IGI Portal: portale web di accesso a risorse Grid e Cloud per le comunità scientifiche

https://portal.italiangrid.it

Marco Bencivenni (IGI/INFN-CNAF)

Workshop GARR Calcolo e Storage Distribuito
Roma- 29/11/2012

Outline

- Introduction
- Authentication/Authorization
- Cloud
- Data management
- Applications

Outline

- Introduction
- Authentication/Authorization
- Cloud
- Data management
- Applications

GRID: distributed computing infrastructure used from many years in the research environment for scientific applications

CLOUD: set of informatics technologies, the most developed by commercial vendor, to use resources delivered as a service over a network

PROBLEMS

- Grid can take some weeks to move the first steps within the Grid: obtain Grid credentials, understanding Grid architecture, learning command sintax
- Grid and Cloud seem separate worlds, but they can coexist and cooperate

GOAL

- **Service for simplifying** the users in **Grid** e **Cloud resources use**, hiding them the difficult of theses systems and making the portal a **single access point** for both

Outline

- Introduction
- Authentication/Authorization
- Cloud
- Data management
- Applications

Authentication/Authorization elements

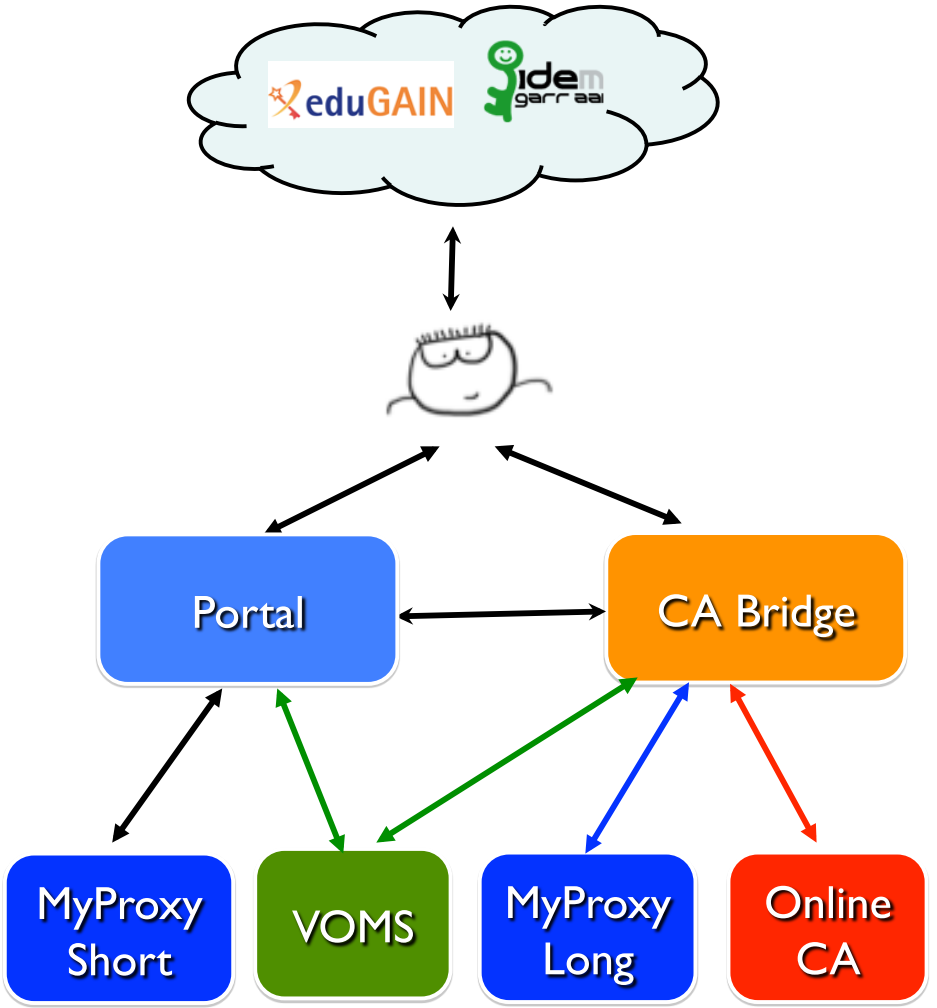
Portal: Grid/Cloud web access point, based on federated authentication

Online CA: Entity that signs certificates, accessible only through CA bridge

CA Bridge: service based on federated access authentication has 3 main tasks: CA access point, Proxy creation and VO registration request

MyProxy: servers where the proxies are stored (long and short proxy)

