Upgrading Web Applications to a GRID Environment

Carmen Maidantchik[1], José Manoel de Seixas[1,2]

[1]Programa de Engenharia Elétrica – Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68504 – 21945-970 – Rio de Janeiro – RJ – Brazil

[2]Escola Politécnica – Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68504 – 21945-970 – Rio de Janeiro – RJ – Brazil

{lodi,seixas}@lps.ufrj.br

Abstract. *Due to the new requirements of the ATLAS detector of CERN, the software applications that are currently being used by the collaboration should be adapted to the context of the Grid computing. In order to support the processing of the enormous amount of data that will be generated by the experiment, large-scale project researchers must share the needed resources. This paper presents three Web applications that are being enhanced to meet the requirements of the new reality. The concepts of distribution and aggregation, security, performance and resources management are being codified and integrated to the CombinedTranfer, Remote Neural Discriminator and Monitor scientific systems.*

Resumo. *Devido aos novos requisitos do detector ATLAS do CERN, as aplicações utilizadas na colaboração deverão ser adaptadas no contexto da computação Grid. Para viabilizar o processamento da enorme quantidade de dados que serão gerados pelo experimento, os pesquisadores deverão compartilhar recursos necessários em projetos de larga escala. Este artigo apresenta três aplicações Web que estão sendo evoluídas para atender a esta nova realidade. Os conceitos de descentralização e agregação, segurança, desempenho e gerência de recursos estão sendo codificados e integrados aos sistemas científicos CombinedTranfer, Remote Neural Discriminator e Monitor.*

1. Introduction

Worldwide scientific collaborations have computing requirements that exceed the capabilities of most computing facilities. The amount of data involved usually reaches the Petabyte scale, which requires merging resources of several institutions to provide compute capability to process the whole data. In many cases, computing processing requires millions of CPU-days, which complexity is enhanced due to the distributed nature of the organizations [M. Mambelli, 2004]. In order to overcome the difficulties of scientific applications that requires high performance in distributed computing platforms, shared Grid infrastructures are already being developed [IVDGL 2004].

Grid computing is emerging as a large computational resource by providing mechanisms for sharing the heterogeneous computing and storage resources of organizations and individuals worldwide distributed to form a massive computing environment through which complex and large scale problems can be processed. The main goal is to offer a large and powerful virtual computer with enormous processing power [A. Ali et al, 2004].

In Switzerland, CERN (European Organization for Nuclear Research) is an example or an international collaboration that requires Grid solutions [CERN 2005]. There, the ATLAS (A Toroidal LHC Apparatus) detector is being constructed to explore the fundamental nature of matter and the basic forces that shape our universe [ATLAS 2005]. ATLAS is the largest collaborative effort ever attempted in physics, getting together 2000 physicists from more than 150 universities and laboratories in 34 countries and will run during at least ten consecutive years. The annual data volume generated by the detector will be equivalent to 1.5 million of 650 Megabytes CDs.

The ATLAS Grid will have to couple a wide variety of geographically distributed computing resources, including storage and computational processing. The software systems are heterogeneous in nature, owned by different individuals and organizations, and have their own policies with different access and cost models. Therefore, the ATLAS Grid computing involves the close coordination of many different sites that offer distinct computational and storage resources to the Grid user community. The resources at each site need to be monitored continuously.

In this context, the current ATLAS software applications need to be adapted to the new reality. The systems that were developed by the collaboration and that are being used require an upgrade to cope with Grid requirements. The Federal University of Rio de Janeiro has developed three important Web systems to support the detector activities such as data access, data analysis, and information storage. These software were used during the performance evaluation of the ATLAS calorimeter, when several generations of prototypes were assembled and submitted to laboratory measurements and experimental particle beam tests, which generated significant amount of data that had to be accessed and analyzed worldwide.

2. The CombinedTransfer Web System

The goal of the CombinedTransfer Web system is to offer an user-friendly Web interface to access all information related to the acquired testbeam data and then, allow the download of the required files to perform physics analysis. The main components that are integrated through Web services are: the TestBeam Analysis database that contains information about all analysis that should be performed, and is placed in the CERN Web Server; the TileInfo database that provides an interface to perform queries in a MySQL server, containing information about all runs; the CASTOR storage management system [CASTOR 2005] and, finally, a sharing resource management that controls resource providers and consumers, defining clearly and carefully what is shared, who is allowed to share, and the conditions under which sharing occurs [I. Foster, C. Kesselman, S. Tuecke 2001].

