



Argonne
NATIONAL
LABORATORY

... for a brighter future



U.S. Department
of Energy

UChicago ►
Argonne_{LLC}



U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC

Database Requirements Updates from LHC Experiments

WLCG Grid Deployment Board Meeting

CERN, Geneva, Switzerland

February 7, 2007

Alexandre Vaniachine (Argonne)

Worldwide LHC Computing Grid
Distributed Production Environment for Physics data Processing

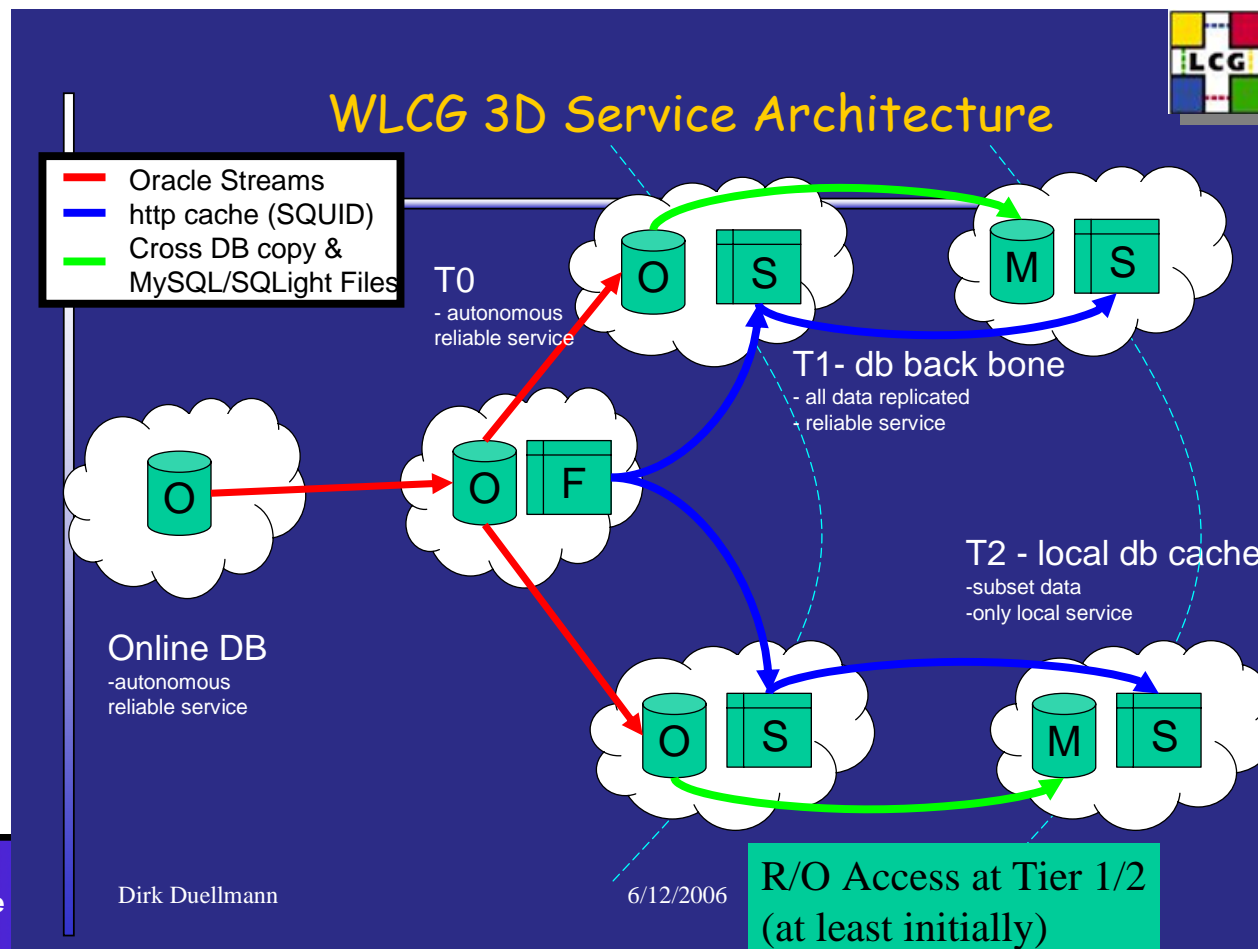


Database Services for the Grid

- Database access is vital for the real data operations
 - Access to calibration data is critical to the event data reconstruction
 - Other use cases also require distributed database access

