

Grid Technologies for Distributed Database Services



3D Project Meeting
CERN, May 19, 2005
A. Vaniachine (ANL)



Outline

- Databases and the Grid
- Expectations and experience
- Grid-enabled database deployment
- Grid-enabled database technology
- Databases for data caching

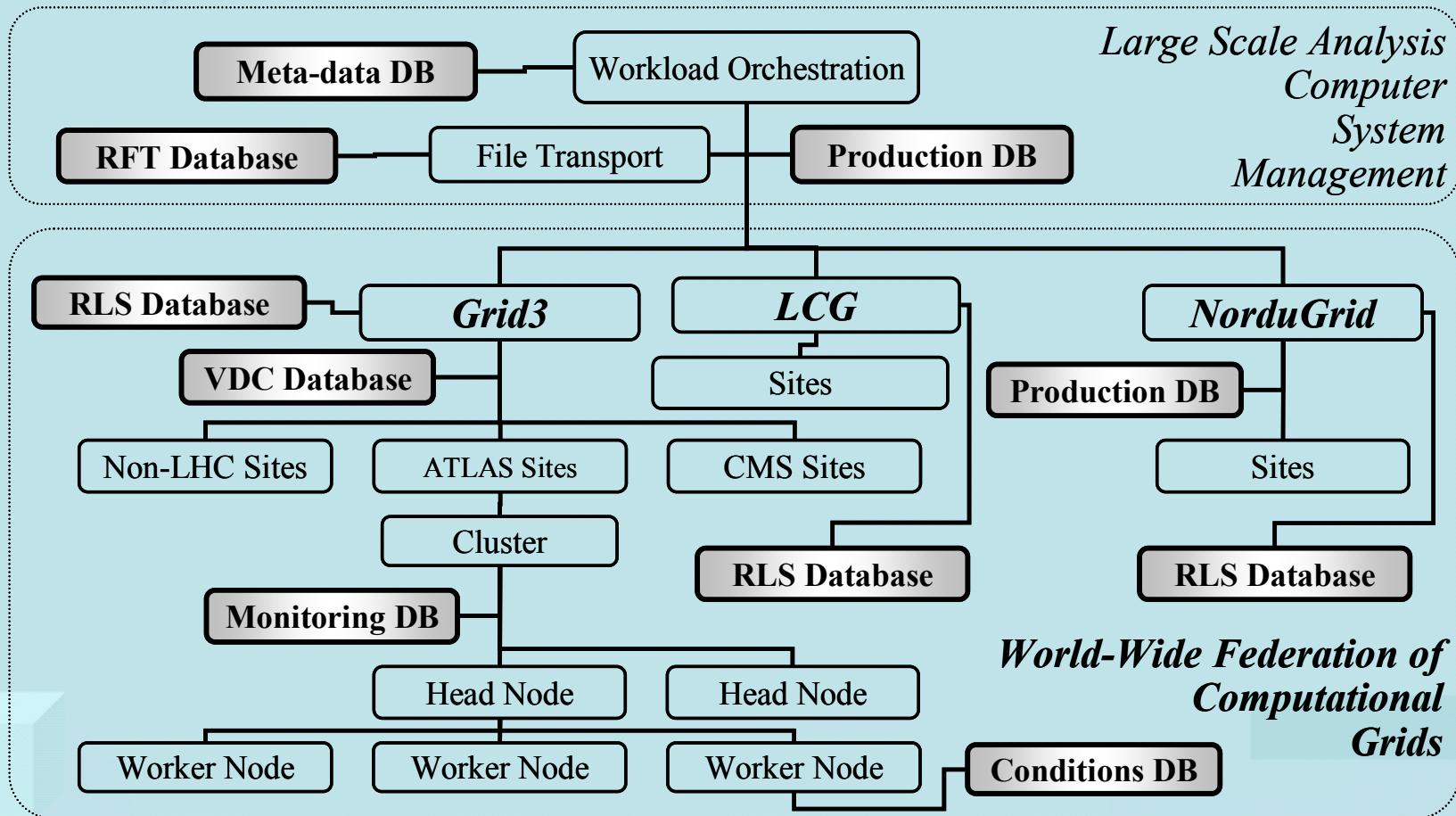


Databases and the Grid

- 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



Emerging Hyperinfrastructure



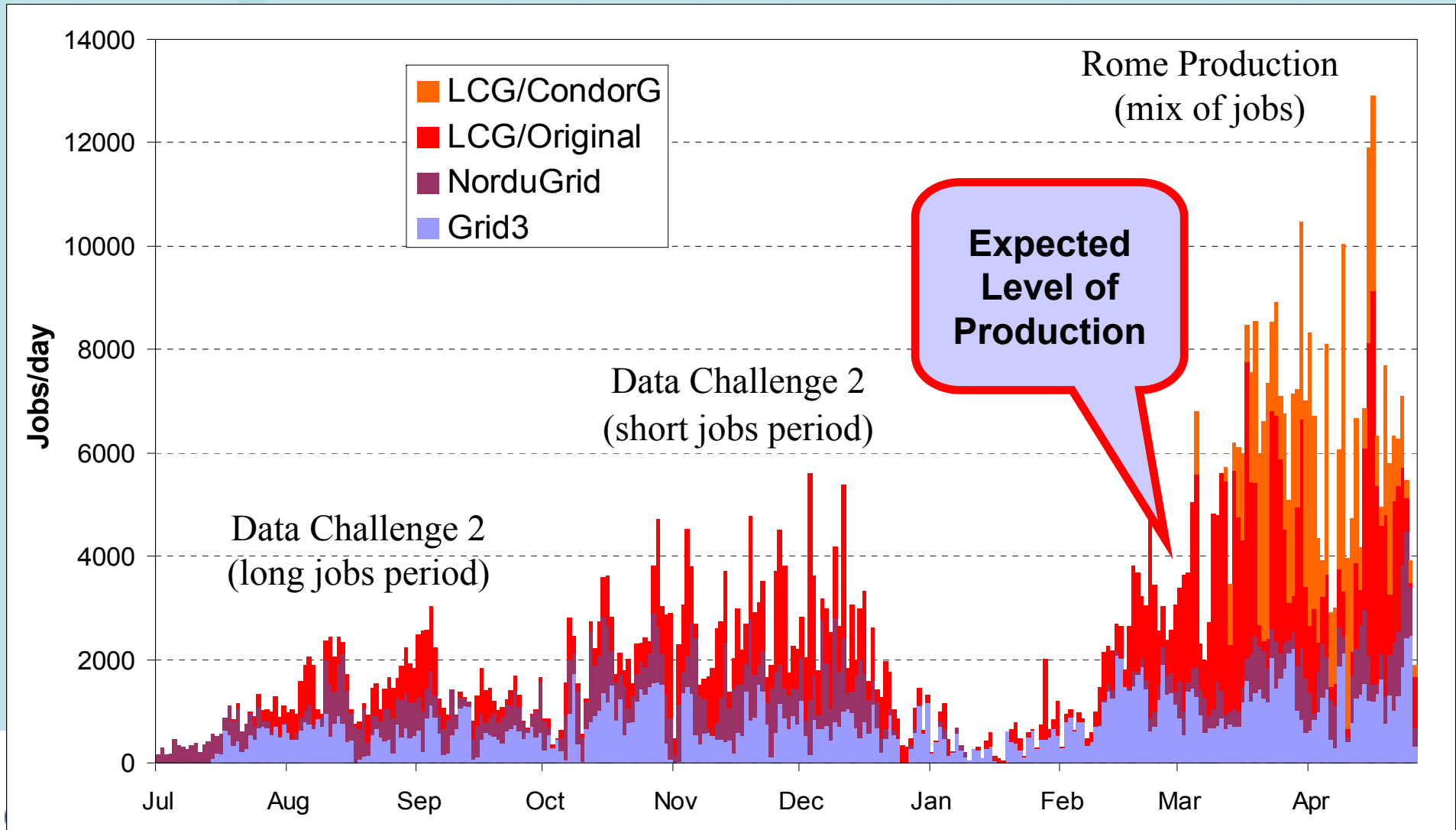


Data Workflow on the Grids

- An emerging hyperinfrastructure of databases on the grid plays the dual role: both as a built-in part of the middleware (monitoring, catalogs, etc.) and as a distributed production system infrastructure orchestrating scatter-gather workflow of applications and data on the grid
- To further detail the database-resident data flow on the grids ATLAS Data Challenges exercise the Computing Model processing and managing data on three different grid flavors



Official Production Being Done Only on Grids





Expectations and Realities

Expectations

- Scalability achieved through replica servers deployment
- Database replicas will be deployed down to each Tier2 center
- To ease an administration burden the database server replica installation was reduced to a one line command

Realities

- Only the centers that has problems with access to the central databases (because of firewalls or geographical remoteness resulting in low data throughput) deployed the replica servers
- On some days the database services were a bottleneck
- **Concerns expressed in regard to replica synchronization with the central servers**



