egee

**Enabling Grids
for E-sciencE**

Replication on the AMGA Metadata Catalogue

Wednesday, 1 March 2006 18:30 (20 minutes)

1. Introduction

Metadata Services play a vital role on Data Grids, primarily as a means of describing and discovering data stored on files but also as a simplified database service. They must, therefore, be accessible to the entire Grid, comprising several thousands of users spread across hundreds of Grid sites geographically distributed. This means they must scale with the number of users, with the amount of data stored and also with geographical distribution, since users in remote locations should have low-latency access to the service. Metadata Services must also be fault-tolerant to ensure high-availability.

To satisfy such requirements, Metadata Services must offer flexible replication and distribution mechanisms especially designed for the Grid environment. They must cope with the heterogeneity and dynamism of a Grid, as well as the typical workloads.

To address these requirements, we are building replication and federation mechanisms into AMGA, the gLite Metadata catalogue. These mechanisms work at the middleware level, providing database independent replication, especially suited for heterogeneous Grids. We use asynchronous replication for scalability on wide-area networks and improved fault-tolerance. Updates are supported on the primary copy, with replicas being read-only. For flexibility, AMGA supports partial replication and federation of independent catalogues, allowing applications to tailor the replication mechanisms to their specific needs.

1. Use Cases

Replication on AMGA is designed to cover a broad range of usage scenarios that are typical of the main user communities of EGEE.

High Energy Physics (HEP) applications are characterised by large amounts of read-only metadata, produced on a single location and accessed by hundreds of physicists spread across many remote sites. By using AMGA replication mechanisms, remote Grid sites can create local replicas of the metadata they require, either of the whole metadata tree or of parts of it. Users at remote sites will experience a much improved performance by accessing a local replica.

For Biomed applications the main concern with metadata is ensuring its security, as it often contains sensitive information about patients that must be protected from unauthorised users. This task is made more difficult by the existence of many grid sites producing metadata, that is, the different hospitals and laboratories where it is generated. Creating copies on remote sites increases the security risk and, therefore, should be avoided. AMGA replication allows the federation of these Grids sites into a single virtual distributed metadata catalogue. Data is kept securely on the site it was generated, but users can access it transparently from any AMGA instance, which discovers where the data is located and redirects the request to that AMGA instance, where it will be executed after the user credentials have been validated.

We believe that partial replication and federation as they are being implemented in AMGA provides the necessary building blocks for the distribution needs of many other applications, while at the same time offering scalability and fault-tolerance.

1. Current Status and Future Work

We have implemented a prototype of the replication mechanisms of AMGA, which is currently undergoing internal testing. Soon we will be ready to start working with the interested communities, with the goal of better evaluating our ideas and of obtaining user feedback to guide us through further development of the replication mechanisms.

A clear user requirement that we will study is the dependability of the system, including mechanisms for detecting failures of replicas and for recovering from those failures. If the failure is on a replica, clients should be redirected transparently to a different replica. If the failure is on the primary copy, then the remaining replicas should elect a new primary copy among themselves. All these mechanisms need an underlying discovery system to allow replicas to locate and query each other, as well as mechanisms for running distributed algorithms among the nodes of the system.

Primary author: DE SOUSA SANTOS, Nuno Filipe (Universidade de Coimbra)

Co-author: KOBLITZ, Birger (CERN)

Presenter: DE SOUSA SANTOS, Nuno Filipe (Universidade de Coimbra)

Session Classification: Poster and Demo session + cocktail

Track Classification: Poster session