IGI Portal architecture and interaction with a CA-online

Abstract

In the framework of the Italian Grid Infrastructure, we are designing a web portal for the grid and cloud services provisioning.
In following this approach, we feel that one of the key point that this kind of application must be able to address is the possibility to hide the complexity of the X509 certificates request and management.
In fact, while the PKI infrastructure is certainly one of the main aspect of the Grid environment, it is also one of the bigger obstacles that very often prevents the new users from approaching it.
In this document, we describe the portal goals, its requirements in term of security and usability and the architecture we have chosen to build it., with particular emphasis on the certificate request and management processes.

Portal Goals and Requirements

The main goals that we want to achieve with the portal are:

- To allow the Grid job submission via web.
- To allow the provisioning of a Cloud environment via web .
- To make easier the request and management of X.509 certificates and the request for a VO membership.
- To minimize the job failure rate

The main requirements of this application are:

- A strong user identification by mean of an accredited identity federation (i.e. IDEM federation). This must represent the external authentication layer.
- The possibility to use a personal X.509 certificate if a user already have one.
- The possibility to specify a VO membership if a user is already a member of a VO
- The possibility to provide a X.509 certificate, through the portal, using an online CA, if a user doesn't have a certificate
- The possibility to ask for VO membership, through the portal, if a user is not a member of any VO

Finally, the portal must be compliant with the EGI policy documents on VO Portal and Site Operations

Portal overall architecture

The portal will be based on the Liferay framework and wil be composed of several modules (portlet) each of them implementing a different service: authentication, job submission, clouds bridge etc.
The portal will interacts with all the external elements as Grid services (VOMS, MyProxy…), IDPs and CA online using Shibboleth protocol and in encrypted form.
An overview of the portal architecture is shown in Figure 1.

