



LCG Service Challenges for 2004

June 1, 2004

Ian Bird

Introduction

This note describes a proposal to put in place a series of service challenges for LCG, based on a set of performance and reliability milestones. The challenges should cover most aspects of the LCG service, both functionally and operationally. It is essential that the LCG service be developed over the remainder of 2004 in order to demonstrate a minimum infrastructure that is capable of handling at least the Tier 0 – Tier 1 interactions expected when LHC switches on. This means that a basic, but extremely robust and reliable data management service that is capable of handling the data rates between Tier 0 and each Tier 1, and between Tier 1's, is put in place and is demonstrated to be capable of delivering the anticipated network bandwidth for periods of weeks or months in a reliable and robust way.

In the same way it is essential that not only the data management service, but other aspects of LCG are set challenges. This is particularly important in bringing the service into a production-like state with continuous (albeit “best-effort”) support, reliability, etc. This means running the service in this way more or less continuously from now on, even outside of the experiment data challenges. Partly this might happen naturally with applications using the grid infrastructure under the aegis of EGEE, but partly might be fulfilled by commissioning job challenges – filling the system with jobs in order to test the limitations of the system and better understand its behaviour under load. In addition, specific milestones on bringing up operational services (monitoring, accounting, support, etc.) are needed. Most of these milestones and services will be independent of the actual middleware deployed and form part of the basic operational and support infrastructure that must be put in place quickly now.

For those developments that affect or rely upon middleware services and functions, it is essential that duplication of effort does not occur unnecessarily. To this end we propose to ensure that needed developments are undertaken as joint projects together with the EGEE middleware development teams, and with other deployment and development projects, in particular the US-LHC computing and grid infrastructure projects as appropriate.

The remainder of this note sets out some suggestions for service challenges and milestones, with some proposals for performance and reliability targets. We propose four sets of service challenges: reliable high performance data transfer; job flooding and filling the resources; security incident response; and interoperability between LCG-2 and Grid2003/OSG and NorduGrid.

The challenges will be overseen and managed by the GDA Steering group, but each set of challenges will require a small dedicated team to be brought together to implement the necessary tools and infrastructure and to manage the challenges themselves.

Proposed Service Challenges



1) A Reliable File Transfer Service

The goal of a reliable file transfer service would be to demonstrate the operation for extended periods of time a service that can be used in a reliable and predictable way, with good performance. The goals that might be foreseen for demonstrating such a service include the following:

- SRM disk pool to SRM disk pool reliable transfers at a rate of 500 MB/s between Tier 0 and a single Tier 1 (FNAL), sustained for a period of 2 weeks. This should include resilience against failure and guarantee that a file will be transferred. Potentially this will include queuing and prioritization of requests from different users. This goal should be achieved by the end of 2004.
- SRM disk pool to SRM disk pool from CERN to each of the other Tier 1 (and large Tier 2) centres, at a substantial data rate (100 MB/s?) sustained for at least 1 day in each case, to demonstrate the existence and basic reliability of the SRM interface to each site.
- SRM tape to SRM tape between CERN and one or two Tier 1 centres, with the goal of understanding total end-end throughput and bottlenecks, sustained over a period of a few days in each case.

To achieve these goals, some of the underlying milestones needed to build up to this level of performance and reliability will include the following:

- Basic network infrastructure between CERN and the Tier 1 centres must be in place. Some of the ideas that are being discussed include specifically routed gridftp traffic in order to make optimum use of the existing network bandwidths, potentially separating the control channel from the data traffic to route around firewalls as needed to avoid firewall bottlenecks. However, these are in some sense implementation issues that must be addressed at each site individually in order to get the best network performance. Once a basic service is up and running these issues can continue to be worked on and refined. Initially the testing will be done with the existing network setups.
- Definition of the Interfaces and APIs to be exposed to higher level services, in order that those higher level services can exist with different evolving implementations of this service. It is proposed that the service be implemented as a simple web service, to be compatible with other middleware and service developments.
- Definition of the service behaviour. The goal is to build a service that guarantees the delivery of a file, but there will of course be situations in which that will be impossible for whatever reason (user error, file exists, destination not available for a long period). How should that information be returned to the user or calling service?
- Definition of the security model for the service. Provision should be made from the beginning for file-level access controls to be put in place. This might require the use of the VOMS service and proxies.

