

ARDA PROTOTYPES

J. Andreeva, *CERN, Geneva, Switzerland*

T. Chen, W. Ueng, *Academia Sinica Computing Centre, Nankang, Taiwan*

J. Herrala*, B. Koblitz†, M. Lamanna, D. Liko, A. Maier, J. T. Moscicki ‡, A. Peters, *CERN, Geneva, Switzerland*

F. Orellana§, *CERN/University of Geneva, Geneva, Switzerland*

V. Pose, *Joint Institute for Nuclear Research, Dubna, Russia*

A. Demichev, *Moscow State University, Moscow, Russia*

D. Feichtinger**, *Swiss Institute of Particle Physics, Switzerland*

Abstract

This paper describes the distributed analysis prototypes for the four LHC experiments ALICE, ATLAS, CMS and LHCb on the EGEE middleware (gLite). Prototypes are currently under development within the ARDA project. Basic components of the prototypes, current status and future plans are covered in the paper.

ARDA PROJECT AND HEP EXPERIMENTS

ARDA stands for “A Realization of Distributed Analysis for LHC”. The goal of the project is to coordinate the activities to prototype the distributed analysis systems for the LHC experiments using gLite.

Experience of the existing Grid systems had proved the importance of the ability to test the functionality and performance of the system by real users at the very early phases and to get users feed back as soon as possible. The ARDA project has to ensure a very good communication channel between the gLite development team and the physics community of the LHC experiments. It has to provide an interface between gLite middleware and existing frameworks and applications of the LHC experiments and to deploy the prototypes supporting analysis tasks in the distributed environment using gLite middleware.

Currently ARDA does not focus on providing the common application layer for all four experiments but rather tries to help every experiment to fulfil its task in a more appropriate way. It is deploying four prototypes, one per experiment.

Since the prototypes are meant for users’ analysis, multi-user environment has a strong impact on the requirements concerning robustness, scalability and performance. On the other side, physicists with different levels of expertise will use the prototype therefore simplicity is an important issue.

By December 2004 ARDA foresees to deploy four end-to-end prototypes allowing users analysis.

GLITE MIDDLEWARE PROTOTYPE

The gLite Middleware prototype [1] became available for the ARDA team on the May 18th. During the first month ARDA users experienced multiple problems related to the stability of the new services, but in the following the prototype evolved with good progress. Currently, the test farm consists of only 2 worker nodes but core services are available:

- File catalogue
- Authentication module
- Task queue
- Meta-data catalogue
- Package manager
- Grid access service

A second site in Madison (Wisconsin USA) was set up by the end of June. Another essential component for the ARDA work was to enable access from the worker nodes to the experiments data files located at CERN Mass Storage System (MSS). This became available only in the middle of August.

The small size of both farms does not allow to try analysis prototypes and middleware performance at a realistic scale. Increasing the number of CPUs and number of sites becomes an urgent need and will be solved in the next weeks.

PROTOTYPES OVERVIEW

Table 1 gives an overview of four prototypes which are currently under development by the ARDA team. Each of them is focussed on a particular aspect, which is mentioned in the second column of the table.

A more detailed description of the prototypes is given below.

Table 1: Prototypes overview

LHC experiment	Main focus	Basic prototype component	Experiment analysis application framework	Middleware prototype
LHCb	GUI to Grid	GANGA [2]	DaVinci [6]	gLite
ALICE	Interactive analysis	PROOF [3] ROOT [4]	AliROOT	gLite
ATLAS	High level Service	DIAL [5]	Athena	gLite
CMS	Use max gLite functionality	Not yet defined by CMS	ORCA [7]	gLite

LHCb PROTOTYPE

The fundamental component of the LHCb prototype is GANGA (Gaudi/Athena aNd Grid Alliance). GANGA is a framework for jobs creating, submitting and monitoring. It can be used to steer different physics applications and to support execution on a variety of backends (local batch systems or grid systems). It is interacting with the experiments book-keeping data base. Users talk to GANGA through graphical user interface. In addition a command line interface is being developed by the ARDA team.

The main task of ARDA is to integrate GANGA with gLite (which in practical terms means to develop a GANGA job-submission handler for gLite) and to be able to run LHCb analysis jobs being submitted through GANGA on gLite middleware. Job-splitting and output merging functionality, retrieving of results and installation and handling of LHCb software on the gLite farm should be also implemented.

ARDA is directly participating in the GANGA project, developing code and handling release management process.

