France Grilles – Proposal to establish a National VO

Authors: Gilles Mathieu, Hélène Cordier, Emmanuel Medernach, David Bouvet, Cécile Barbier

Last modified: 21/02/2011

# Starting point

## Identified needs and goals

The definition of a resource allocation strategy for France Grilles highlights the need of a better frame for:

* providing central services to the whole NGI and its users
* integrating new users and new communities to the grid more easily
* assessing the amount of resources and services provided to the French community
* limiting the amount of work put on site to integrate new users and new communities

One important consideration is that the process needs to allow resource owners to decide on how to allocate these resources, while providing a nationwide frame for supporting the French user community.

## Proposal

The implementation we propose to address those needs is to establish a national VO. The purpose of this document is to list all the elements involved in the definition of such a VO, in order to build precise specifications and a concrete deployment plan.

### How does this address our needs?

We address the need for an easier integration of new users by establishing a VO supported nationwide and open to all. The services provided through this VO are made available to all its users, and accounting for this VO can be used to assess the amount of provided/used resources. From a site’s point of view, it is much easier to deal with the needs of a single VO than with multiple demands from small communities fragmented in small VOs.

### What are the problems it is not trying to solve?

The overall resource allocation strategy includes resource allocation to well established communities such as international VOs. Establishing a national VO doesn’t address this.

# Definition

## Requirements on the VO itself

### Prerequisites

If a national VO is established, this VO should be:

* **Catch-all and multi-disciplinary**: it should accept users from all scientific communities
* **Widely supported**: It only makes sense if the VO has resources on as many sites as possible
* **Properly managed**: rules for this national VO and its sub-groups should be the same as other VO and could be derived of biomed experience (see annex), or as a matter of fact of any other established and experimented VO. This includes people responsible namely for :
  + User management & Usage management
  + Resource management
  + Resource follow-up
* **Properly monitored** at sites :
  + Dedicated resources, network and dedicated services availability are to be specifically monitored according to this VO credentials.
  + Operational Level Agreement should be enforced at the same level as for any large VO ( ensure data replication when e.g. SE are upgraded, decommissioned)
  + Sites to enforce the metrics defined on:
    - pilot jobs usage/submission, with a rate to be agreed upon
    - applications support (service catalogue is then to be carefully dimensioned with sites)

### VO membership and access to resources

Access policy to the resources of the national VO follows the resource allocation strategy. In short, a user could get resources below a first threshold (example of first thresholds could be a few CPU hours over a few months time). In most cases, an *a posteriori* analysis on the usage of these resources is necessary, although it can be lightweight.

A community with large requirement and whose need would rise above a second threshold would have to request the scientific committee guidance to have access to this, after an *a priori* analysis of the project. An *a posteriori* assessment would have to be done especially on the return on investment from the communities using resources whose requirements go above this second threshold.

## Proposed rules

### Providing services to the VO

A new VO will need central services such as LFC and VOMS. An element of the proposal is that those central services could be shared and sites’ local VOs could be referenced as other subgroups of this national VO, for which the large sites could foster one central service.

### Providing resources to the VO

Requirements on the set of resources and configuration offered by sites to the National VO should be minimal unless the NGI provides resources, but operational agreements should be stringent in any case.

The list of resources made available to the VO should be built from bottom to top: The VO will have what resource providers can offer. Indeed, the list will adapt to the site’s size and expertise.

Generic/necessary services will be the first to be deployed, while services dedicated to a given project can be added later (possibly by asking concerned project to take responsibility of those services).

If we consider the proposal of integrating local VOs as sub-groups of the national VO, a model could be for each site to support their own sub-group associated to their local VOs and allow a percentage of resources to the rest of the national VO.

It is mentioned at this stage of the implementation proposal that sites will offer resources on a “best effort” basis. The initial list of available resources and services will certainly be small but with the ability to grow in the future.

In terms of quality of service, it is also mentioned that resource usage only has to be guaranteed if job submission is uniform.

### Who does what and who decides?

It is foreseen that local site administrators would have full responsibility on their local VOs administration – i.e. the relevant subgroups of the national VO. But this VO would be handled like a production VO: having rights and rules to follow regarding sites – cf § requirements.

