



# Advanced Technologies for Scalable ATLAS Conditions Database Access on the Grid

A. Vaniachine (Argonne)

P. Nevski (Brookhaven)

R. Basset, L. Canali, M. Girone, R. Hawkings, A. Valassi, F. Viegas (CERN)

G. Dimitrov (DESY)

R. Walker (LMU Munich)

A. Wong (TRIUMF)

# Outline

- Introduction
  - Conditions Data
  - Access patterns
- Problems
- Solutions
  - Staggered job starts
  - Pilot queries
  - SQLite
  - FroNTier
- Future work and Conclusions

# Conditions Data

- Time dependent data on the "condition" of the detector
  - Detector Control System
    - Voltages, temperatures, magnetic field map, ...
  - Calibration and Alignment
    - Prompt calibrations, and ones added later
- DCS Interval Of Validity(IOV) ranges from few seconds to few days
  - 1-2GB/day
- Calibrations typically valid for a day or so
  - 100MB/day
- Processing tasks over a large run range need access to many different data, eg. ~500GB/year

# Access Patterns

- Reconstruction/reprocessing of real data
  - run by run pass through all data
  - 1 job processes few minutes of data
    - loads ALL conditions data on first event
    - 5mins Db access per 8hrs jobs
  - concurrently running jobs  $O(2000)$  [T1 cpu]
- Alignment, e.g. ID
  - every event in large run range. Highly parallel.
  - very high Db query load
- Analysis
  - little experience so far
  - may skim over large run range
    - access Db for every event!

# Db Deployment

- ATLAS Conditions Data stored in Oracle Database
  - RAC for redundancy and scaling
- T0 RAC used for primary reconstruction and CAF access
- Each of 10 Tier1s has Oracle RAC replica: WLCG 3D group
  - replication via Streams
    - overloaded Db can affect Streams
  - T1 RAC used for reprocessing and analysis
    - 1<sup>st</sup> level of scaling: not all using T0
    - must check usage scales withing T1 cloud



# Finding Oracle Replica on the Grid



- Job landing on WN needs to know which Db to use
- LCG InfoSys does not have this information
- Use mapping from WN domain name, e.g.
  - .de .pl .cz FZK
- Private networks, missing mappings - override with env
  - `export ATLAS_CONDDB=gatekeeper.hostname`

# Testing

- To validate database performance at peak loads, we tested database scalability with very high numbers of concurrent jobs
- This has been achieved through coordinated database stress tests performed in series of *ATLAS* reprocessing exercises at the Tier-1 sites
- The goal of database stress tests is to detect scalability limits of the hardware deployed at the Tier-1 sites, so that the server overload conditions can be safely avoided in a production environment
- And to test potential solutions

# Problems

- During massive data reprocessing operations, an ATLAS Conditions Database application must support concurrent access from numerous ATLAS data processing jobs running on the Grid
  - observation: some T1 RACs could not cope
  - peak loads are much higher than average access rates
    - size(\$\$) RAC for average.
  - required job start rate, i.e. concurrent connections, is modest
    - 5min Db access in 8hrs with 2000cpus ~ 20 concurrent accesses
    - but little control over concurrency of job starts
- ID alignment task much heavier than reprocessing
- Access from T2 limited by WAN round-trip time (RTT)
  - multiple RTTs per query



# Stagger job starts in Panda

- Panda is the Production system pull-method WMS
  - Pilot job pulls real payload directly to WN
  - activated->sent->starting->running
- Add sleep delay as function of number of job sent
  - delay proportional to number of jobs in 'sent' state
  - job stays in 'sent' state for delay seconds.
  - tune to get N starts per minute(only need ~4 to keep 2000 cpu full)
- T0 primary reconstruction also staggers job starts @ 2/min

