SECURE GRID DATA MANAGEMENT TECHNOLOGIES IN ATLAS

Miguel Branco[#], CERN, Geneva, Switzerland
D. Malon, A. Vaniachine, ANL, Argonne, IL 60439, USA

Abstract

In a resource-sharing environment on the grid both grid users and grid production managers call for security and data protection from unauthorized access. To secure data management several grid technologies were introduced in ATLAS data management. Our presentation will review new grid technologies introduced in HEP production environment for database access through the Grid Security Infrastructure (GSI): secure GSI channel mechanisms for database services delivery for reconstruction on grid clusters behind closed firewalls; grid certificate authorization technologies for production database access control and scalable locking technologies for the chaotic 'on-demand' production mode. We address the separation of file transfer process from the file catalog interaction process (file location registration, file medadata querying, etc.), database transactions capturing data integrity and the high availability fault-tolerant database solutions for the core data management tasks. We discuss the complementarities of the security model for the online and the offline computing environments; best practices (and realities) of the database users' roles: administrators, developers, data writers, data replicators and data readers, need for elimination of the clear-text passwords; stateless and stateful protocols for the binary data transfers over secure grid data transport channels in heterogeneous grids. We present the security policies and technologies integrated in the ATLAS Production Data Management System - Don Quijote (GSI-enabled services oriented architecture, GSI proxy certificate delegation) and approaches for seamless integration of Don Quijote with POOL event collections and tag databases - while making the system non-intrusive to end-users.

INTRODUCTION

LHC experiments are facing an unprecedented multi-petabyte event data processing task. To address that challenge LHC computing is adopting emerging grid computing technologies. The resource-sharing paradigm of grid computing presents technical challenges for security in the LHC data management task.

ATLAS Data Challenges

ATLAS decided to undertake a series of Data Challenges in order to validate its Computing Model, its software, its data model.

ATLAS Data Challenges 2 (DC2) started during the summer of 2004. In DC2 a new ATLAS Automatic Production System [1] was introduced. Its goal is to manage all aspects of DC2 operation in a fully unsupervised and automatic mode.

The DC2 production is distributed across many computing sites which collaborate with the ATLAS experiment. These sites are spread across three different Grid Computing resources - LCG-2, US Grid3 and NorduGrid.

FILE ACCESS ON THE GRID

In order to prepare for the Data Challenges, a listing of security requirements from the perspective of the production managers was collected.

The user(s) in charge of managing the production are mostly concerned about producing data. They do not intend to focus on the underlying grid middleware security or site security issues. These should be dealt by the grid production service and development teams as well as the site managers. Availability of computing resources and data is likely to be the top priority from the perspective of the production manager. Given the magnitude of a Data Challenges this is typically the most critical area. Also, given the cost associated with producing data (CPU hours, storage space, …) not loosing or allowing data to be corrupted is obviously a concern. Due to the fact that there are many collaborating sites providing computing resources it is important for ATLAS not to impose heavy security constraints om sites wishing to join the ATLAS DC. In addition it is important not to expose sites to vulnerabilities that might be derived from running ATLAS software. Overall the production manager wishes to be in charge of all grid production at all times and of all the produced data. Auditing the usage and access to data is not only desirable but mandatory.

As for end-users it is important not to overload them with a coarse security infrastructure. The High Energy Physics (HEP) community wishes to focus on their analysis work without having to deal with grid security. In HEP, especially in comparison to other sciences such as Bio-medics, having no end-user visible security is usually regarded as an advantage. It is therefore important to provide simple and easy to use security mechanisms for end-users.

The ATLAS Automatic Production System

With DC2 the new ATLAS Automatic Production System was introduced. This production framework is composed of different modules and is the responsible for running the Data Challenges by interacting with all available grid computing resources. Therefore it is the central point for implementing a security infrastructure.

[#]Miguel.Branco@cern.ch

The automatic production system is made of four major components: the production database and Windmill [1], the ATLAS automatic production supervisor, Job Executors [2] responsible for dispatching jobs to the underlying grid middleware and a common data management system: Don Quijote [3].

The production database is where the Data Challenges tasks are stored. This database is hosted in a centrally maintained server using Oracle and located at CERN.

