LCG Operations Workshop

LCG VOBox Operations Recommendations and Questionnaire

Date:	**30 September 2005**
EDMS Reference:	https://edms.cern.ch/document/655277
Internal Version:	**0.2**
Status:	**Draft**
Author:	**Steve Traylen**

Document Log			
Issue	**Date**	**Author**	**Comment**
0.1	30-Sep-05	Steve Traylen	• Initial version
0.2	06-Oct-05	Steve Traylen	• Greater details about what should be backed up and disaster recovery techniques. • Details for sites where they are expected to monitor the VO services. • Details for service intervention. • Shared VO boxes across sites. • Time period for return to production service. • More that one VO box could be possible. • Clarify needs for special requirements such as large data throughput. • Consequence of sites being unable or unwilling to provide VObox. • Correct EDMS number in the header.
0.3	09-Oct-05	Steve Traylen	• Split question in 2 in question 2 and 3. • Try and clear up what can be expected under LCG MOU. • Add section that service can be terminated. • Add operating system question. • Software certification.
0.4	11-Oct-05	Steve Traylen	• Choose and OS similar to EGEE middleware. • Expand network deployment question to a more general question. • Make a note that LCG MOU not relevant at all sites.

Table of Contents

References

[1] LCG/EGEE Security Policy Documents' VO-Box-Security-Policy
https://edms.cern.ch/document/639856/

1 Introduction

The concept of the VO dedicated systems or VO Boxes has introduced new operation scenarios due to the split responsibility and knowledge about what may be a critical service at a particular site.

This document describes recommendations to Site Managers from an operational point of view to implement a VO Box at a particular site.

For Virtual Organisations proposing to use VO Boxes it lists recommendations and procedures that must be adhered to, as well as details of what information to make available to the Site Managers in order to assess the best method of operation for a VO Box. Once the operational procedures and level of service has been specified to a site by a VO then sites are in position to decide if they are able to support a VO Box. If a site is unable or unwilling to provide this level of service then it may choose not to support a particular VO Box and subsequently the VO if a VO Box is a requirement. It MUST not be assumed that providing the details as defined below will result in a VO Box at sites being offered as a service. For example many sites have no back up facilities.

2 Definitions

The term *VO Box* designates a system provided by a site for use by one or more designated Virtual Organisations (VOs) for interactive access and running persistent processes. The terms MUST, SHOULD, MAY and MAY NOT must be interpreted according to RFC 2119.

3 Operational Security Concerns

Before any VO Box enters operation at any site the recommendations of the Joint Security Policy Group for VO Boxes will be adhered to as they are described in [1].

4 Hardware Resources

A clear indication from the VO of the hardware resources MUST be given. In particular these should include the obvious requirements of memory and disk space. In the case of disk space should this space be on the local hardware or is some networked file system acceptable? Some indication as to the CPU saturation that the service may use when running normally should be supplied. Is spare CPU capacity required in order to ensure a low latency in any service response time?

The site may like to co-locate VO service boxes on a single piece of hardware in order to make better use of resources. This may be done using hardware virtualisation techniques, a secondary virtual IP interface or just a single VO Box enabled for and accepting VO services from multiple VOs. For each of these co-locating methods the VO must give technical reasons why any of these co-locating solutions could not satisfy their needs for a VO box. The sites should provide dedicated physical hardware where the load from any one VO is such that it degrades the performance of other VOs. Requests for more than one VO box at a site for any single VO will be considered on case by case basis.

Some sites that work closely together may like to pool resources and run one VO box to be utilised by several sites. The VO SHOULD indicate is this is an acceptable situation.

5 Operating System

Sites MUST run an operating system that the VO can use. VOs MUST provide a preference for which operating system is preferred or could be used. Choosing an operating system supported by EGEE/LCG/EDG middleware is generally the sites preference. The site MUST apply all security updates and MAY apply bug fixes as supplied by the operating system vendor.

6 Software Certification

Sites MUST require that all software to be run in a production environment has been certified to work at some level. The VO MUST provide details of what certification has taken place for the VO services. Updates to the VO services MUST happen with a versioned release of VO middleware including release notes for the changes

between versions. These release notes should at least provide details of how the VO service fits into the overall grid and VO workflow.

