ATLAS Frontier/Squid Status Update

Douglas A. Smith
For the Frontier Group
Of the ATLAS Collaboration

Frontier Overview

- A server that provides database connection management and translation of sql statements into HTTP protocal.
- Solution for Oracle access problems for Conditions data to Tier-2/3 sites.
  - Not a replacement for Oracle streams to Tier-1 Oracle RAC servers.
  - Database access to Tier-2/3 from Tier-1 has scaling and latency issues.
- Job access part of ATLAS software.

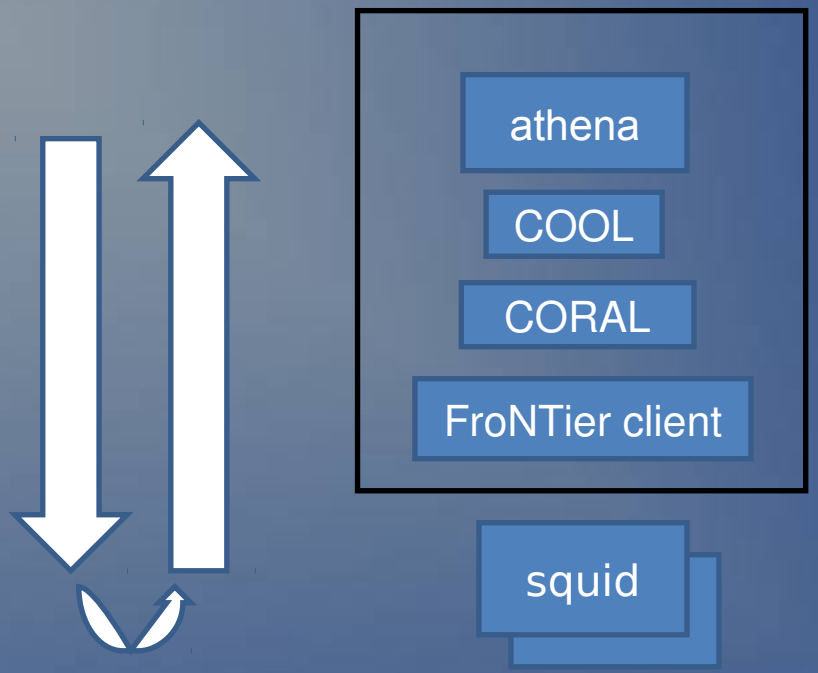
Squid Overview

- A server that provides HTTP data caching, for local cache of data at the Tier-2/3 site.
- Recommendation: one Squid setup per 500 analysis queue slots at a site.
  - Usually at least one squid per Tier-2, some have 2.
  - Smaller Tier-3 site could use nearby Tier-2 squids.
- Squid installs now from ATLAS maintained RPMS.

Frontier/Squid software RPMs

- Installation and configuration of Frontier and Squid software can be performed using rpm distributions provided by F. Donno (CERN/IT-GS)
  - Relocatable rpms available for both SLC4 and SLC5, with post installation script.
  - Squid rpm in use, new release today, Frontier rpm still changing, new release for next week.
  - Common distribution in use for ATLAS and CMS (and other experiments if needed).