VOMS: service for the VO registration



Grid Credentials Provisioning

- Grid Credentials: Personal Certificate and VO membership

Personal Certificate

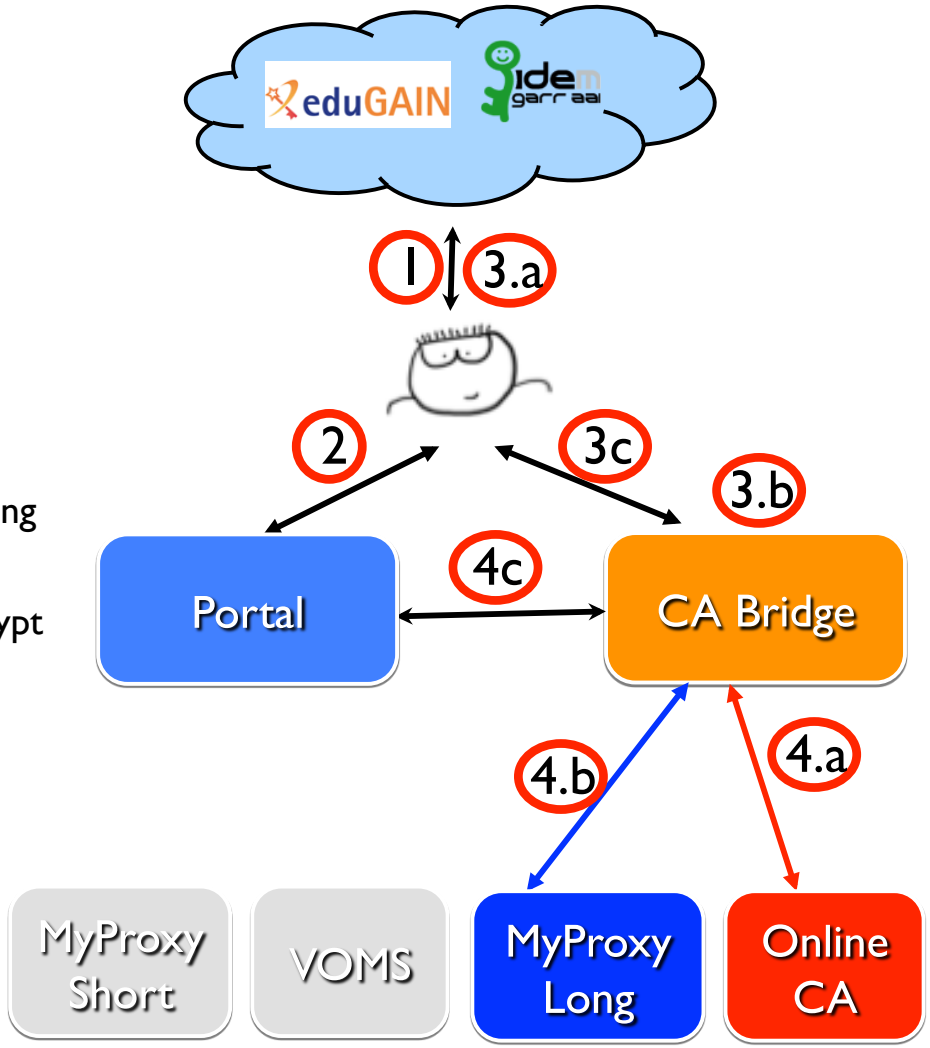
- Upload it if you already have one
  http://www.eugridpma.org/guidelines/pkp/pk-protection-1.1-20100921.pdf
- Request a new certificate using an online CA

VO membership

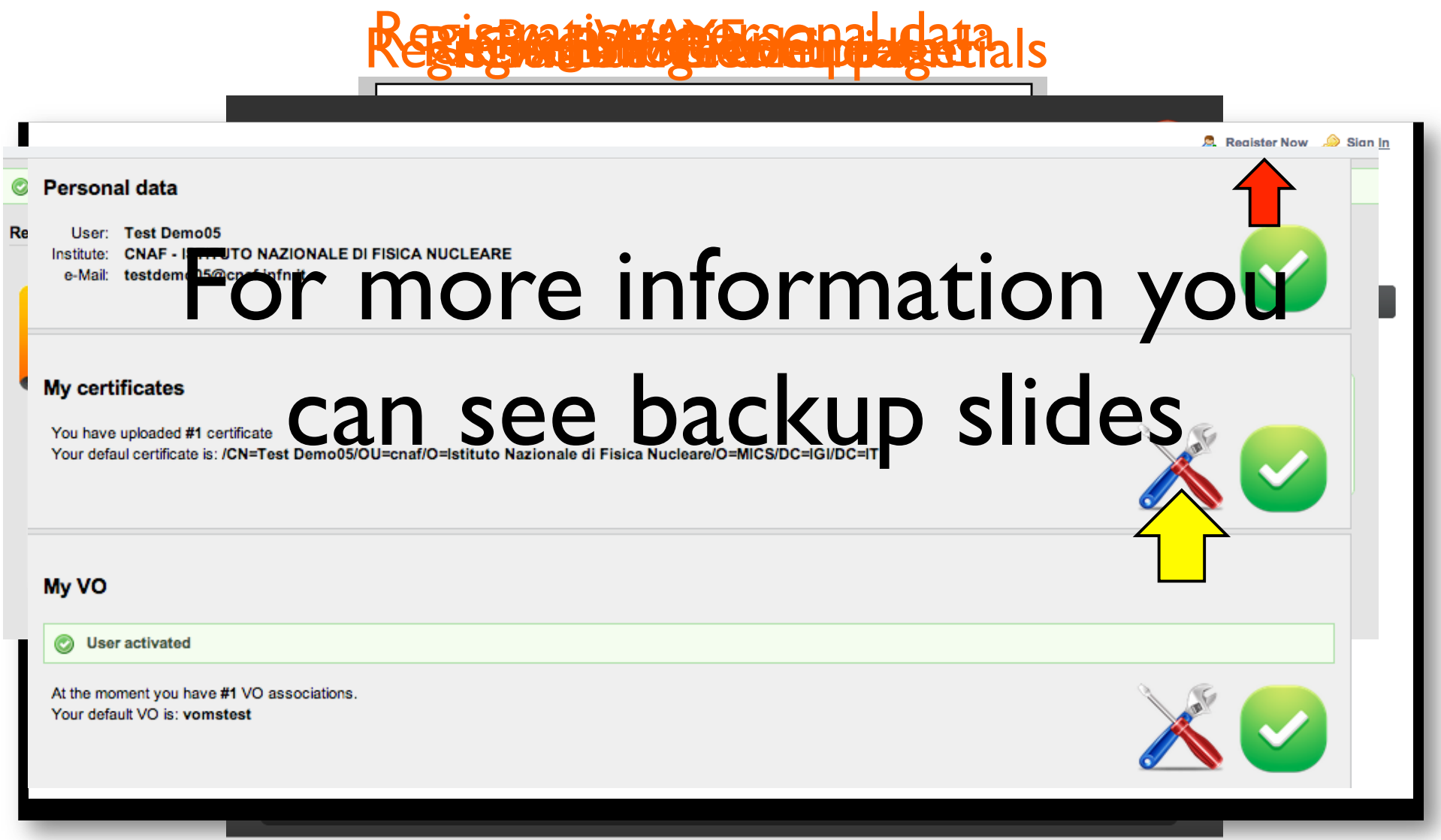
- Set your role/group for the VO you are already member of (if needed)
- Automatically registered in a Catch all VO with access to a limited set of resources