The following criteria have been gathered by site administrators. Their definition could be done by the relevant bodies defining the two thresholds, rather than a priori by sites (they need to be reasonable and adapted to the situation):

* **National VO configuration specifics definition**: Job-slots numbers, FairShare scheduling parameters, amount of TB required or network bandwidth to be defined carefully
* **National VO service set-up** like cream-ce, nearline storage, vobox, and storage type like xrootd and irods to be defined carefully.

### Further possibilities

The national VO could be a way to bring new resources within France-Grilles. In particular, well-sized projects could be asked to contribute with some resources under the form of new sites dedicated to this VO. Financial contributions could also be asked.

## Proposed tools

Discussions have highlighted the importance of having tools, for both users and sites, to assess, check and validate what the exact needs are and what the service offered is.

In particular, having a tool to deal with resource reservation is essential, both for long term resource allocation and for planning and anticipating temporary activity peaks. Different tools are available including PL-Grid’s solution (Bazaar): choosing the tool and studying the feasibility of its usage in our context will have to be a dedicated task in our implementation/deployment plan.

# Validation

### How does this idea fit within the overall EGI environment and the NGI strategic goals?

Establishing a national VO will simplify access to the grid for new users and new communities, which is a key element of France Grilles’ global strategy.

There is little to be said of the international context, since this solution is by essence national.

### What are the pros and cons of this solution? Feasibility study to be completed

Pros:

* Sharing of central national VO services management
* Traceability of French resources utilization by the communities
* Flexibility to users’ needs and sites
* Easier management of users’ needs from a site’s point of view
* Clear Implementation of our scientific strategy: building a multi-disciplinary grid
* Gain in visibility for the sites supporting the national VO
* Gain of a political and financial support to sites through a proof of concept
* If a job submission framework with a centralized submission queue is used, management of priorities between groups/projects can also be done at VO level. This would simplify management tasks for sites.

Cons:

* Some initial technical work would be needed for storage access
* Some technical workaround – transparent! - might be needed to account for local users

Modalities of thresholds definition can be tricky

Further needs have to be assessed.

### Modus operandi and proposed actions

* Validate the principle of establishing a national VO in the scope of France Grilles global resource allocation strategy
* Define precise VO specifications and initial requirements
* Gather feedback from sites about who might want to:
  + Provide central services to the national VO
  + Provide resources to the national VO
* Complete the feasibility study and define a deployment plan, especially with regards to:
  + Defining the management structure of the VO, and who will be in charge of what
  + Trying, choosing and deploying a resource allocation or “reservation” tool for simplifying negotiations between the VO, France Grilles and sites
  + Deploying central services
  + Implementing the whole workflow, from new users requesting to join to projects integrating the national VO.

# Annex - VO management tasks with 2 FTEs

T. Glatard, for the biomed VO managers - 14/10/2010

The following tasks are ordered by decreasing priority. Manpower estimations are probably a bit underestimated (assumes a skilled, experienced VO manager).

## User management and coordination (0.1 FTE; 0.5 day per week)

* analyze VO membership requests (scientific relevance)
* maintain an updated list of VO members (periodical polls, de-registration of users)
* manage VO groups (projects, thematic, countries)
* list ported available applications and foster collaboration
* list application-level middleware and foster collaboration
* provide first-line VO information (entry page, wiki, etc)
* organize LFC structure (home directories and permissions)

## Resource management (0.1 FTE; 0.5 day per week)

* evaluate VO needs
* recruit sites supporting the VO (CEs and SEs + core services LFC, VOMS)
* establish technical agreements
* collect decommissions

## Monitoring (0.5 FTE)

* (0.1 FTE ; 0.5 day per week) develop and deploy probes and monitoring system
* (0.2 FTE ; 1 day per week) react to detected problems (submit GGUS tickets, apply incident procedures)
* (0.1 FTE ; 0.5 day per week) improve QoS (design/adjust operational procedures ; core service dimensioning)
* (0.1 FTE ; 0.5 day per week) flag and cleanup unused data (unregistered users, temporary data) ; flag and replicate critical data

## Application support (0.3 FTE; 1.5 day per week)

* specific services operation and consolidation (e.g. Dashboard, Hydra, biological databases, portals, MOTEUR)
* specific user support on these services (!= GGUS)

## Progress beyond state of the art (1 FTE)

* Dissemination, training; attract new users
* Develop new services (data migration service, LFC interface, user management DB)