Evolution of grid-wide access to database resident information in ATLAS using Frontier

Alastair Dewhurst,
on behalf of the ATLAS collaboration

Outline

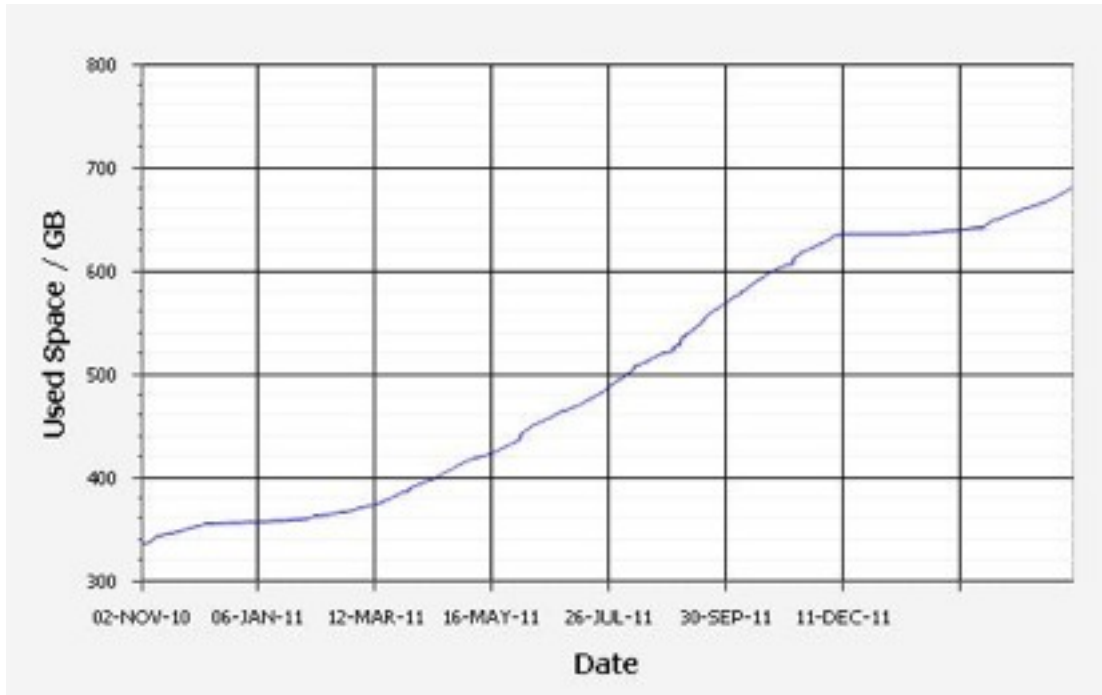
- What is Condition data and how is it accessed through Frontier
- Squid performance and site evolution
- Development of AGIS and monitoring

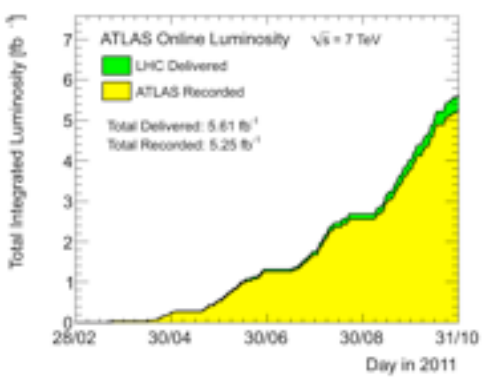
What is conditions data

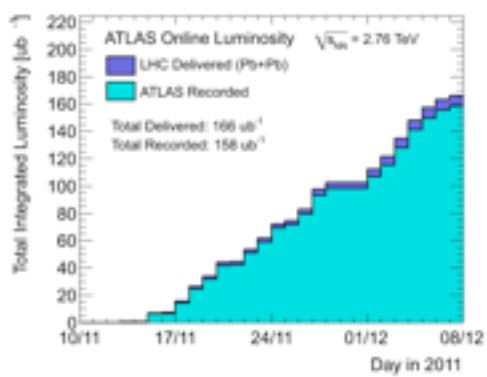
- Conditions data is defined as event independent time varying data
- In ATLAS there are two elements to conditions database access:
  - Access to the relational database
  - Access to POOL files storing large calibration objects
- The relational database can either be:
  - Oracle database
  - SQLite file
- Frontier allows access to the Oracle database

