

ATLAS Feedback

Alastair Dewhurst on behalf of the ATLAS
Collaboration

Outline

- ATLAS use cases
- CVMFS
 - Operational issues. (John Hover, Asoka de Silva, Alessandro de Salvo)
 - AFS migration. (Graeme Stewart)
- CernVM
 - How CernVM is used. (Randall Sobie)
- Covered in other talks
 - Rod Walker spoke about HPC [1].
 - Jose Caballero will speak about Stratum 1 replication[2].

[1] <https://indico.cern.ch/event/469775/contributions/2148340/attachments/1285298/1911421/cvmfs4MUC.pdf>

[2] https://indico.cern.ch/event/469775/contributions/2168632/attachments/1282500/1907664/cvmfsreplica_CVMFS_workshop_RAL.pdf



ATLAS use cases

- CVMFS is now used by all aspects of ATLAS offline computing.
 - Grid and batch jobs as well as interactive user analysis.
- Multiple repositories:
 - `/cvmfs/atlas.cern.ch` – ATLAS offline software releases, site configuration information, user interface software.
 - `/cvmfs/atlas-cond.cern.ch` – Conditions data
 - `/cvmfs/atlas-nightlies.cern.ch` – Nightly software releases



Inode Problem

- If the CVMFS inode counter goes above the 32 bit limit, 32 bit applications not using 64 bit file system calls will break.
 - Happens at some point in most ATLAS Grid Jobs.
- Requires CVMFS continuously mounted for ~80 days
 - In WN with lots of job slots this will happen
- Often requires experts to debug.
 - Once identified is trivial to fix.
- Would love to see it fixed in software.
 - Several sites have developed their own automated checks followed by fixes. Could this be distributed centrally?



Maximum Object size

- Some objects in /cvmfs/atlas-cond.cern.ch are larger than 1GB.
- Squids have configuration for maximum object size to Cache.
 - By default (?) this is set to 1GB so lots of sites use it.
 - File is still stored on cache on WN but this is shared by other repositories.
- One of the reasons CVMFS works so well because of the effective caching
 - Who worries about this?
 - Are we sleep walking into a significant problem?



Other issues

- Taiwan's Stratum I has poor availability.
 - It is an annoyance for ATLAS shifters.
- Some Stratum Is are not on GeolP.
- Some ATLAS sites mount CVMFS over NFS. We have found a few rouge repositories.
 - <https://its.cern.ch/jira/browse/AGENE-1134>
- Method for adding content is “brittle and risky”
 - Would like to see a safer programmatic interface to add content to CVMFS.
 - OSG/OASIS have a more protected arrangement.
 - BNL RACF is about to setup local repo server (contact John Hover).



ATLAS Nightlies

- ATLAS would like to gently move towards having all nightly builds published onto CVMFS.
- Long term aim is to deprecate the use of AFS, which has been our traditional way of making nightly builds available to developers
- CVMFS is not a drop in replacement for AFS.
- ATLAS is trying to reduce the number of full software builds compared to today.

Build For...	Notes	Opt	Debug	N_files
Tier-0	Very stable reconstruction	5GB	20GB	100k
Trigger	Critical release for online farm	5GB	20GB	100k
Development	Next release — evolves rapidly, less stable	5GB	20GB	100k
Simulation (light)	Does simulation only, maximally portable	0.5GB	2.5GB	15k
Analysis	Actually multi-flavoured: stable, development	0.5GB	2.5GB	15k



Build Architectures

- We are becoming more heterogeneous again in CPU architectures
 - Currently targeting ports to ARM64 and PowerPC.
 - Evidently x86_64 remains our workhorse.
- Most nightly builds will only be kept for 1 week, then deleted
- Some we may want to keep for more extensive testing
 - Release candidates for physics validation.

Build For...	x86_64	AArch64	PowerPC
Tier-0	✓		✓
Trigger	✓		
Development	✓	✓	✓
Simulation (light)	✓	✓	✓
Analysis	✓	✓	

Take this with a large pinch of salt...