ATLAS testbeams were performed during several periods along the last years, generating data that is organized in ntuples and stored in a hierarchical storage management system developed at CERN. The size of a single ntuple can achieve hundreds of Megabytes and the target physics analysis may involve several ntuples. Files in CASTOR are accessed using rfio (Remote File Input/Output) protocols either at the command level or, for C programs, via function calls. CASTOR is used for the transfer of raw data from experimental areas to central storage. The migration of CASTOR files between tape storage and disk is managed by a stager daemon process. The stager manages one or more disk pools or groups of one or more UNIX filesystems, residing in one or more disk servers.

The CombinedTransfer integrates the TileInfo database and the access to the files stored at CASTOR by hypertext links. Each line of the resulting query is associated with a program that, just when the user selects a link, executes CASTOR commands. This program controls all necessary steps starting at staging out the file to a spool area until copying the ntuple to a scratch directory and then starting the transfer process. In 2005, the CombinedTransfer was used during the first combined testbeam that gathered several components of the ATLAS detector.

2.1. Transferring Data Across the Network

In order to optimize file transmission, the system is integrated with a central repository that stores information of the latest accesses. The aim is that once a user connects to the CombinedTransfer system, the client machine can also become a file server to other users. When the system is not busy, it looks through the repository updating information about the spread servers, verifying whether the specific computer is on, the ntuples are still placed there, the network speed has changed and so on. The merging of two different types of databases, one containing the data description and the other with information about the owner of the file, makes it possible to achieve real helpful solutions. And in order to speed up the download time, at different servers, the selected file is split into several pieces. Each piece is sent from one server in parallel to the others and such fragments are built up together at the final destination.

Spiders' techniques were applied to rank most suitable file providers among the whole group of CombinedTransfer users who have the target data available. The elected server for data provision depends on several factors such as network speed and stability. The nearest server is not necessarily the one that will provide the ntuple to a requesting client. Each volunteer is associated with a score based on the network characteristics and the reliability of the server. In case the first chosen provider is not available, the framework automatically switches to the next qualified one.

3. The Remote Neural Discriminator Web System

One of the ATLAS subsystems is the hadronic calorimeter that measures the energy of particles. The information provided by the calorimeter is used to distinguish the different type of particles. Particle discriminating problem can be redefined as pattern recognition problem and neural networks techniques can be used to solve such kind of problems. In order to provide the possibility of using a neural network approach by the ATLAS researchers around the world, the techniques were put available via a Web browser. The

Remote Neural Discriminator (RND) Web system was developed to perform a complete analysis as required by the neural network application that runs in a CERN server.

The application builds different Web pages depending on the user choices. The system provides different neural network training options, such as default or advanced. The system manages the neural network package processing. The package reads ntuples produced by the data acquisition system for the hadron calorimeter testbeam. The results of the network are then, analyzed by the Physics Analysis Workstation (PAW) program developed by CERN specialists in data analysis. PAW is integrated to the Remote Neural Discriminator Web system.

The Web system maintains a database of users, identified by username and password. This identification leads to a specific user directory where all the files produced by the user analysis are placed. This includes data files from the neural network, PostScript and ntuples files produced by the user analysis. This feature allows users to continue their analysis performed in a previous section. The database also stores the user level, which can be Beginner, Advanced or Expert. The Beginner user only follows a default analysis and is not allowed to change parameters of the neural network processing. This user can see all the graphic results of predefined analysis. The Advanced user can change the network parameters, retrain the neural network and verify the graphic results of the neural network analysis with the predefined analysis tools. The Expert users can also provide their own PAW script to analyze the data. Therefore, the system is fully programmable. New data files can be produced and used in analysis together with the results of the neural network. All the files produced by any analysis can be downloaded to the user remote account.

3.1. Analyzing Data Across the Network

Due to the large number of users and the huge size of the data files, the neural network processing should be distributed. A central repository can store information about users, their last connections, tasks and a rank the most suitable computing resources provider.

In order to speed up the download time, at different servers, the selected file is split into several pieces. Each piece is sent from one server in parallel to the others and such fragments are built up together at the final destination. Therefore, the information must be continuously updated among spread servers.

