



Status of ATLAS Cloud Activities

John Hover <jhover @bnl.gov>

Outline



Organization

Rationale

Proximate Goals

Projects

- Workload Management
- Cloud Analysis Clusters
- Private T1/T2 Clouds
- Cloud Storage
- Virtualization

Next Steps

Organization



ATLAS Cloud R&D

- Bi-weekly phone conf led by Dan Van der Ster, Fernando Barreiro
- Proposal Document: Goals and rationales
<https://indico.cern.ch/getFile.py/access?resId=1&materialId=0&confId=13675>
- Wiki
<https://twiki.cern.ch/twiki/bin/viewauth/Atlas/CloudcomputingRnD>
- Collaborative. Groups pursue projects, share progress and ideas, try to avoid duplication.

US ATLAS effort coordinated with OSG Technology Investigations

- Bi-weekly phone conf. BNL and T2 representation.

Overall Rationale



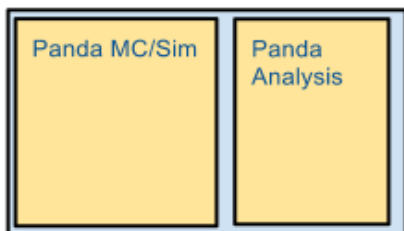
Uniform access to virtualized computing allows dynamic repartitioning of total resources.

- Improved utilization of all resources.
- Tier 1s can provide dynamic Tier 2/3 clusters
- Tier 2s can provide temporary single-user analysis clusters and specialized analysis clusters.
- Work can be easily extended to public scientific and/or commercial clouds as justified.
- Greater possibility of opportunistic usage for small-data MC/Simulation.
- Total expansion of resources during crunch periods.

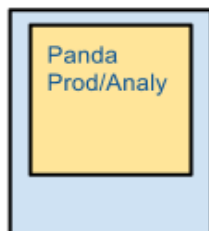
Dynamic Resource Partitioning



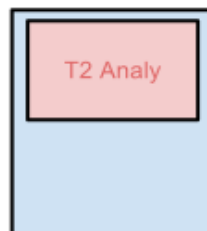
From



Tier 1

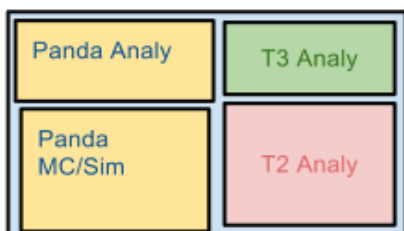


Tier 2

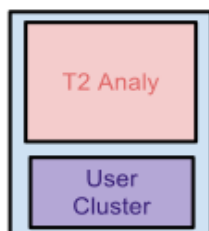


Tier 2

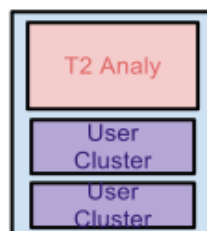
To..



Tier 1

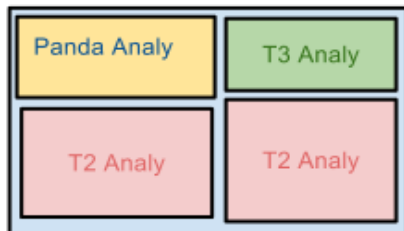


Tier 2

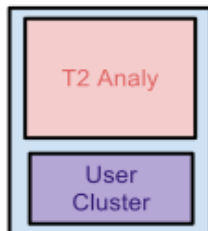


Tier 2

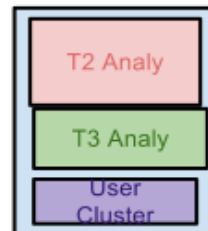
Or...



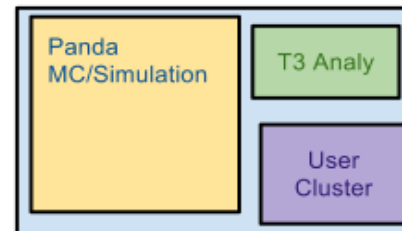
Tier 1



Tier 2



Tier 2



EC2/Scientific Clouds

Primary Goals



Type	Duration	Managed By	Typical Cloud Target	Rationale
Dynamic extension of static resources. (Cloud Panda)	Duration of extra need. HOURS-DAYS	Tier 1 Admins, on Tier 1 resources.	T1 Local cloud + Scientific Clouds	Clears MC off of T1 resources. Freeing them for data-intensive.
Tier 2/3 User Analysis Resource	WEEKS-MONTHS (although exact size dynamic)	Tier 2/3 Admins, on Tier 1 resources.	Tier 1 Private Cloud. Tier 2?	Utilizes T1/2 resources efficiently. Shared across T2s.
End-user Ad-Hoc Analysis (Cluster in the cloud)	Duration of urgent project, DAYS-WEEKS	By end-user, using tools provided by VO.	Local private clouds, ATLAS private clouds, EC2	Alternative to / expansion of T2/T3.