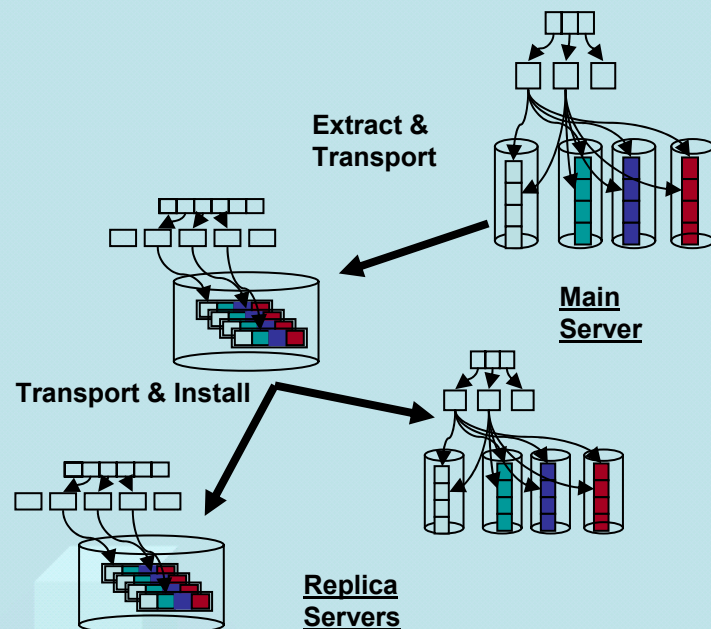
Database Deployment Options

- Central Deployment:
 - Most of ATLAS current experience in production
 - Scalability problem
 - Remote site firewall problem
 - Remote site server-side timeouts problem
- Replica deployment on worker node:
 - Extensive experience in ATLAS Data Challenge 1 (next slide)
 - Replica update problem
 - Fault-tolerance cleanup problem
- Replica deployment on the head node (gatekeeper):
 - Proof-of-the-principle deployment performed
 - Work-in-progress – main theme of this presentation



Worker Node Deployment

Extract-Transport-Install



- MySQL simplified the delivery of the extract-transport-install components of ATLAS database architecture to provide database services needed for the Data Challenges for sites with Grid Compute Elements behind closed firewalls (some sites on Grid3 and NorduGrid)



Head Node Deployment

Proof-of-the-principle demonstrated at the SMU site of the Grid3 production testbed

- in collaboration with Yuri Smirnov, iVDGL

1. globus-job-run command to shutdown the old mysql server replica (that was previously installed by the site administrator)
2. globus-url-copy command to transfer the mysql replica rpm to the site
3. globus-job-run command to install the relocatable rpm into the atlas data area on the site.
4. globus-url-copy command to transfer to the site the modified mysql configuration file for the main server instance
5. globus-job-run command to launch the main mysql server replica instance
6. globus-job-run command to verify the main mysql server replica instance launch
7. globus-url-copy command to transfer to the site the modified mysql configuration file for the main server instance.
8. globus-url-copy command to transfer to the site the modified mysql configuration file for the ANSI-compatible mysql server instance
9. globus-job-run command to launch the ANSI-compatible mysql server replica



Work in Progress

Our work discussed at the recent GriPhyN meeting:

http://www.interactions.org/sgtw/learnmore/GriPhyN_more.html

- Discussions findings
 - Head Node deployment fits well in the grid architecture
 - Head Node deployment of service tunnels is an option
 - Worker node deployment used in other sciences (SQL Server)
 - The newer version of the GRAM protocol provides an option for the process to be automatically restarted upon failure
- Work is now in progress to
 - Merge all or most of deployment steps into one job
 - Develop deployment procedures for the OSG testbed (globus 3) (Grid3 production testbed is based on globus 2.4 technology)
 - Use new GRAM protocol options



Database Authentication

Two models and their data transport implications

- A separate layer does the grid authentication:
 - Spitfire (EDG WP2) – SOAP/XML text-only data transport
 - DAI (IBM UK) – Spitfire technologies + XML binary extensions
 - Perl DBI database proxy (ALICE) – SQL data transport
 - Oracle (separate grid authentication layer)
- Authentication is integrated in database server:
 - Instead of surrounding database with external secure layers the safety features are embedded inside of the code
 - By pushing secure authorization into the database engine the inefficient data transfer bottlenecks are eliminated



Grid-enabling Databases

- A small business with a long record of collaboration with Fermilab and experience in Oracle developed:



Database Grid Solution

Technology enabling queries of distributed, heterogeneous databases, all through one simple, Grid enabled interface.

- DOE funds are now awarded for embedded grid-enabled database authentication
 - New collaboration project with ANL



FroNTier for Data Caching

- We consider the technology for the conditions/calibrations data caching at the remote site as critical
- We consider FroNTier as a very promising technology that does address this critical need
- We would like to learn more on the requirements for FroNTier deployment in the grid environment
- We would like to learn more on the resolution of the cache invalidation problem
 - How do we guarantee that the job running somewhere on the grid does get access to the very latest calibration data instead of some old obsolete data that were cached in the FroNTier web cache before?



Databases for Data Caching

- “Cache validation is a ‘big problem’”
Peter Yared, *Founder and CEO*
ActiveGrid, a commercial open source company
- ActiveGrid develops a data caching solution for mid-tier
 - An equivalent to our data caching deployment on grid Head Node
- We are interested in collaboration with ActiveGrid on mid-tier data caching technologies
- We believe that the data caching technology based on MySQL database capabilities will resolve the cache validation problem



Summary

Integration of grid technologies to build the distributed database services hyperinfrastructure

- Efficient data transport for grid-enabled MySQL
 - Project funded
- Grid-enabled database deployment solutions
 - Proof-of-the-principle demonstrated
- Mid-tier data caching
 - Planning evaluation of technologies