The first steps towards understanding the issues are starting now, with the setup of simple disk-disk transfers of files between CERN and FNAL (and other Tier 1 sites), to get an idea of the existing baseline performance and bottlenecks. This work is being coordinated by Bernd Panzer at CERN, together with the CERN networking group.

The goal of the transfer service proposed here, however, is to build on that basic infrastructure to provide a service that can underpin higher level services (such as replica



management, etc.), and which can be used as a framework and test-bed within which to address all the detailed system level issues that will arise. These issues include:

- Understanding the failure modes of each of the components like gridftp, and putting in place mechanisms and strategies by which the service can recover.
- Understanding how to optimise the parameters in the transfer itself in order to maximise the throughput. Studies on number of parallel streams, TCP window sizes, etc. have been done before, but here we want to encapsulate that knowledge in this service.
- Understand the influence and impacts of using different sized files.
- Understanding how to deal with firewalls...
- Understanding how to build a scalable system with load-balancing transfer services at each end.
- Provide a simple but sufficient management interface, not only to be used to understand what the problems are, but to provide information to service users on the state of the service, transfers, queues, priorities, and statistics and accounting information.
- Given a service with well defined interfaces and API's, it would then be simple to replace the actual underlying transport mechanism with others for direct comparisons of performance.
- Behaviour in case of full disk caches.
- Understanding how to provide a layer providing workflow management (queuing, prioritisation), scheduling of transfers and eventual recovery from problems.

Once a basic service is in place, it must be extended to understand how to communicate with standard interfaces (SRM). A question might be potentially different behaviour given a MSS as an endpoint rather than a disk cache.

There should be a set of milestones with gradually increasing functionality, transfer rates, and sustainability.

Subsidiary work needed to operate such a challenge would include building a generator to provide the needed data and files for the tests. For a sustained test at the rate suggested, using 2GB files would require some 250K files. Policies for handling files at the end-points (at which point are they discarded?) will need to be thought out.

The existing LCG replica manager and other data management tools would be interfaced to this service.

2) Job Flooding/Grid Exercise Challenge

The idea of this challenge is to fill the entire system with jobs in order to make some baseline measurements of performance, robustness, and scalability, as well as to understand some of the limitations of the system. It is understood that these types of measurements will certainly differ depending on the middleware used to provide the services. However, it will be valuable to have such measurements based on the existing system in order to compare with future upgrades, either of individual services or eventually of the entire system with the EGEE/gLite developments.

In the 2004 data challenges, some problems with very long jobs became apparent. It would therefore be appropriate to try and use a mixture of jobs of different lengths. The test jobs



should really exercise the full range of grid services including in particular the data management system.

It would be expected that this challenge and the security response challenge (below) would be initiated, monitored, and managed by the grid operations centres. Eventually the set of tests might become part of a suite that the GOCs use for regular testing and verification of the infrastructure.

The testing framework currently in use by the LCG certification team to launch “job storms” etc. on the certification test-bed can be re-used with suitable modification as a tool to manage this challenge.

The jobs themselves can either be simple “hello world” type jobs that are constructed to exercise appropriate services, or can be real applications that provide some real benefit. An example of the latter might be the LHC “sixtrack” application.

The results of these tests would be used as a guide to improving service stability where appropriate on the existing infrastructure, and as input to ongoing middleware development projects.

Scheduling

Since this challenge could be potentially disruptive to ongoing production work, it will be essential that careful scheduling is done with the application communities around their data challenges and other work. It is expected that after adequate testing we would try and run such a challenge for around a week, to measure and observe the system. In the final stages, destructive testing (such as that described below) where services are deliberately restarted or killed, or network connections disrupted, would be attempted. This should certainly be done when it is clear no other important work is ongoing.

Typical goals of this challenge would be:

- Small scale tests to demonstrate capability to submit sufficient jobs to fill the system. This will involve using enough resource brokers to make sure that the submission rate is adequate.
- Run for a week in November and keep the entire system filled with jobs, and measure:
 - Job throughput
 - I/O throughput (where?)
 - How many of the available CPU can actually be used
 - etc....

A significant effort must be devoted to providing sufficient and adequate monitoring tools to be able to provide the appropriate performance measurements, and debugging of problems. It is expected to make use of existing tools such as MonaLisa, R-GMA with suitable sensors, and Gridlce. A secondary benefit of this kind of exercise will be to get a better feel for the coverage of these tools and where effort needs to be expended to improve them.