Other contributions to the LHCb software were done by the ARDA team:

- Integrating of GANGA with Dirac [8] (LHCb production system),
- Integrating of GANGA with Condor
- Stress testing of the LHCb metadata catalogue

Current status and short-term plans

The GANGA job-submission handler for gLite is being developed. LHCb user analysis jobs (DaVinci application) submitted to gLite through GANGA successfully ran on the gLite testbed.

Currently ARDA is focussed on the implementation of the job-splitting and output-merging functionality as well as using the gLite package manager for handling the LHCb software.

ALICE PROTOTYPE

Main focus of the ARDA project for ALICE experiment is to enable interactive analysis, which relies on ROOT/PROOF, emphasising the stability and performance.

The ARDA group is also involved in the gLite-related development, which is of big importance not only for ALICE, but also for the whole gLite user community. ARDA has developed C++ gLite access library and C library for Posix-like input/output. Libraries provide the security layer and the APIs to access gLite services from user applications.

Alice follows a rather pragmatic approach for the batch analysis model. The experiment provides a physics application (AliROOT), the user interface is ROOT, all other functionality is considered a middleware task.

Interactive analysis model [9] (See Fig.1) is based on PROOF.

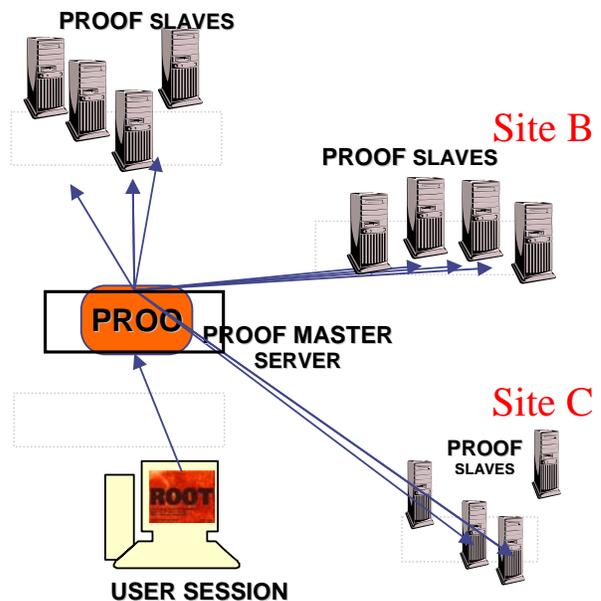


Figure 1 Interactive analysis model based on PROOF

The user starts an interactive local session by using the ROOT user interface. He gets connected to the PROOF master server which initiates slave nodes at the remote sites where the data files belonging to the user defined collection are located. The task gets splitted in sub-jobs (one per slave) which are processed in parallel. This results in a fast response for the user.

The main effort is currently to improve the robustness and error recovery of PROOF, to make the system independent of the grid flavour and to achieve a better performance. For example the parallelization of the start-up of the PROOF slaves was implemented just recently.

Current status and short-term plans

By the beginning of November both batch and interactive analysis should be possible with the gLite middleware.

ATLAS PROTOTYPE

The ATLAS prototype is based on DIAL (Distributed Analysis of Large datasets). DIAL is a high level service supporting user analysis on different types of environments.

A job in DIAL is composed of a dataset which specifies the input of the job, the application and the task describing the algorithm for data processing. The job is passed to the DIAL scheduler, which splits the job and submits sub-jobs to a given batch system or a Grid job queue. The scheduler collects the results of the accomplished jobs. All complexity of the task is hidden from the user.

The main focus for ARDA is to provide an integration of DIAL with the gLite middleware and to enable ATLAS analysis jobs (Athena application) submitted through DIAL to run on the gLite middleware.

ARDA is also performing the task of integrating the ATLAS data management system (Don Quijote) with gLite. Since ATLAS is currently using three Grid systems (Grid3, LCG2, NorduGrid) exchanging data between them is an important issue. Don Quijote provides the possibility for accessing files regardless of the Grid flavour used for data files production and registration. In this respect integration of Don Quijote with gLite would ensure the access from gLite to the full range of ATLAS data independently of data provenance.

Current status and short-term plans

The DIAL server has been adapted to CERN environment and installed at CERN. The first implementation of the DIAL scheduler for gLite has been developed. The DIAL scheduler has been tested to run Athena jobs on the gLite middleware. Currently the implementation is still depending on the shared file system for inter-job communication, this should be changed in the next version.

The gLite–Don Quijote integration is well under way.

