Program of work for Cloud development

John Hover

jhover@bnl.gov

Outline

Why cloud?

The State of the Art

- CERNVM project

- ATLAS / ADC Work

- Panda Team Work

- LBNL Work and Proposal

BNL / RACF / OSG Work and Plans

- Computing Models for Cloud Utilization

Key Issues/Challenges

Questions/Discussion

Why?

Why are we (DOE HEP program, ATLAS) interested?

- Economics of central, possibly commercial, cloud resources vs. dedicated computing centers.

- VMs offer uniform environment vs. Grid.

- If Cloud is mandated by funding agencies, ATLAS must be able to smoothly transition existing model to a cloud-compatible one.

- Evolving data model already more adapted to non-locality anyway (direct I/O, federated xrootd, caching/Squid, FronTier).

- Other reasons? If so, it should influence decisions.

CERNVM Project

General-pupose LHC VM image

- Relatively small (~500MB download)

- CMVFS global filesystem included

- Standard contextualization

Challenges

- ATLAS-specific pieces

- EGI/glite -centric grid tools

New info?

ATLAS ADC Cloud work

Kick Off Workshop May 2011

- https://indico.cern.ch/conferenceDisplay.py?confId=136751

Cloud Panda Experiments (van Der Stern)

- Cvm tool -> CernVM instance via EC2 API on LxCloud.
- Software in CVMFS.
- Reads input from Castor via xrootd.
- Output to SRM (CERN-PROD_SCRATCHDISK)
- Cloud performance roughly equivalent to standard batch/grid.

He notes 2 options:

- Directly managed cloud VM instantiation (cvm+tools)
- Cloud-behind-Condor approach (BaBar CS).

Ref

Panda Team Work

Panda Server (Panitkin)

- Successfully ran on EC2 and Magellan.
- Used cloud for Panda server, development, data archiving.
- Ran trivial (non-ATLAS) job via EC2 Panda.

VM Worker Node (Alden—immediately preceeding)

- ??

Probably more I just don't know about...

LBNL Work and Proposal

Cloud CRV

- Dynamic, automated cluster-in-a-box. Deploys VMs with Condor CM and startds.

- Generalized node type management model. Can dynamically create arbitrary node types and associate them.

- Includes monitoring and startup-shutdown management.

- Web-based GUI for control and monitoring.

Current Proposal for New Work on the table:

- Automatic/dynamic scaling.

- GUI for low-level configuration.

- Data handling.

BNL / RACF / OSG

Organization:

– Organized as part of OSG Technology Group.

– Working with Brian B., Ashu Guru, Ryan Lee at UNL.

– Jose Caballero, John Destefano, John Hover, Tom Wlodek, and Xin Zhao at BNL. Official effort started Oct 1.

Goals:

– Gain admin expertise within our group at BNL.

– Easily deployable, release-able, components.

– Automated service and VM setup (RPM, Puppet).

– RPM repository via YUM. VM repository.

– Documentation sufficient for T2/T3 deployment where useful.

– Refined enough for eventual inclusion in OSG stack.

BNL / RACF / OSG Plan

1a. Condor VM Universe

– Run virtualized at site. Full ATLAS prod, analysis jobs.

– Testbed for VM refinement and testing.

– Wean ourselves from NFS, site locality.

– Initially avoid any latency/throughput issues.

1b. VM Creation and Testing

– CERNVM + CVMFS

– Possible use of APF on VMs as integrated pilot-runner. c.f. custom runpanda script by ADC.

– Scalable, public VM download repository.

BNL / RACF / OSG Plan (cont.)

2. Elastic local cluster

- BaBar CloudScheduler.

- Simple node model (one static VM image).

- Dynamically scalable. Full lifecycle management.

- Based on Condor/Condor-G, avoids re-invention. Supports all provider APIs/quirks.

- Fits within existing model.
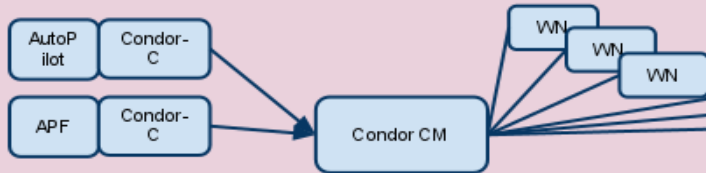
3. Cloud provider infrastructure

- Installed locally.