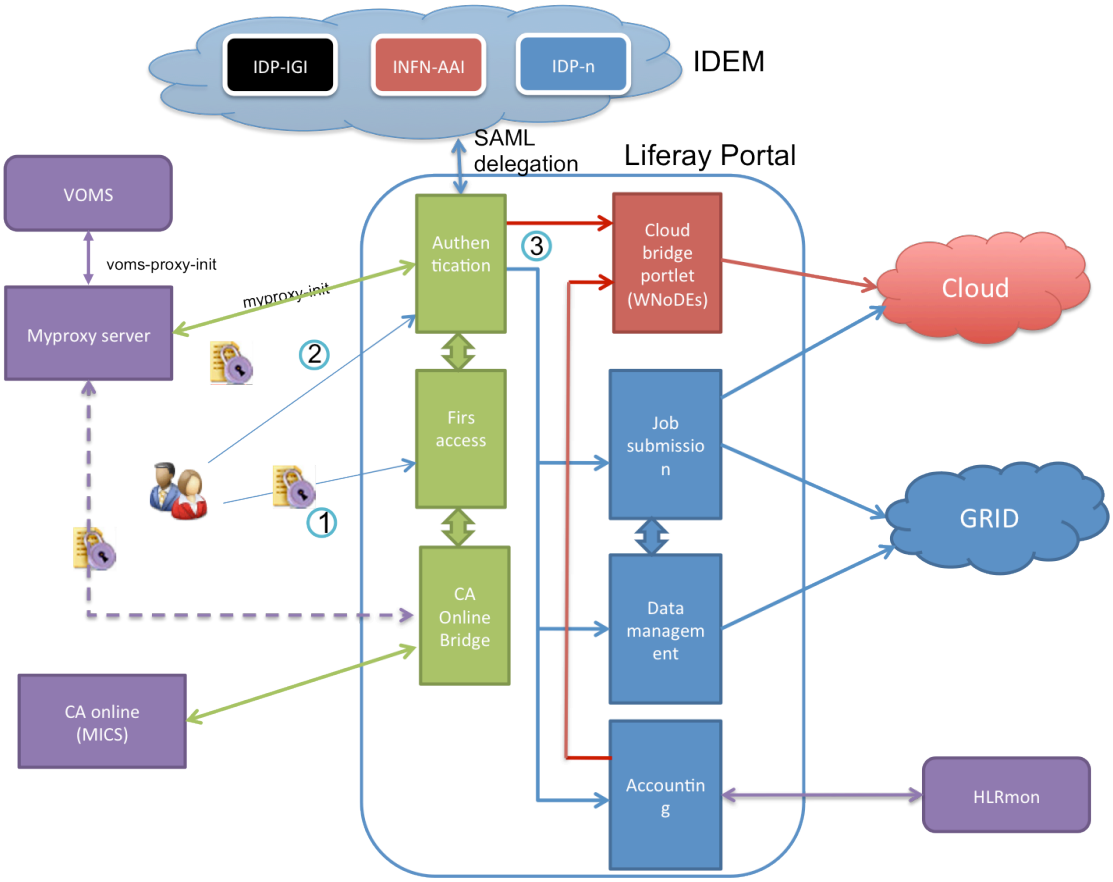

Figure 1 – Portal Architecture

The portal is composed by several elements (portlets), a brief description of each of them:

- Authentication: a module of login to access at the portal. This module redirect the user at the WAYF service in order to choose his IDP. More details are in the next sections.
- First access: a module utilized to collect some user information and to understand if the user has all the instruments to use the grid. More details are in the next sections.

- CA-online Bridge: a module to interface the portal to a CA-online in order to request certificate on behalf of the portal user.
- Job submission: a module where the user can build workflow to submit in a grid environment.
- Data Management: a module where the user can manage file in a grid environment.
- Accounting: a module to collect data about job submission and data management executed through the portal. Moreover it can interact with external grid accounting system as HLRmon.
- Cloud Bridge: a module where the user can compile a form in order to request a specific cloud environment.

And a brief description of the elements that interact with the portal

- CA-Online: CA online MICS, IGTF accredited.
- MyProxy Server: the server where the long-term proxies are saved.
- HLRmon: Accounting information system about CPU resource usage for both Grid and "local" jobs is recorded in the Grid.
- IDEM: Italian identity federation of universities and research institutes for authentication and authorization

Portal Usage

In the typical usage scenario of the portal, we can distinguish three different phases:

1. A first-access phase: in which the user has to identify itself against an accredited IdP, eventually provide some additional information and get valid X509 credentials to use the Grid infrastructure.
   Some of these information (non related to the X509 certificate) will be stored locally in the portal database, so that, in the future, the portal will automatically prepare the right environment for that user.
   The details of this phase, and especially of the X509 credential provisioning, will be outlined in the following section.

After the first-access phase, each time the user wants to operate on the portal, he will go through:

2. An authentication/authorization phase: in which the user has to authenticate himself against an external IDP federated in IDEM and get a valid proxy to operate on the Grid.
3. Operation phase: in which the user will use all the services which is allowed to use according to his privileges

It is important to point out that under no circumstances, PKI credentials and activation data will be stored locally on the portal.

For the scope of this document the phase 1 is the most important and will be described in details in the following section

First access phase

The two main aspects of this phase are those of the user identification and the X509 credential management.

Step 1: Identification and personal data collection.

The portal will check if the user is registered in some IdP of the IDEM federation. A Where Are You From (WAYF) service will provide the possibility to authenticate users against a IDEM federated IdP. If the user is not registered in a IDEM IdP, he will be automatically redirected to a static page where he can find some useful information about how to make the registration in IDEM.
After that, the user has to fill a form in order to collect some personal data. Some information are retrieved from the portal-SP interaction (Name, Surname and email) but other information (Organization, Institute, Country …) could be useful in order to filter accounting data and portal usage statistics.
Each external communication will be done in encrypted form.