Authorization: Portal and online-CA Interaction

1. Federated authentication for using the portal
2. During the registration on the portal it is possible to request for a new certificate
3. A pop up window (CA Bridge web page) will appear
   a. user are authenticated again (transparently thanks to the browser's cookies)
   b. the certificate's Distinguish Name fields are filled according to the user's attributes retrieved from the IDP
   c. Users are asked for a password that will be used to encrypt the proxy certificate
4. In a few seconds:
   a. A certificate is generated and signed
   b. A long encrypted proxy (13 months) is generated and stored in MyProxy Server Long
   c. The certificate is deleted and the portal notified



Authorization steps



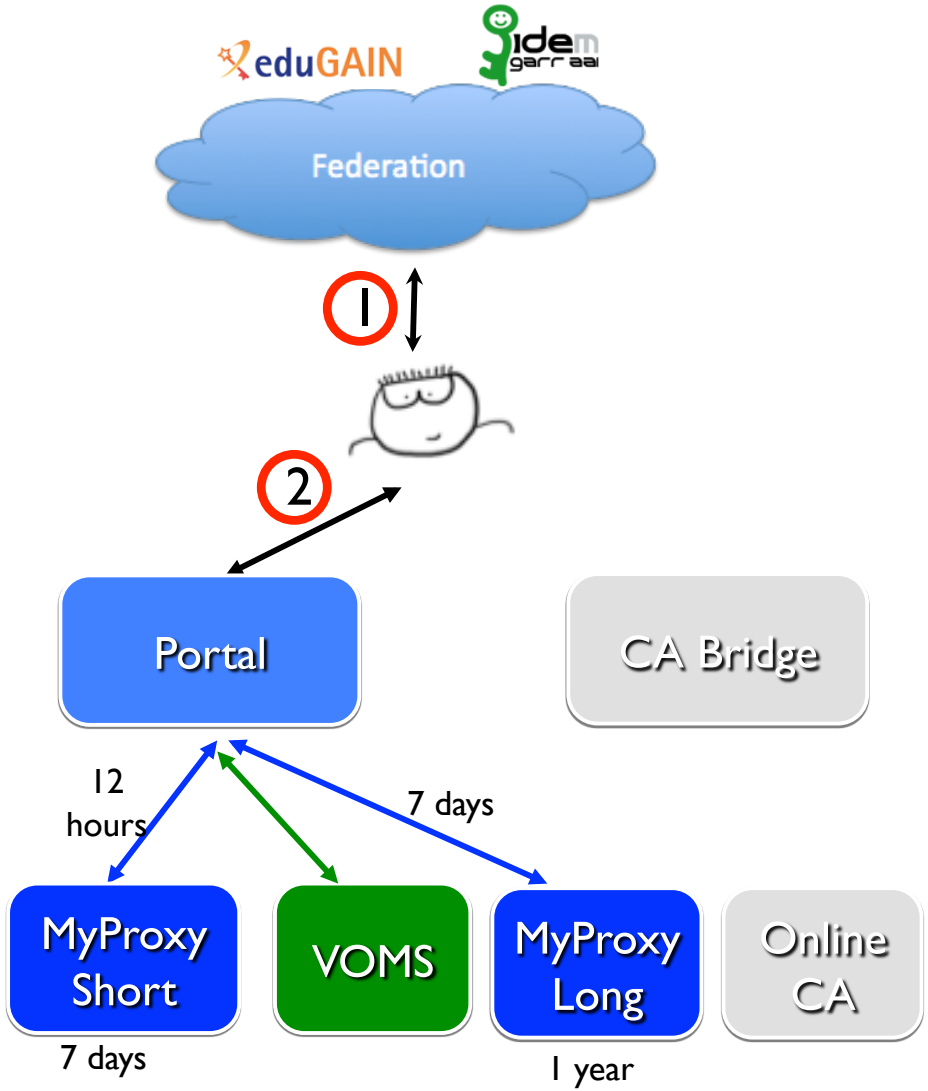
For more information you can see backup slides

After the first time…

From then on, users can access the grid by 2 simple steps:

1. They authenticate themselves on the portal using the federated credentials
2. They retrieve their proxy certificate (protected by the password set during registration) and it is copied on the proxy server short (7 days)

