Frontier use in ATLAS

D. A. Smith*, D. Barberis, J. DeStefano, A. Dewhurst, F. Donno, D. Dykstra, D. Front, E. Gallas, R. Hawkings, F. Luehring, R Walker
On behalf of the ATLAS collaboration

Analysis of ATLAS data can require on the order of 10.000 simultaneous jobs running on the GRID. Each of these jobs need to read detector conditions from a database, and no database server system can handle this number of concurrent connections. Analysis tests at large centers showed problems on Oracle servers when over 150 simultaneous jobs were started.

Frontier was developed for the CMS experiment. The system manages database connections, and translates information requests into the HTTP protocol. These HTTP interactions are cached by Squid web cache instances, hosted on servers near the jobs. ATLAS can scale up to the number of required database interactions by using multiple Frontier server sites, each providing data to Squid caches at multiple sites.

Latency of interactions can be a large issue for remote sites. Jobs required to span long distances to get data directly from Oracle servers can take 10 to 30 min to start, and for some distant sites much longer. By placing an easily-maintained Squid server close to job sites, latency has been lowered dramatically: the same jobs now start within 20 to 50 seconds.

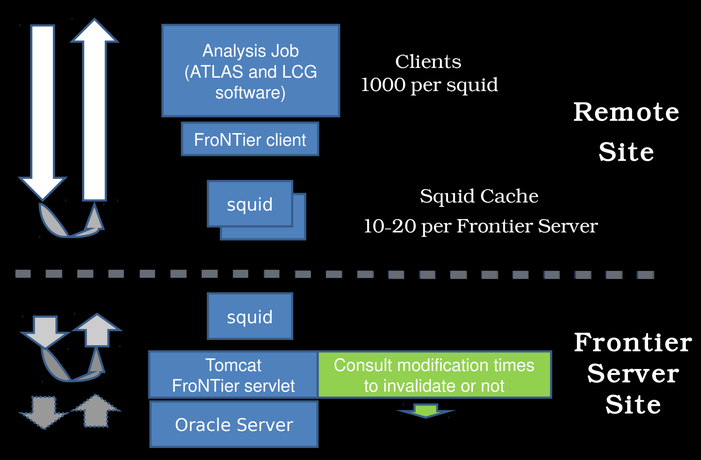
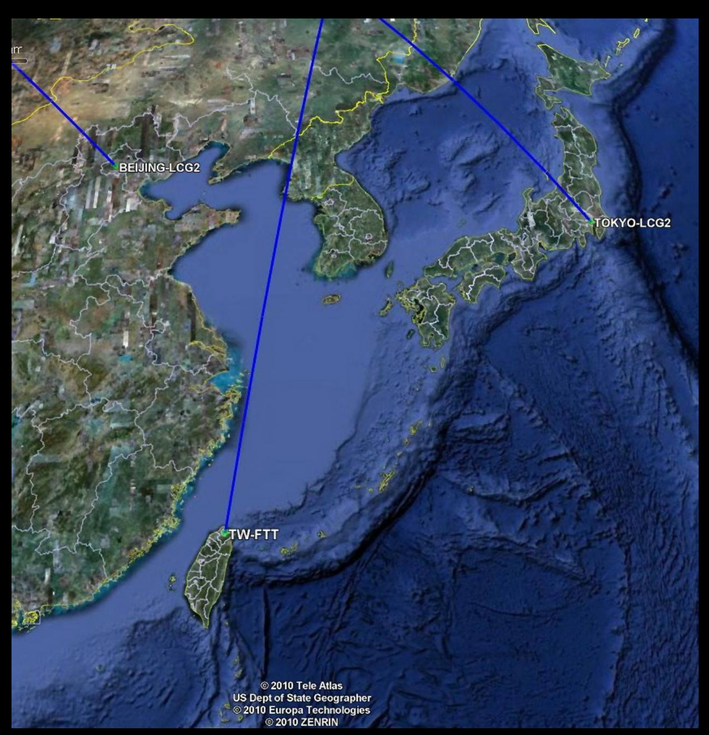


Figure 1: Diagram of system, with Oracle server, Frontier server, and remote Squid servers. Jobs mostly interact with local squid and cached data, retrieving data as needed from remote central site.

Number of test jobs	50	100	150
Throughput [hz] (# Jobs read/7200)	2.5	3	4
Number Jobs Read [k]	17	24	29
Load on Frontier Server [percent]	6.5	7	10
NW Utilization of Frontier Server [percent]	54	83	100

Figure 2: Scaling test average results showing CPU and bandwidth use of special test jobs. On average, analysis jobs will use 1/7 of their run time in database access, estimated number of analysis will 100 times greater than the listed number of test jobs.

Scaling tests have been performed on resources at CERN. These tests are comprised of special jobs that read conditions database data without performing analysis. The volume of conditions data used by these test jobs was purposely excessive: normal analysis jobs require much less data. These tests proved that the servers could scale to the service levels necessary for analysis with very little load on CPU and disk, with network bandwidth being the only limiting factor with regard to scaling.



