LCG PEB Applications Area Meeting

MySQL Service Plans and Needs in ATLAS

November 20, 2002

Alexandre Vaniachine (ANL)

Database Needs

- Current ATLAS software needs for relational databases are coming from three domains
- Data Challenges (support of ATLAS Grid tools):
  - Metadata catalog: AMI
  - Replica catalog: Magda
  - Virtual Data Catalog: VDC
  - Grid portal: ATCOM
- Offline Software (framework services, simulations)
  - primary numbers, conditions
- On-line/Testbeam
  - LAr, tilecal,...

Observations

ATLAS:

- Three domains - three different schedules and priorities lists
- Same database technology was deployed by different software developers groups everywhere
- ATLAS gained experience in use of MySQL

LHC/EDG:

- Other LHC experiments are using MySQL for many different needs
- LCG persistency project POOL selected MySQL for implementation of FileCatalog and Collections
- EDG project uses MySQL for Replica Location Service

Considerations

- Oracle is used in some LHC experiments for certain needs
- Oracle support is available from IT/DB
- In a comprehensive benchmark of the latest available versions of five server databases done by the eWEEK Labs/PC Labs both Oracle 9i and MySQL had the best performance and scalability. Both Oracle and MySQL come on top of other databases (IBM DB2, Microsoft SQL Server and Sybase ACE) scaling up to a 1000 simultaneous connections, with Oracle 9i just very slightly ahead of MySQL for most of the run.

  http://www.eweek.com/article2/0,3959,293,00.asp

- LCG persistency project POOL plan to use both Oracle and MySQL (BlueprintPlan.xls)

Experience

- MySQL functionality is sufficient
- Users would like faster database response
- Single server model is not convenient:
  - When developed databases started to be used in production this prevented server shutdown for maintenance, upgrades
  - Security upgrades done without preliminary studies of their potential side effects resulted in service disruption for many users
- Current authorization procedures are inadequate:
  - clear-text database passwords are used in scripts or distributed via e-mails

Outcome

- MySQL server was deployed at CERN for evaluation
  (thanks to Gordon Lee from IT/DB and RD Shaffer)
- MySQL services working group was established:
  Christian Arnault , Nectarios Benekos, Jerome Fulachier,
  Jorge Lima, Antoine Perus, Alex Undrus, Sasha Vaniachine
  People who want to participate are welcome to join
- Group will define the requirements for MySQL services, and propose a phased deployment plan
- These requirements will provide input to the LCG infrastructure effort
- During discussion with MySQL services group members and a meeting at LAL on October 11 a proposal for a multi-tier server model was developed

Three Tiers for Database Services

Production servers
- deployed for use in production data processing
- stable and tested server versions are used
- made highly secured and reliable (backed up, archived, replicated, protected)

Software development servers
- used to develop databases and test the software before its deployment on the production servers
- validated and tested server version is used
- used during release tests
- used for tutorials, demos

Database development servers
- needed for evaluation of the new server releases
- needed for testing effects of the new security upgrades
- needed for testing effects of the operating system upgrades
- may reside on the same hardware as the development server, although this will prevent evaluation of the new operating systems and the effects of network security upgrades

Four User Roles

Access rights will be granted according to the roles:

- Administrators
  - manage the installation of database servers
  - give access rights to other users
- Developers
  - develop database for some software application, e.g. online conditions database, calibration database, tag collector database etc.
  - have full access rights to that database on the server
- Data writers
  - these are user applications or scripts that fill the data to the database
  - have only the rights to add the data, not to replace or delete
- Data readers
  - these are user applications or scripts that read data from the database
  - no password is needed for read data access

It is expected that only a limited number of users will have administrative rights to on the server

For database development servers, more users may have full access rights

Long-terms solutions for the data writer password management issues have to be investigated

Support Issues

- We think that for production servers, CERN resources should be first considered
- However, considering that this solution is submitted to quite wide preconditions (relationships with computing plans, COCOTIME, etc.) it is probably wise to consider short term solutions
  - using private Atlas resources at CERN (atlasdev1)
  - using offered resources from external sites (BNL, Grenoble are current examples)
- In case other computing centers would offer similar services, it may be appropriate to undertake agreements on the level of service (Similarly to software agreements)

Web Services

- One central aspect of database services is the web access to those services. These include management services (such as phpMyAdmin) or Web applications accessing the databases (e.g. the NOVA browser). It is noticed that the providing Web access to databases is a different issue than managing the database server and that the Web server where applications run may be different than the SQL server. The constraints are different (different technologies are required, such as php, Java, etc.)
- It is expected that several servers will provide at least a php resource
- So far we have CERN home service and a possibility with the LCG web server (not operational these days). We need strong support for these servers, since it's likely that Web services may become part of normal software usage

Short-term Solutions

Production servers

- Master server - atlasdev1.cern.ch
- Replica server 1 - db2.usatlas.bnl.gov (to be deployed)
  - Data backup and archiving by BNL Tier 1 Facility
- Replica server for AMI database - isngrid1.in2p3.fr
  - Data backup support by Grenoble
- Web access service provided by BNL and Grenoble
- Databases to be deployed on production servers:
  - AMI, Magda, VDC, NOVA

Development servers

- Master server - db1.usatlas.bnl.gov
- Replica server - atlassw1.phy.bnl.gov
- Web access service - atlassw1.phy.bnl.gov
- back-up support by BNL Tier 1 Facility

Evaluation server

- lxshare0701d.cern.ch (provided by LCG POOL project)

Future Developments

- Blueprint RTAG recommendation (laptop user case)

=> MySQL Embedded Server Technology

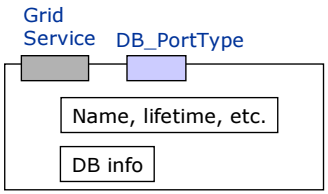
- Deployment model should provide for future Grid integration issues (OGSA, UDDI, Grid database services architecture)

=> MySQL X509 Certificate Authorization Technology

Grid example by Kate Keahey

Grid Service Example: Database Service

- A DBaccess Grid service will support at least two portTypes
  - GridService
  - Database_PortType
- Each has service data
  - GridService: basic introspection information, lifetime, …
  - DB info: database type, query languages supported, current load, …, …



Long-term Evolution

The three ATLAS software database clients:

- Data Challenges (AMI, Magda, ATCOM, VDC)
- Core Software (primary numbers, conditions)
- On-line/Testbeam

In the long run they should be supported by different sets of servers

Extreme range of requirements for on-line database and the end-user laptop

Need for two more tiers?

⇒ Ultimate robustness for the Tier 0 (on-line DB)

⇒ Least cost for the Tier 4 (end-user laptop)

Oracle Deployment

Can we have an Oracle on every user laptop?
Can we have embedded Oracle?
Can we have Oracle Grid certificate authorization?

These issues call for two technologies in
a multi-tier model (thanks to Tom LeCompte)

Technologies ranging from
Oracle on a production tier server
(ultimately robust service level)
to
MySQL on an end-user laptop
("do-it-yourself" service level)

Potential IT/DB Roles

Provide support for production-level tiers of database servers

Develop and implement strategies for
- fault-tolerant database services
- data backup and archiving
- crash recoveries

Give advice for evaluation tier:
MySQL versions, optimizations, hardware platforms

Investigate value of purchasing support from MySQL AB

Partnership agreement with MySQL AB to develop code providing
- better support for binary data delivery (Conditions DB)
  binary data transfer through TCP/IP sockets
  direct client access to data (a la Objectivity or some RDBMS)
- support for non-X509 certificates (grid proxies)
- …