The proxy on the portal has 12 hours validity

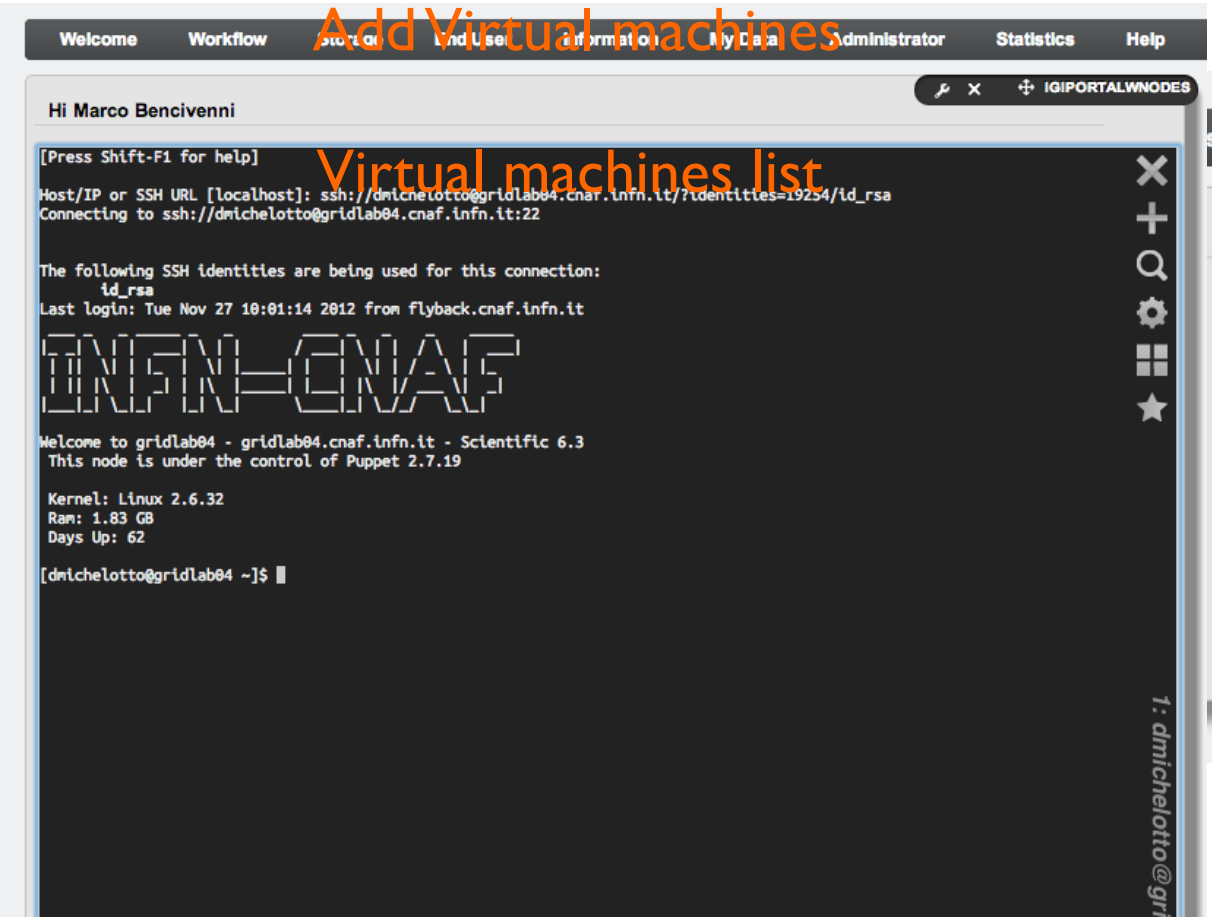


Outline

- Introduction
- Authentication/Authorization
- Cloud
- Data management
- Applications

Virtual Machine on demand

- WNoDeS integration
- Possibility to obtain virtual machine on demand
- Authorization based on Certificate and VO membership
- Shell terminal integrated in the portal web page
- Portal manages also the key pair for ssh password-less