Access Layers

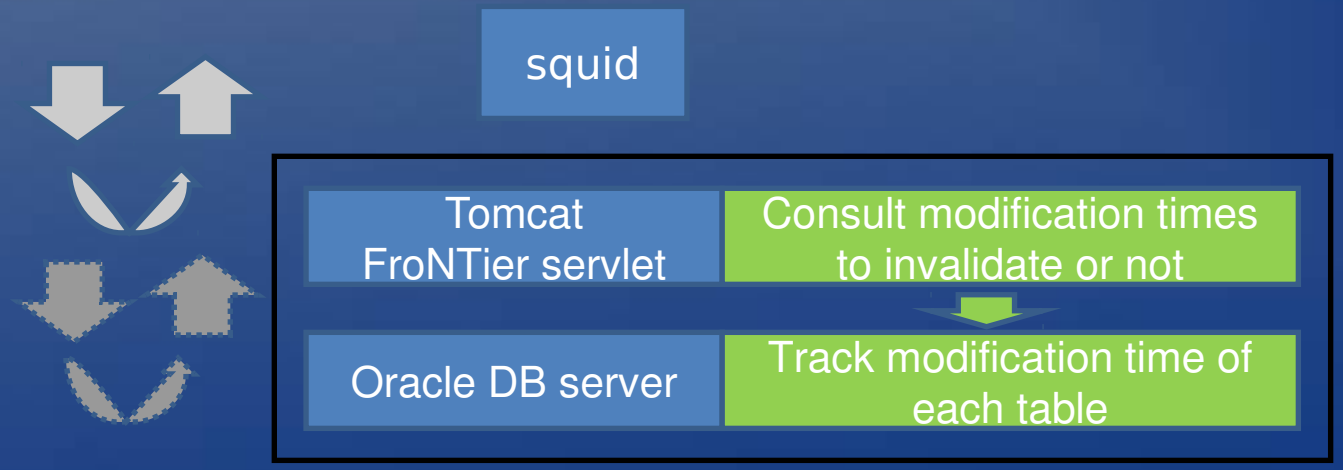


Clients
500 per squid

Remote Site

Squid Cache
10-20 per Frontier Server

WAN



Frontier Server Site

Atlas-wide Deployment

- Early testing of Frontier and Squid access in Canada and US clouds.
  - Individual jobs, scaling tests, latency issues
- Soon after German cloud setup.
  - Site setup issues, job control, central testing
- At ATLAS week in Barcelona, Oct. 5-9, 2009 decision to deploy ATLAS wide
  - Deployment effort lead by Rod Walker and John DeStefano

