Setting up VOs in EGI: An inventory of basic and general services required

I. Blanquer, A. Carrion[1]
J. Gomes, G. Borges[2]
UPV, Valencia, Spain[1]
LIP, Lisbon, Portugal[2]

After the approval of a new VO, there are several services that need to be provided for. Those services are general ones and could be hosted on resources shared by different VOs. This document tries to identify the critical VO-services and to analyse the cost in setting-up and operating them, from the point of view of the provider (which can be either EGI or the VO), the VO community and NA3 people.

Services are described into tables. Each service has the following entries:

Service Name		
Objective	Purpose for the service	
Need	Need of the service for a VO to work.	
Hw Resources	Hardware resources to host the service.	
Issues	Comments on specific difficulties.	
	tasks	*cost*
Infrastructure operation cost (VO or EGI)	Task that the centre physically hosting the service should perform.	Costs that such operation tasks will have for the centre.
VO operation cost	Tasks that the members of the VO should perform with the service	Costs for the VO members.
NA3 cost	Tasks that the EGI NA3 people should assume to set up the service and to maintain it	Costs from the EGI-NA3 point of view.

VOMS Service		
Objective	Manage membership and authorization on a VO.	
Need	Compulsory for the use of any resource in gLite.	
Availability	http://glite.web.cern.ch/glite/packages/R3.2/sl5_x86_64/deployment/glite-VOMS_mysql/glite-VOMS_mysql.asp	
Hw Resources	Not defined but it is a lightweight software. The resource can be shared with other services and VOs.	
Issues	It is a critically service regarding availability, although it requires low bandwidth, disk and CPU consume. Replication is recommendable.	
	tasks	*cost*
Infrastructure operation cost (VO or EGI)	• Ensuring availability through automatic monitoring. • Periodic backups needed.	• Low cost for setting up a new VO. • Low operation costs but need for quick reaction (on duty operation).
VO operation cost	• Remote (web) operation for creating and deleting users. • Answer requests for information on users misusing resources.	• Low operation cost but on a daily basis.
NA3 cost	• Interface for requesting new VOs for the VOMS server. • In the case of misusing of resources, NA3 could play an active role.	• Low cost. Interaction at the beginning when setting up a new VO. • Punctual intervention when problems.

MyProxy Service

Objective	Stores medium-term proxies and renews short-living ones (including VOMS attributes).	
Need	Important for jobs living in the system (including queues) more than 24h. This will depend on the availability of resources and the nature of the jobs in the VO.	
Availability	http://grid.ncsa.illinois.edu/myproxy/	
Hw Resources	Not defined but it is a lightweight software. The resource can be shared with other services and VOs.	
Issues	It requires low bandwidth, disk and CPU consume. Replication is easy and can reduce effort on availability.	
	tasks	***cost***
Infrastructure operation cost	• Ensuring availability through automatic monitoring. • No backup needed. • Reconfiguration with the WMS supported is needed.	• Low cost for setting up a new VO. • Low operation costs.
VO operation cost	• Basic monitoring can be included in a dashboard.	• Basically no operation cost.
NA3 cost	• Compilation of usage information and tests.	• Low cost mainly at the setup (negotiation with WMS and MyProxy providers, preparation of use cases).

LFC Service	
Objective	Manages logical directories and replicas of files stored in the Grid. It enables defining Access Control Lists to prevent and grant access to data files.
Need	Critical for an efficient use of the Grid. Most of the VOs require large bulk files accessed from many geographical locations.
Availability	http://glite.web.cern.ch/glite/packages/R3.2/sl5_x86_64/deployment/glite-LFC_mysql/glite-LFC_mysql.asp
Hw Resources	A dual-core computer with 1-2 Gb RAM also hosting the database on the same or separate host is recommended. The resource can be shared with other services and VOs, depending on its capability.
Issues	It is a critically service regarding availability and it requires high bandwidth, but average disk and CPU. Replication is only possible for read-only, so it becomes a bottleneck in many cases. It should be provided by the VRC/VO as the usage demand increases.

	tasks	*cost*
Infrastructure operation cost (either EGI or VO)	- Daily monitoring. - Reports are often requested (e.g. when SEs of a VO are decomissed, users leave from a supported VO, SEs are becoming full, etc.). - Daily backups needed.	- Medium cost for setting up a new VO. - Medium operation costs and short reaction time (on duty operation).
VO operation cost	Interactions with LFC manager when issues appear.	No specific operation from the VO point of view.
NA3 cost	- Interface for requesting new VOs for the LFC server. - Negotiation to request SEs to support the VO in the infrastructure.	Medium cost on the set-up. Weekly operation to monitor the usage of resources.