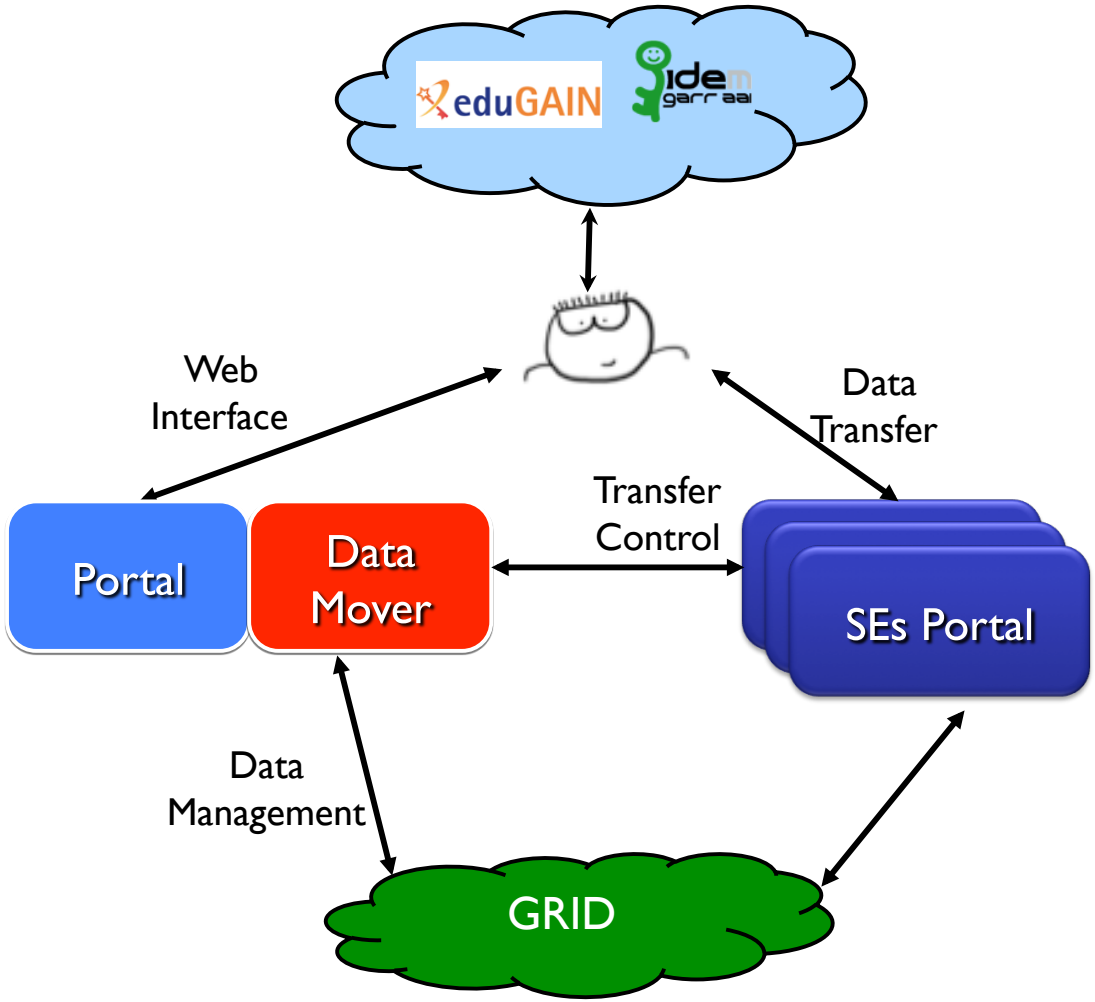


Outline

- Introduction
- Authentication/Authorization
- Cloud
- Data management
- Applications

Data Management: Architecture

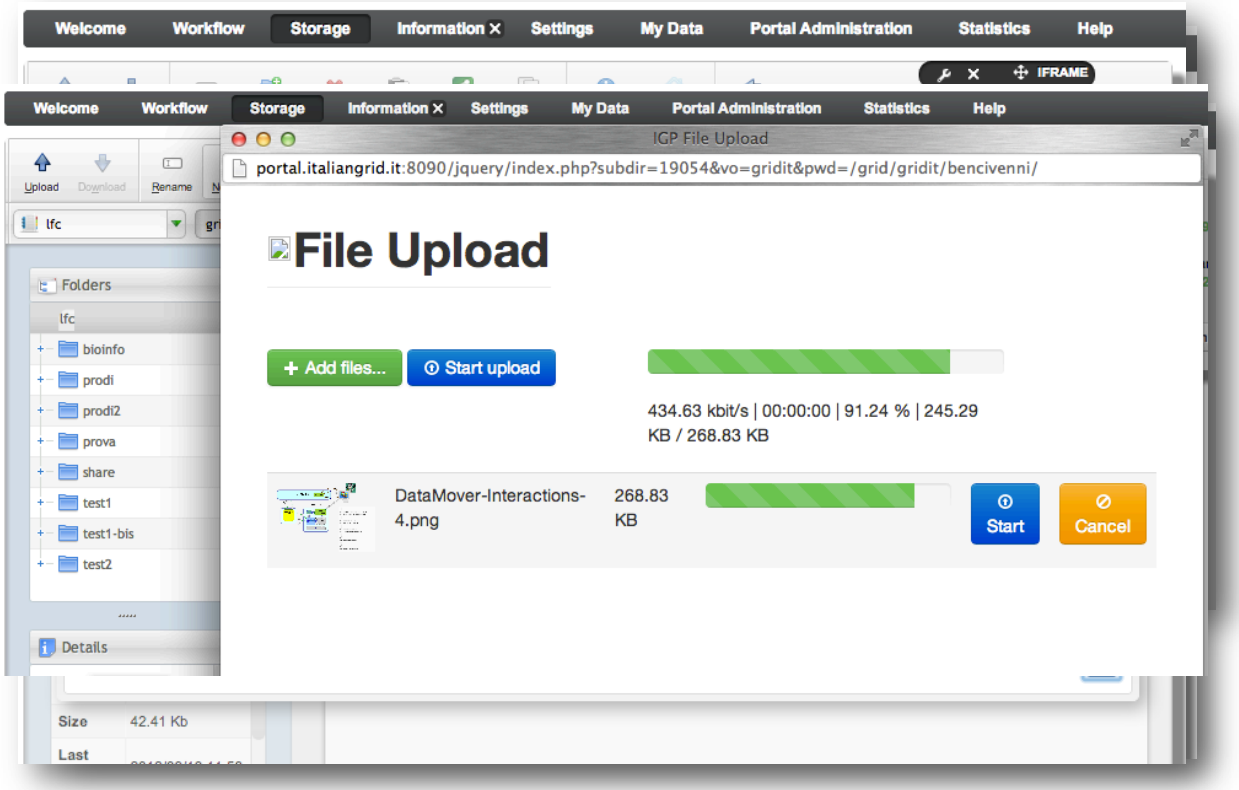
- To avoid bottlenecks for the data movement, files will be moved through an external service
- A Data Mover service and a battery of SEs (portal SEs) are the main components of this architecture
- The **Portal SEs** act as cache memory until the file has not been transferred to the Grid SE (upload phase) or downloaded by the user (download phase)
- The **Data Mover** controls and manages every step of the transfer



Data Management: Interface

- **A new data** management interface has been developed and it is now available
- Complete integration with existing grid storage infrastructure: LFC and SRM protocols. Ready to support other protocols as WebDAV, FTP, MySQL ...