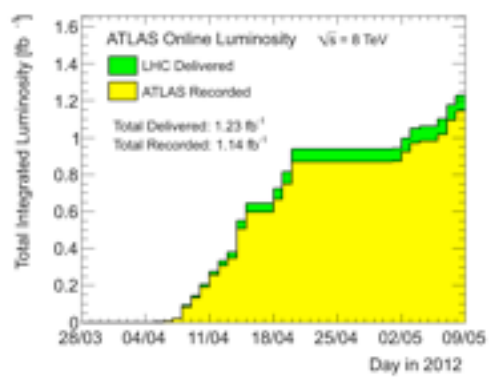
Conditions data size

Plot showing how the Oracle conditions database has grown in 2011



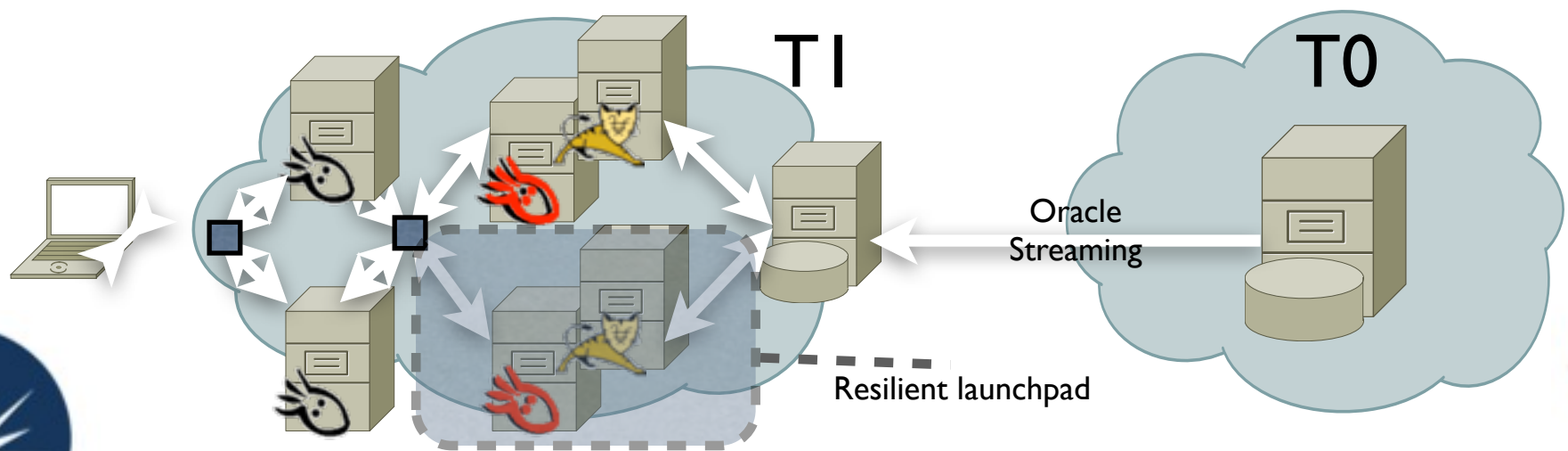






What is Frontier

- The Frontier system distributes data from central databases that are read by many client systems around the world
- The Frontier setup relies on two levels of caching.
  - Squid cache at every site
  - Squid cache in front of every Frontier accelerator