Windmill is the supervisor of which there are multiple instances running worldwide at any given time. Its communication and resource discovery framework is based on Jabber. Work is on-going regarding the implementation of a GSI-enabled Jabber server. Windmill fetches tasks from the production database and dispatches them to a job executor. Jabber has proven to be useful for resource discovery and communication between Windmill and each grid job executor.

The job executors interact with the native grid middleware translating the high-level job definitions taken from the production database to the grid middleware job definition language. Production managers start a job executor which is typically sitting on a grid user interface machine, with the grid certificate of the person in charge of running part of the DC production.

To allow for transparent data access within or across the several grid flavors used by ATLAS, a data management interface was also developed. Don Quijote (DQ) is an access service for grid file-based data. It was built as a high-level interface for grid data management for the ATLAS Automatic Production System allowing transparent registration and movement of replicas between all grid “flavors”, using a services oriented architecture.

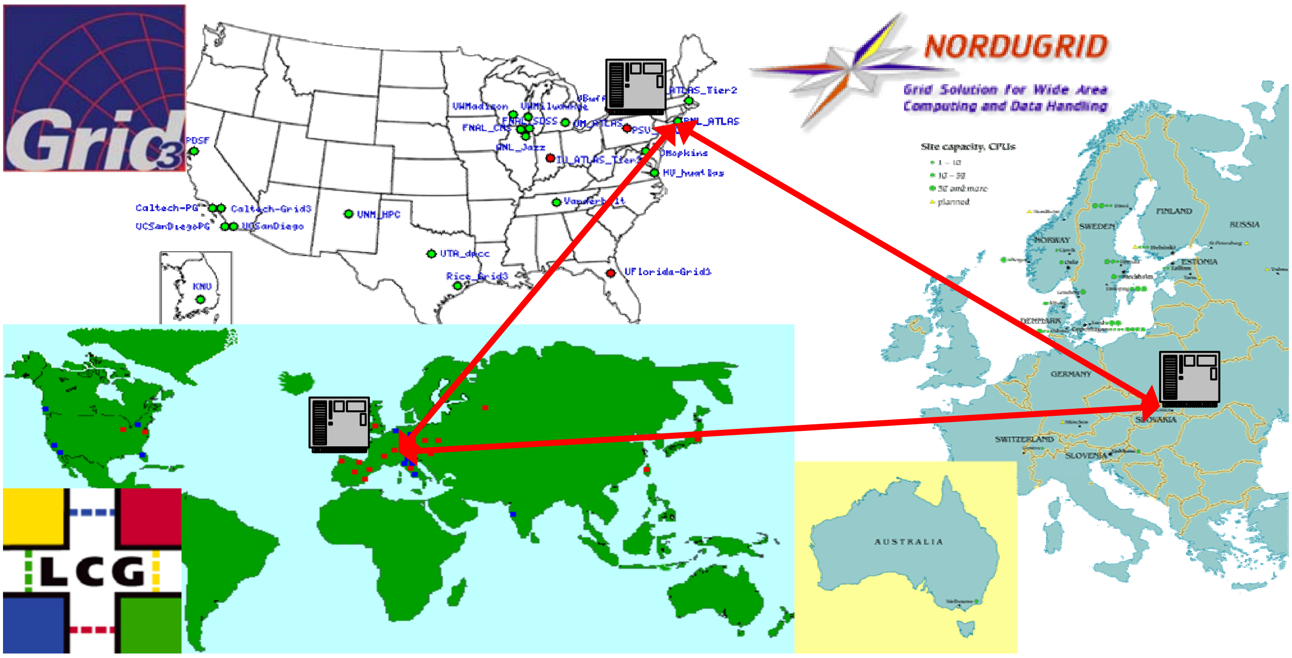

Figure 1: In Data Challenge 2 ATLAS deployed data services and exercised processing and managing data on a federation of computational grids.

Don Quijote provides both secure and insecure versions of the servers (with different subsets of the functionality). There is a DQ secure client (using GSI) and insecure client. The DQ server can forward the user grid certificate if the user accessed the server using the secure version of the client. If the user didn’t use the secure client, the server is able to act on behalf of the user by using instead a grid service certificate. Typically only search requests (read-only access) is allowed from the insecure version.

During DC production, it became clear that access to all ATLAS resources was not being properly propagated to users. This was due to a combination of factors: evolving or missing grid security middleware features, slowness in establishing effectively a common ATLAS VO matching the complete ATLAS Collaboration. In addition, most end-users had not yet been through the (occasionally) painful process of requesting grid certificates.