Step2: Certificate management and VO membership

If the user has his personal certificate, the portal will give the possibility to upload it, through a protect network, to handle future grid credentials.
The user personal certificate will be used only to store a long-term proxy (the certificate duration long) on a remote myproxy server and then immediately removed from the portal.
The passphrase to activate the certificate will be requested on-line, kept in a non-swappable memory area for few minutes and then deleted.
If the user doesn't have a personal certificate, the portal should be able to provide one in a way that is transparent to him.
We will now try to highlight how we think to implement this phase.
The key point here is that the portal must be delegated to request the user certificate to the online CA.
One possibility is to make use of the SAML delegation mechanism that is available on the new Shibboleth SP and IdP services.
The portal, while authentication the user to the IdP, will receive also a delegation token. Using this token, the CA bridge module of the portal will contact the online CA and will request a certificate on behalf of the user. The use will only have to enter his passphrase for the private key encryption.
Again, the certificate is used to store a long-term proxy on a myproxy server and once the proxy has been created the private key encrypted will be conserved on the portal, on the contrary the passphrase will be not conserved.
The passphrase edited by the user remains in clear text only for few seconds: the time to use it to encrypt the private key.

A user can ask for a certificate only in this phase; this means that if a user has already a certificate he can upload it on the portal and conclude the registration, but he will not be able to ask for a certificate through the portal in the future. On the other hand a user that during this phase asks for a certificate, he will be able to ask a certificate renew through the portal.

Finally, if a user is already a member of a VO he can select the VO he belongs to, otherwise, if a user is not already a member of a VO, he can ask for VO membership, through the portal. In any case the user, to use the portal, has to belong to a VO. It's worth noting that, in the second case, the mechanism to hide the request for VO membership to a VOMS server will be based on VOMS admin API using.

Step 3: Terms and conditions of approval

At last the user will be prompted with a box containing both terms and conditions, such as privacy rules and VO AUP, to be approved.

The step 1 and 2 of first-access process is described in Figure 2

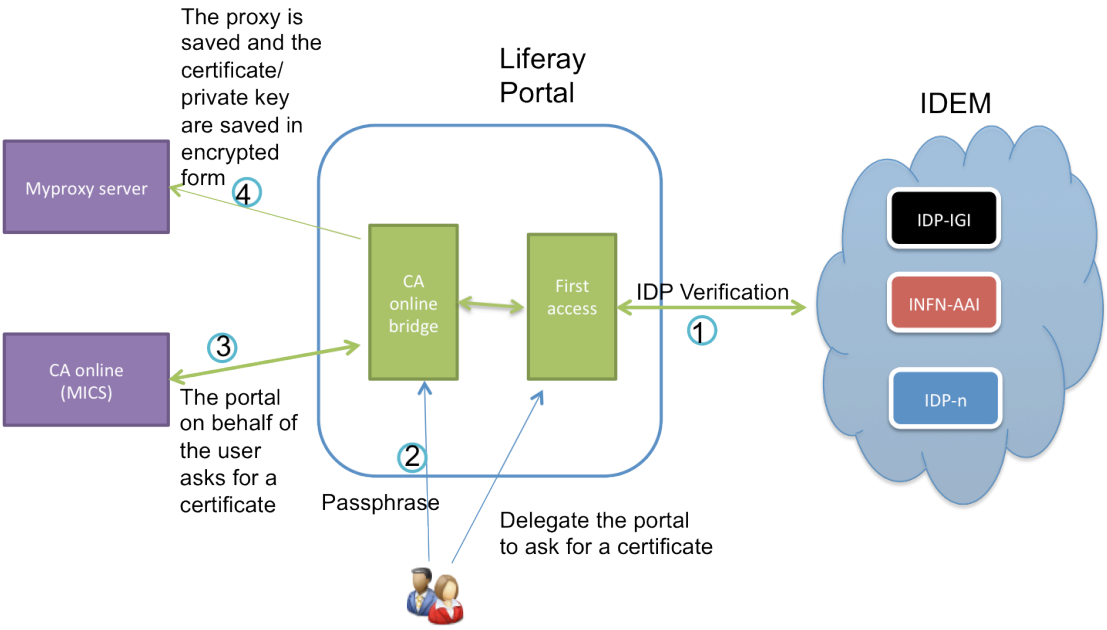

Figure 2 - Step 1 and 2 of first-access process

In view of this explanation it is clearer the reason we prefer to interface the portal to CA online MICS instead of SLCS solution. In this way it is easier and more transparent the certificate management because the user can be notified one month before his proxy will be expired in order to renew the certificate through the portal using the SAML delegation. Using the CA online SLCS the mechanism is more complex, especially for the users that submit jobs that last more than 11 days. In this case the portal has to keep track if the user has one or