Currently the work is focussing on enabling the gLite DIAL scheduler to interact directly with the underlying middleware services.

CMS PROTOTYPE

The detailed CMS prototype is still under discussion . Nevertheless ARDA started preparatory activities:

- Running CMS analysis jobs (ORCA application) on the gLite test-bed
- Populating the gLite catalog with CMS data collections, residing at the CERN MSS
- Trying the gLite package manager for installing and handling the CMS software

- Adopting gLite job-splitting service for the CMS use-cases

At the same time the ARDA group has started to work on the CMS prototype following strategy to use as much of the native gLite functionality as possible and only in case of very CMS specific tasks to construct something on top of existing middleware. This approach can prove that gLite functionality is complete and generic enough

And provide useful information for both CMS and the gLite team. On the other hand it can help to minimise time spent on integration with tool(s) CMS will finally choose.

Data management tasks are vital for CMS. Access to CMS data collections requires a POOL [10] catalogue and a set of COBRA (CMS reconstruction framework) [11] META files corresponding to a given collection. A system for publishing of POOL XML catalogues and COBRA meta files for user access was missing in the CMS data management structure, and this had a bad impact on the user analysis during CMS data challenge of 2004. The ARDA group therefore is participating in the development of the distributed data bases for publication (PubDB). A first prototype is already working at CERN and several CMS regional centres.

ARDA is also contributing to the redesign of the CMS reference data base (RefDB, CMS metadata catalogue) [12].

Stress testing of the RefDB has been performed by the ARDA team to estimate performance in case RefDB would be used for analysis purposes .

Current status and short-term plans

A first version of the prototype is under development.

The prototype is based on a workflow planner with a gLite back-end and a command line interface. It is interacting with RefDB/PubDB to get a required POOL catalogue and COBRA META files. It registers extracted files in the gLite catalogue, creates and submits jobs to gLite and retrieves the results. The tool is using the job-splitting and job-tracing functionality provided by gLite.

The first version of the prototype should be available to the users by the end of October at the same time when the gLite middleware should be deployed on a larger facility capable to support more users.

CONCLUSIONS

The ARDA team is working on the first version of the LHC analysis prototypes based on the gLite middleware providing a feedback to the gLite development team, directly influencing middleware evolution.

At the same time, ARDA is substantially contributing to evolve experiments software to match the requirements for the LHC distributed analysis.

To meet the milestone for the prototypes deployment in December 2004, ARDA needs more real users to test already available functionality and a considerable

increase of the hardware resources, both in terms of CPUs and in terms of sites.

ACKNOWLEDGEMENTS

We would like to acknowledge the gLite development team for their support and collaboration.

REFERENCES

- [1] B. Koblitz , “First Experience with the EGEE Middleware”, CHEP04 , Interlaken, Switzerland, 2004
- [2] A. Soroko , “*The Ganga User Interface – interface for physics analysis on distributed resources*”, CHEP04 , Interlaken, Switzerland, 2004
- [3] M. Ballintijn , “*Super scaling PROOF to very large clusters*”, CHEP04, Interlaken, Switzerland, 2004
- [4] ROOT Project, <http://root.cern.ch>
- [5] D. Adams , “*ATLAS Distributed Analysis*”, CHEP04, Interlaken, Switzerland, 2004
- [6] Da Vinci Project , <http://lhcb-comp.web.cern.ch/lhcb-comp/Analysis/default.htm>
- [7] ORCA Project , <http://cmsdoc.cern.ch/orca>
- [8] A. Tsaregorodtsev , “*DIRAC - The Distributed MC Production and Analysis for LHCb*”, CHEP04, Interlaken, Switzerland, 2004
- [9] F. Radenmakers, “*Global Distributed Parallel Analysis using Proof and Alien*”, CHEP04, Interlaken, Switzerland, 2004
- [10] POOL Project , <http://lcgapp.cern.ch/project/persist>
- [11] COBRA Project, <http://cobra.home.cern.ch/cobra>
- [12] V. Lefebure, J. Andreeva , “*RefDB: The Reference Database for CMS Monte Carlo Production*”, CHEP03, La Jolla, California 2003

* Funded by Federal Ministry of Education and Research, Berlin, Germany

† Funded by Federal Ministry of Education and Research, Berlin, Germany

‡ Funded by Particle Physics and Astronomy Research Council , UK

§ Financial support from the Paul Scherrer Institute and the European COST Action Nr.283

** Financial support by the European COST Action Nr.283