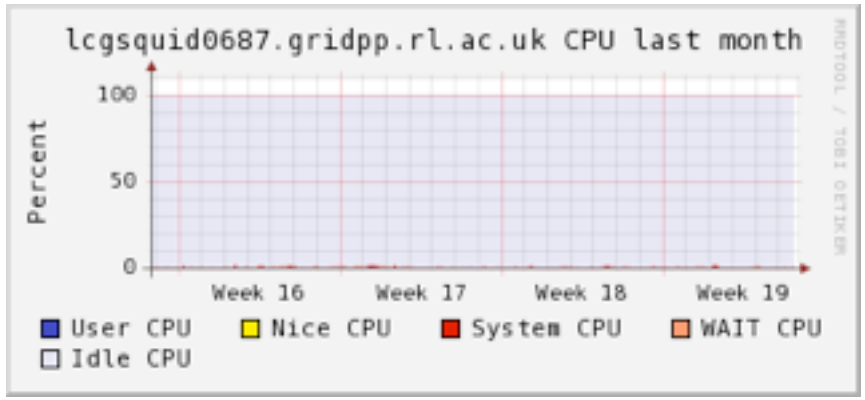
Squid Load

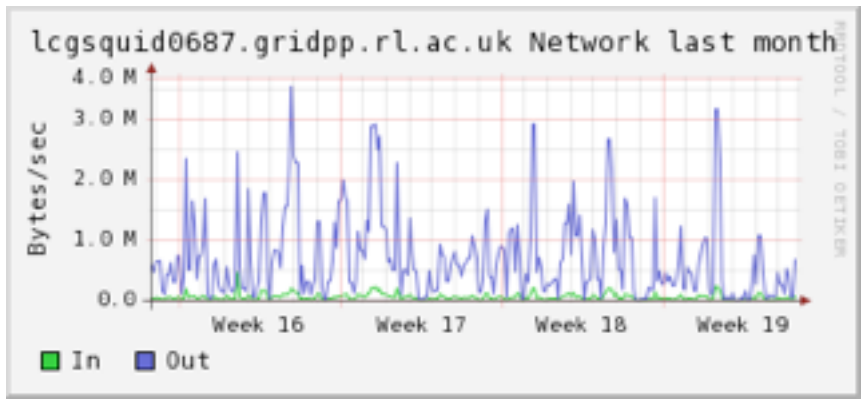
Plots show load on one of the squid servers at RAL in the last month

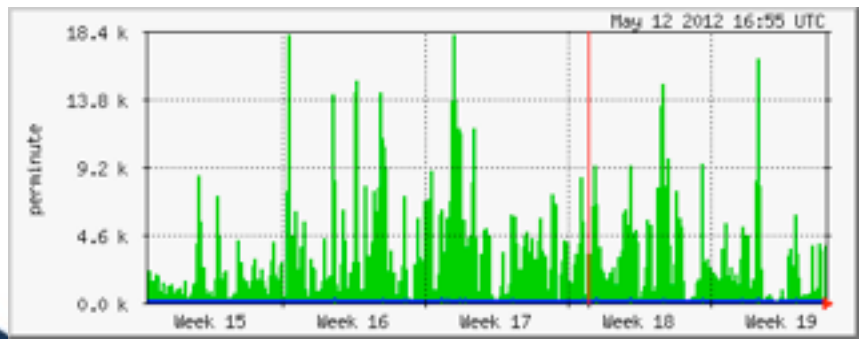
RAL has 2 squid, ~6000 job slots, ~45% allocated to ATLAS