mores job in running state when the certificate is near to the expiration date and contact the user (i.e. via mail) in order he ask for a new certificate through the portal using the SAML delegation. The problem is that if the user doesn't ask for a new certificate in time the jobs fail. So using the SLCS solution the portal doesn't target completely two of the main goals: to make easier the request and management of X.509 certificates and minimize the job failure rate.

Authentication/Authorization phase.

After the first-access phase, the user is enabled to use the grid resources, because he is registered in a IDEM federated IdP and has a valid long-term proxy stored in the myproxy server.
The next times the user makes the login he will choose his IDP and the portal redirect him to the appropriate IdP login page.
Once the proper IDP has authenticated the user he will be automatically logged into the portal as well (Point 1 in the figure 3).
For the grid authorization purpose, the portal will ask him the passphrase in order to retrieve the proxy from myproxy server (Point 2 in the figure 3).
The Proxy server allows or denies access based on the CRL entry for that particular user. CRL update is every 6 hours.
The portal, having in his database the DN of the user and momentarily also the passphrase, can retrieve the proxy from proxy-server and at the same time contact the VOMS server in order to sign the proxy with VO extension (Point 3 in the figure 3).
At this point the user has all the credential to submit job in a grid environment.

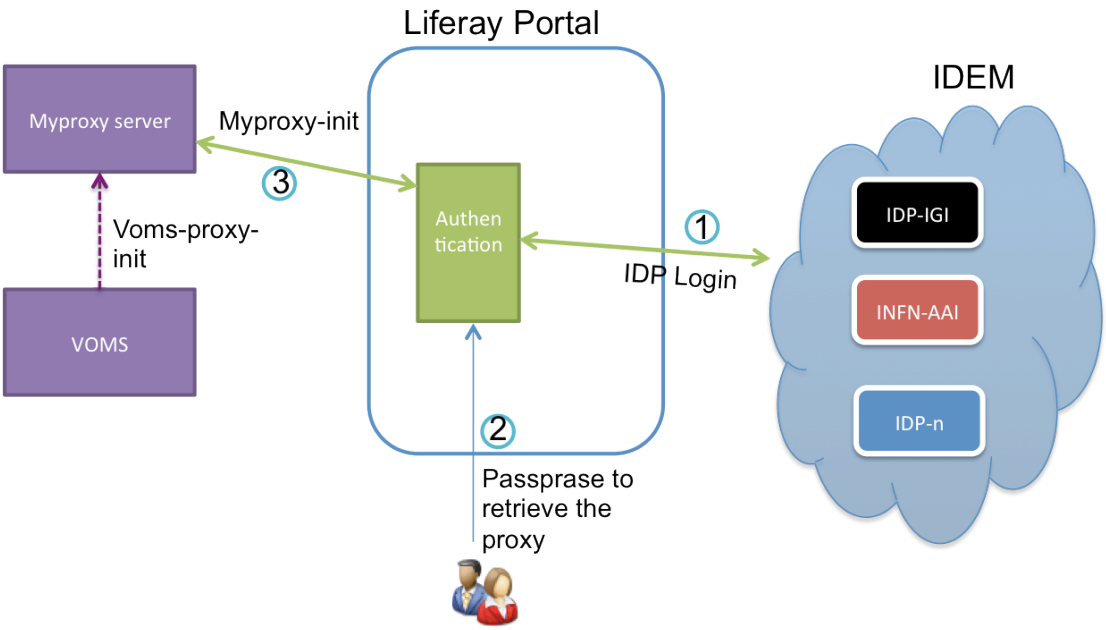

Figure 3 – Authentication/Authorization workflow