The Frontier system was installed at ATLAS sites at the end of 2009 and beginning of 2010. Since it was shown to work for analysis jobs, and the installation of Squid servers was a simple matter for remote sites, use in production happened fairly quickly, and with few problems. The system has been providing stable conditions data access for analysis jobs for many months now, and users have been quite pleased with the performance and use.

Figure 3: World map showing ATLAS sites in Europe and Asia. Conditions data are produced at CERN and are transferred to remote sites hosting Frontier servers via Oracle streams. These data transfers are shown in red. These Frontier servers provide conditions data for analysis to Squid instances at their own site, and others. These data transfers are shown in blue.

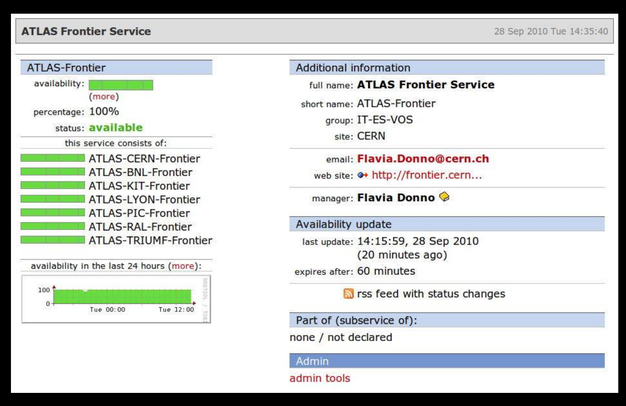


Figure 4: Standard monitoring of the Frontier Servers installed. This monitoring shows the availability, and sets alarms on when services are down.

Monitoring of sites is performed centrally from CERN. The Frontier servers are monitored with AWStats, and information is gathered into the Service Level Status (SLS) system to provide availability information. The Squid servers at analysis sites are monitored with MRTG, which collects information in 5 min intervals. Fallback servers are defined for conditions data access for jobs, so analysis will still work even in periods of downtime for local services.

Server administration has proven to be quite simple for sites. Frontier and Squid server code is packaged into rpms for distribution, and to provide simple installations and upgrades. After an initial period of setup and configuration, servers have been quite stable, and have not shown significant problems with downtimes and outages.

Analysis use of Frontier has been stable for many months, and is providing trouble-free database access for jobs. At the larger sites average levels of 10k requests per minute are supplied by local Squid servers, with peak levels at up to 100k per min. Job failures due to access problems have not been seen, and server load is usually quite low, at less than 10% cpu load. For scaling and reliability purposes, a standard of one Squid instance per 500 to 1000 analysis job slots has been recommended.

Job configuration is controlled by environment variables, set at the site at run-time, defining the Squid and Frontier servers used, along with fallback servers. The variable values are controlled centrally and are not set by the sites themselves, although users could override. In this way, changes to server URLs can be made centrally and propagated quickly to other sites for use by all jobs. As a result jobs are submitted by users without a need to know which site or servers will be used for data access.

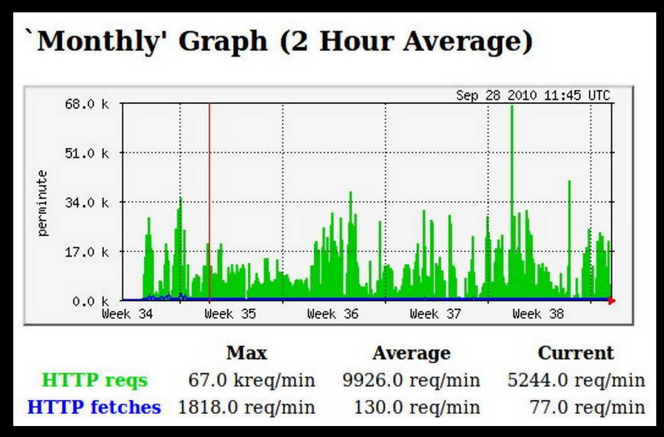


Figure 6: Squid instance use at BNL, showing steady analysis use for weeks. In green are the data requests served by the Squid cache, in blue are the data requests served by the Frontier server, showing the dramatic scaling provided by the use of Squid caches.

