LCG Data Management: From EDG to EGEE

Jean-Philippe Baud[2], James Casey[2], Sophie Lemaitre[2],
Caitriana Nicholson[1], **Graeme Stewart**[1]

[1] University of Glasgow, Glasgow, G12 8QQ, Scotland

[2] CERN, European Organisation for
Nuclear Research, 1211 Geneva, Switzerland

1 Introduction

The Large Hadron Collider (LHC) at CERN, the European Organisation for Nuclear Research, will produce unprecedented volumes of data when it starts operation in 2007. To provide for its computational needs, the LHC Computing Grid (LCG) is being deployed as a worldwide computational grid service, providing the middleware upon which the physics analysis for the LHC will be carried out.

Data management middleware will be a key component of LCG, enabling users to analyse their data without reference to the complex details of the storage and computational environment.

In this paper we review the performance tests of the LFC in comparison with other data management catalogs. We will also survey the deployment status of the LFC within the LCG.

2 EDG Replica Location Service

The European Data Grid (EDG) produced data management middleware consisting of the Replica Metadata Catalog (RMC) and the Local Replica Catalog (LRC). Together these formed the Replica Location Service (RLS), which was deployed in the LCG in 2003-4.

The 2004 series of LHC experiment data challenges were the first to use the LCG-2 set of middleware tools in a realistic environment [1]. Feedback from the experiment groups after the data challenges highlighted various problems and limitations, as well as differences between the expected and actual usage patterns.

During these challenges it became apparent that the file catalog infrastructure was too slow both for inserts and for queries [2]. Queries involving both the LRC and RMC were particularly slow [3]. Missing functionality identified included lack of support for bulk operations and transactions. It also became clear that queries were generally based on metadata attributes and were not simple lookups of a file's physical location. On the other hand, users did not use the web services approach in the way which had been anticipated when the EDG components were developed. They were implemented such that a remote procedure call (RPC) was performed for each low-level operation; users, however, wanted to send higher-level or multiple commands in a single RPC. As this was not available, the cumulative overheads from a large number of low-level RPCs led to considerable loss of performance. Also, although a C++ API was available, command-line tools were available only in Java which led to added loss of performance due to the overhead in starting up the Java Virtual Machine with each call.

3 EGEE LCG File Catalog

To address the problem of the EDG RLS the EGEE (Enabling Grids for E-sciencE) project had designed a new data management component, the LCG File Catalog (LFC). The LFC moves away from the Replica Location Service model used in previous LCG releases, towards a hierarchical filesystem model and adds missing functionality which was requested by the experiments.

The LFC has a completely different architecture from the RLS framework. Like the EDG

catalog, it contains a GUID (Globally Unique Identifier) as an identifier for a logical file, but unlike the EDG catalog it stores both logical and physical mappings for the file in the same database. This speeds up operations which span both sets of mappings [4]. It also treats all entities as files in a UNIX-like filesystem. The API is designed to mimic a UNIX filesystem API, with calls which are intuitive to the user, such as `creat`, `mkdir` and `chown`.

The main entities of the LFC design are shown in Figure 1. There is a global hierar-

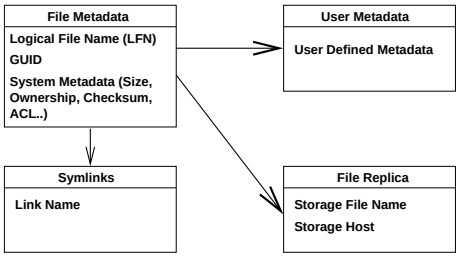

Figure 1: Components of the LFC.

chical namespace of Logical File Names (LFNs) which are mapped to the GUIDs. GUIDs are mapped to the physical locations of file replicas in storage (Storage File Names or SFNs). System attributes of the replicas (such as creation time, last access time, file size and checksum) are stored as attributes on the LFN, but user-defined metadata is restricted to one field. Multiple LFNs per GUID are allowed as symbolic links.

Bulk operations are supported, with transactions, and cursors for handling large query results. As there is only one catalog, transactions are possible across both LFN and SFN operations, which was impossible with the EDG RLS. In case of momentary loss of connection to the catalog, timeouts and retries are supported. Authentication is by Grid Security Infrastructure (GSI), which will allow single sign-on to the catalog with users' Grid certificates.

4 Performance

The performance of the EDG RLS and other catalogs are compared. In particular we concentrate on the performance tests we believe best model the end user usage patterns expected from the LHC experiments.

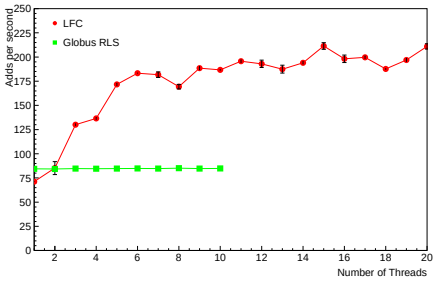

Figure 2: Entry insertion rate in LFC and Globus RLS.

5 Conclusions

The LCG File Catalog is a replacement for the EDG RLS system. It offers the features required by users but unavailable in other replica management systems.

Performance measurements show the robustness and scalability of the LFC up to many millions of entries and hundreds of client threads.

The deployment status of data management middleware in the LCG will be reviewed.

References

[1] A. Delgado Peris, *et al.* LCG-2 User Guide. Technical Report CERN-LCG-GEDIS-454439, CERN, Geneva, Switzerland, September 2004.

[2] A. Fanfani *et al.* Distributed Computing Experiences in CMS DC04. In *Computing in High Energy and Nuclear Physics (CHEP)*, Interlaken, Switzerland, September 2004.

[3] M. Girone et al. Experience with POOL in the LCG Data Challenges of Three LHC Experiments. In *Computing in High Energy and Nuclear Physics (CHEP)*, Interlaken, Switzerland, September 2004.

[4] J-P. Baud and J. Casey. Evolution of LCG-2 Data Management. In *Computing in High Energy and Nuclear Physics (CHEP)*, Interlaken, Switzerland, September 2004.