- Rename file/directory or Create new directory
- Upload/Download files to/from Grid (Resumable upload for large files)
- Share files with: other portal users or entire VO members
- Replicate files on others SEs

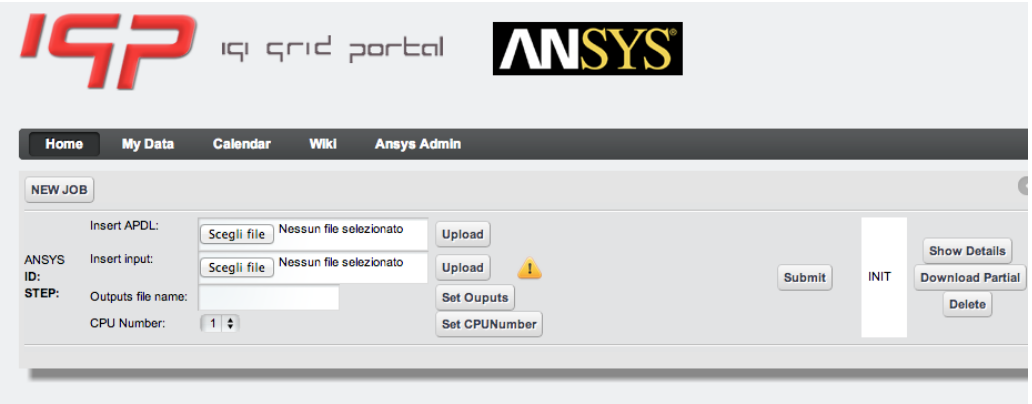


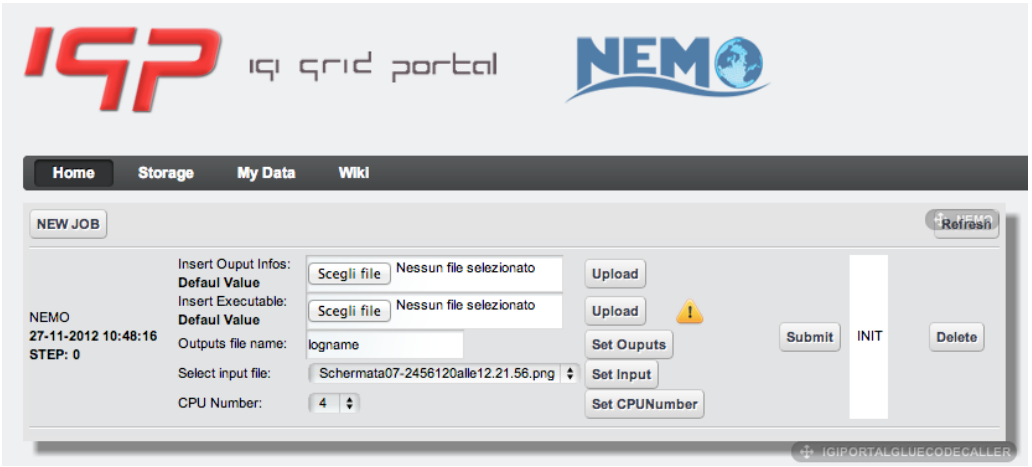
Outline

- Introduction
- Authentication/Authorization
- Cloud
- Data management
- Applications

Applications

- Several applications have been ported in Grid
- Applications used for different disciplines: bioinformatics, geophysics, chemistry ecc
- NAMD, Gaussian, Nemo, Ansys, Blast
- For some of these applications a specific and simplified but powerful interface has been developed

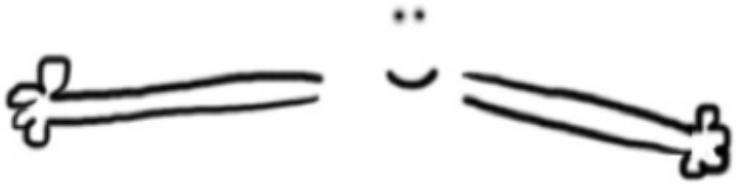




Work in progress

- Authorization: online-CA accreditation: CA and other components almost ready, documents necessary (CP/CPS) will be submitted in Genuary
- Cloud: continuing of test, the component will be in production for February (EMI 3)
- Data Management : Integration of Cloud Storage in Data Management Component in order to move data easily from Grid to Cloud and vicecersa (we are investigating if it is possible to use GarrBox)
- Applications: Several others applications are waiting for being imported in Grid

Questions?



