight than, the Shibboleth/SAML [19][20] WebSSO solution. The ESG team performed an in-depth evaluation of both solutions and concluded that OpenID would be a viable WebSSO solution if a number of concerns were addressed, which resulted in a definition of an OpenID secure usage profile for the ESG federation. This profile required enforcement of a white list of IdPs, that ensured that only a trusted set of IdPs, agreed upon by the ESG federation are accepted by the SPs. Furthermore all communication between the IdP-SP must be protected by a mutually authenticated TLS-channel[25].

ESG leverages OpenID4Java[17], an open source Java implementation of the OpenID protocol. We have actively participated in the open source effort by contributing code for the features and patches that address the perceived shortcomings. Those contributions are:

- **IdP Whitelisting Extension**. This extension is designed to allow developers to plug-in their own IdP validation during the discovery phase. A well-defined IdP Validator interface is provided that allows arbitrary parameters to be configured. Developers can leverage this interface to define their own validation algorithm. Similarly, a configuration interface allows deployers to choose one or more IdP validation mechanisms to be used for a given deployment. Both as an example and to meet the ESG use case, an implementation of the IdP Validator interface that takes a plain text file with a list of IdP endpoints to trust, is provided. This validator implementation will check if the IdP used for asserting the user's identity is a member of the IdP-whitelist. Another contributed Validator implementation provides a whitelist for individual identities that can be asserted only by a defined set of IdPs.
- **Attribute Provider extension**. This IdP-extension provides a pluggable interface to obtain attributes from external sources and communicates those to the SPs as part of the OpenID protocol communication. An Attribute Provider interface has been defined that can be configured with arbitrary parameters and used to obtain the attributes for a given identity. To meet the ESG use case, an implementation that extracts the attributes from a back-end database has been provided as an example.

3.2 Short-Lived Credential Services.

The issuing of long-lived X.509 public key end-entity certificates to individual users is notoriously heavy-weight and would put a substantial extra burden on ESG's user management. For that reason, ESG has chosen a short-lived credential services (SLCS) solution [13] based on the deployment of an online-CA that issues short-lived X.509 EE-Certificates derived from a pluggable primary authentication mechanism. The implementation choice for the online-CA is MyProxy[12], which leverages the standard Pluggable Authentication Module (PAM) for the primary authentication mechanism. This allows the ESG infrastructure to plug-in a primary local authentication service that is used by both an OpenID IdP as well as the backend to an Online CA. This setup allows for the sharing of a single username-password database by the OpenID IdP and the MyProxy-CA.

In addition, the MyProxy server can be configured to consume attributes from arbitrary sources and embed them as non-critical extensions in the X.509 EE-Certificate. These attributes will be transparently communicated to a relying party during the authentication process Note that this feature is equivalent to the before mentioned OpenID attribute provider extension, where the attributes can be obtained by the OpenID IdP and communicated to the SPs. Both the Open IdP and the MyProxy online-CA will obtain the attributes from the same source.

3.3 Autoprovisioning.

To ease the burden of the system administration to maintain the security configurations on all clients and servers, ESG is deploying the autoprovisioning feature of the MyProxy service for both clients and servers.

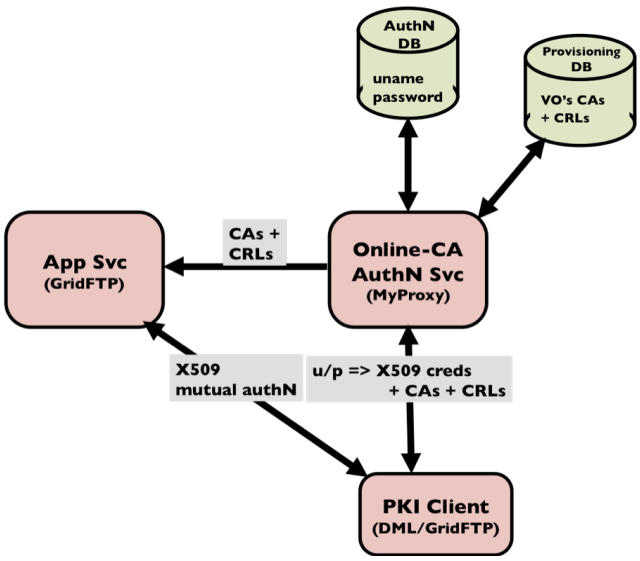

Figure 2. MyProxy Autoprovisioning Integrated with Login

After successful authentication, in addition to the short-lived X.509 credentials, a client also receives security configuration information from MyProxy. In this way, the system administrator can centrally maintain the correct and updated security configurations for the clients with MyProxy, and the clients will receive the updated information with every login. In addition, ESG has

contributed code to the MyProxy effort that enhances that service such that servers can obtain the trust-root information in a similar fashion, which again facilitates the system administration of the server's configuration.

