PROCEEDINGS OF SCIENCE

Grid Operational Supports for Middleware Deployment and User Administration

Eisaku Sakane[*]
National Institute of Informatics
E-mail: sakane@nii.ac.jp

Kento Aida
National Institute of Informatics/Tokyo Institute of Technology
E-mail: aida@nii.ac.jp

Manabu Higashida
Osaka University
E-mail: manabu@higashida.net

Taizo Kobayashi
Kyushu Univeristy
E-mail: tkoba@cc.kyushu-u.ac.jp

Hirofumi Amano
Kyushu Univeristy
E-mail: amano@cc.kyushu-u.ac.jp

Mutsumi Aoyagi
Kyushu Univeristy
E-mail: aoyagi@cc.kyushu-u.ac.jp

In this paper, we present our experience of grid operational supports in the inter-university grid infrastructure in Japan focusing on grid middleware deployment and user administration. The installation tools enable administrators in resource provider sites to install the grid middleware and customize middleware settings. The user administration tools help administrators to register user accounts and maintain grid-mapfiles in multiple resource provider sites.

The International Symposium on Grids and Clouds and the Open Grid Forum - ISGC2011,
March 25-30, 2011
Academia Sinica, Taipei, Taiwan

[*] Speaker.

1. Introduction

E-Science is a new scientific research methodology to process multidisciplinary data using advanced information technology in order to achieve new scientific discoveries. Utilizing various distributed resources, e.g. computers and storages, is required for e-Science, and the grid is a key technology to federate resources for e-Science. The grid technology has been developed in the last two decades, and many grid infrastructures are currently operated in the world, e.g. TeraGrid [1], Open Science Grid [2], EGI [3] and NAREGI [4, 5].

An operation of a production level grid infrastructure is not an easy task. Although a lot of grid middleware is available, deploying middleware for the production level operation is not straightforward. For instance, grid middleware in multiple sites needs to be configured consistently, and administrators need to configure settings properly communicating with administrators in multiple sites. Procedures for these settings are often complicated and error-prone. User administration in the grid infrastructure also requires complicated procedures for administrators. The grid security infrastructure (GSI) [6, 7], which is the de facto standard of grid security mechanism, assumes that each user has local accounts in all computing resources in the grid infrastructure. It forces each user to apply user accounts to multiple sites. Furthermore, administrators need to share and maintain mapping information between users' client certificates and local accounts in computing resources, e.g. grid-mapfile.

In this paper, we present our experience of grid operational supports in the inter-university grid infrastructure in Japan focusing on the grid middleware deployment and the user administration. We developed tools for grid middleware installation and user administration. The installation tools enable administrators in resource provider sites to install the NAREGI grid middleware [4, 5] and customize middleware settings, e.g. configurations for network and resources information. Our user administration tools help administrators to register users' local accounts and maintain grid-mapfiles in multiple resource provider sites.

The rest of this paper is organized as follows: Section 2 introduces an overview of the inter-university grid infrastructure. Section 3 presents tools to deploy grid middleware among resource providers, and Section 4 shows tools for user administration. Section 5 summarizes the work presented in this paper and outlines the future work.

2. Inter-university Grid Infrastructure

This section presents an overview of the inter-university grid infrastructure in Japan. The grid infrastructure is organized by supercomputer centers in nine universities and an operation center in the National Institute of Informatics (NII).

2.1 Grid Infrastructure

The supercomputer centers, or resource providers, operate computing resources presented in Table 1; and NII, or an operational center, operates servers for grid services, e.g. the information service, the job brokering service and the grid portal. The NAREGI grid middleware Ver.1.1 [5] is used to operate the grid infrastructure. The operation center also operates the certificate authority that is approved by the Asia Pacific Grid Policy Management Authority, ApGrid PMA [8].

Table 1: Computer systems in the inter-university grid infrastructure

site	hardware	#cores/node	memory[GB]/node	#nodes
Hokkaido U.	DELL PowerEdge R200 Hitachi HA8000/110W	2	2/4	27
Tohoku U.	NEC SX-9	16	1000	4
U. Tsukuba	Appro XtremeServer-X3	16	32	4
U. Tokyo	Hitachi HA8000-tc/RS425	16	32	8
Tokyo Tech.	HP ProLiant SL390s	12	54/96	375
Nagoya U.	Fujitsu PRIMERGY RX200	2	2	6
	Fujitsu HX600	16	64	16
Kyoto U.	Fujitsu HX600	16	32	4
Osaka U.	NEC SX-8R	8	64/256	8
	NEC SX-9	16	1000	8
	NEC Express 5800/120Rg-1	4	16	32
Kyushu U.	Fujitsu PRIMERGY RX200S3	4	8	12

The typical scenario of running an application on the grid infrastructure is as follows: The user first signs on the grid infrastructure through the grid portal. The Grid Security Infrastructure (GSI) [6, 7] and the Virtual Organization Membership Service (VOMS) [9] are used for the authentication and the VO management, respectively. Then, the user edits and submits a workflow of the user's application on the portal. The submitted workflow is decomposed into jobs by the workflow engine and they are assigned to computing resources through the job brokering service, or the super scheduler. The user can monitor the running jobs through the portal. Status of computing resources are periodically collected and archived in the information service. Finally, the user retrieves the computed results through the portal.

2.2 Grid Operation

Resource providers in the grid infrastructure offer their resources for users in their local sites following their local policies. However, they need to operate their resources following not only the local policies but also an operational policy of the grid infrastructure. The operational policy of the grid infrastructure needs to be agreed by administrators of all resource providers and the operation center. We organized an administrative team, the grid deployment and operation task force, to operate the grid infrastructure. The task force is organized by researchers and administrators of nine resource providers and the operation center. The task force holds periodical meetings to discuss about grid middleware deployment, development of administration tools, operational procedures and operation policies.

Table 2 presents the milestone of the taskforce. First, the task force started to deploy middleware for computing service in FY 2008 - 2009. It also started experimental operation for user administration using Grid-Pack, which is the user registration procedure for single sign-on service on the grid infrastructure, in FY 2009. (See Section 4 for the details.). In FY 2010, middleware for data grid service is deployed among two sites and experimental operation of the authentication

Table 2: Milestone of the grid deployment and operation

FY	goal	operation
2008-2009	deployment of grid middleware for computing service user administration using Grid-Pack	experimental
2010	deployment of middleware for data grid service user administration using Grid-Pack with Shibboleth	experimental
2011	user administration using Grid-Pack with Shibboleth	production
2012	user administration computing service data grid service	production

system using Shibboleth is started. (See Section 4 for the details.) Operation in FY 2009 and FY 2010 was in experimental phase, but the task force plans to start production level operation for user administration in FY 2011 followed by computing and data grid service in FY 2012.

3. Grid Middleware Deployment

This section presents tools to deploy grid middleware among resource providers in the grid infrastructure. The NAREGI grid middleware consists of multiple service components. Thus, we need to deploy suitable components to both resource providers and the operation center. We developed grid middleware installation tools to deploy the NAREGI grid middleware.

3.1 Middleware Installation Tools

Our installation tools enable administrators of both resource providers and the operation center to install suitable components in their sites. Table 3 summarizes components installed in each site. The installation tools are developed based on the hierarchical management concept [10] in order to minimize the information of the configurations.

For resource providers, installation tools install and configure a component of the computing service (GridVM) and a component of the information service (IS-CDAS). GridVM runs on a gateway node of computing resources in each site, and it accepts jobs assigned to the site and submits the jobs to computing nodes through the local batch scheduler. IS-CDAS collects resource information in the resource provider site and reports the collected information to the information service (IS-NAS) running in the operation center.

The operation center runs services for the grid portal (Portal), the information service (IS-NAS), the job brokering service (SS) and the security service. The installation tools install components for these services and configure the settings. Portal is a user interface to access resources in the grid infrastructure. The user first signs on Portal; then the user can make a workflow of the user's application, submit the workflow and monitor the submitted workflow (or jobs in the workflow) through Portal. IS-NAS collects resource information from IS-CDASes running on resource provider sites and archives the collected information. SS dispatches jobs in the submitted workflow to suitable computing resources following scheduling policies. We use GSI and VOMS for user authentication and VO management, respectively. The security service is performed by the user

Table 3: Software component installed on the grid infrastructure

service	operation center	resource provider
computing	grid portal service (Portal) Information service (IS-NAS) job brokering service (SS) security service (UMS, VOMS)	computing service (GridVM) Information service (IS-CDAS)
data grid	metadata server (gfmd)	file system node (gfsd)

management service (UMS) and VOMS. Client certificate of users are archived in the certificate repository on UMS. When the user signs on the portal, UMS create the user's proxy certificates with VO attributes using the information in VOMS.

Data grid service is also important service for e-Science applications. We operate a distributed file system in the grid infrastructure using the Gfarm File System (Gfarm) [11]. Gfarm enables users to access files on storages geographically distributed over the grid infrastructure. The Gfarm file system is organized by a metadata server and file system nodes. Files are saved on storages in file system nodes distributed over the grid infrastructure, and the metadata server manages metadata of distributed files. In our operation, the operation center operates the metadata server and resource providers operate file system nodes. Installation tools install components, both metadata server (gfmd) and file system nodes (gfsd), on suitable sites and configure the settings.

3.2 Middleware Deployment Procedures

After installation of middleware components, administrators of both resource providers and the operation center need to exchange information of middleware setting in each site and configure consistent middleware settings. We established the procedure to exchange information between resource providers and the operation center.

Figure 1 shows an example of the procedure to add/delete computing nodes in a resource provider. First, the resource provider sends an application form, including information of computing nodes, to the operation center. Then the operation center configures middleware settings of grid services and makes a configuration file including information about servers for grid services. The configuration file is sent to the resource provider, and the administrator of the resource provider can configure middleware settings by running the installation tools with the configuration file. We also have similar procedures for configuring the data grid service.

The deployment procedures and the installation tools help administrators of resource providers to configure middleware settings. For instance, roles of administrators are sending/receiving files to/from the operation center and running the installation tools. Thus, the administrators can configures middleware setting by using the installation tools without detailed knowledge about the middleware.

4. User Administration

This section presents tools for user administration. The user administration tools help administrators to register users' local accounts and to maintain grid-mapfiles in multiple resource provider

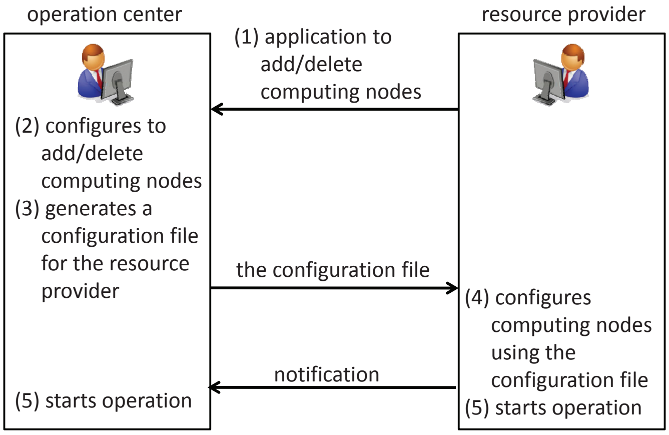
operation center
resource provider
(1) application to add/delete computing nodes
(2) configures to add/delete computing nodes
(3) generates a configuration file for the resource provider
the configuration file
(4) configures computing nodes using the configuration file
notification
(5) starts operation
(5) starts operation

Figure 1: An example of installation procedure

sites.

4.1 Federated User Registration Procedures (Grid-Pack)

We established the user registration procedure called "Grid-Pack", which enable a user to apply both local accounts in multiple resource providers and a client certificate by submitting an application to the local registration authority (LRA) in the user's local site.

Applying Grid-Pack Figure 2 presents a diagram to issue a client certificate in Grid-Pack. The user submits an application for Grid-Pack at LRA in the primary resource provider, or the resource provider A in the figure. LRA in the primary resource provider is in charge of customer service for users in the local site, e.g., users in the university A apply Grid-Pack at LRA in the supercomputer center located in the university A. The LRA operator confirms identity of the applicant through a face-to-face interview and assigns the Grid-Pack ID (GPID), which is a unique ID assigned to a user in the grid infrastructure, to the applicant. Then, the LRA operator requests the LRA administrator in the operation center to issue the client certificate for the confirmed applicant by sending the request form and the applicant's GPID.

Issuing a Client Certificates A client certificate is issued by an online procedure by the applicant through the portal. The LRA administrator sends the account information to access the portal (UMS acct and password in the figure) and the license ID to the applicant. When the applicant requests a client certificate on the portal, the authentication of the applicant is conducted not only by the password but also by the license ID. There are two options to save an issued certificate: saving the certificate in the certificate repository on UMS and downloading the certificate to the user's machine. In Grid-Pack, we chose the former option considering security issues.

Maintaining grid-mapfile The LRA administrator maintains the GPID-DN list, which contains a pair of GPID and a Distinguished Name (DN) of a client certificate for each user. The GPID-DN list is published to administrators of resource providers through the subversion (SVN) repository [12]. The user's local account is created in each resource provider following the procedure illustrated

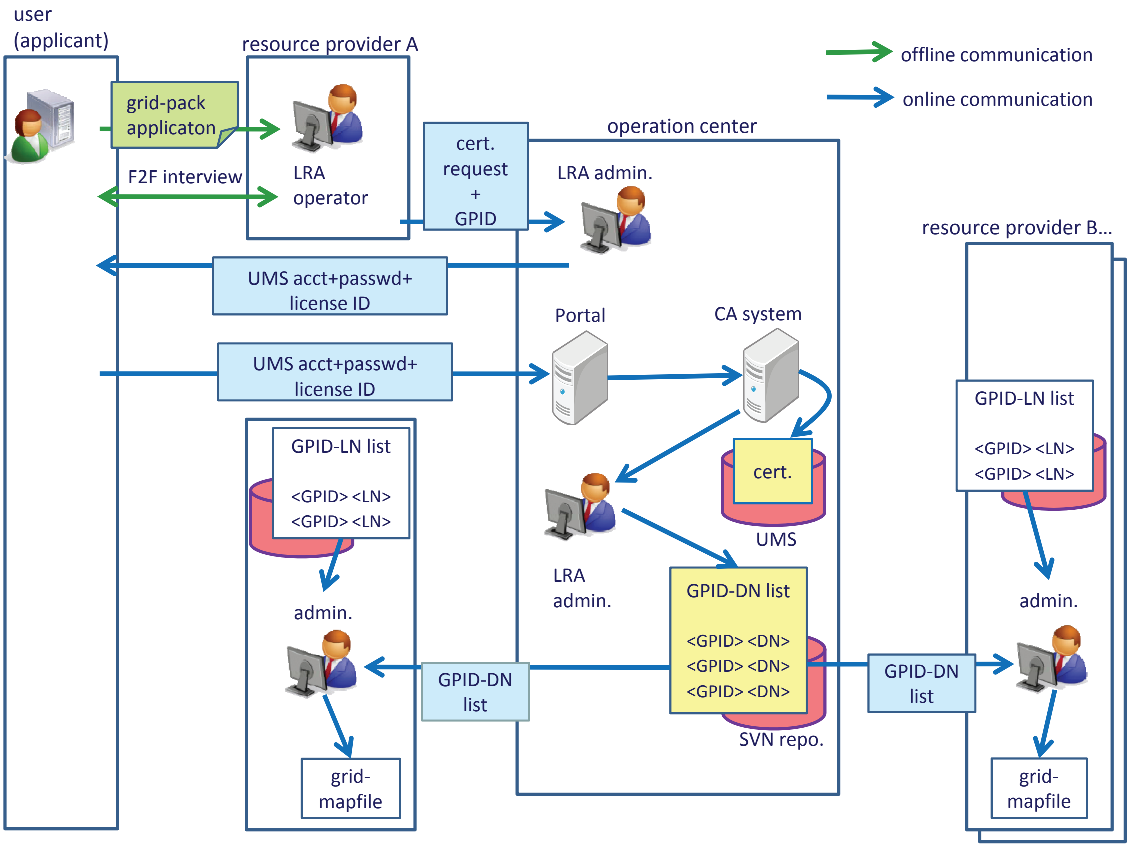

Figure 2: A diagram to issue a client certificate in Grid-Pack

in Figure 3. The administrator of each resource provider maintains the GPID-LN list, which contains a pair of GPID and the local account name (LN) for each user. Then, the administrator of each resource provider makes a grid-mapfile by running the grid-mapfile generation tool. The tool downloads the GPID-DN list on the SVN repository and generates a grid-mapfile by retrieving information both from the GPID-DN list and the GPID-LN list.

4.2 Federated Authentication System with Shibboleth

Grid-Pack reduces cost for user administration in the grid infrastructure both for users and administrators. However, there still exists a bottleneck in administration of the operation center. For instance, all requests from users are gathered at the LRA administrator in the operation center. Furthermore, the user needs to manage an account for the portal, or UMS acct in the figure, to sign-on the portal.

We plan to operate the authentication system using Grid-Pack with Shibboleth [13, 14]. Figure 4 illustrates software architecture of the authentication system. In this system, each resource provider operates an identity provider (IdP), which manage an account DB of the local site. The portal running on the operation center works as SP, and it enable a user to sign-on the portal using his/her local account of the primary resource provider.

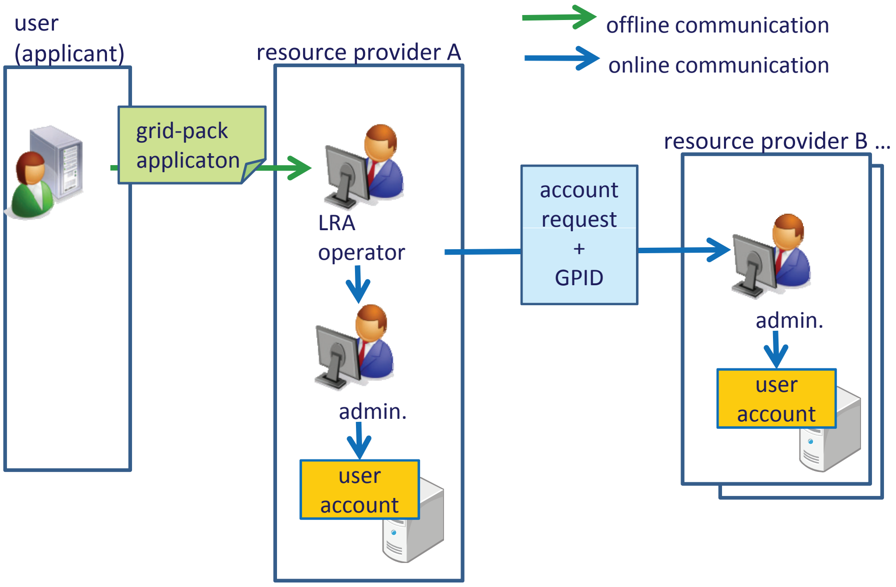

Figure 3: A diagram to register local accounts in Grid-Pack

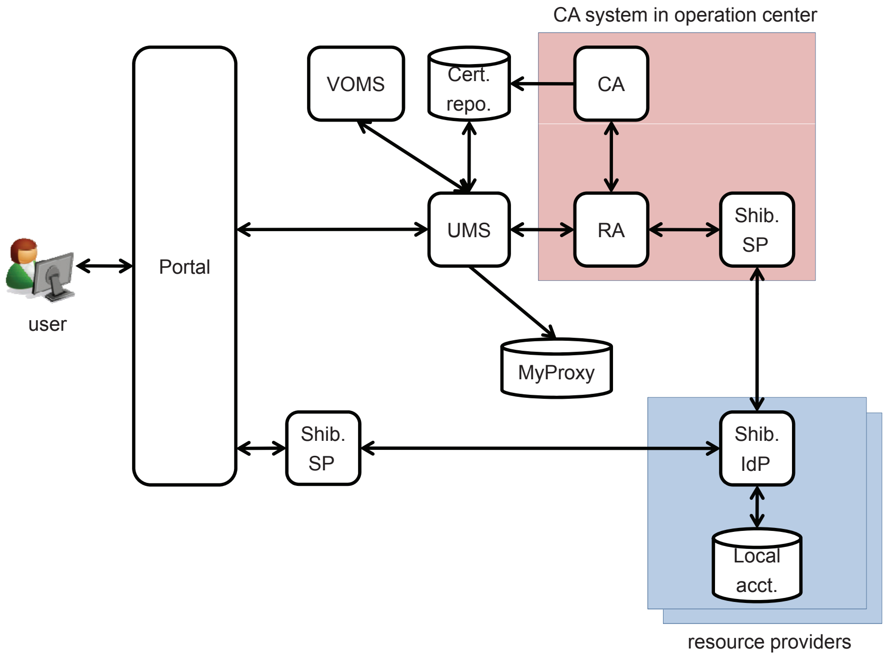

Figure 4: An authentication system using GIS and Shibboleth

Issuing Client Certificate The user can request issuing a client certificate through the portal. First, the user signs on the portal and requests issuing the client certificate. For the authentication on the portal, the user inputs the local account issued from the primary resource provider [1]. The certificate request is sent to the RA server via UMS. The RA server authenticates the user by communicating IdP in the primary resource provider. After the user's identity is authenticated, the RA server requests the CA server to issue the client certificate. Then, the CA server issues the client certificate and save the certificate in the certificate repository (Cert. repo. in the figure).

Single Sign-on We use GSI and VOMS to enable single sign-on and VO management, respectively. When the user sings on the portal, the proxy certificate is generated by UMS. Then, UMS add the VO attributes of the user by using information in VOMS. The proxy certificate with the VO attribute is saved in MyProxy [15, 16]. The user can access resources in the grid infrastructure using the proxy certificate in MyProxy.

The system reduces the bottleneck in the operation center and reduces the administration cost. The client certificate is issued through online procedures on the portal. The current testbed is organized by the operation center (SP) and two resource providers (IdPs). We are now discussing the detailed procedure for user administration using the authentication system.

5. Conclusions

In this paper, we presented our experience of grid operational supports in the inter-university grid infrastructure in Japan focusing on grid middleware deployment and user administration. The grid middleware installation tools enable administrators in resource provider to install and configure grid middleware without detailed knowledge of the middleware. The user administration tools enable users to apply accounts to use the grid infrastructure with a single application, and the tools also help administrators to register user accounts and maintain grid-mapfiles in multiple resource providers. We also established procedures for the middleware deployment and the user administration.

Currently, we plan to extend the testbed for the authentication system using GSI and Shibboleth in order to start operation among nine resource providers. The goal is to start the production level operation of the user administration in FY 2011.

References

[1] TeraGrid. https://www.teragrid.org/.

[2] Open Science Grid. http://www.opensciencegrid.org/.

[3] European Grid Infrastructure. http://www.egi.eu/.

[4] K. Miura. Toward Cyber Science Infrastructure - NAREGI Grid Middleware and Beyond -. In *Proc. of the International Symposium on Grid Computing 2010 (ISGC2010)*, 2010.

[5] NAREGI. http://www.naregi.org/.

[1] We assume that the user's identity is confirmed in proper way when the user apply the local account.

[6] V. Welch, F. Siebenlist, I. Foster, J. Bresnahan, K. Czajkowski, J. Gawor, C. Kesselman, S. Meder, L. Pearlman, and S. Tuecke. Security for Grid Services. In *Proc. of the 12th IEEE International Symposium on High Performance Distributed Computing*, 2003.

[7] V. Welch. Globus Toolkit Version 4 Grid Security Infrastructure: A Standards Perspective. In *http://www.globus.org/toolkit/docs/4.0/security/GT4-GSI-Overview.pdf*, 2005.

[8] Asia Pacific Grid Plicy Management Authority. http://www.apgridpma.org/.

[9] R. Alfieri, R. Cecchini, V. Ciascini, L. dell'Agnello, A. Frohner, A. Gianoli, K. Lorentey, and F. Spataro. VOMS: an Authorization System for Virtual System for Virtual Organization. In *Proc. of the 1st European Across Grids Conference*, 2003.

[10] T. Kobayashi, J. Ooba R. Mibu, T. Takami, J. Maki, R. Nogita, and M. Aoyagi. A New Concept for Constructing HPC environment - packaging the NAREGI grid middleware by apt-rpm -. In *Proc. of the 9th International Conference on High Performance Computing, Grid and e-Science in Asia Pacific Region (HPC Asia 2007)*, 2007.

[11] O. Tatebe, N. Soda, Y. Morita, S. Matsuoka, and S. Sekiguchi. Gfarm v2: A Grid file system that supports high-performance distributed and parallel data computing. In *Proc. of the 2004 Computing in High Energy and Nuclear Physics (CHEP04)*, 2004.

[12] Apache Subversion. http://subversion.apache.org/.

[13] Shibboleth. http://shibboleth.internet2.edu/.

[14] R. L. Morgan, S. Cantor, S. Carmody, W. Hoehn, and K. Klingenstein. Federated Security: The Shibboleth Approach. *EDUCAUSE Quarterly*, 27(4), 2004.

[15] MyProxy: Credential Management Service. http://grid.ncsa.illinois.edu/myproxy/.

[16] J. Novotny, S. Tuecke, and V. Welch. Initial Experiences with an Online Certificate Repository for the Grid: Myproxy. In *Proc. of the Tenth International Symposium on High Performance Distributed Computing (HPDC-10)*, 2001.