Thanks

portal.italiangrid.it

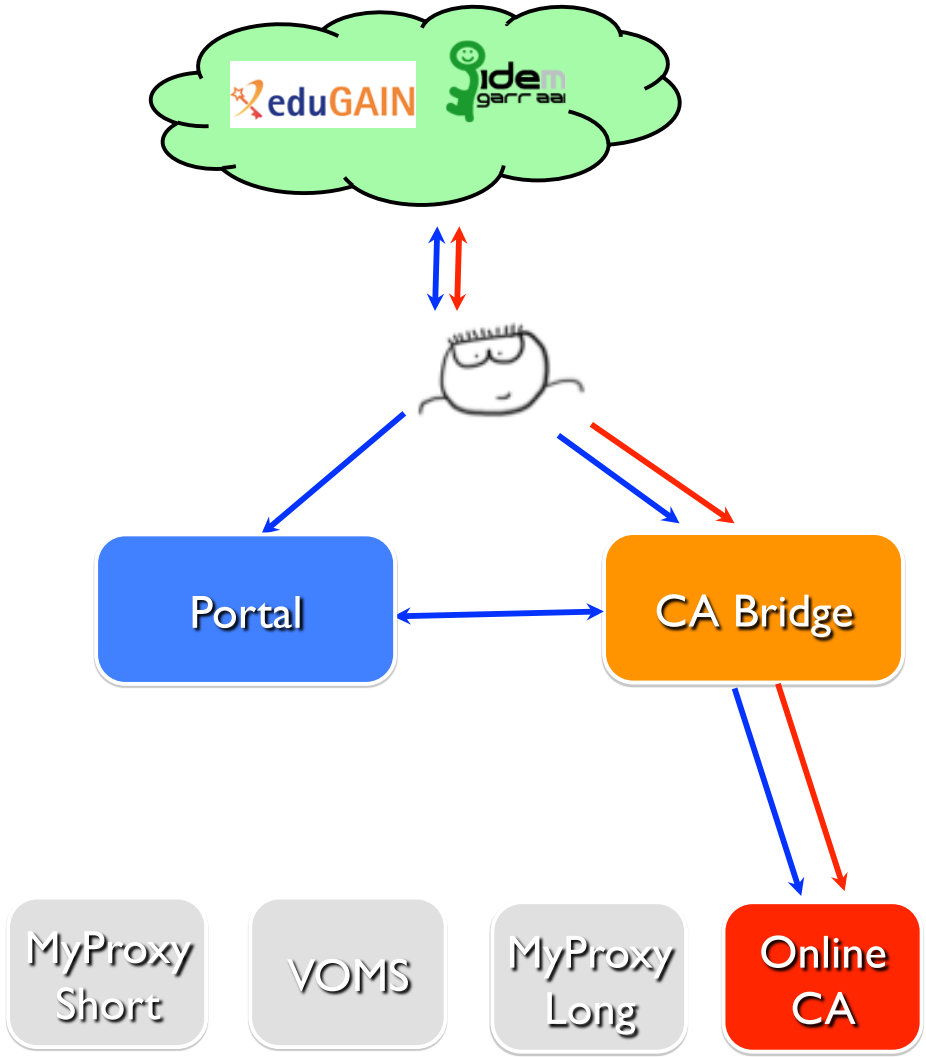
igi-portal-admin@lists.italiangrid.it

Backup Slides

Online CA

Online CA relies on federated authentication.
Certificates can be requested:

- directly by users → at the end of the procedure they will have a certificate installed in the browser
- by the portal → users don't have to manage a certificate directly



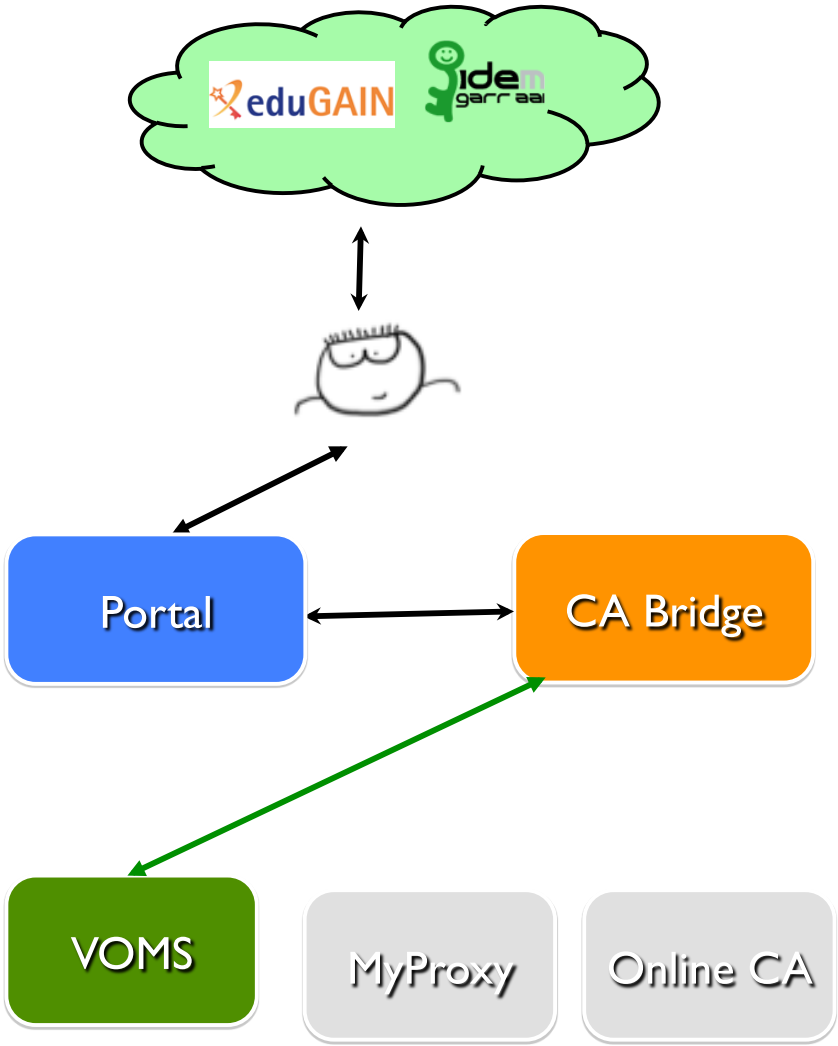
- Software: EJBCA (http://www.ejbca.org/)
- Hardware: THALES NSHIELD PCI 500E F3
- Functional schema already discussed and approved by EUGridPMA