4. CURRENT AND FUTURE WORK

We have several activities planned or under way to enhance ESG.

4.1 Leverage Trust Framework Initiatives.

The different member organizations of the ESG-Federation have to agree on the identity providers that they will trust. We are starting the process of formalizing the requirements for the inclusion of such IdPs, which would express terms like the minimal operational guarantees concerning the server deployment and account&password management that are acceptable for ESG's end-user communities and resource providers. Having a formal description of those requirements becomes even more important for the establishment of the future international federations.

We expect to be able to leverage the work and results of a number of recent trust framework initiatives, like InCommon[9], the OpenID Foundation's Open Trust Framework[18], and NIST's Electronic Authentication Guideline[13].

4.2 The Americas Grid Policy Management Authority (TAGPMA) Accreditation[24].

In order for clients to use their ESG-issued short-lived X.509 credentials to submit jobs to, for example, TG's compute services, ESG's online-CA signing policy must conform to the resource owner's, i.e. TG's, policy. We are investigating how to obtain the requires TAGPMA-accreditation.

4.3 Attribute Assertions.

Currently, we are leveraging OpenID's mechanism to communicate attribute information with the authentication assertion from IdP to SP. However, we recognize the issue that the IdP does not always has the option to access and communicate the relevant VO-associated attributes. An extra, optional attribute service call-out may be needed, and we are looking to leverage GridShib's [6] experience in that area. We plan to deploy some of MyProxy/GridShib's mechanisms to embed attributes in the issued X.509 EE-certificates. Furthermore, we're investigating the use of SAML-formatted attribute statements for OpenID's attribute values. The latter would allow us to deploy a single attribute assertion format that can optionally be signed.

4.4 Autoprovisioning of WebSSO trust roots.

We expect the number of service providers within the ESG to grow from tens to potentially hundreds of hosts. Each of those SPs will have a property file with a white-list of IdPs that includes those identity providers trusted by the ESG-Federation. We also expect the list of trusted IdPs to change and grow over time, as the membership of ESG's virtual organization changes and planned (international) federations are established.

In order to cater for those changes and to ease the administrative burden of maintaining the up-to-date trusted-IdP list with all SPs throughout the ESG, we see the need for an autoprovisioning service for the IdP-white-list configuration as shown in Figure 3.

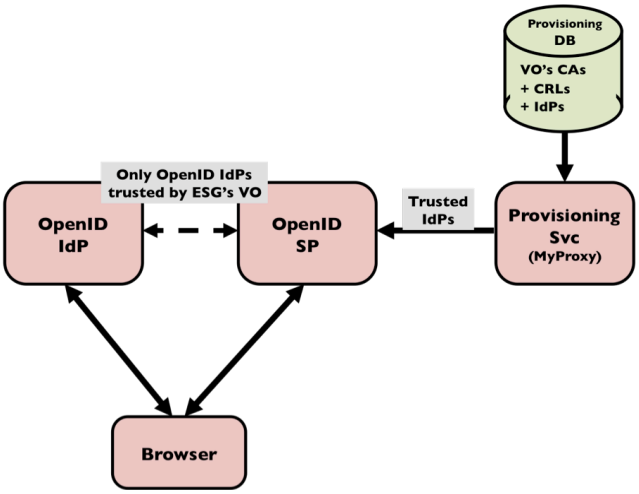

Figure 3. Provisioning of VO-trusted IdPs

The current implementation of MyProxy's autoprovisioning scheme for trusted CAs, CRLs, and such, is based on a simple file-system directory lay-out. After our initial investigation, we believe that this approach could also accommodate the IdP white-list property file. The SPs could use the MyProxy autoprovisioning service to regularly download the most current trusted-IdP-list with the trusted CAs and such. This process allows for a centralized administration of the trusted-IdP list, and its distribution can be automated without the need for administrator intervention.