WLCG 3D Project:

- Focus on Oracle streaming to Tier 1
 - ATLAS
 - LHCb
- FronTier technology for database data caching at Tier 1 and beyond
 - CMS
 - ATLAS (evaluating)
- Regular updates of DB requirements



WLCG Collaboration Workshop (Tier0/Tier1/Tier2)

mini workshop (26 January 2007) - Windows Internet Explorer

http://indico.cern.ch/conferenceDisplay.py?confid=10132

Database mini workshop (26 January 2007)

Database mini workshop

Friday 26 January 2007
from 09:00 to 18:00
at 160 1-009

Description: To join the phone conference, do one of the following: a. Dial +41 2276 76000 and enter access code 0103408, or b. To have the system call you, click here: <https://audioconf.cern.ch/call/0103408>

[Friday 26 January 2007](#) | [top↑](#)

Friday 26 January 2007

12:35

lunch break

13:40->17:40 Experiment Experience and Plans

- | | | | |
|-------|---|------------|--|
| 14:00 | Database authentication in CORAL and COOL (20') | (Slides) | Giacomo Govi (CERN) |
| 14:20 | COOL status and plans (20') | (Slides) | Andrea Valassi (CERN) |
| 14:40 | COOL tests with FroNTier in ATLAS (20') | (Slides) | David Front (Weizmann Institute) |
| 15:00 | CMS online/offline replication studies (20') | (Slides) | Marcin Bogusz |
| 15:20 | coffee break | | |
| 15:50 | Status and plans for online database setups in ALICE, ATLAS, CMS and LHCb (30') | (Slides) | Frank Glege (CERN) , Hans von der Schmitt (MPI for Physics, Munich) , Beat Jost (CERN) , Peter Chochula (CERN) |
| 16:20 | ATLAS experience, plans and resource request update for T0 and T1 (15') | (Slides) | Alexandre Vaniachine (ANL) |
| 16:35 | CMS experience, plans and resource request update for T0, T1 and T2 (15') | (Slides) | |
| 16:50 | LHCb experience, plans and resource request update for T0 and T1 (15') | (Slides) | |
| 17:05 | wrap up session (20') | | |

Experiments' Session

- Separate updates of online and offline requirements
- Collect the main resource requirements pointing out the proposed changes for the next six months
 - present them to the sites
 - **this GDB Meeting**



CMS Requirement Update and Plans

Lee Lueking

3D Workshop
26 January 2006



LCG 3D Workshop
26 January 2007

LHCb experience, plans and resource request update

Marco Clemencic
Marco.Clemencic@cern.ch



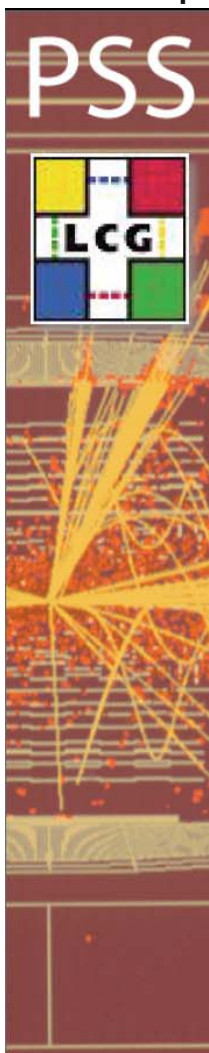
ATLAS experience, plans and resource request update for T0 and T1

WLCG Collaboration Workshop (Tier0/Tier1/Tier2)
CERN, Geneva, Switzerland
January 22-26, 2007
Alexandre Vaniachine (Argonne)



Baseline Requirements

- Presented at the previous GDB Meeting:





PSS Experiment Requirements

CERN IT Department

- Baseline (unchanged since November '05 GDB)
 - ATLAS T1
 - 3 node db server, 300 GB usable DB space
 - LHCb T1
 - 2 node db server, 100 GB usable DB space
 - CMS T1+2
 - 2 squid server nodes, 100 GB cache space per node
- ATLAS
 - R. Hawkings collected COOL requirements and proposed a model for total volume and client accesses - preliminary number: some 1.7 GB /day
 - J. Cranshaw gave first numbers for TAGS
 - some 4 TB /year (not including indices), access not clear yet
- CMS and LHCb
 - Currently no major upgrades requested