7 Network Connectivity

When deploying a VO Box sites should consider the most suitable location within their network topology for deploying such a service. While a VO's answers to the questionnaire included in [1] will provide a complete set of network connections to and from any VO Box this will not provide any data rates for these connections. VOs SHOULD provide as much information as possible about special requirements for services, such as the need to transport large quantities of data. Special routes exist between some T1s and the T0 which require host level configuration to be added.

Sites are encouraged to use the network port lists provided by [1] to enforce the required level of access. Tools such as IPTables to restrict protocol/source and destination of network connections and SELinux can be used to restrict particular users and binaries to particular ports.

8 Monitoring the Service

Since the VO Box and the services running inside are managed distinctly by both the local site administrators and the VO service managers there must be a clear distinction between the responsibilities of the two groups. The site is responsible for the operating system and for all services that are provided by a VO Box before a VO accesses and installs any of its VO services on the VO box. The site will monitor a VO Box for common failures such as kernel crashes, disk i/o errors, disk space, etc. In certain cases the site may then believe that a full disk or other failure is caused by a particular VO service and so will notify the VO service manager who is then responsible for correcting the situation. The working operation of the VO service must be monitored and maintained by the VO. It may be sensible for any VO service to be monitored by the site at some level so that effective use of a site's infrastructure can be used such as monitoring consoles, operators and call out facilities. In this case the VO SHOULD provide details of how a service's health can be monitored and when incidents should be triggered. Once an incident has occurred any details of subsequent action should be well defined. For example the action could be that a VO MUST be notified by email or by telephone immediately, during working hours but not on public holidays.

9 Business Impact of the Service

Sites should assign a level of service to VO boxes after considering the impact of a VO Box being unavailable. The VO is best placed to understand how any loss of a VO Box is likely to impact the VO as a whole. For instance, will the site be unable to run jobs? Will the site drop out a particular class of new job submissions such as analysis jobs? The VO should provide this information.

In the case where the service has been completely lost due to a complete hardware disk failure or similar then as a minimum the site SHOULD return an empty VO box into production in a time period that reflects their LCG MOU service commitments if they exist and are relevant at a site. The VO services then installed or recovered by the VO is not part of any sites commitment. Additional services such as a nightly backup should not be assumed. The VO should state in advance if anything other than a reinstallation of the VO Box is required when there has been a complete service loss. For these extra services clear agreement MUST be in place between the site and VO.

In the case where a service state is to be recovered from a backup then the site should have a plan that returns the service in to a state that the VO is expecting. This plan should detail which directories or databases are going to be restored. The VO MUST provide guidelines as to how this can be achieved. Lists of directories and methods for storing consistent databases should be given. If the downtime of the service requires a synchronisation with other services or events changing state during downtime then details must be given by the VO for this. For instance the service may have to be synchronised with jobs that completed during the downtime and have registered files.

10 Service Interventions

At certain times sites will have to intervene with any service for planned maintenance such as hardware failures or operating system upgrades. These would normally be planned and published in advance but may happen immediately where security implications or unforeseen operational circumstances are present. Clear agreed lines

of contact with the VO must be put in place either by named individuals or standard routes such as GlueSchema flags or broadcasts to VOs. The method must be defined and agreed in advance. If a site must perform steps before or after the outage such as a drain then this must be made clear by the VOs to the site.

11 Service Termination

As will all services ran at a site the site can at any time terminate a VO service or VO Box if they have good reason for doing so. This could be for instance if a VO service is doing something different to what is expected when compared to the service information provided by the VO. Other reasons for a service termination include the service impacting other services in a detrimental way or there could be some security concerns. In all cases the site MUST inform the VO of its intentions or actions. It MUST also inform other sites or security contacts if the behaviour might be expected to be grid wide.

.

An Operational assessment questionnaire for VOs

The answers to this questionnaire must be available to the *Site Managers* at all times. The information provided here MAY be stored in an on-line repository, accessible to all *site managers*. Specific *Site*s MAY require the information below and (at their discretion any additional information) to be sent explicitly to them.

By submitting this questionnaire or by using any service provided by a VO box service system, the VO maintainer agree that their personal information will be shared with all *Sites* and shall be used for administrative and operational purposes only.

Name of the Virtual Organisation:

ATLAS

Name and contact details of the maintainer:

The VO BOX is maintained jointly by ATLAS site contacts and ATLAS Operations team.