Secondary Goals



Goal	Duration	Technology	Rationale
Cloud-based Data Caching	Short term	Squid, HDFS, xrootd	Minimize WAN transfers for efficiency and economy.
Full DDM in the cloud.	Medium/Long term	Via DQ2 and/or XrootD ?	

Workload Management



CloudFactory:

- Dan Van der Ster, Fernando Barreiro (CERN IT)
- *cf* wraps *cvm* to invoke VM. custom wrapper for Panda pilot. EC2 compatible.
- Ran Panda jobs on CloudSigma. Stage-in/out to EOS at CERN.

AutoPyFactory:

- John Hover, Jose Caballero (BNL)
- General-purpose job factory. Also designed to invoke EC2 VMs in response to Panda queue status.
- Condor startd VMs manually tested and ran Panda jobs successfully. Data from/to BNL SRM.

Workload Management



Condor CloudScheduler

- Ian Gable, Ryan Taylor, Mike Paterson (UVic)
- Creates Condor cluster by invoking startd VMs (associated with static Master).
- Triggered by presence of jobs in local pool.

StratusLab demo:

- Rod Walker (LMU)
- Custom VM w/ Condor startd associated with static Master.
- Invoked manually.
- Ran ATLAS Panda jobs. Data to/from LRZ.

Cloud Analysis Clusters



CloudCRV:

- Val Hendrix, Doug Benjamin, Sergey Panitkin
- Puppet-based cloud appliance framework for defining, invoking, and managing heterogenous multi-node systems in Clouds.
- Also demo'd Starcluster and Scalr. Successfully ran analysis.

StarCluster PROOF/XrootD:

- Sergey Panitkin
- Added root and xrootd to existing cluster framework. Tested ATLAS analysis performance and cost on Amazon EC2. ~15VMs.
- Now collaborating with Doug and Val w/ CloudCRV.

Private T1/2 Clouds



LxCloud

- Ulrich Schwickerath, Belmiro Moreira (CERN IT)
- 430 VMs, OpenNebula based. Investigating OpenStack.

OpenStack at Fresno

- Cui Lin, et al.
- Initial setup working, 13T storage. Coordinating with LBL group.

OpenStack at BNL

- Xin Zhao, Tom Wlodek, with OSG TI
- Diablo (v3) installed. Register image, run, shutdown OK.
- Intended to be reproducible in OSG. Puppetized setup.



Current ATLAS-accessible Clouds

Cloud	Type	API	Max VMs	Status
LxCloud/CERN	OpenNebula 3	EC2 (partial)	~450	Production
StratusLab	OpenNebula 1.4	EC2?	~256	Production
BNL	OpenStack 3	EC2	~15	Initial
Fresno	OpenStack 3?	EC2	~100s?	Initial+
Nebraska	OpenStack 2	EC2	~50?	Dev
CloudSigma		?	~100s	Public/Commercial
Fermicloud?	OpenNebula		~130?	Production

Others?

- Helix Nebula European Cloud Europe-only?!
-

Virtualization



CERNVM:

- Used by many groups. General purpose. EGI middleware. Configured via Conary, Hepix contextualization..

Hepix Contextualization project

- EC2 Userdata + standard CDRM-based schema.

Custom images:

- Varying degrees of repeatability/automation.
- Customized Amazon AMIs.
- Ad-hoc custom images by others.
- BNL working on reproducible libvirt/Puppet/Hepix VM generation workflow. Investigating Boxgrinder.

Cloud Storage



XrootD Cluster in the Cloud

- Doug Benjamin (Duke), Sergey Panitkin (PAS)
- Dynamic creation. Performance testing.

Others?

Summary/ Next Steps



Ongoing Progress

- Most proof-of-concepts performed, and function.
- Now time to 'productize', scale up, and do wider deployments.
- CloudFactory: Testing on Amazon, personal pilots.
- AutoPyFactory: Run on OpenStack invoked VMs automatically.
- Private cloud infrastructure/Cloud APIs for grid sites.

Data model?

- So far tied to static site storage. Something between full DDM site in cloud and one-site static association?

VMs?

- CERNVM+Hepix only current standard. Could be role for OSG-managed image generation and publication.