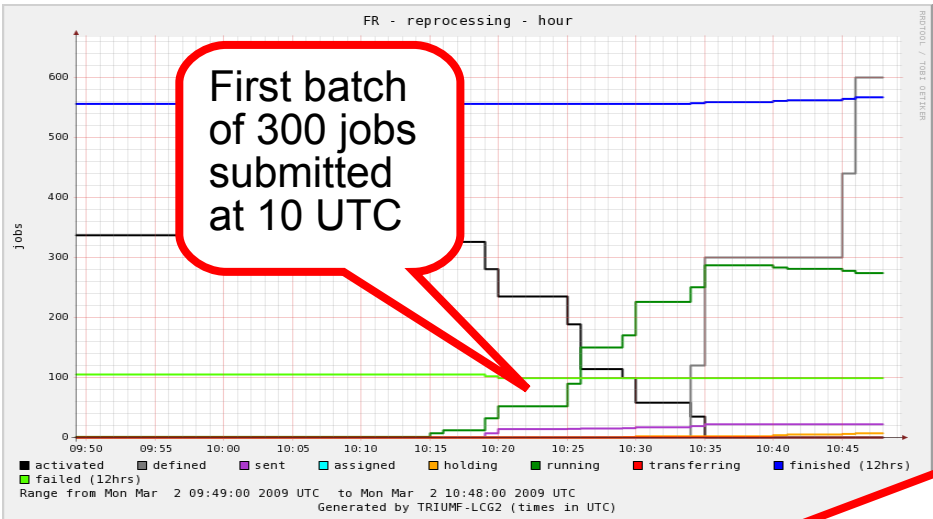
# Pilot Query Approach

- Active staggering of starts possible if the job knows the Db is busy
- T1 RAC can be configured to respond to query from WN
  - OK or notOK for job to start, based on Db load
  - notOK makes job sleep and retry(with limits)
- Pilot query approach requires WLCG 3D Services approval and DBA privileges for installation at the Tier-1 site servers
  - The 3D Services alternative to the pilot query approach is to buy more hardware to support peak database loads
- Effective protection against Db overload
- Version 2 under test to give even more control
  - Track requests and give permission for them to start



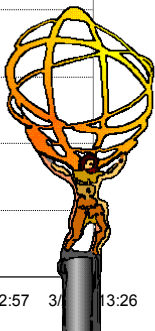
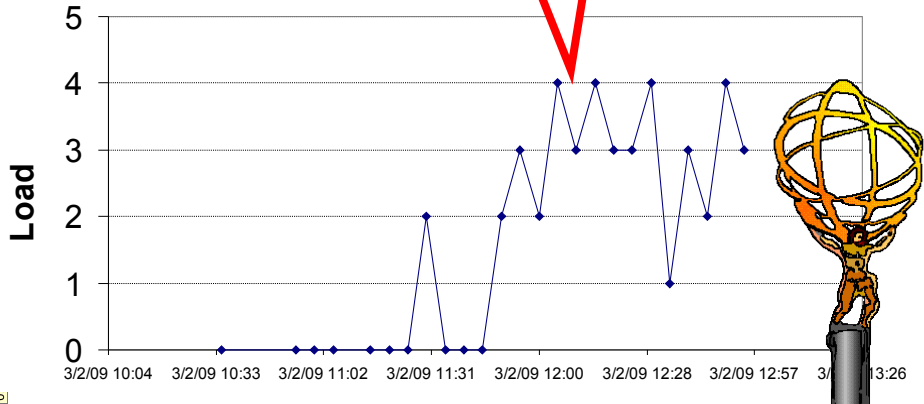
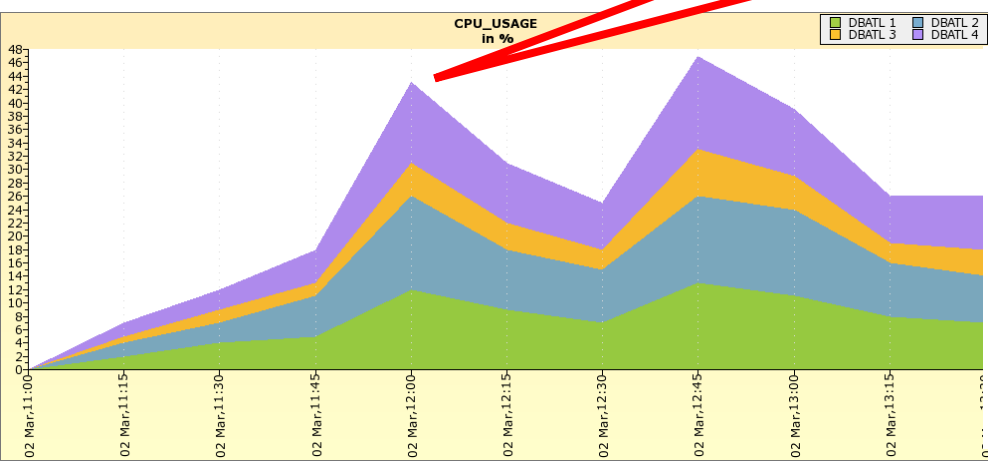
# CHEP'09 - 23 March 2009