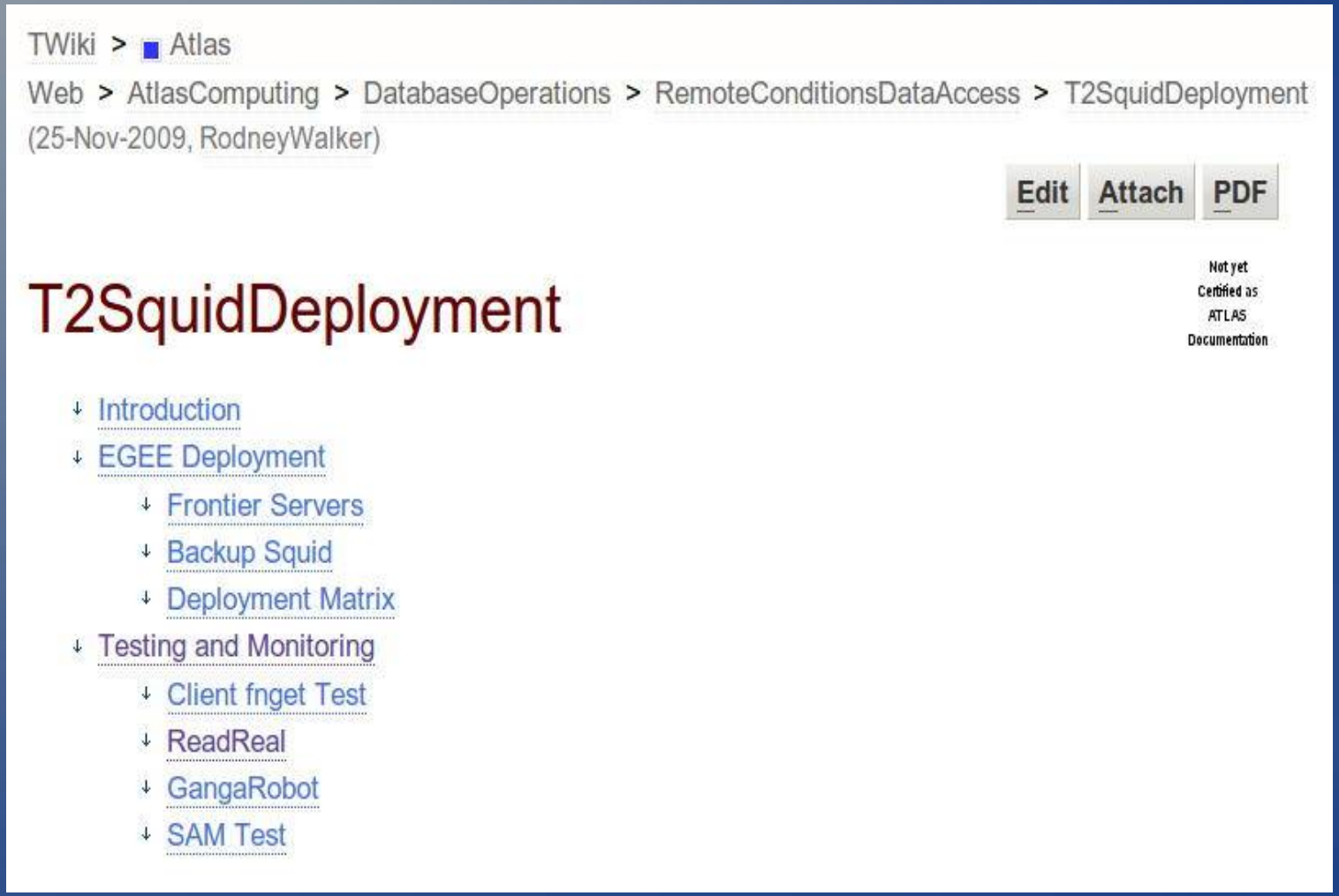
North American Deployment

European, Pacific Deployment

Current Deployment

- 5 Production Frontier servers (BNL, FZK, TRIUMF, PIC, RAL).
- 1 test Frontier server at CERN, also for backup.
- 51 Atlas sites now have a squid installed.
- 19 sites will use a near by squid from another site.
- 70% of Atlas site now have a local squid.
- A number of Tier-3 installs, and plans for Tier-3 use of nearby Tier-2 squids.

Wiki page for tracking


TWiki > Atlas

Web > AtlasComputing > DatabaseOperations > RemoteConditionsDataAccess > T2SquidDeployment

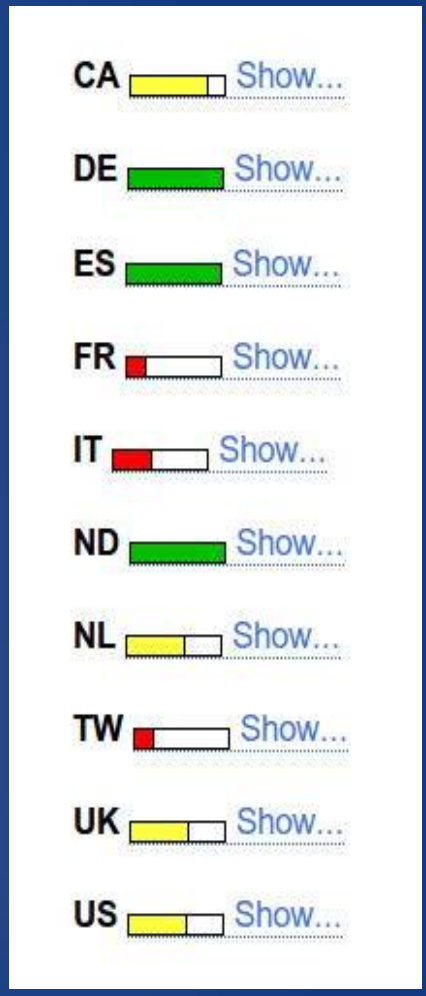
(25-Nov-2009, RodneyWalker)

Edit Attach PDF

Not yet Certified as ATLAS Documentation

T2SquidDeployment

- Introduction
- EGEE Deployment
  - Frontier Servers
  - Backup Squid
  - Deployment Matrix
- Testing and Monitoring
  - Client fnget Test
  - ReadReal
  - GangaRobot
  - SAM Test



https://twiki.cern.ch/twiki/bin/view/Atlas/T2SquidDeployment

Job Control

- Need to tell jobs server address:
  - Happens at run time, jobs at each site need to know local squid for site.
  - Control through env. variable set for ATLAS jobs at site.
  - Env. variable tells jobs where local squid is, and which Frontier server to use.
  - Part of Release since 15.5.1, in use a couple months now.