Data analysis can be a collaborative activity. For that reason, it is mandatory that the system offers Web publication features and an open storage tool with data categorization according to the performed analysis. The system must also provide search mechanisms. Another requirement is the possibility to integrate to different analysis package, such as ROOT, or even to other particle discrimination techniques.

4. The Monitor Web System

The ATLAS Database stores a huge amount of data about the ATLAS construction. Retrieving data is a hard task due to the diversity of data types and search requirements. The Monitor Web system offers a parametric search according to the kind of data that is being searched. The description for each document type is automatically described in XML, and then, used to generate the search interface.

The concept of this application is related to a universal description process. Any relational database can be described through a single hierarchical description file and the representation method is unique. Consequently, a general algorithm may browse the database structure and provides its representation through a markup language.

The current version allows the possibility to browse existing search interfaces or to create new search interfaces. When browsing, one can edit or delete an existing interface. To create a new interface, the user selects a schema name, a table name that belongs to the previous selected schema, and the system automatically generates the XML that describes the interface. This markup description contains the column names and types of a selected table. Therefore, the interface guides the search informing the users whether the data type is string, real, etc. Moreover, scripts automatically verify the correctness of input parameters. Expert users may edit the XML description adding further information, integrating with help files or modifying the column name in order to increase the user understanding.

The Monitor Web system also represents the relationship among tables and allows the combination of "and/or" searches. Another feature of the software is the customization of searches. A user of a group of users can define their own search interfaces that have priority over the common ones. Since the equipment maintenance team has different attributions from the physics analysis team, the former group will have access to a specific set of search interfaces that will be diverse from the other group.

Finally, the Monitor Web system provides the definition of different abstract layers. The goal is to simulate the navigation through the ATLAS detector structure that is composed by an overlay of diverse sections. Search engines are also applied to the layers, creating a bi-dimensional exploration of the whole storage data.

4.1. Searching Data Across the Network

The storage and cataloging of all users data, including the search interface descriptions and the layers definition, can be placed in different servers. The entire ATLAS database itself should be divided and distributed in a way to improve query performance, optimize disk space, and avoid information lost in case of a repository rupture.

When there are a large number of results, the system should manage their display to users and avoid transferring the whole set to avoid network traffic. Therefore, the user section should be saved, managed and restored. This supervision is also necessary in case of any connection problems.

The application has also to be independent from the database in a way that any user can use the software to search data in any collaboration database located anywhere.

5. Conclusions

For each Web application we presented some considerations about the requirements they have to respect in order to properly run in a worldwide collaboration, taking into account the complexity of the computing tasks raised from several factors. High energy physics systems are characterized by the huge amount of data, enormous file size, analysis processing, number of users, spread resources, diversity of configurations, knowledge, etc. In the specific case of the ATLAS detector, the longevity of the project is also an important matter, since it is directly related to the development, user, training, and maintenance of data and software applications.

The Web systems were designed to be scalable systems, serving additional users or transactions without fundamentally altering the application's architecture or program structure. It is possible to identify common features in all Web applications as the data transferring which guides developers to implement a single program suitable for different purposes.

The Grid is supposed to be capable of including various kinds of devices or resources such as computers, networks, data, software, storage, etc. In the context of ATLAS, the coordination of multiple remote system entities is required. Monitoring systems successfully integrate existing tools to a reliable Grid infrastructure. The Web applications architecture is now being enhanced by integrating and revising existing Grid software components, chosen by the ATLAS community.

References

M. Mambelli (2004), “Grid2003 Monitoring, Metrics, and Grid Cataloging System”, Computing in High Energy Physics Conference (CHEP), Interlaken, Switzerland.

IVDGL (2004), International Virtual Data Grid Laboratory, http://www.ivdgl.org/

A. Ali et. Al. (2004), “Job Monitoring In An Interactive Grid Analysis Environment” Computing in High Energy Physics Conference (CHEP), Interlaken, Switzerland.

CERN (2005) - European Organization for Nuclear Research. http://www.cern.ch

ATLAS (2005) - A Toroidal LHC ApparatuS. http://atlas.web.cern.ch/Atlas

CASTOR (2005) - CERN Advanced STORage Manager. http://castor.web.cern.ch/castor/

I. Foster, C. Kesselman, S. Tuecke (2001), “The Anatomy of the Grid: Enabling Scalable Virtual Organizations”, International Journal of High Performance Computing Applications; Vol. 15, No. 3, pp. 200-222.