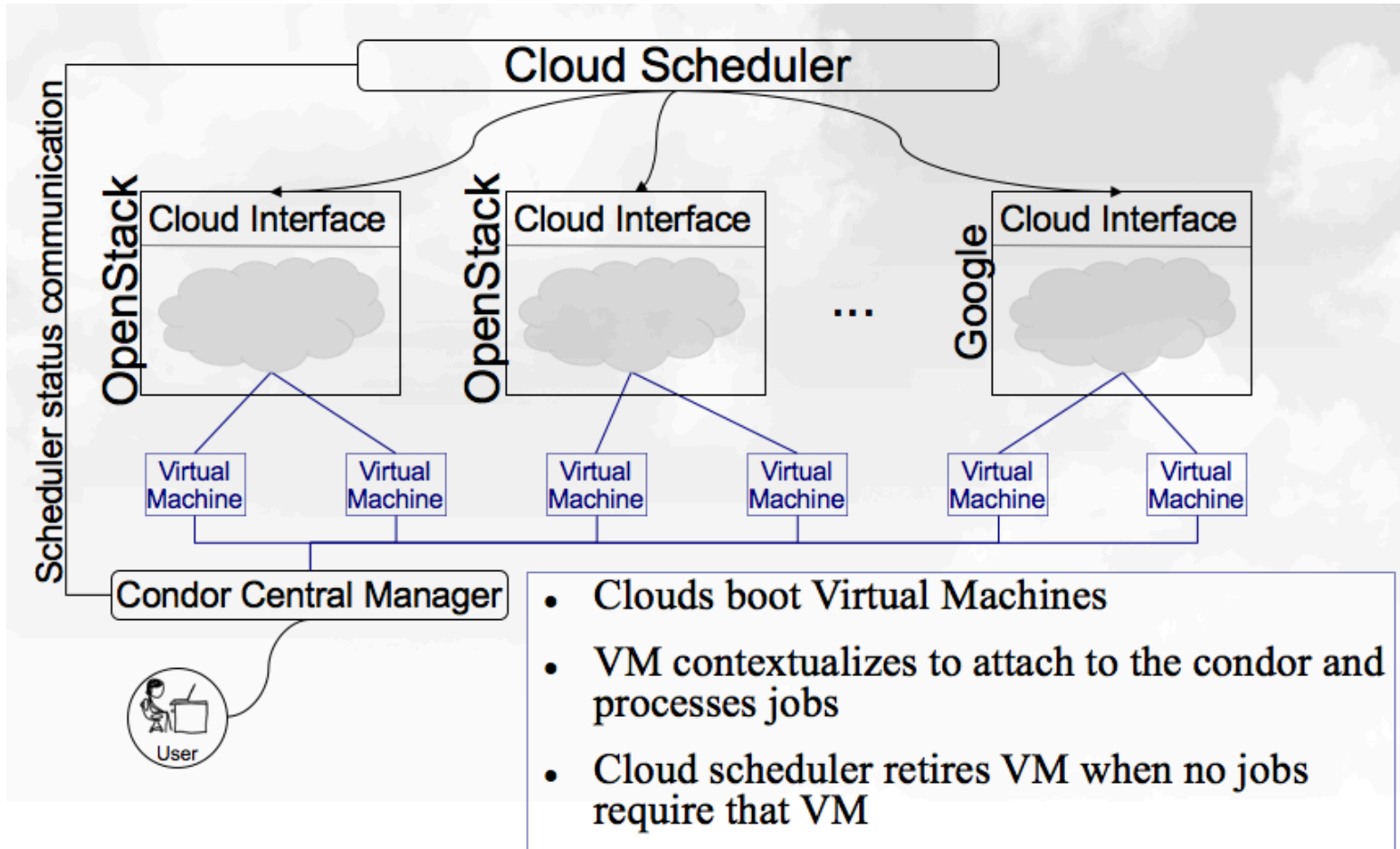


Estimates

- 10 builds a day (release and architecture)
 - x 2 for flavours (Opt and Debug)
 - 20 publications
- 175GB and 7 million files per day to publish
 - 2MB/s continuously
- Same amount to delete per day
 - No problem if deletes and garbage collection happens less frequently.
- No build should be delayed by more than 30 minutes to get onto CVMFS
 - People can't work until build is published.
- Total volume 6TB – 10 rolling releases + 20 flagged for keeping.
- These figures are hugely speculative, but it would be surprising if they are more than x2-3 out...



Cloud Computing



CernVM use

11

- Micro-CernVM used for ATLAS (and Belle-II)
 - Typical use 2000-4000 cores using 8-core VMs on private and commercial clouds in Europe and North America
 - All clouds using micro-cernvm version 2.4-6 include EC2 and Azure
- Site Customization
 - cloud_init using multi-mime YAML
 - Using Shoal (Squid discovery service) to find nearest HTTP Squid for CVMFS based on GeoIP
 - Condor master and startd are started
 - Add ssh-keys, condor-users (for condor slots)
 - Hot fixes (eg. DNS additions)
- CERN CVMFS puppet module used on Amazon
- Issues
 - Need better transition process from development to production
 - Contextuational issues (need further discussion)



Conclusions

- ATLAS rely on CVMFS for software distribution and it works very well.
- Plan to migrate away AFS to atlas-nightlies.
- CernVM is used extensively in cloud computing.