Dirk.Duellmann@cern.ch



LCG 3D Project Status - 4

- CondDB
 - Validation of hardware setup
 - Validation of software libraries (COOL and CORAL)
 - Add missing Tier-1s (March)
 - Switch to production systems
 - LHCb Alignment Challenge (April)
- LFC
 - Get all Tier-1s in production

- CondDB (per site)
 - Before start-up (no change)
 - 2 nodes Oracle RAC
 - 100 GB user space
 - Start-up
 - Possible CPU upgrade
 - More disk space (200GB)

- LCF (per site)
 - Before start-up
 - 2 nodes Oracle RAC
 - 2 r-o LFC servers
 - 200 GB user space
 - Start-up
 - Possible CPU upgrade
 - More disk space (0.5TB)



Rough Estimates of CMS DB Resources

March 2007 through August 2007

Area (CMS contact)	Disk Maximum (GB)	Concurrent Users Peak usage	Transactions Peak usage (Hz)
Online P5	500	20	10
Offline Conditions Tier-0 (CMSR)	500 (DB) 100 (per Squid) 2-3 Squids	10 (DB)* 10 (Squid) * Incl. On2off Xfer	10 (DB)* 10 (Squid) * Incl. On2off Xfer
Offline Conditions Tier-1 (each site)	100 (per Squid) 2-3 Squids/site	10(per Squid) >100 Σ all sites	10 per (Squid) >100 Σ all sites
Offline DBS Tier-0 (CMSR)	20	10 (currently ~2)	10 (currently ~5)



Production Hardware Needs

- **Offline Database:** The RAC (CMSR) is sufficient and support has been very good. Upgrade to 8-nodes is planned and will be used.
- **Frontier: cmsfrontier1,2,3**
 - System is now in place with 3 machines, load balancing and failover via DNS Round Robin.
 - Limited by 100 Mbps ethernet.
- Will install 2 additional Squid servers that are well connected, network-wise, to processing resources at CERN, like CMS T0 (prompt reco) farm, Grid CE's at CERN, et cetera.
- **Dataset Bookkeeping Service (DBS):**
 - Lxgate40 is current production server.
 - Will request additional server to load balance w/ lxgate40.
 - Will request additional server dedicated to Tier-0 farm.



CMS Squid Deployment

- Squids deployed at 30 Tier 1 and 2 sites.
 - Tier 1: LCG: ASCC, IN2P3, PIC, RAL, CNAF, FZK, CERN and OSG: FNAL
 - Tier 2: LCG: Belgium, Legnaro, Bari, CIEMAT, DESY, Estonia, CSCS, Pisa, Taiwan, GRIF, Rome, OSG: UCSD, Purdue, Caltech, Nebraska, Florida, Wisconsin, MIT +
- Will install additional Tier-2, and begin installing at Tier-3 sites as needed.
 - Possibly 20 or 30 additional sites by August
 - Including: Budapest, Imperial, NCU/NTU, ITEP, SINP, JINR, IHEP, KNU,

ATLAS Schedule

- Preparation for ATLAS CDC in Spring:
 - Install SRM 2.2 and commission: March/mid-April
 - **3D services for CDC production: February**
- Readiness for the final phase of ATLAS FDR in September-October:
 - 3D services stress testing: June/July
 - FDR initial phase: July/August
 - **Max-out T1 3D capacities: August/September**
 - *most T1s now have only two nodes for 3D*
 - *ATLAS DB will need several TB per year*
 - 3D Production running: October-December

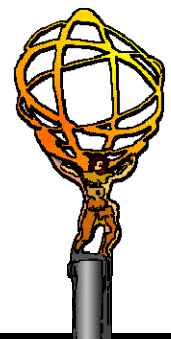


Request to max out T1 3D capacities

- We have to max out (fully utilize) the available resources:
- We expect to be memory bound, not CPU bound
 - Since mostly jobs request the same data

Max out request:

