PHENIX

Data Management for Physics Analysis in PHENIX (BNL, RHIC)

Evaluation of Grid architecture components in PHENIX context

Barbara Jacak, Roy Lacey,
Saskia Mioduszewski,
Dave Morrison, Zhiping Qui,
Andrey Shevel, Irina Sourikova

CHEP03

PHENIX

Overview

- Collaboration
- Data volumes
- Aims, conditions, expectations, schemes
- Main needs: File/Replica Catalog, Job submission, data moving
- Known systems under development
- Working prototype setup
- Conclusion and Future plans

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

PHENIX

Brief info on PHENIX

- Large, widely-spread collaboration (same scale as CDF and D0)
- ~ 400 Collaborators
- 12 Nations
- 57 Institutions
- 11 U.S. Universities
- Currently in third year of data-taking

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

PHENIX

Distributed Data

- 100 TB/yr of raw data
- 100 TB/yr of reconstructed output
- ~ 25 TB/yr microDST (analyzed output)
- Primary event reconstruction occurs at BNL RCF (RHIC Computing Facility)
- Complete copy of reconstructed output at CC-J (Computing Center in Japan) and most of reconstructed output at CC-F (France)
- Multi-cluster environment and distributed data is our reality.

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

PHENIX

Conditions and Aims

- PHENIX is running experiment. That means it is not easy to test new program tools. We try to test new program environment at remote university – SUNY. We consider SUNY as typical example for remote site. Hopefully our experience could be used at all other remote midsized teams.
- To install, evaluate and tune existing advanced GRID program tools to make robust and flexible distributed computing platform for physics analysis. We plan to keep it with minimum of maintenance efforts.

Our Expectation on distributed computing

- By distributing the data we hope to decrease the latency (response time) for remote users for simulation and data analysis.
- Remote end users (physicists) will able to get more computing power in multi-cluster environment.

General scheme: jobs are going where data are and to less loaded clusters

SUNY RAM

Partial Data Replica

Main Data Repository

RCF

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

Replica/File catalog

- Needs to maintain information about significant files in different locations (at BNL and remote sites). Expected total number of files is about 10**5 at remote site and is about 10**6 at central PHENIX location. The consistency of the catalog is the requirement.
- At PHENIX there was created catalog ARGO http://replicator.phenix.bnl.gov/~replicator/fileCatalog.html based on DBMS Postgres (presented at CHEP03). At remote site (SUNY) there were tested replication of ARGO and adapted version of MAGDA (presented at CHEP03).

PHENIX

Computing Resource Status and job submission

- We need for simple and reliable tool to see current status of available computing resources in multi-cluster environment (GUI and CLI).

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

Known systems under development

- "**Chimera**: A Virtual Data System for Representing, Querying, and Automating Data Derivation" *http://www.griphyn.org/chimera/*
- **Clarens:** "The Clarens Remote Dataserver is a wide-area network system for remote analysis of data generated by the CMS" *http://clarens.sourceforge.net/*
- **AliEn** *http://alien.cern.ch/* AliEn is a GRID prototype created by the Alice Offline Group for Alice Environment.

PHENIX

Need for prototype

- All mentioned systems are not trivial, they include many components.
- The underlying structure (Globus) is not trivial as well.
- To be sure for base structure we have to test it first.
- Prototyping with minimum of manpower - a balance between potential profit and achievable results.

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

Job Submission: Initial Configuration

In our case we used two computing clusters which are available for us

- At SUNY (RAM: *http://nucwww.chem.sunysb.edu/ramdatal*)
- At BNL PHENIX (RCF: *http://www.rhic.bnl.gov/RCF/*)

- **Globus gateways are used in both cases.**

Reporter: Andrey Shevel@mail.chem.sunysb.edu

Job Submission and retrieval Commands

- Submission to SUNY - Submision to RCF	- Job queue status at SUNY - Job queue status at RCF
-Submission to less loaded cluster gsub script	-Get job standard output + standard error output gget job-id
-Submission where the file is located gsub-data script file	-Job status gstat job-id

Reporter: Andrey Shevel@mail.chem.sunysb.edu

PHENIX

Data moving

- **Our Conditions**
  - From time to time we need to transfer a group of files (from about 10**2 to 10**4 files) in between different locations (in between SUNY and BNL). Apparently we need to keep newly copied files in Replica/File Catalog. Trace of all data transfers is required as well.
  - Now it is realized with our MAGDA distribution (with GRIDftp). To see SUNY data catalog based on our MAGDA distribution you can use "*http://ram3.chem.sunysb.edu/Main*"

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

Requirements to deploy the prototype

- **To deploy the prototype of computing infrastructure for physics analysis we tested somebody needs:**
  - PC, Linux 7.2/7.3 (it was tested);
  - Globus Tools 2.2.3;
  - To get two tarballs with scripts (including SUNY distribution for MAGDA): magda-client.tar.gz and gsuny.tar.gz. Dedicated site can be seen *http://nucwww.chem.sunysb.edu/ramdata/docs/globus.htmlx*
  - MySql server (if required) + MySql client + Postgres + C++ interfaces + perl interfaces;
  - To get Globus certificates.

PHENIX

CONCLUSION

- Transition to GRID architecture could only follow the understanding in GRID computing model of all involved people . It is not business for one person.
- GRID tools have to be available and supported on centralized computing resources.

Reporter: Andrey
Shevel@mail.chem.sunysb.edu

PHENIX

Future plans

- To add another computing cluster to our prototype.
- To evaluate more sophisticated and advanced middleware components.

Reporter: Andrey
Shevel@mail.chem.sunysb.edu