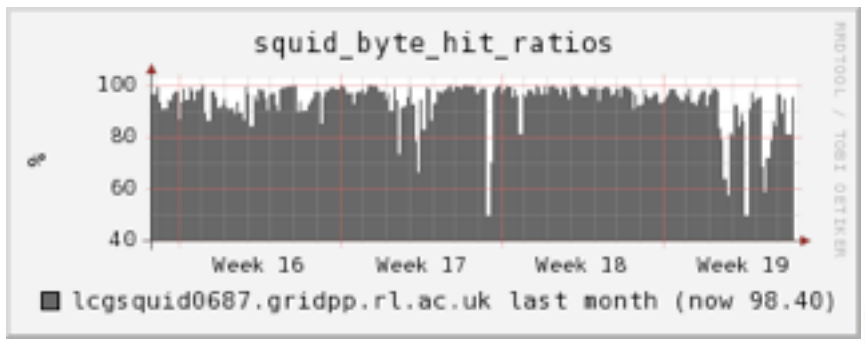
Tiny load on CPU and bandwidth

High hit ratio



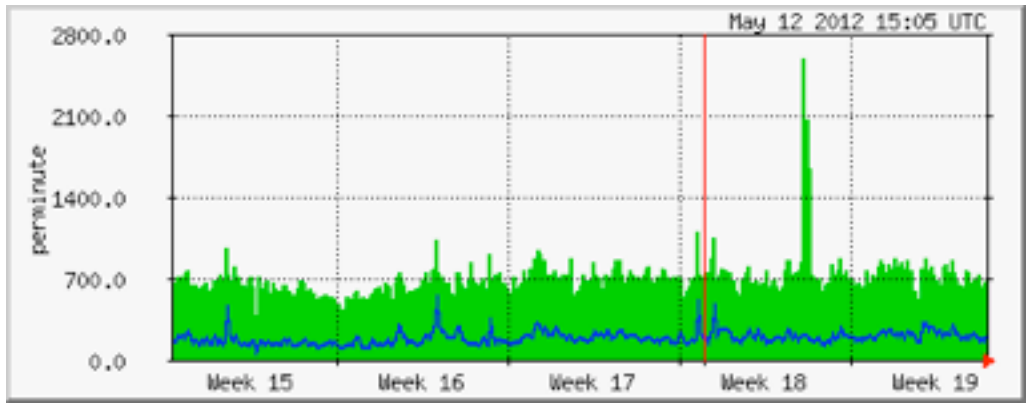






Frontier load

- The amount of caching at the Frontier level is significantly less than at the site squid level.
- A site squid will have possibly thousands of different jobs accessing the same data where as a Frontier should only have to deal with tens of site squids
- Second level of caching does still help and protects the Oracle database should things failover.

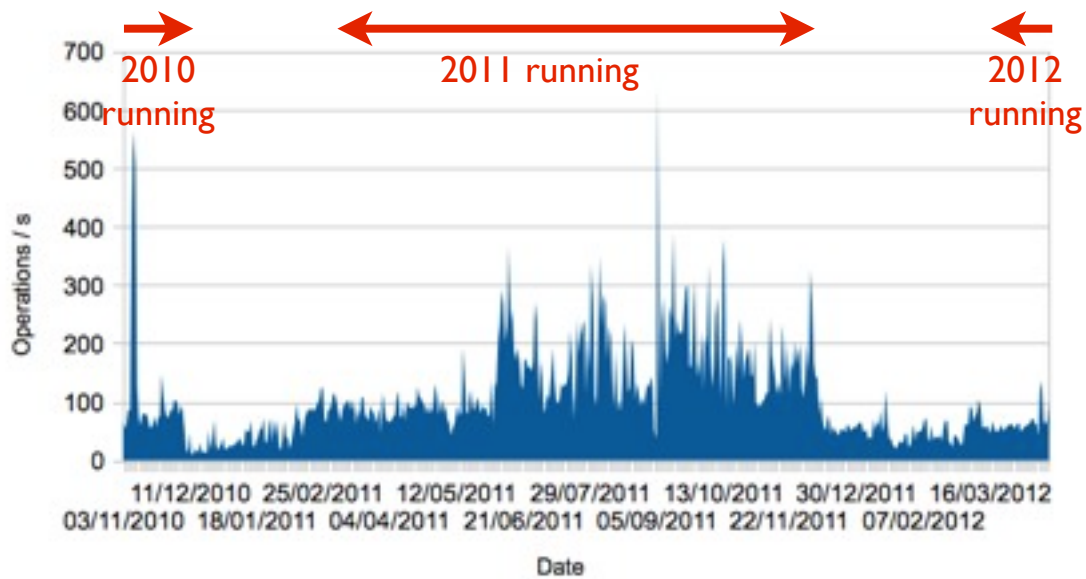


Oracle Evolution

- It has been demonstrated that as long as your site has a squid cache you do not need to be near a Frontier
- Since 2010 Oracle database at Taiwan and PIC have been decommissioned
- In mid 2011 IN2P3 switched from direct Oracle access at their Tier 1 to using Frontier.

Average number of operations /s per day for the RAL database

Operations is both reads and writes, so that include the streaming from CERN.



Site evolution

- It has been demonstrated that it is not essential to be near a Frontier launchpad
- RPMs have been developed. Default settings more sensible for Frontier use case
- There has been a large increase in the number of site squids in the last year.
- Squids being used frequently for other services such as CVMFS