4.5 OpenID-enable the OnLine-CA.

For the GridShib project, the so-called GridShib-CA has been developed[7]. This GridShib-CA allows a user to authenticate using Shibboleth, and then creates and installs a short-lived Grid credential based on that authentication on the user's local system. This credential is suitable for use with Grid Tools such as GridFTP. We are planning to work with the GridShib team to modify their GridShib-CA code such that it will also allow OpenID authentication. The process flow to obtain Grid-credentials from an OpenID-enabled Online-CA is depicted in Figure 4. The Online-CA is provided with an OpenID-SP frontend. If the client's web browser requests Grid-credentials from the Online-CA for

the first time (1), the user is redirected to her OpenID IdP (2) and redirected back after successful authentication. The trust relationship between the Online-CA SP and the IdP is such that the IdP's OpenID assertion for the user allows the Online-CA to issue X.509 Grid-credentials for that user. Those credentials have to be obtained thru a certificate request protocol exchange with the Online-CA and subsequently stored on the user's local file system. A Java Web Start application is used for that purpose[10], which is kicked off by the Online-CA SP web application (3). This Java application will generate a key-pair, send a certificate request to the Online-CA (4). The Online-CA will validate the request, ensure that it is associated with the OpenID context for that user, and issue a X.509 EE-certificate that is stored with the other parts of the credential in the client's local file system (5). Those credentials will empower any client Grid-application to access services like GridFTP (6).

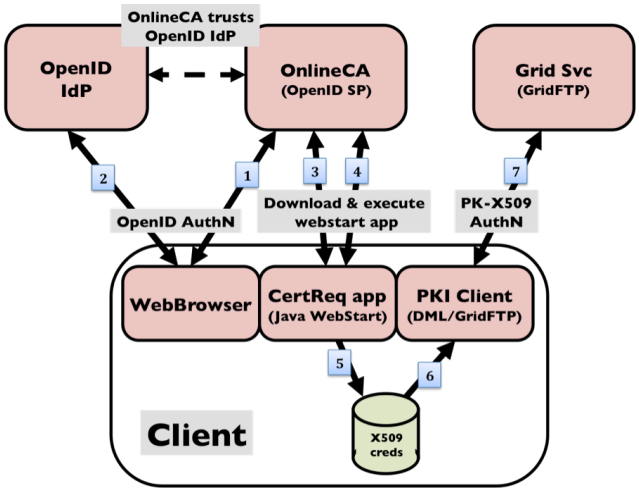

Figure 4. OpenID-enable the Online-CA

As explained, our current SSO solution is implemented by sharing the username/password database between the OpenID IdP and the MyProxy Online-CA servers. This coupling is very tight and it will only allow us to provide this integration when the OpenID IdP and the MyProxy servers are deployed by the same organization. Through the described OpenID-enabled Online-CA, we will have only a one-way dependency, which would facilitate the general deployment of this SSO solution.

4.6 SAML/Shibboleth Federation.

Although we do not have the requirement currently, future collaborating organizations may have a WebSSO infrastructure based on SAML/Shibboleth.

The Internet2 project recently kicked-off a development effort to address the Shibboleth-OpenID interoperability. The project page states that "Work is underway to add native OpenID support to the Shibboleth 2.x Identity Provider. The current goal is for a working deliverable by the end of the 2009 calendar year." [21]. We are following these technological developments concerning federation of the different WebSSO mechanisms with great interest.

5. SUMMARY

In this paper, we present an update of ESG's development and implementation efforts concerning its authentication infrastructure. ESG's requirements are to leverage existing primary authentication mechanisms, to deploy a light-weight but secure WebSSO, to deploy a light-weight X.509-PKI, and to ease the burden of security configuration management. We believe that our choice of OpenID, Short Lived Credential Services, and autoprovisioning meets those requirements. Furthermore, we have identified a number of possible future enhancements that will make the deployment even easier, like OpenID enabling the Online-CA and the autoprovisioning of the trusted IdP-list.