For practical reasons service certificates were used on some of the ATLAS Production System components, notably on the Don Quijote. This enabled users who do not posses a valid grid credential to execute a very limited subset of the requests – the most common being a request to search for file replicas in the Grids replica catalogs. The ATLAS Production System team had to pragmatically consider this option in order to provide data availability for end-users.

The ATLAS Production System will be reviewed taken security implications into account more extensively. The need to have a working system on a short timescale delayed the process. With the DC2 experience the reviewed Production System will include better integration of security mechanism: e.g, delegation of grid credentials between the supervisor, job executors and data management system.

Nevertheless security "gaps" derive mostly from the existing grid middleware. One example is the lack of fine-grained access to grid storage and grid replica catalogs. In the meantime the usage of client tools allowed some security measures to be put in place by ATLAS to compensate for the missing features in the existing middleware.

DATABASE ACCESS ON THE GRID

Until recently LHC computing models focus has been limited to the problem of managing the petabyte-scale event data that are considered now to be in a traditional file-based data store. The grid security model and efficient data transport mechanisms are particularly well suited for handling the file-based data. The same is true for the database-resident file cataloguing and the file-level metadata that are well served by the grid-based RLS and meta-data catalogs supplied through the grid middleware infrastructure.

In addition to file-based event data, LHC data processing applications traditionally require access to large amounts of valuable non-event data (detector conditions, calibrations, etc.) stored in relational databases. In contrast to the file-based data, this database-resident data flow has to be detailed further [4, 5]. For that purpose ATLAS Data Challenges exercise Computing Model processing and managing data on a federation of LCG, Grid3 and Nordugrid computational grids (Figure 1).

To secure database-resident data ATLAS is evaluating several technologies: secure grid query engine technologies federating heterogeneous databases on the grid [6], methods utilizing GSI data-transport channel for database services delivery and grid certificate authorization technologies for database access control.

Penetrating Firewalls

A standard approach for delivery of database-resident data to applications on dedicated ATLAS computing resources requires open TCP/IP channels to the database servers. To harness Grid commuting resources that are not dedicated to ATLAS one must address the problem of data delivery to the computing nodes on the clusters with stricter security policies of commercial grids or other grid clusters behind closed firewalls. As practical solutions in ATLAS DC2 we used both database server replica deployment on a dedicated cluster node behind the firewall and the network address translation (NAT) techniques providing TCP/IP conduits to the listed database servers addresses and ports. Both these methods were successfully deployed in the production environment of ATLAS Data Challenge 2 increasing availability of opportunistic resources for ATLAS grid by many hundreds of computing nodes.

GSI Transport for Database Services

Both the replica and the NAT methods for database services delivery require considerable involvement of the cluster support personnel. An alternative approach utilizes existing Grid GSI data-transfer channels and does not require modification of the cluster configuration.

This method simplified the delivery of the extract-transport-install components of ATLAS database architecture (Figure 2) to provide database services needed for the Data Challenges for sites with Grid with worker nodes behind closed firewalls (several Grid3 and Nordugrid sites). Both during DC1 and DC2 we have benefited from the use of MySQL database servers for deployment of the server replica to each worker node for the duration of the data processing job.

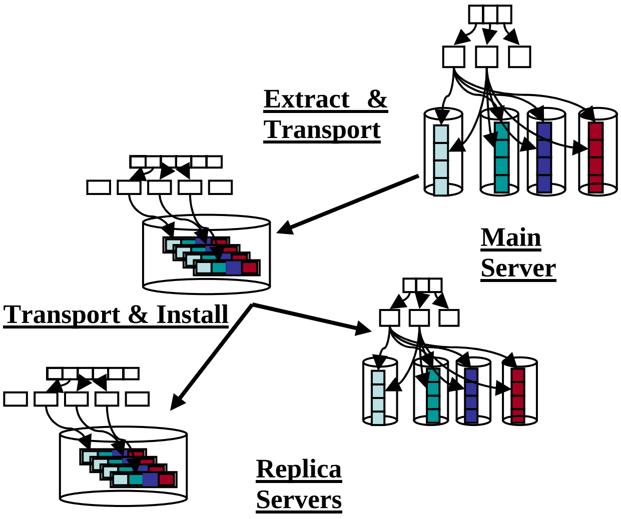

Figure 2: Extract-Transport-Install concept of ATLAS database architecture.