3) Incident Response Challenge

The proposal below is extracted from: <https://edms.cern.ch/document/478367>.



Non-intrusive service challenges

The following challenges pose a small set of simplistic problems. They are not intended to simulate a real security incident but have the characteristics which can be usefully but relatively harmlessly used to test procedures.

1. **Compute resource usage trace to owner**

The GOC causes a job to be run by a known user at a known site. The job sends an email (*is email enabled from all workers?*) to an external mailbox and execute some other trivial command (e.g. *hostname*). The user, location of the issuing User Interface and job commands should be unknown to the site(s).

Challenge: "An email was received by ... at ... from ... What user initiated the job, from where was it initiated and what else did it do?"

Extension: "What is the contact information for the user?"

2. **Storage resource trace to owner**

It is not yet clear whether this is technically possible with the current middleware audit.

The GOC should cause a job to be run by a known user at a known site. The job should create a known file in a known Storage Element.

Challenge: "File ... has been found on storage at ... When was it created, who created it and from where."

Intrusive service challenges

Test programs should be developed to stress network control channels and interfaces to individual grid services. Tests should include invalid and incomplete inputs, dropped connections and "fire hose" activity where random noise is directed at control channels. Use of such programs in parallel with production grid use is necessarily disruptive so careful management control of the tests will be necessary.

Challenge: "Service ... has been subject to disruptive activity. How did the service react and what were the wider effects on grid operation."

This challenge would be run as part of the job flooding exercises.

It is expected that such tests could be developed in conjunction with middleware testing activities.

Team:

Joint Security Group and people they might co-opt.

4) Interoperability

The fourth type of challenge that is foreseen is that of demonstrating interoperability between the grid services being used by the experiments (and indeed other applications). The most important of these are LCG-2, Grid2003/Open Science Grid, and NorduGrid. The idea would be to address a set of different interoperability topics and to provide solutions (or understand why a solution is impractical) that should be implemented in a lasting way – i.e. not solely for the purpose of the challenge but can be used in an ongoing way.

With the new version of the LCG-2 Resource Broker that can be configured to use only information from the BDII there could be a path to make GRID3, NorduGrid and LCG-2 interoperate.



We could use a modified version of the LCG-2 BDII (MDS) that translates the information collected from the GRID3 and Nordugrid sites into the Glue Schema used by LCG-2. Since the RB would see this information only through the BDII it could select sites for submitting jobs. The job-managers in LCG and GRID3 should be compatible thanks to using VDT-based software.

Authentication and authorization should not be a big problem for the experiments that have access to those resources. The mapping to a local account is done based on the DN of the certificate. The core CAs are recognized by all projects.

To make use of the replica manager tools is a bit more difficult, but since these tools are not services, but client tools this problem could be overcome rather easily.

The idea is not to do a full integration, but allow experiments to use all their resources in a transparent way. In this proposed first step there are still two RLS systems that are not synchronized.

The very first step to try this is to test running a simple job to one of the Grid3/NorduGrid CEs.

We had a brief look at the NorduGrid schema and it is clear that it will take some work to convert the information into the LCG-2 schema. For the GRID3 we need to understand better what the differences are. A sample CE and SE to query would be useful.

The aspects of interoperability that must be addressed are the following:

- VO management and user registration
 - It is important that each grid infrastructure accepts users from the other infrastructures. This condition is probably already fulfilled but some work must be done (with the Joint Security Group) to address all the details and make sure that a long term arrangement is in place.
- Information system.
 - LCG and OSG both use MDS and the GLUE schema. However, different semantic interpretation of the schema is possible (and has happened!), as well as different extensions assumed by each of the projects. Work to implement filters between OSG and LCG Information Systems will be done to bring them together. This should allow at least basic job submission, and should lead to updates of the GLUE schema based on common understanding.
 - NorduGrid, while also using MDS, however uses a very different information schema. It has to be understood how closely the GLUE and NG schema can be brought together, and indeed whether they can at all.
 - A script is required for each grid that takes output from an ldap search on a GIIS and converts this to the (common?) schema format. Problems might arise where information not published is needed in the submitting grid.
- Job submission
 - Foresee 2 steps here. Initially, using for example, central file catalogues
- File catalogues
- Data management
 - MSS



- Policies and expectations

Team:

- GLUE schema people
- Grid3 (Ian Fisk)
- LCG (Laurence Field – IS)