- There are six Oracle licenses per T1 (for ATLAS use)
 - If you have not used all your Oracle licenses – use it
 - *Deploy the third node*
- If your Oracle nodes do not have 4 GB of memory
 - Upgrade to 4 GB of memory
- If your Oracle nodes do not have 64-bit Linux
 - Upgrade to 64-bit Linux
 - *In preparation for memory upgrade beyond 4 GB*
- Since this is a rather modest upgrade request, the upgrade can be accomplished by August/September



Summary of ATLAS Request for Resources

Readiness for Calibration Data Challenge (by March):

- Upgrade 3D services to production level

Readiness for Final Dress Rehearsal:

- Max out request (by August/September):
 - Deploy the third node for ATLAS (for sites without third node)
 - Upgrade to 4 GB of memory (for sites with less than 4 GB of memory)
 - Upgrade to 64-bit Linux (for sites without 64-bit Linux)
 - In preparation for memory upgrade beyond 4GB

Readiness for Real Data:

- Disk space upgrade (by December)
 - T1 sites with Oracle disk storage less than 1 TB for ATLAS user tablespace are requested to double their capacities



ATLAS Disk Schedule for T1 Databases

- Gancho and Florbela made detailed estimates for ATLAS nominal year (including indexes overhead)
 - COOL storage **0.8 TB per year**
 - TAGS storage **6 TB per year**(200 Hz during active data taking, $2 \cdot 10^9$ events per year)
- **We need to ramp up T1 Oracle capacities to these numbers**
 - The priority is storage for COOL
 - *ATLAS TAGS can be stored outside of Oracle, in ROOT files*

Ramping up to ATLAS nominal running conditions:

- 2008 will be 40% of a nominal year (starting in June)
 - 0.2 TB COOL + 1.4 TB TAGS
- 2009 will be 60% of a nominal year
 - 0.5 TB COOL + 3.7 TB TAGS



Baseline Requirements Update towards LHC turn-on

- ATLAS T1
 - 3 node db server, 0.3 TB usable DB space **ramp-up to 0.6 TB**
- LHCb T1
 - 2 node db server, 0.1 TB usable DB space **ramp-up to 0.2 TB**
- CMS T1+2
 - 2-3 squid server nodes, 0.1 TB cache space per node

Towards LHC turn-on:

- ATLAS and LHCb
 - For the LHC turn-on **doubling of the storage** requested for 2006
- CMS
 - Currently no major upgrades requested

and beyond:

- ATLAS - first LHC operations ramp-up estimate
 - 2008: 1.6 TB for Oracle user tablespace
 - *Not including site-specific storage for backup and mirroring*





Argonne
NATIONAL
LABORATORY

... for a brighter future



U.S. Department
of Energy

UChicago ►
Argonne_{LLC}



A U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC

Backup Slides

Worldwide LHC Computing Grid
Distributed Production Environment for Physics data Processing



Summary of Online Requirements



Status and plans for on line database set ups in ALICE, ATLAS, CMS and LHCb

Database mini workshop

Friday, 26.01.07

Frank Glege

- Not covered in this (offline) summary



ATLAS Disk Storage Request

- Current capacities for ATLAS (collected by Gancho and Florbela)

T1 site	Disks (TB)	Nodes	CPU/node	Memory (GB)	64-bit	Comment
IN2P3	0.5	2	4		yes	Shared with LHCb
CNAF	1.0	2	4			
RAL	0.4	2	2			
Gridka	0.6	2	4			
BNL	0.4	2	4			ATLAS site
ASGC	1.6	2	4		yes	Expandable to 10TB
TRIUMF	3.1	2	2		yes	
NDGF						3D Phase II T1 site
PIC						3D Phase II T1 site
SARA						3D Phase II T1 site

- T1 sites with Oracle disk storage less than 1 TB are requested to double their capacities by December



ATLAS Requirements for FroNTier

- We can deal with FroNTier cache consistency largely by policies
 - e.g only accessing the data by frozen CondDB tag
 - *so the cache result for this query will not go stale)*
 - and/or by having the cached data expire after a certain time
 - *as shown by CMS in their FroNTier presentation this week*
 - which is quite encouraging

- If ATLAS choose the FroNTier alternative the major issue will be to get the T2 sites to deploy squid caches
 - in principle, this is straightforward
 - *as shown by the CMS deployment at ~20 T2 sites now*
 - but of course requires some hardware resources and some support at the T2s