## Controlling Oracle RAC Load at CC-IN2P3 Lyon



WLCG 3D monitoring show Oracle RAC load limited by ATLAS PILOT

Because we set ATLAS application-specific Oracle load limit at 4





# Give up

## Database operations pilot at LYON

pilot detected status NOGO - Load: 02.80 Sessions: 420 Threshold:2  
 Mon, 02 Mar 2009 12:21:26 avoiding load of 02.80 at 420 concurrent COOL sessions  
 pilot detected status NOGO - Load: 02.40 Sessions: 471 Threshold:2  
 Mon, 02 Mar 2009 12:28:26 avoiding load of 02.40 at 471 concurrent COOL sessions  
 pilot detected status NOGO - Load: 02.50 Sessions: 486 Threshold:2  
 Mon, 02 Mar 2009 12:44:09 avoiding load of 02.50 at 486 concurrent COOL sessions  
 pilot detected status NOGO - Load: 02.90 Sessions: 776 Threshold:2  
 Mon, 02 Mar 2009 13:08:27 avoiding load of 02.90 at 776 concurrent COOL sessions  
 pilot detected status NOGO - Load: 03.40 Sessions: 701 Threshold:2  
 Mon, 02 Mar 2009 14:00:11 avoiding load of 03.40 at 701 concurrent COOL sessions  
 FATAL: Killing job to avoid Oracle overload

## Panda job information

Jobs: 27791616

[Click for help](#)

Showing 1 jobs modified from 2009-03-02 14:59:36 to 2009-03-02 14:59:36

Jobs:

PandaID, Owner	Job	Status	Created	Time to start	Duration	Ended/Modified	Cloud/Site, Type
<a href="#">27791616</a> <a href="mailto:Rodney.Walker@physik.uni-muenchen.de">Rodney.Walker@physik.uni-muenchen.de</a>	<a href="#">data08_cvalid.00088425.physics_MBTS_BCM_LUCID.recon.o4_r625_tid042225_00177.job</a> #1	failed	03-02 10:30	0:37:54	3:50:43	03-02 14:59	FR/LYON, production
<b>Error details:</b> trans: Athena core dump or timeout, or conddb DB connect exception							
In: <a href="#">data08_cosmag.00088425.physics_MBTS_BCM_LUCID.daq.RAW.o4</a>							
Out: <a href="#">data08_cvalid.00088425.physics_MBTS_BCM_LUCID.recon.HIST.o4_r625_tid042225</a>							