WMS Service		
Objective	Schedules jobs on the different resources and manages their input and output. Manages all the submission procedures and gathers all the information from the status of a job.	
Need	Critical for executing any job in the grid.	
Availability	http://glite.web.cern.ch/glite/packages/R3.1/deployment/glite-WMS/glite-WMS.asp	
Hw Resources	A quad-core processor is recommended to better handle parallel matchmaking and all the different sub-services running on a WMS, with a minimum of 4 GB RAM memory and a minimum disk space that depends on the load and type of jobs submitted (Minimum 30-40 GB) The resource can be shared with other services and VOs, depending on its capability. It should be provided by the VRC/VO as the usage demand increases.	
Issues	It is a critically service regarding availability and it requires high bandwidth and CPU. Replication is simple and recommendable, since it stores only local information.	
	tasks	***cost***
Infrastructure operation cost	• Intensive monitoring. No need for backups.	• Medium cost for setting up a new VO. • Medium operation costs and short reaction time (on duty operation).
VO operation cost	• Monitoring through Dashboard.	• No operation from the VO point of view.
NA3 cost	• Interface for requesting the support of new VOs for the WMS server. • Requesting CEs from the system in the start-up. • Periodic intervention to assist on the increase / decrease of resources according to the usage.	• Medium cost on the set-up. • Low cost after the operation.

Dashboard Service

Objective	Eases gathering information and accessing the resources of a specific VO in a Grid. Performs a monitoring of resources, jobs and data in a grid from a VO point of view.	
Need	It is not a critical service for the use of the resources, although it can strongly help the monitoring of the resources on a VO. In some cases, a resource could work for the operations VO but not for a specific VO, and this can only be possible from a user of the VO.	
Availability	Not yet defined.	
Hw Resources	Depending on the final features, but considering ARDA dashboard or other generic solutions, it requires a web server on a conventional computer.	
Issues	Many different solutions, but there is no official release yet.	
	Tasks	*cost*
Infrastructure operation cost	• Setting-up the general and specific tests for each VO. • Ensuring the availability of the service.	• Mainly at the first deployment. • Minor actions from then on.
VO operation cost	• Configuration of VO specific tests. • Daily supervision of the tests. • Monitoring for checking availability of the core services (VOMS, LFC, MyProxy, WMS). • Checking the availability of SEs from the specific VO. • Checking the availability of CEs from the specific VO. • Periodic monitoring of the usage of the resources by users.	• High daily costs in real-world operation. It can involve a few hours per day in a medium-sized VO to check alarms, detect failure patterns, rise and follow-on tickets.
NA3 cost	• Assist on the deployment of new sites. • Compilation of usage information of CEs and SEs to assist on the scaling (up and down) of the resources of the VO.	• Low operation costs.

Wiki Service

Objective	Provide a portal and a schema to centralize and share information for a VO. Large VOs will prefer using their own portals, but new or small VOs will benefit from having a portal and a schema where all the information a VO user may need is available and customized for his/her community.	
Need	Very important to centralize information on a VO.	
Availability	Not yet defined.	
Hw Resources	Lightweight. A web server and a wiki site.	
Issues	Definition of a general framework suitable for different communities.	
	Tasks	*cost*
Infrastructure operation cost	• Deploying a new site and configuring the authorization. • Basic checking of the availability.	• Low costs only on the start up. • Minor maintenance.
VO operation cost	• Filling-in and management of content. • Updating information on the particularities of the VO.	• Medium cost to keep information up to date.
NA3 cost	• Initial definition of structure and deployment. • Periodic review of contents. • Update of general information (review and report when changes on services, certificates, documentation urls, etc.).	• Medium cost in order to ensure the quality of the contents and the update of general information.