CA online – Security Aspects

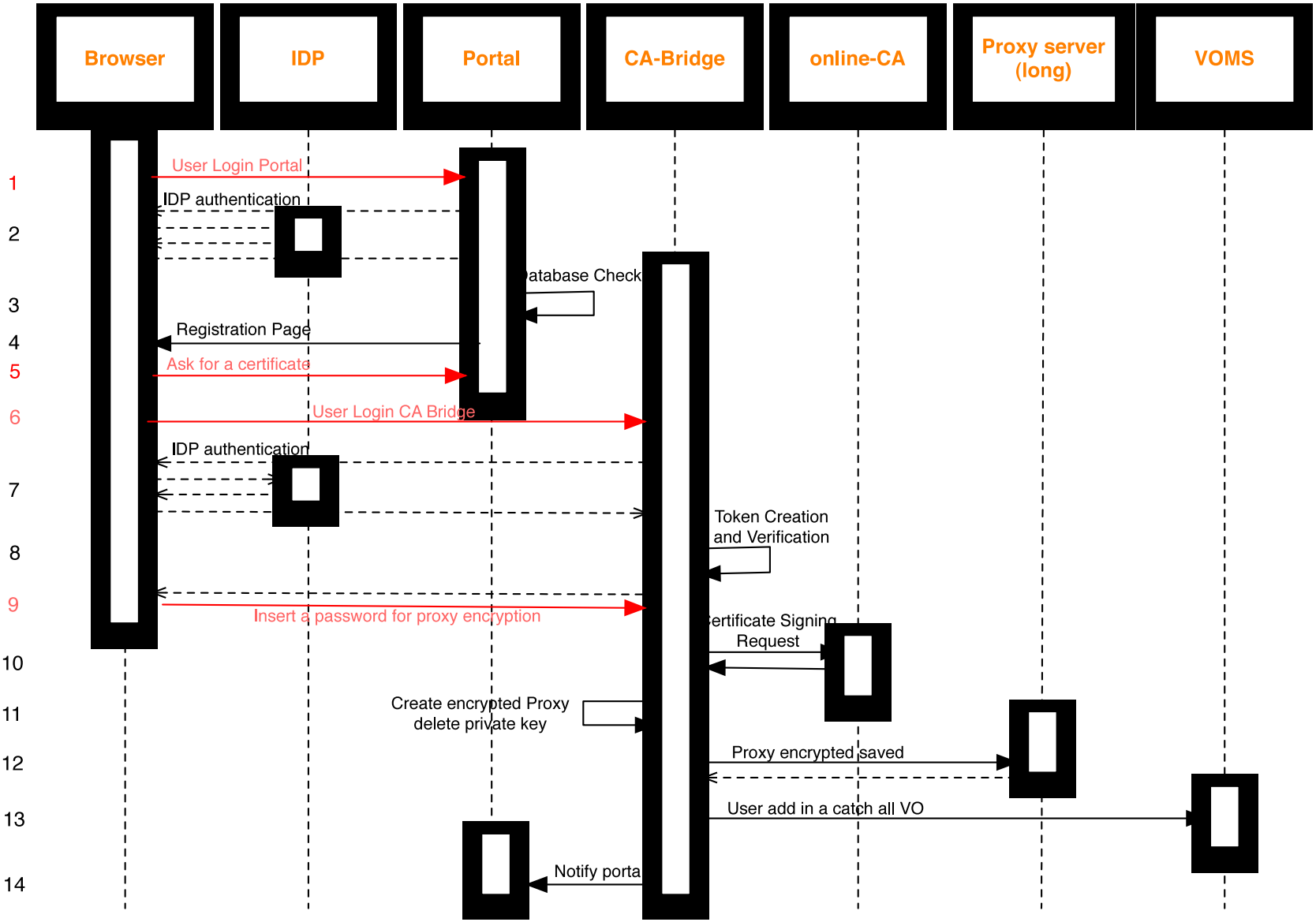
- 2 tokens are created: one in the Portal and one in the CA bridge → **if and only if** these tokens match, users can continue the certificate request procedure
- **TOTP Mechanism** used to create these tokens (Time-Based One Time Password Algorithm - rfc6238)
  - Token=func(secretKey, time, IDP attribute)
  - The **secretKey** is shared between Portal and CA Bridge and refreshed at regular intervals → to check if the request really comes from the portal
  - one or more **IDP attributes** can be chosen → this grants that the user is the same who is logged in the portal
  - The token has **2 minutes valid time** → limited time interval to conclude the operation
- All the connections are SSL/TSL encrypted.
- All the web authentications are shibboleth based.

VO Membership

- To be registered to a VO users must have a valid certificate
- Users with a certificate issued by the online CA are automatically registered to a catch-all VO
  - They can use a limited set of resources
  - They can ask for a new VO membership (EMI-3)
- Users with a personal certificate and VOs membership can set their roles/ groups for each VO and choose which one to use through the portal



CA online – Flow Diagram



Data Management – Flow Diagram