Additional Data Access

- Some Condition data is outside database.
  - Stored in POOL data files.
  - Files in datasets in DDM, and distributed to sites as other ATLAS data.
  - Files pointed to by Cond. Data in database.
- Mapping to local storage through PFC
  - Local file catalog setup to tell jobs how to access local data from storage.
  - Multiple copies of these files in local storage for scalable access to 100's of jobs.

Status of setup

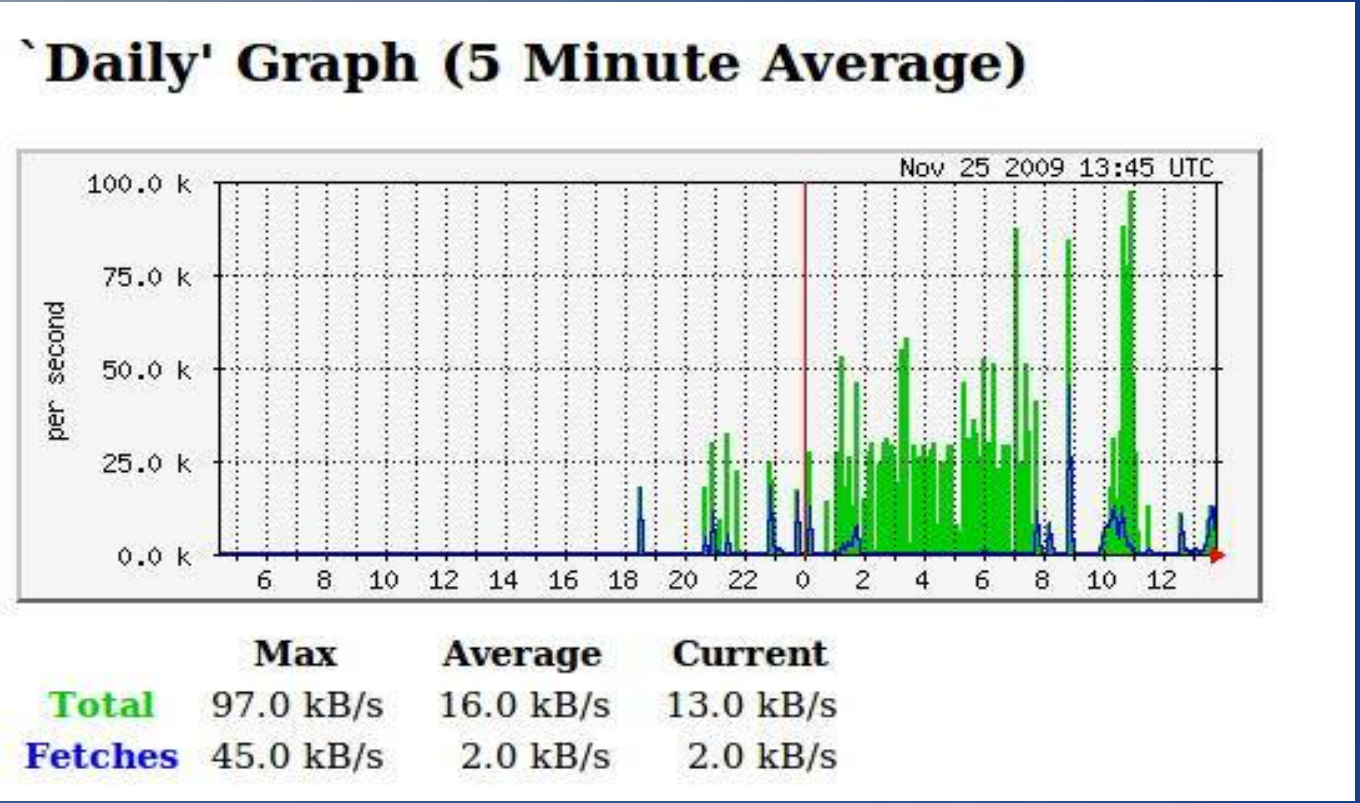
- Setting of env. done for most sites.
  - Still sites getting setup for this, changing quickly.
- Setup of local PFC still going on for sites.
  - Some bugs for certain storage types.
  - Getting worked out as more sites setup.
  - Certain data access bugs found for certain interfaces, current problems fixed in current releases.