6. ACKNOWLEDGMENTS

The work of ESG is supported by the SciDAC program of the U.S. Department of Energy through its Office of Science Advanced Scientific Computing Research and Biological and Environmental Research. Affiliations of participating ESG team members include the following institutions: Argonne National Laboratory (ANL) is managed by University of Chicago Argonne LLC under Contract DE-AC02-06CH11357. Information Sciences Institute (ISI) is a research institute of the Viterbi School of Engineering at the University of Southern California (USC). Lawrence Berkeley National Laboratory (LBNL) is managed by the University of California for the U.S. Department of Energy under Contract DE-AC02-05CH11231. Lawrence Livermore National Laboratory is managed by the Lawrence Livermore National Security, LLC for the U.S. Department of Energy under Contract DE-AC52-07NA27344. Los Alamos National Laboratory (LANL) is managed by Los Alamos National Security, LLC for the U.S. Department of Energy under the Contract DE-AC52-06NA25396. National Center for Atmospheric Research (NCAR) is managed by the University Corporation for Atmospheric Research. Oak Ridge National Laboratory (ORNL) is managed by University of Tennessee—Battelle, LLC for the U.S. Department of Energy under Contract DE-AC-05-00OR22725. Pacific Marine Environment Laboratory (PMEL) is under the National Oceanic and Atmospheric Administration's line office of Ocean and Atmosphere Research, lies within the U.S. Department of Commerce. This work was partially funded by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Dept. of Energy, under Contract DE-AC02-06CH11357.

7. REFERENCES

[1] D N Williams, R Ananthakrishnan, D E Bernholdt, S Bharathi, D Brown, M Chen, A L Chervenak, L Cinquini, R Drach, I T Foster, P Fox, D Fraser, S

Hankin, P Jones, C Kesselman, D E Middleton, J Schwidder, R Schweitzer, R Schuler, A Shoshani, F Siebenlist, A Sim, W G Strand, N. Wilhelmi, "The Earth System Grid: Enabling Access to Multi-Model Climate Simulation Data" Bulletin of the American Meteorological Society (BAMS), 2008

[2] Data Mover Light (DML), http://www.earthsystemgrid.org/about/dmlPage.do

[3] Earth System Grid (ESG), http://www.earthsystemgrid.org/

[4] Globus Toolkit, http://www.globus.org/toolkit/

[5] GridFTP, http://en.wikipedia.org/wiki/GridFTP, http://dev.globus.org/wiki/GridFTP

[6] GridShib: A Policy Controlled Attribute Framework, http://gridshib.globus.org/

[7] GridShib-CA: GridShib Certification Authority, http://gridshib.globus.org/docs/gridshib-ca-1.0.0/

[8] GT Security (GSI), http://www.globus.org/toolkit/security/

[9] InCommon: InCommon Federation, http://www.incommonfederation.org/

[10] Java Web Start: Java Web Start Overview, http://java.sun.com/javase/technologies/desktop/javawebstart/overview.html

[11] Live Access Server (LAS), http://www.ferret.noaa.gov/LAS

[12] MyProxy Credential Management Service, http://grid.ncsa.uiuc.edu/myproxy/

[13] NIST Electronic Authentication Guideline: NIST Special Publication 800-63 Version 1.0.2, http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf

[14] OPeNDAP: Open-source Project for a Network Data Access Protocol, http://www.opendap.org/

[15] OpenID, http://openid.net/

[16] OpenID Specifications, http://openid.net/developers/specs/

[17] OpenID4Java, http://code.google.com/p/openid4java/

[18] Open Trust Frameworks for Open Government: Enabling Citizen Involvement through Open Identity Technologies: http://openid.net/docs/Open_Trust_Frameworks_for_Govts.pdf

[19] Security Assertion Markup Language (SAML), OASIS Security Services (SAML) TC, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security

[20] Shibboleth, http://shibboleth.internet2.edu/

[21] Shibboleth OpenID Support, https://spaces.internet2.edu/display/SHIB2/OpenIDSupport

[22] Short Lived Credential Services X.509 Public Key Certification Authorities (SLCA AP), Profile for SLCS - X.509 Public Key Certification Authorities with Secured Infrastructure, http://www.tagpma.org/files/SLCS-2.1b.pdf

[23] TeraGrid, http://www.teragrid.org/

[24] The Americas Grid Policy Management Authority (TAGPMA), http://www.tagpma.org/

[25] Transport Layer Security (TLS), http://en.wikipedia.org/wiki/Secure_Sockets_Layer

The following government licenses should be removed before publication: