

Integrating VOMS with SAM-Grid

Date: May 02, 2005

Contents

1.	<i>Introduction</i>	1
2.	<i>VO Management in SAM-Grid today</i>	2
2.1.	<i>VO Management for the users</i>	2
2.2.	<i>VO Management for SAM-Grid Services</i>	3
2.3.	<i>Management of the trusted Certificate Authorities</i>	3
3.	<i>Motivations for the migration to standard VO management tools</i>	4
4.	<i>Integrating VOMS/VOMRS in SAM-Grid</i>	5
4.1.	<i>PHASE I: Populating VOMS with VO members and using VOMS client</i>	5
4.2.	<i>PHASE II: Using VOMRS for registration</i>	6
4.2.1.	<i>Challenges and show stoppers</i>	7
4.3.	<i>PHASE III: Interfacing SAM and VOMS/VOMRS</i>	9
4.3.1.	<i>Challenges and show stoppers</i>	9
4.4.	<i>PHASE IV: Using VO specific feature in VOMS</i>	10
5.	<i>Estimated Timeline and Manpower</i>	10
6.	<i>References</i>	11

1. Introduction

This document discusses Virtual Organization (VO) management for DZero in the context of SAM-Grid. It describes the mechanisms used for VO management as of today and it proposes new mechanisms that take advantage of standard Grid VO Management tools, such as VOMS and VOMRS.

A Virtual Organization (VO) is a collection of individuals and institutions that agree upon resource sharing on the Grid. The VO is responsible for establishing agreements between the resources providers and the resource users. In particular, the VO is responsible for maintaining the lists of users and services authorized to use the resources of the organization. A scientific collaboration, such as the DZero experiment, is generally considered a VO.

In this document we explain the current VO Management process in the SAM-Grid and our motivations to integrate and interface new standard VO management tools. In the SAM-Grid VO Management is done with tools developed in-house that are specifically customized to the SAM-Grid's architecture. These tools have fewer acceptances than the standard tools. Site administrators are vary about running these tools periodically via cron jobs. In addition, integrating these tools with the SAM-Grid would make it easier to interoperate with other grids that already adopt these standards.

This document also discusses VO registration services. These services allow the implementation of VO-specific registration policies for users and services. The current

registration process for users submitting grid jobs using the SAM-Grid is cumbersome and can be streamlined. This work is an opportunity to redesign the registration process, using standard tools, such as VOMRS, which are naturally interfaced to the standard VO management services.

We conclude by discussing a plan to integrate VOMS/VOMRS in the SAM-Grid, using a multi-phase approach. We also put forward the various challenges and shortcomings we foresee, followed by the estimated timeline and manpower required for the project.

2. VO Management in SAM-Grid today

2.1. VO Management for the users

Main purpose of VO management in SAM-Grid is authorizing users and services for data and job management. In order to access SAM-Grid services, the user needs to register at two locations: the SAM database for data movement and the sam_gsi_config repository for job submission. See

Figure 1 for a diagram of the virtual organization management and registration mechanism.

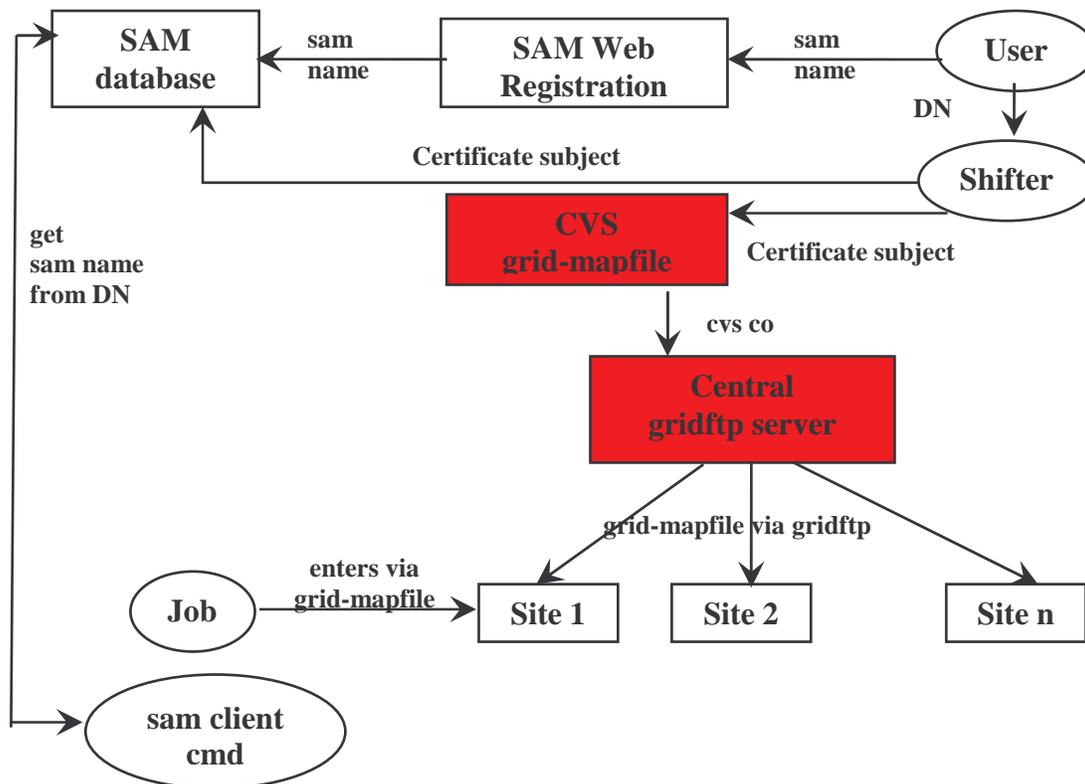


Figure 1: The SAM-Grid Virtual Organization management and registration system. The boxes in red are going to be replaced with a VOM server.

Registration to the SAM database grants access to the SAM data handling system. The user can register over the web, providing 2 pieces of information

1. Fermilab Kerberos principal: the system checks that the principal exists. No further identity certification is performed. The principal becomes the identifier i.e. the *sam-user* name, for the user in the SAM system. The SAM commands implement accounting using the *sam-user* name.
2. A set of SAM “groups”: these are used by the system to implement fair share of storage resources.

The *sam_gsi_config* repository maintains the list of users that want to access resources for job submission via grid mechanisms. The list is maintained as a file in the *sam_gsi_config* CVS repository. To register, the user sends his/her DN (“Distinguished Name”, also called “certificate subject”) to the *sam_admin* mailing list. Since the shifter only gets the DN he/she is required to certify the identity of the user. After certification, shifters insert the DN in the *sam_gsi_config* repository. The list is made available hourly to a central GridFTP server, using the *cvs checkout* command. Remote sites download periodically the official list of DN from the GridFTP server, using *sam_gsi_config* utilities. Users that submit jobs to the sites via grid mechanisms are granted access to the resources if their DN is present in the local user list (*grid-mapfile*).

SAM-Grid shifters also register the user DN with the SAM database. The DN is used to associate a person’s identity with a *sam-user*, when executing *sam* commands. The *sam* command extracts the DN from the user certificate proxy, if present. Otherwise, a simple comparison between local unix user name and *sam-user* is performed. This mechanism guarantees backward compatibility for users that do not use grid certificates for identification.

This mechanism respects the registration policy of the DZero experiment for the SAM-Grid system: simple Kerberos principal name comparison for access to the SAM system (data movement); identity certification from the DZero Virtual Organization, via the SAM Shifters, for access to computing resources.

2.2. VO Management for SAM-Grid Services

Some of the services in the SAM-Grid use certificates to establish security contexts. A typical service that needs such certificates is the SAM station, which moves files on behalf of the users. Multiple users, in fact, may want to access the same data, which is moved for the collectivity using the station identity.

The VO management infrastructure of the SAM-Grid must be able to manage lists of services, in addition to list of users. These lists are used to authorize access of services to resources, such as storage elements. Currently these lists are maintained in the *sam_gsi_config* package by the shifters. Typically, new DN for the services are added when new SAM stations are installed.

2.3. Management of the trusted Certificate Authorities

Certificate Authorities (CA) are organizations that certify the identity of users and services. After the certification, the CA releases a certificate of identity to the user/service. Users

and services present credentials derived from this certificate to other services. If these services trust the CA, they can authenticate the user/service identity.

Currently, the SAM-Grid team maintains the list of CAs that are trusted by the SAM-Grid services. This list consists of a set of files for each CA. The most important file in the set is the CA certificate, which univocally identifies the CA. The list of CA certificates needs to be maintained, as each CA certificate has a validity limited in time.

There are two main categories of services in the SAM-Grid: data handling and job management services. Each site has its own security policy and may allow or deny honoring different CAs. So the set of CAs trusted by job management is more restrictive than the set trusted by the data handling. For example, Fermilab had a policy for allowing users who use Fermilab's KCA to run their jobs. This has changed recently and now Fermilab allows users with certificates from a bigger subset of CAs.

In the new VO management scheme, we need to investigate whether to have the SAM-Grid team maintain the list of CAs or to delegate the responsibility to other entities, such as the Virtual Data Toolkit.

3. Motivations for the migration to standard VO management tools

The current VO management and registration system is adequate for the current needs of the experiment. On the other hand, it presents various drawbacks:

1. Site Administrators do not trust the `sam_gsi_config` tools: `sam_gsi_config` provides a set of tools to download and install the VO membership lists at a site. The download should be performed periodically as root. In our experience, administrators tend not to run these tools as they are not standard. They prefer to be personally contacted and add new DNs by hand. This is inconvenient and non-scalable. The migration to standard tools will alleviate this problem.
2. The registration process should be optimized: the current scheme requires that a user registers twice (*sam-user* name and DN independently) and that a shifter adds the user's DN to both the SAM database and the `sam_gsi_config` repository. This is driven in part by the policies of the experiment, in part by the technological choices. The migration to standard VO management tools is an occasion to revisit the policies and the technologies used to implement them.
3. The SAM database acts as a VO repository: SAM commands access the SAM database to retrieve VO information and perform user authentication. This is a duplication of information (DN is in multiple places) that ideally should be avoided. Also, the database does not currently offer secure communication contexts.
4. Maintenance: all the `sam_gsi_config` tools are developed in-house by the SAM-Grid team. The maintenance accounts for a 2 FTE-days / year.

5. Extensibility: the standard tools support advanced functionalities, such as the extension of user credentials with VO-defined roles. These functionalities are used by authorization services, such as fine grain authorization to resources. The migration to the standard tools makes it possible for the system to integrate new security services.

We therefore believe that it is time to investigate the migration to new technologies. We propose to investigate the use of VOMS for VO management and of VOMRS for user registration. Since this is a considerable change to how we manage users, it would be carried over in phases as explained in the next section.

4. Integrating VOMS/VOMRS in SAM-Grid

A similar project of migration to VO management services from the traditional gridmap approach was done by the EDG group in CERN [2]. There is a significant correlation to the overall VO management done at CERN with that present in SAM-Grid with the significant difference being the customized tools used for the VO management. EDG used LDAP while SAM-Grid currently uses `sam_gsi_config` tool described in the previous section. Shifting to standard VO management scheme is a significant change and these changes are easier carried in phases. This allows for a smoother introduction of changes in the project with minimal/easy transition to VO management for the users.

4.1. PHASE I: Populating VOMS/VOMRS with existing VO members

In the first phase we intend to perform the following tasks –

1. Setup a VOMS and VOMRS server test-bed for the development activities.
2. Create initial VO membership list for DZero for users and services: this will be based on the users registered on the SAM database. Check certificates currently in the SAM database for consistency and add them to VOMS/VOMRS. Understand the administrative tools available for VOMRS and develop new tools to help us in migration. We currently do not have any agreements on the use of VOMS/VOMRS for CDF and we will need to investigate their interests/plans.
3. Integrate the VOMS client tools in the SAM-Grid products and devise a strategy for the deployment. We will use these tools to create the authorization lists at the remote sites. We plan to use VDT to distribute the client tools and configure SAM-Grid to use them.
4. Establish a production VOMS/VOMRS and migrate the content of the development VOMS/VOMRS to production.
5. Identify people responsible for server availability, infrastructure maintenance, VO membership administration.
6. Write documentation on how to use the tools.
7. Understand the compatibility with the European VOMS.

4.2. PHASE II: Using VOMRS for new user registration

The second phase of the project is defining an integrated registration process for data and job management. We need to automate/ease the tasks for both users and shifters. We want to achieve this goal by using VO registration tools such as VOMRS. We choose VOMRS because it automates the registration process, providing services such as notification of a registration request to VO administrators, convenient tools for approval/denial of the registration, automatic interfacing to VOMS, etc. [4]. The diagram in **Figure 2** outlines at a high level the registration flow. The diagram is explained hereby.

1. User registers using the VOMRS service giving the required details like VO, group of the VO, DN, etc. We hope to be able to eliminate the double registration requirement for data and job management.
2. VO Administrators/Shifters get an email about the new request. Administrators certify that the user belongs to the VO using VOMRS interfaces. The user information is then automatically entered in VOMS and the SAM database.
3. On the execution site VOMS client tools will be used to generate the gridmap file. This process is most likely a cron-based command using standard tools.

In this scheme, the users and services are distinguished based on the concept of VO groups. This concept is implemented by VOMS and it is the equivalent of a metadata field in the user/service entry. DZero users and services will be organized in two different groups.

The registration process for service certificates is not completely defined at the moment. We identify the following main options:

1. Service administrators load the service certificate in their browser. The registration service (VOMRS) is told via a web form whether the certificate is for a user or for a service. We alternatively envision writing code that will automatically discriminate between the two. This approach is more involving for the service administrator and requires the development of the registration form.
2. The service administrator sends the identity of the service (DN) by email and the VO administrator/shifter registers it to VOMS directly. This approach does not have the benefits of automating the registration process.

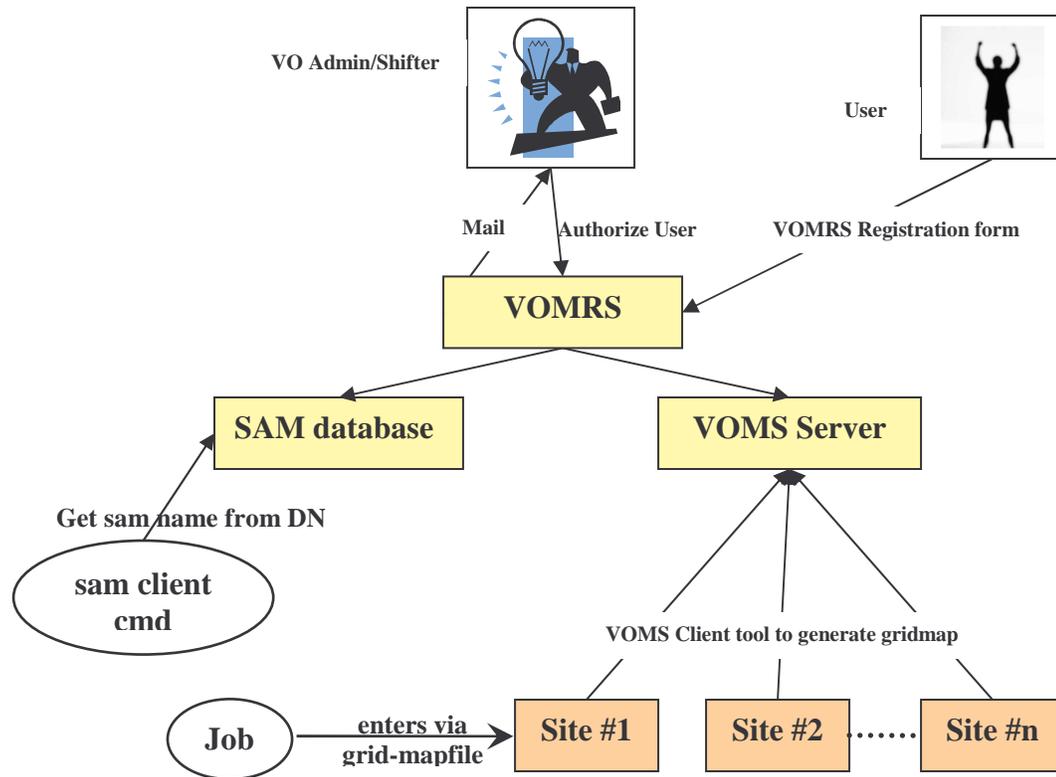


Figure 2: Proposed SAM-Grid Virtual Organization management and registration system.

4.2.1. Challenges and show stoppers

1. The diagram in **Figure 2** assumes that the administrator approves or disapproves the user requests interacting only with VOMRS. VOMRS is then responsible for the user registration with the SAM database and VOMS. To implement this information flow, VOMRS will need to support registration with multiple entities.
2. VOMRS requires user to load his/her certificate in the browser. There are SAM users that issue SAM client commands to access data handling services and are not interested in accessing the grid to run their jobs. Today, the registration process does not require that a user have a certificate. The current registration workflow is shown in **Figure 3** we can envision maintaining this registration process by having VOMRS redirect these users to the current registration web page. This page only requests that a user enters his/her Fermilab id. The major drawback in this scheme is that there is no explicit certification of the user identity, as no certificate/DN verification is involved. The system accepts the identity of the user on the basis of his/her claim. This work is an opportunity to re-discuss the registration procedure. These are the options that we envision:
 - a. The experiment is satisfied with the “claim-to-be” registration policy. We follow the process as shown in **Figure 3**. This procedure is the least secure.

Pros:

- § No changes are required to the registration process of the users who do not use the grid.

Cons:

- § There is no explicit checking of user identity. This policy cannot be used for the grid. We need to maintain two different registration processes.

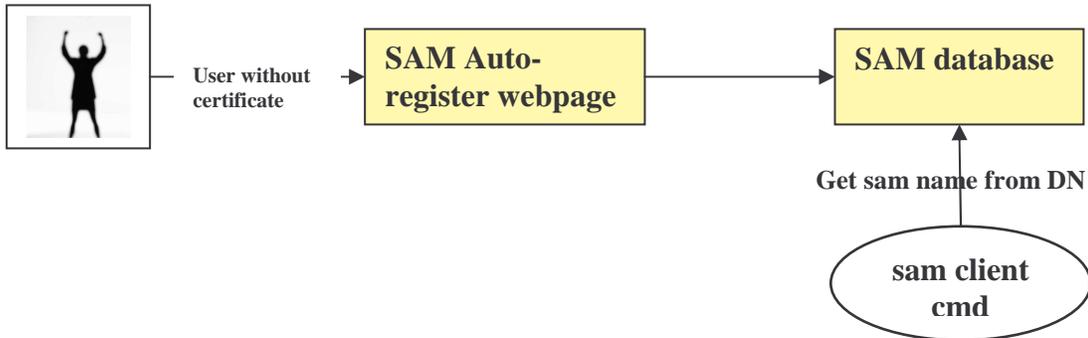


Figure 3: Current registration system for user without certificates.

- b. The experiment mandates that the user load the certificate in his/her browser and register only via VOMRS. This procedure is more involving of the user. To facilitate the user, we envision providing easy to use tools to generate certificate from the Kerberos CA and to load the resulting certificate into the user's browser.

Pros:

- § This procedure is secure as we confirm the identity of the registering user.
- § It is a single registration point for the grid and SAM (one stop shopping)
- § It is more forward looking

Cons:

- § This policy is more involving for the user as s/he needs to load the certificate to his/her browser. We can provide tools to facilitate this.

- c. Users email the information to the VO Administrator. The Administrator certifies the identity of the user and adds the information to the SAM database and VOMS. This procedure is more involving of the Administrator.

Pros:

- § This procedure is secure, as the administrator confirms the identity of the registering user.
- § This procedure can be used for SAM and the grid.

Cons:

- § This procedure is more involving of the Administrator.
- § Administrator needs to certify the identity of the user again. In policy 2, VOMRS does that using the user certificate.

3. Devise a plan for the migration of the registration procedure.

4.3. PHASE III: Interfacing SAM and VOMS/VOMRS

Today, the SAM database stores the DN of the users and services. This represents a duplication of information, as the DNs are also stored in the VO management system (`sam_gsi_config` or VOMS). We propose to investigate whether VOMS can replace the SAM database to map the user DN to the *sam-user* identity. If this is possible, the SAM client code will need to be modified to access VOMS instead of the SAM database.

4.3.1. Challenges and show stoppers

- ∅ VOMS is currently not flexible enough to allow VO-specific metadata associated with the registration entry like *sam-user* discussed above. There are two possible alternatives to this problem
 1. Obtain the information when the user registers with VOMRS and store the VO-specific metadata in VOMRS database. VOMRS needs to sustain high number of repetitive requests to provide this information. We deem this is as a potential problem to this alternative.
 2. Obtain the information when the user registers with VOMRS and store it as *capability* feature in VOMS. SAM clients can extract the information from the extended proxy. *Capability* feature of VOMRS has not been tested in production and it is not clear if it works. This solution needs design and code modifications to VOMRS. SAM also needs to be modified to use this extended proxy and extract the required information.

Another such important VO specific metadata is SAM group membership. Every SAM user is part of one or more data-handling groups. Group membership information is required by SAM to implement various data-access related features like fair share.

We propose following alternatives to tackle this problem –

1. Create an equivalent groups structure in VOMS and make the user part of the corresponding groups in VOMS as well. This membership information is only used by SAM. Putting the group's information in VOMS might not be right design decision. Also, VOMRS does not allow users to select group(s) he/she needs to be part of. Only the VO admins can assign the users the group membership.
2. Modify the VOMRS user registration page to allow the user to input the group information. This information can be collected as a part of the registration process. A SAM interface can be implemented and register with VOMRS, so the groups information can be stored in SAM database. However this requires modifications to VOMRS so that group information of already registered users can be obtained from the SAM database.

Ø The scalability of VOMS/VOMRS to serve a high number of concurrent requests must be investigated.

4.4. PHASE IV: Using VO specific feature in VOMS

VOMS is integrated with advanced security services. These services include fine-grain authorization to resource. Integrating the SAM-Grid with these services would allow the definition of role-based authorization policies.

Current tools/technologies involved are PRIMA, GUMS. Please refer to **Figure 4** [1]. This phase will also need a transition to GT web services i.e. GT 3.x+ or GT 4.x+.

There are not enough elements at this time to devise a more detailed plan.

5. Estimated Timeline and Manpower

Assuming that the project will not face any major show stoppers (see previous sections), we estimate the following time lines and man power requirements for the project.

Phase I : 2 to 3 weeks, 50% of 1 FTE

Some of the tasks for Phase I have already started. We have a production VOMS server running the “dzero” VO on fermigrd2.fnal.gov. It will take around 4 to 6 hours to setup another VOMS/VOMRS server as a test-bed. We also have some basic tools to extract the DNs from the SAM database. We will add features to these tools to extract the information from the SAM database and register them with the VOMRS. We estimate an approximate period of 20 to 30 hours to develop and document these tools. Another important task is to use VDT (used for SAM-Grid deployment) to distribute the VOMS/VOMRS tools and the CA certificates. This task should take around 10 to 15 hours. We estimate remaining period for understanding the European VOMS. Some of the tasks could be done in parallel.

Phase II : 5 to 6 weeks, 50% of 1 FTE

This phase includes revamping the workflow of VO registration and management. This might involve communicating with several people and incorporating their feedback. This could be iterative process and we estimate around 3 - 4 weeks for this task. We also need to write tools for the users to load their certificate into the browser and document detailed instructions on our new registration policy. We also need to write tools to instruct VOMRS to store the information in SAM database as well as in VOMS. This needs some development to be done in VOMRS. We estimate a period of around 3 - 4 weeks for this process. Some of the tasks can be done in parallel.

Phase III : Undetermined

Phase IV : Undetermined

Authorization Architecture Compute Node Functionality for OSG-0

FNAL Privilege Project

Version 4 - 2005-01-09
mlorch@fnal.gov

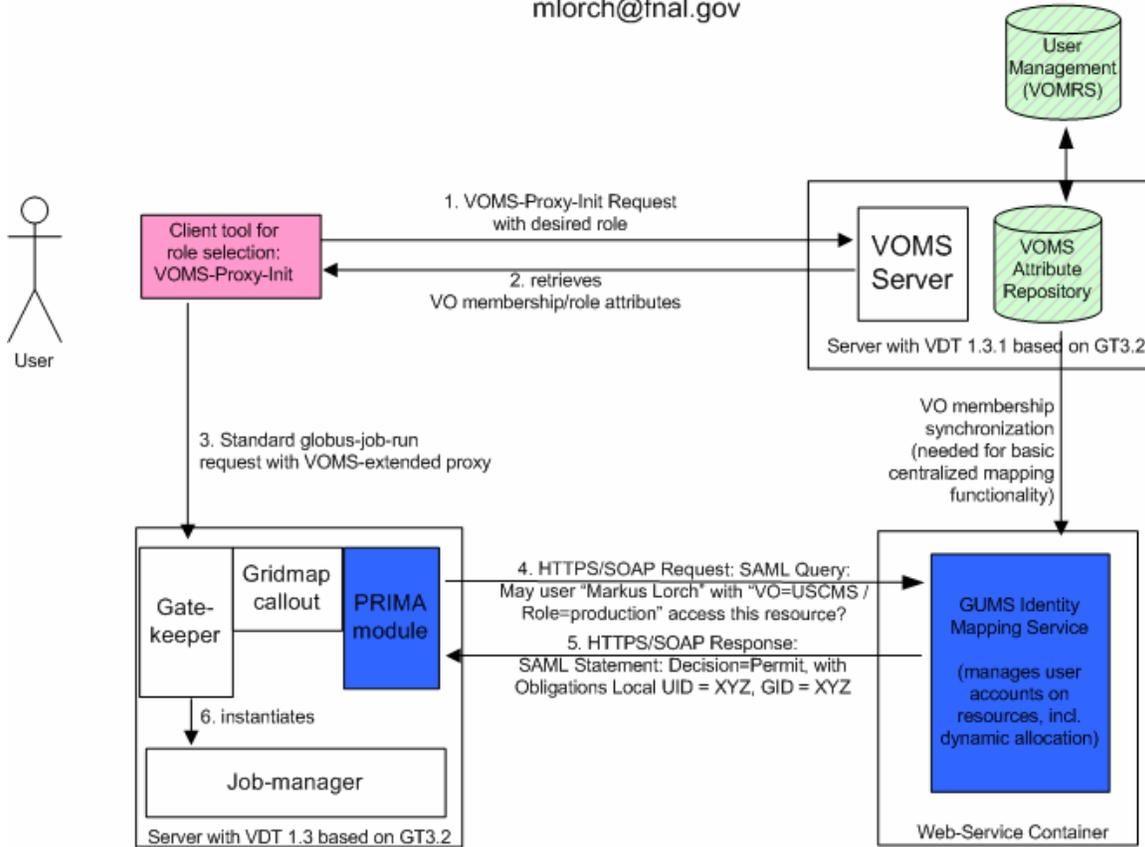


Figure 4: Interfacing Gatekeeper to VOMS through PRIMA and GUMS.

6. References

- [1] <http://computing.fnal.gov/docs/products/voprivilege/documents/transition-to-privilege.html>
- [2] <http://edg-wp2.web.cern.ch/edg-wp2/security/voms/voms-migration.png>
- [3] http://grid-auth.infn.it/docs/VOMS_stress_tests.html
- [4] <http://computing.fnal.gov/docs/products/vomrs/>