# SQLite Access

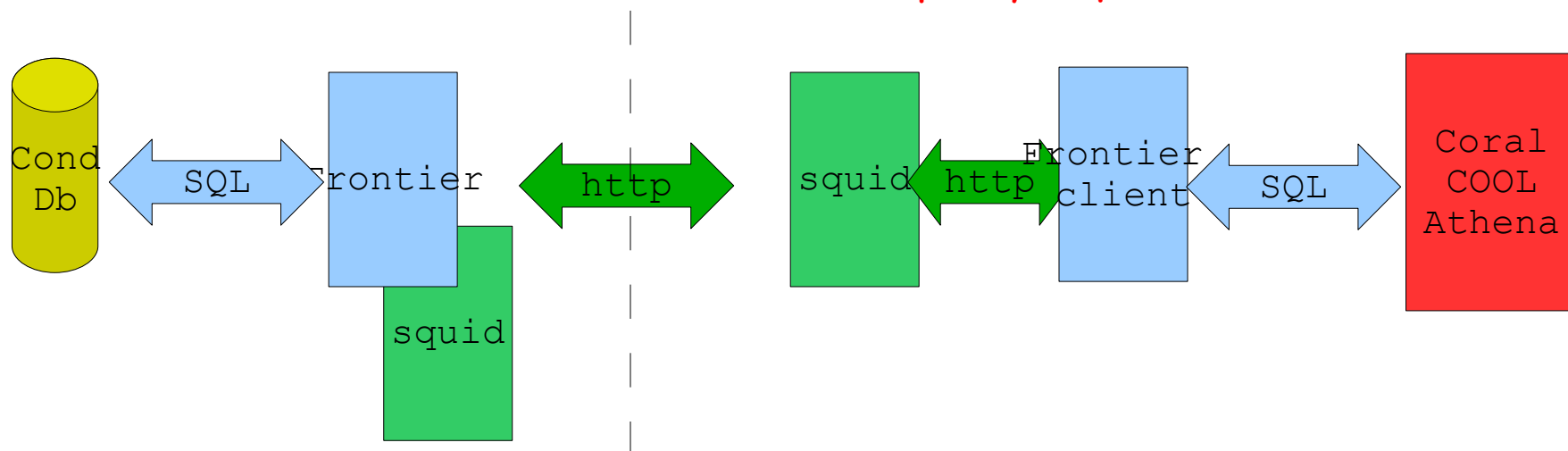
- Conditions Db can be extracted to SQLite files and accessed as flatfile
  - completely avoids Db overload, and scales perfectly (see MC prod)
- Costs
  - 2GB files per day must be created and distributed
  - useful for reprocessing, not skimming large run range
    - cannot create SQLite files for every use-case
- Production jobs during ATLAS 2008 data reprocessing campaign, and March 2009, accessed bulk of Conditions data via SQLite replicas, plus some via Oracle
  - 2008 - 1 big file for whole run period
  - 2009 - 1 files per run (~100)

# SQLite: Hot File

- Every job in 2008 reprocessing downloaded the same 2GB file
- Usually only one disk replica available
  - bandwidth restriction can lead to timeouts and failures
  - can replicate file: result depends on disk/network infrastructure
- 2009 repro has conditions data tar ball per run and not one for all 100 runs together as we had over Christmas
  - so there will be fewer jobs at the same time trying to access these data
- Exposes general problem/inefficiency
  - cache popular files on WN to reduce SRM load
  - test the "pcache" solution where the input data left in WN cache after the job has finished
  - if another job on that node needs the same data it will just use it and not bring a fresh copy in

# Frontier

- US CMS tool to access Oracle via caching web proxies
- Queries encoded to http url.
- Frontier web server, close to Db, decodes and runs query
  - removes WAN RTT from queries
  - compressed result returned to client. Application gets result.
- Identical query = identical url
  - can cache the result in normal web proxy, Squid



# Further Improvements

- For the next release 15, changes in Athena database code to further improve scalability of Oracle access:
  - Single Oracle session at a time, reduced sessions idle time
  - Aggressive look-ahead auto-configured caching strategy
  - Collection of statistics on data volumes, spot unnecessary data
  - IOV alignment (e.g. on 10-min boundaries) for FroNTier caching
    - more identical queries



# Conclusions

- Extensive pre-testing of scalability enabled problems to be addressed before the Xmas reprocessing campaign
  - bulk access via SQLite used
  - frees Db for chaotic usage
- Direct RAC access works with new solutions
  - staggered job starts
  - pilot query approach works as expected
- User/Group analysis, chaotic access not yet solved
  - Frontier under test and is promising
    - protect the T1 RAC using proxy
    - enable T2 access. Remove RTT and cache
- Db access is in good shape for reprocessing real data