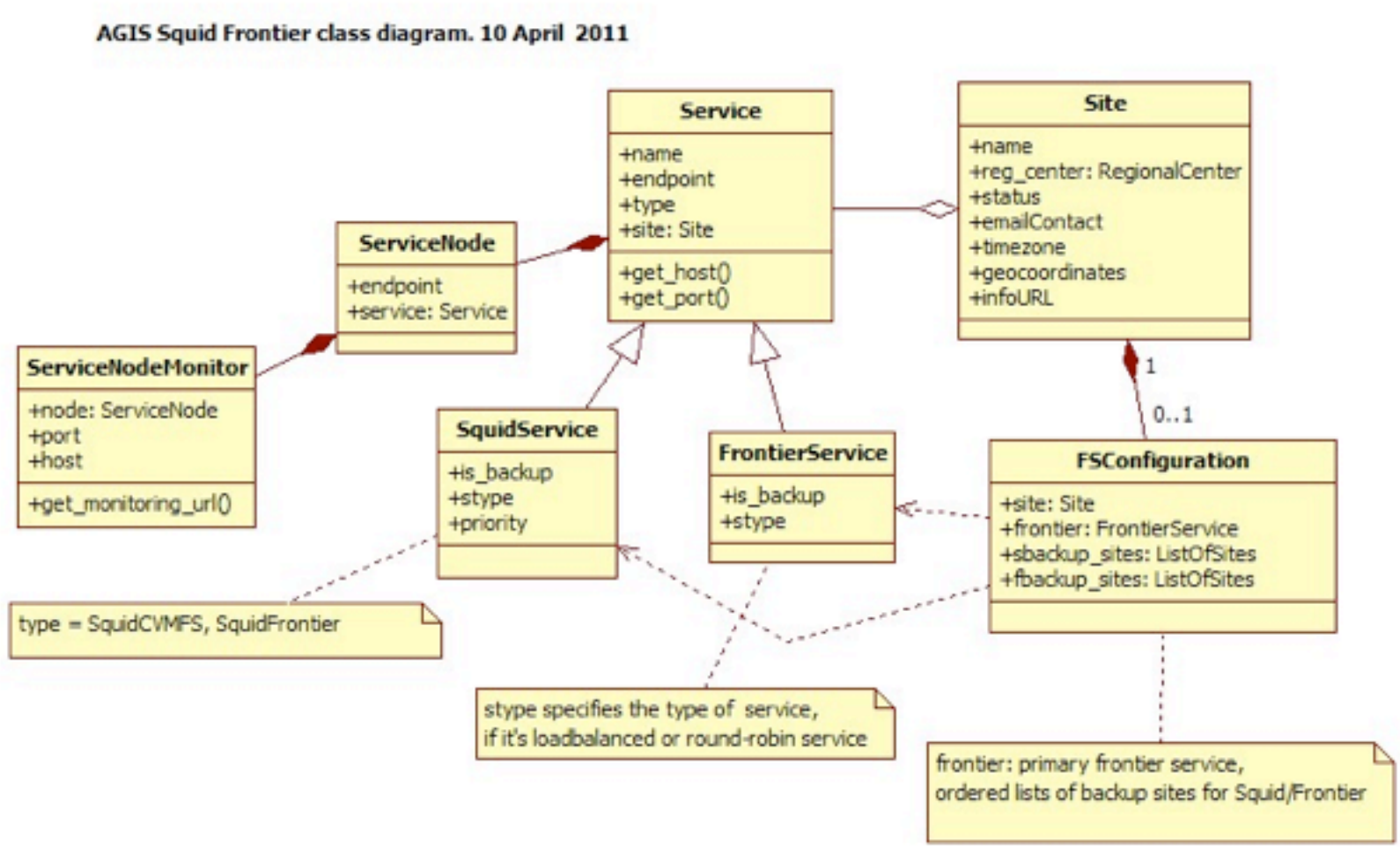
Failover policy

- If a squid breaks and a request fails the client can try other combinations of squids and Frontier launchpads to try and get the data.
- ATLAS decided to ask sites to failover to nearby "friendly" sites.
- The problem with automatic failover is that because jobs won't fail the problem is not obvious until it's a much larger one.
- An easy to use monitoring system needed to be developed.

AGIS use case

- At the start of 2011 Frontier configuration for ATLAS was controlled via a single text file "Tiersofatlas.py"
- Configuration changes were slow to make, a single typo could break everything.
- It was not possible in Tiersofatlas to configure monitoring.
- The ATLAS Grid Information Service was being developed at the time. Ideal for Frontier use.