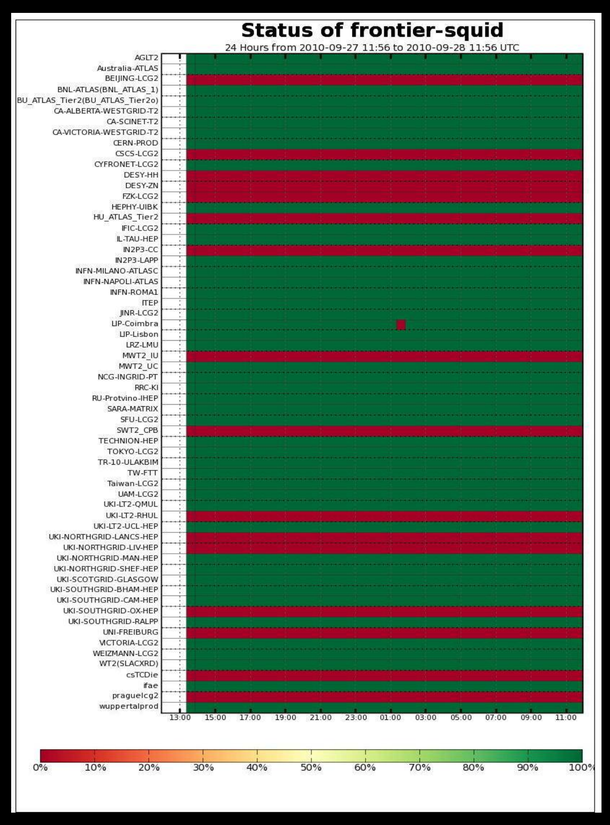
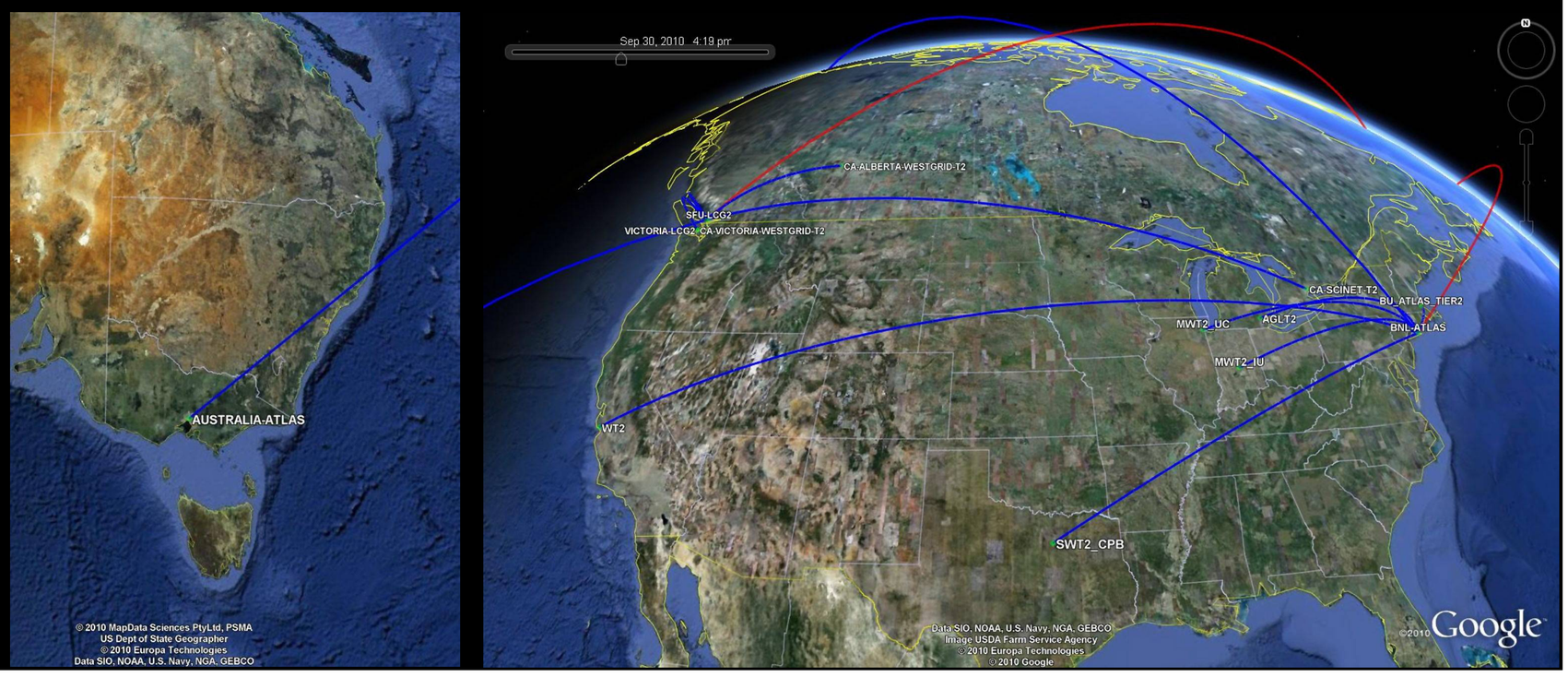


Figure 5: Monitoring of installed squid servers, one for each analysis site used in ATLAS. Some are red because of missing installs, or problems with monitor setup, some sites are still coming on-line.

Figure 7: Frontier server, and Squid server installations in North America and Australia. For this figure, and Figure 3, the transfer lines shown are the default sources, in cases of downtime or failure there are fallback sources defined, which are not shown here.



ATLAS has been quite happy with the Frontier installation and use, and has been an important component of the success of ATLAS analysis use of GRID resources. We expect the system be scale to future use and provide stable database access for jobs. For further technical details on the Frontier system please see: http://frontier.cern.ch

(*) Contact information: Douglas A. Smith -- Douglas.Smith@cern.ch
For more information: https://www.racf.bnl.gov/docs/services/frontier