GRID-ENABLING DATABASES

There are two different models in providing securing database access on the grid. In a traditional approach a separate security layer (a separate server) does the grid authorization. In an alternative approach instead of surrounding database with external secure layers the safety features are embedded inside of server.

External Security

Providing security in a separate layer is a proven technique used in the Spitfire [6] and OGSA-DAI [7] projects. It leaves the traditionally weak database authorization techniques behind the secure layer, the clear-text passwords embedded in the deployed configurations, limited control over the secure transport channel, cryptographic handshake for every SOAP message, and requires protocol extensions (XML with binary attachments) for efficient data transfers.

Embedded Security

In an alternative approach the grid authorization is integrated in database server, which is possible with the the open-source databases or e.g. through the IBM DB2 loadable security modules techniques. This innovative approach is listed among the top ten innovations in security [9]. In addition to the elimination of the clear-text passwords through the deployment of the same grid security model cross-cutting all data flow channels the inefficient data transfer bottlenecks are eliminated.

Grid-enabling Databases

To overcome concerns that pushing secure authorization into the database engine result in a rigid system that can be brittle we deployed the grid-enabled MySQL server in ATLAS on the database development server tier and used the technology extensively in ATLAS DC2 pre-production on Grid3. For efficiency, the secure layer can be used only for authentication, with the data transfer channel operated in the un-encrypted mode (similar to the GridFTP).

Both the grid-proxy (x509up) the grid (x509) certificate authorization were supported by the database server to enable two different user roles for the holder of the same certificate. Ii addition, use of certificate credentials provided capabilities for efficient locking mechanism to support chaotic mode of distributed parallel submission of jobs on the Grid. The grid-proxy certificate authorization was used successfully in the processing of more than 7K of jobs by several ATLAS production operators [10].

CONCLUSIONS

Grid middleware developments have a strong focus on securing file-based data. Nevertheless additional security features are necessary to respond to the HEP security requirements. Many promising projects are currently addressing these issues. In the meantime ATLAS as part of the Production System developed its own security layer taking into consideration current limitations. This has proved to work successfully for the ATLAS Data Challenges 2. Future work will complement the existing system in two main areas. One is enabling additional finer-grained access control to file-based data. The other is providing proper credential delegation within the production chain, from the job definition step to the final job execution at a given grid site, passing through all ATLAS Production System components.

In addition, to overcome database access limitations one must to go beyond the existing grid infrastructure. In preparation for future challenges we are evaluating the technologies laying a foundation of a new hyperinfrastructure:

- secure grid query engine technologies federating heterogeneous databases on the grid;
- methods utilizing Grid Security Infrastructure data-transport channel for database services delivery to the grid clusters behind closed firewalls;
- grid certificate authorization technologies for database access control where the safety features are pushed into the database engine code.

We have tested these technologies in a production environment of ATLAS Data Challenges on emerging computational grids.

It is our belief that to implement a proper security infrastructure a strong collaboration must be set up between the teams developing grid middleware and the experiment framework developers. Only by embedding the security mechanism throughout the entire chain - from the experiment production framework to the underlying grid middleware - can security be effectively put in place. This is an area in which some additional work and collaborations have to be put in place.

ACKNOWLEDGEMENTS

We thank all of our ATLAS collaborators and, in particular, developers from database, grid and data challenges domains. We also wish to thank Jack Cranshaw, Rob Gardner, Lee Lueking, Yuri Smirnov and Matthew Vranicar for numerous discussions and interest in this work. Argonne National Laboratory's work was supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under contract W-31-109-Eng-38.

REFERENCES

[1] L. Goossens [501], this conference.
[2] M. Mambelli [503], D. Rebatto [364], O. Smirnova [499], this conference.
[3] M. Branco [142], this conference.
[4] P. Watson, Databases and the Grid, UKeS-2002-01.
[5] A. Vaniachine, D. Malon, M. Vranicar, DPF’2004, Riverside, USA, 2004, paper 334.
[6] http://www.piocon.com/databasegrid.html
[7] http://edg-wp2.web.cern.ch/edg-wp2/spitfire
[8] http://www.ogsadai.org.uk
[9] http://www.battelle.org/forecasts/defense.stm
[10] http://griddev.uchicago.edu/swhome/atgce