The questionnaire has been answered by:

Name: Miguel Branco

Email address: Miguel.branco@cern.ch

Phone number(s): +41 22 76 71268

1. **What are the hardware requirements for your VO Box?** (*Example – A node comparable to your site's fastest batch workers is sufficient. Twenty GigaBytes of possible network attached disk space is required.*)

A worker node class machine is sufficient.

2. **What operating system is preferred or would be acceptable?** (*Example – We prefer RHEL3 derivatives such as CERN Scientific Linux 3, Scientific Linux 3, CentOS 3. RHEL4 would be acceptable)*

RHEL3 e.g. SLC3

3. **What certification has been performed of the VO services? (*Examples – The software has been used within the EGEE pre-production service. The software has been used within the ARDA prototype testbed. The software has been test deployed at 3 Tier1 centres. The software is yet to be written or has had no testing.*)**

The s/w has been used in production on all ATLAS Tier1s. It is first deployed on a small testbed constituted of volunteering sites.

4. **Can the VO Box for your VO be co-located on the same physical hardware as another VO's VO Box?** (*Example – This is not possible since we require our services to bind to a particular port that cannot be altered or can not bind to a particular network interface..*)

In principle yes, although there may be conflicts in Apache and MySQL installations. It is possible to reconfigure our site services to not conflict with other s/w versions.

5. **Can one VO box be shared across multiple sites to be used by the VO?** *(Example – A VO box cannot be shared across sites since it requires access to the NFS shared software area not typically available on the WAN.)*

   In principle yes, although the performance implications must be understood. Some VO BOX services perform bulk requests to LCG site services and these may suffer from network round-trip time.

6. **Should any VO services be monitored by the site? If so please provide technical details of how this can be achieved. In the case where incidents are triggered please provide details of the action required**? (*Examples – Please use a TCP Connect check on port 8123 to check service X is listening. If this test fails please reboot the node and subsequently telephone us immediately at any time of night or day.*)

All services are centrally monitored and do not require site intervention.

7. **How should the VO box be recovered after a complete loss? If more than a return of the vanilla VO box service is required please include a list of directories to be backed up and if any provision should be made for preserving and restoring consistent databases. Once the service is restored to previous state are any subsequent actions required by a site such as synchronisation with job outputs that completed during the downtime?** (*Examples – A reinstallation of a vanilla VO Box with no VO additions is acceptable. A recovery from a nightly backup is required. All areas under the home directory should be recovered after a backup. MySQL tables should be write locked during a backup.*.)

The reinstallation of a vanilla VO BOX with no VO additions is acceptable.

8. **What is the time period in which a production service can be restored by the VO? (***Production service will be restored by the VO the next working day.)*

The VO will restore the Production service the next working day or earlier.

9. **Are there any special requirements for the service which might influence the VO boxes initial deployment? For instance are there special network requirements of the VO box? Please provide details of how a VO service fits in to the overall workflow for the experiment.** (*Examples – Only command and control operations will take place so just the standard production network is required. All data access will be channelled through the VO box. Please see a particular document for a description of the VO service in the workflow.*)

No special requirements. The VO BOX only triggers and monitors requests to the normal production infrastructure provided by LCG (e.g. LFC, SRM, FTS, …)

10. **What is the business impact for the failure of VO Box?** (*Examples – All jobs at the site will be lost. We will be unable to submit new work to this particular site. We will be unable to access any existing data at this site.*)

Data transfers to the site are affected. Data transfers from the site may be affected, depending on when the machine crashed (if it is down for a substantial period of time these are likely to be affected).

In case the VO BOX fails completely, the site will no longer receive any new datasets (no new files are copied onto the site). Depending on the failure, it may also stop providing new datasets to other sites.In case of temporary failures (e.g. network fails for a few minutes), no significant impact on the data transfers is likely to occur.

11. **How should the VO be contacted for service interventions?** (*Examples – Via a mailing list. Via an EGEE broadcast. Via status tags in the GlueSchema.*)

Mailing list:

(ATLAS DQ2 Support) <atlas-dq2-support@cern.ch>

(ATLAS Computing Operations) <atlas-comp-oper@cern.ch>

(in addition to the VO BOX contacts provided in a separate questionnaire)

12. **Are any steps required before or after the service outage?** (Examples - *Please drain all work from your farm belonging to VO X before the VO box is rebooted.)*

No. If service is failing, it is responsibility of the VO to take all required steps to resume it (refer to answer 8)