SAM tests on the grid

Sitename	Service Type	Service Name	atlas-frontier-squid
ALBERTA-LCG2	CE	lcgce01.cpp.ualberta.ca	warn
Australia-ATLAS	CE	agh2.atlas.unimelb.edu.au	error
BEIJING-LCG2	CE	lcg002.ihep.ac.cn	warn
CSCS-LCG2	CE	ce01.lcg.cscs.ch	ok
CYFRONET-LCG2	CE	ce.cyf-kr.edu.pl	warn
DESY-HH	CE	grid-ce4.desy.de	ok
		grid-ce5.desy.de	ok
DESY-ZN	CE	lcg-ce0.ifh.de	ok
GRIF	CE	grid10.lal.in2p3.fr	ok
		lpnce.in2p3.fr	error
		node07.datagrid.cea.fr	error
GoeGrid	CE	ce-goegrid.gwdg.de	ok
HEPHY-UIBK	CE	grid.uibk.ac.at	ok
IFIC-LCG2	CE	ce01.ific.uv.es	ok
		ce04.ific.uv.es	ok
		lcg2ce.ific.uv.es	ok
IN2P3-LAPP	CE	lapp-ce01.in2p3.fr	ok

- Standard tests for ATLAS sites.
- http://dashb-sam-atlas.cern.ch/dashboard/request.py/latestresultssmry
  - Service type: CE, Test type: CE-ATLAS-sft-Frontier-Squid

Centralized Monitoring



- Provide access stats.
- Setup for BNL, moving out to other sites.
- http://frontier.cern.ch/squidstats/indexatlas.html

Stress Tests at BNL – Setup

- Tests were done at BNL to stress Oracle, Frontier / Squid, and dCache.
  - 5 sets of 650 histogramming jobs were submitted using pathena.
    - First test used Frontier/Squid
    - Second configuration use direct Oracle (done twice!)
  - Each job had ~4 GB of pre-staged event data (2008 cosmic) plus the conditions data in Oracle and Pool files.
  - Each job took 10-15 minutes to read the entire input data set and fill histograms. (Typical tasks a user will do to monitor collision data.)
  - 900 new cores were used, with submission speed tuned.
    - The new cores were unoccupied, it was possible to start large numbers of jobs simultaneously (increasing staging and Oracle or Squid/Frontier load).

BNL Stress Test – Results

- Only one job out of ~10k jobs failed.
  - dCache I/O rates reached 3.5 GB/s when running Oracle tests and 6 GB/s when running Squid/Frontier.
  - Oracle node CPU utilization was a few percent using Squid/Frontier but ~80% using direct Oracle access.
  - During the direct Oracle tests someone in Beijing used Frontier/Squid to read conditions data for a large number of jobs apparently without issues…
    - Also suring testing a number of other sites (Melbourne (via Taiwan), SLAC, Japan, CERN) were also using the BNL Frontier. People are using Frontier!

Open Issues

- Service access policies
  - Who can access Frontier service?
  - Who can access local Squid service?
  - Limits to what sort of data on Squid server?
- Service level agreements
  - What are the acceptable uptimes for a Frontier server? for a Squid server?
- Failover policies
  - When squid is out, failover to another Squid? Same with Frontier server? Allow failover to direct Oracle access?

Acknowledgements

- ATLAS setup of Frontier use is a fairly large diverse team
  - Lots of help from Dave Dykstra of CMS.
  - Many people from many sites working on coordination and details of development.
  - Many, many people at Tier-1/2/3 sites for server setup and maintenance
- Many thanks to all!!!

Summary

- ATLAS wide use of Frontier decided early Oct.
- Currently 5 Frontier servers installed, 51 local squid cache servers installed, 70% of sites.
- Local setup for job control and file access getting in place.
- Bugs in ATLAS job control have been found, and are getting fixed quickly, moving into ATLAS release for full support.
- Standard testing and monitoring frameworks still getting setup, and shift reporting.