- OpenStack

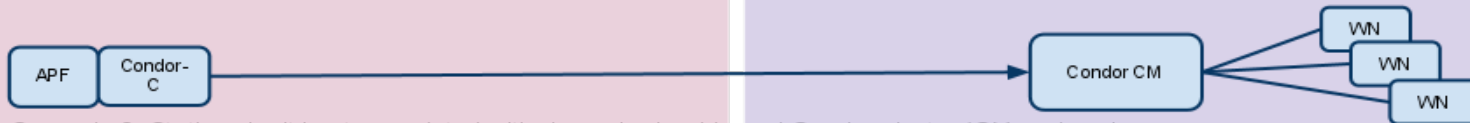- Redhat DeltaCloud (feeds OpenStack and Condor-VM)

Computing Models

BNL Cloud Implementation Roadmap.


Static/Local

Cloud/Dynamic

Scenario 1: Dynamic Cloud-based expansion of existing Condor-C cluster. Initially with AutoPilot and then APF.

Scenario 2: Static submit host associated with dynamic cloud-based Condor cluster (CM and worker nodes). This could utilize the CloudCRV developed by Yushu.

Scenario 3: Condor-G submission to static CE, which is associated with dynamic cloud-based Condor cluster. Possibly a good time to move to CERNVM and CERNVMFS for release, conditions, and data access.
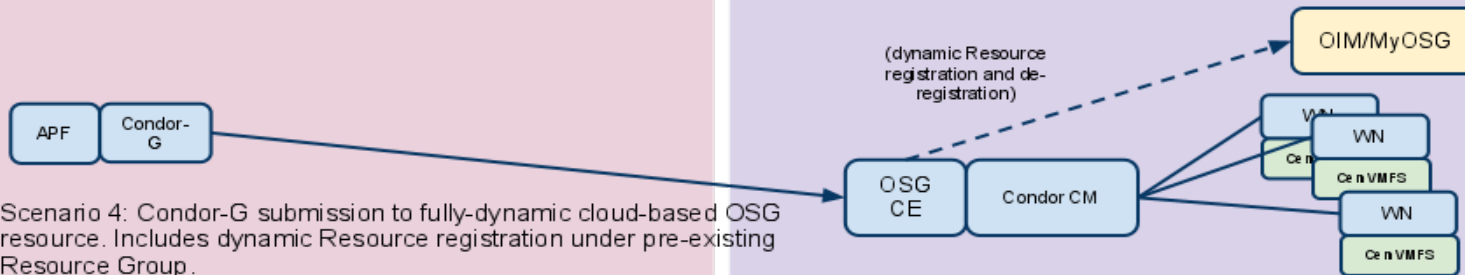
Scenario 4: Condor-G submission to fully-dynamic cloud-based OSG resource. Includes dynamic Resource registration under pre-existing Resource Group.

(dynamic Resource registration and de-registration)

John Hover, RACF, BNL

BNL / RACF / OSG Plan (cont.)

4. OSG supporting infrastructure

- WN monitoring, network throughput metrics

5. Full dynamic grid/cluster sites in cloud?

- Scenario 3 and 4 in diagram.
- CloudCRV likely mechanism.

6. Sites as cloud providers? Replace GRAM/CREAM w/ EC2 API?

Challenges

General Challenges

- Cloud provider APIs. EC2 default standard, but...

- Simple, cloud WNs tied to static sites and services vs. fully "Cloud-ified" model (complex services + processing in cloud).

ATLAS Challenges

- Data input access. Direct I/O? Cloud-based caching?

- Job brokering. Cloud is "siteless", how does Panda decide what jobs to send?

- Data stage-out.

- How to integrate Cloud-based processing with existing DDM model?

Questions/Discussion

What do we *need*?

– Assuming it works and scales, would the CloudScheduler satisfy our needs? Or is cloud-based complex processing (PROOF?) required?

What do we *want*?

Anything else OSG / Tier 1 should investigate?

What don't we know yet?

Extra Slides/Refs

What is cloud?

– Grid = Queued provisioning of reciprocately shared, non-uniform resources via uniform interface(s).

– Cloud = On-demand provisioning of uniform computing resources via uniform interface(s).

Van Der Stern's Talk

– https://indico.cern.ch/conferenceDisplay.py?confId=154168

WLCG Cloud Meeting

– https://indico.cern.ch/conferenceDisplay.py?confId=136751