Frontier setup in AGIS

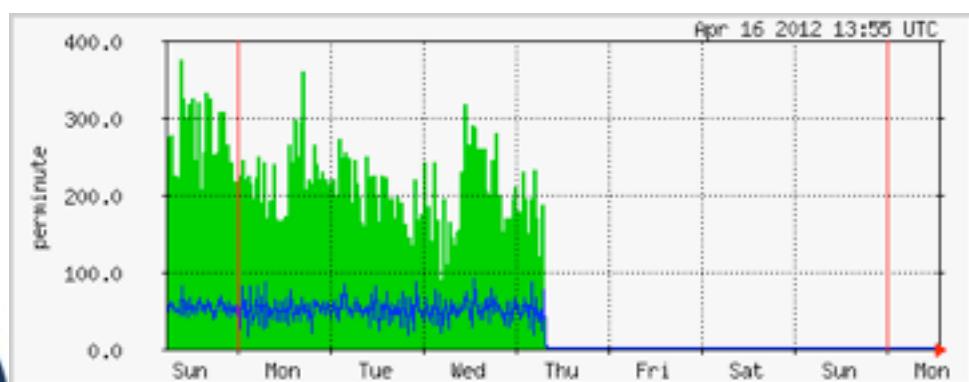


Monitoring squids

Show 200 entries | Copy | Print | Save | View: Frontier_Squid | Search...

Site Name	Frontier	frontier-squid
CSCS-LCG2	n/a	OK
csTCDie	n/a	OK
CYFRONET-LCG2	n/a	OK
DESY-HH	n/a	OK
DESY-ZN	n/a	OK
FZK-LCG2	100	OK
GoeGrid	n/a	NotAvailable
HEPHY-UIBK	n/a	OK
ifae	n/a	OK
IFIC-LCG2	n/a	OK
IL-TAU-HEP	n/a	down
IllinoisHEP	n/a	OK
IN2P3-CC	100	OK
IN2P3-LAPP	n/a	OK
INFN-COSENZA	n/a	down
INFN-FRASCATI	n/a	OK

Simple interface for shifters, linking to more detailed information



Sitename	Flavour	Hosts	1
AGLT2	OSG-CE	gate01.aglt2.org	OK
		gate02.grid.umich.edu	C
Sitename	**Flavour**	**Hosts**	**1**
AM-04-YERPHI	CREAM-CE	ce.yerphi-cluster.grid.am	C
Sitename	**Flavour**	**Hosts**	**1**
Australia-ATLAS	CREAM-CE	agh5.atlas.unimelb.edu.au	OK
Sitename	**Flavour**	**Hosts**	**1**
BEIJING-LCG2	CE	lcg002.ihep.ac.cn	W
			1
	CREAM-CE	cce.ihep.ac.cn	W
Sitename	**Flavour**	**Hosts**	**1**
BNL-ATLAS	OSG-CE	gridgk01.racf.bnl.gov	OK
		gridgk02.racf.bnl.gov	OK
		gridgk03.racf.bnl.gov	OK
		gridgk04.racf.bnl.gov	OK
		gridgk05.racf.bnl.gov	OK
		gridgk06.racf.bnl.gov	OK

Failover testing

- Recommendation for squids is to restrict incoming requests to a local subnet.
- Also option to allow destination sites to be restricted to 6 Frontier launchpads.
- Need to regularly check each site allows failover correctly.
- Simple test job run on each site testing all combintations of primary and backup squids and reporting warning if some fail and error if all fail.

AWstats

http://frontier.cern.ch/awstatsatlas.html

Visitors domains/countries (Top 10) - Full list

Domains/Countries		Pages	Hits	Bandwidth
USA Educational	edu	8206618	8206618	53.23 GB
USA Government	gov	5776678	5776678	75.78 GB
Non-Profit Organizations	org	4492962	4492962	28.95 GB
Taiwan	tw	1089097	1089097	8.21 GB
Canada	ca	15410	15410	264.15 MB
Unknown	ip	8972	8972	103.91 MB
Switzerland	ch	8640	8640	6.13 MB
Australia	au	4606	4606	105.64 MB
South Africa	za	20	20	336.03 KB
Others		0	0	0

Awstats monitoring of Frontier launchpads allows debugging of most problems.

	Frequency[/s]	Hits	Percent
0-44	6.33216	16412814	83.6 %
44-100			
100-500	0.00184	4766	0 %
500-1K	0.07751	200906	1 %
1K-2K	0.35872	929797	4.7 %
2K-5K	0.03153	81716	0.4 %
5K+	0.76546	1984060	10.1 %

Conclusion

- ATLAS have used Frontier to successfully access their conditions data throughout 2011
- System is scaleable and capable of handling the extra load in 2012 and beyond
- Developing detailed monitoring allows ATLAS to choose the best failover strategy for each site
- Tier 0 has started using Frontier and work is starting to move all conditions access to Frontier