Agile Infrastructure: an updated overview of IaaS at CERN

Luis FERNANDEZ ALVAREZ
on behalf of Cloud Infrastructure Team
luis.fernandez.alvarez@cern.ch
HEPiX Spring 2013
Agenda

• Introduction
• Goals & Use Cases
• CERN Contribution
• IaaS Overview
  – OpenStack deployment
  – Work-in-progress
• Current Status & Timeline
• Outlook
Agile Infrastructure - IaaS

- **Rationale**
  - Need to manage twice the servers as today
  - No increase in staff numbers
  - ...we have to manage the infrastructure efficiently

- **Infrastructure as a Service (IaaS)**

---

### Traditional IT
- Applications
- Data
- Runtime
- Middleware
- Operating System
- Virtualization
- Servers
- Storage
- Networking

### Infrastructure (as a Service)
- Applications
- Data
- Runtime
- Middleware
- Operating System
- Virtualization
- Servers
- Storage
- Networking

### Platform (as a Service)
- Applications
- Data
- Runtime
- Middleware
- Operating System
- Virtualization
- Servers
- Storage
- Networking

### Software (as a Service)
- Applications
- Data
- Runtime
- Middleware
- Operating System
- Virtualization
- Servers
- Storage
- Networking
What is OpenStack?

“OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter, all managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface.”
Goals

- **Objectives**
  - 90% of hardware virtualized
  - ~15k hypervisors needed over next two years
- **Following industry reference structure**
  - Infrastructure, Platform and Applications
- **Deploy multi-site**
  - Extend to 2\textsuperscript{nd} data center in Hungary
  - Enable remote management
- **More efficient use of our hardware**
- **Better tracking of usage**
- **Consolidate support to a single device**
- **Support potential new use cases (PaaS)**
Use cases

Pets are given names like pussinboots.cern.ch
They are unique, lovingly hand raised and cared for
When they get ill, you nurse them back to health

Cattle are given numbers like vm0042.cern.ch
They are almost identical to other cattle
When they get ill, you get another one

Future application architectures tend towards Cattle but Pet support is needed for some specific zones of the cloud
Collaborating

• CERN is contributing, the Open Source way:
  – OpenStack, candidates testing, bug fixing,…
    • Keystone, nova,…
    • Hyper-V Team
  – Fedora, active participation during OpenStack Test Day,…
  – PuppetLabs, Foreman, Aeolus Oz,…

• Exploit the collaboration tools benefits:
  – Mailing lists, IRC, bugzilla, Launchpad, GitHub,…

• Collaborations already starting around
  – BNL, IN2P3, NECTaR (Australia), IHEP (China), ATLAS/CMS
    Trigger teams,…
Infrastructure Overview

CERN

Account mgmt. system
Microsoft Active Directory

CERN Block Storage provider
CERN Network Database

Cinder
Keystone
Horizon
Glance

Compute
Network
Scheduler

openstack™

CERN DB on Demand
Nova

Cloud computing fabric controller

- Network manager modified for CERN
  - Integration with network database
  - These problems are specific to our case
  - No modifications pushed upstream
- Nova Compute aware of CERN DNS and AD service
- Multiple availability zones
  - Special zone for Hyper-V hypervisors
  - Scheduler has an additional filter based on image distribution metadata
- Other small tweaks…
  - Renaming while launching multiple instances…
Glance

Services for discovering, registering, and retrieving vm images

• No major modifications in this module
  – New image meta-data for CERN integration

• Aim for automated image creation and updating
  – Design common process for Linux & Windows images.
  – Use of common tools, basically Aeolus Oz
    • Contribution for improving Windows support in Oz
  – CERN tool for communicate Oz and Glance API

• AI will provide images for CERN supported OS:
  – SLC5, SLC6, Windows {2008 R2, 7, 2012}
  – Users can upload their own images

• Deployed images are contextualized with cloud-init
  – New cloud-init for Windows implemented by Cloudbase
Identity service: authentication, authorization and service catalog

- Full integration with Active Directory via LDAP
  - CERN's AD: 44,000 users & 29,000 groups
  - Uses the LDAP backend with some particular configuration settings
  - Aim for minimal changes to Active Directory
  - Patches submitted around hard coded values, additional filtering and pagination

- Account Mgmt. System integration for project creation/deletion
  - Implemented generic connector to interact with keystone

- Connections are made through secure channels (SSL)

- Now in use in our pre-production instance
  - Map project roles (admins, members) to groups
  - Documentation in the OpenStack.Org wiki
Horizon (I)

Provides a graphical user interface for managing OpenStack services

- Secure connections through SSL
- Working on integration with CERN SSO system
  - Shibboleth authentication
- Customization for end-users:
  - CERN specific
  - Make it easier for the end user: select OS, flavor, launch!
  - Hide details of security groups, floating IPs, snapshots, etc.
  - Make use of Horizon modularity.
## Overview

Select a month to query its usage:

<table>
<thead>
<tr>
<th>Month</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>April</td>
<td>2013</td>
</tr>
</tbody>
</table>

Active Instances: 365  Active RAM: 2TB  This Month's VCPU-Hours: 70775.81  This Month's GB-Hours: 3978503.25

### Usage Summary

<table>
<thead>
<tr>
<th>Project Name</th>
<th>VCPUs</th>
<th>Disk</th>
<th>RAM</th>
<th>VCPU Hours</th>
<th>Disk GB Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS Victoria</td>
<td>408</td>
<td>9180</td>
<td>816GB</td>
<td>5189.63</td>
<td>896002.59</td>
</tr>
<tr>
<td>ATLAS</td>
<td>56</td>
<td>1120</td>
<td>112GB</td>
<td>2794.98</td>
<td>223598.42</td>
</tr>
<tr>
<td>IT SVN</td>
<td>52</td>
<td>1040</td>
<td>104GB</td>
<td>1445.59</td>
<td>115647.37</td>
</tr>
<tr>
<td>IT Batch</td>
<td>49</td>
<td>980</td>
<td>98GB</td>
<td>886.13</td>
<td>28934.22</td>
</tr>
<tr>
<td>Personal straylen</td>
<td>48</td>
<td>960</td>
<td>96GB</td>
<td>5973.34</td>
<td>238933.69</td>
</tr>
<tr>
<td>PH-SFT build</td>
<td>44</td>
<td>660</td>
<td>88GB</td>
<td>2455.78</td>
<td>147346.55</td>
</tr>
<tr>
<td>NA61</td>
<td>43</td>
<td>860</td>
<td>86GB</td>
<td>1562.77</td>
<td>191997.01</td>
</tr>
<tr>
<td>CMS</td>
<td>41</td>
<td>960</td>
<td>82GB</td>
<td>2796.92</td>
<td>193565.16</td>
</tr>
</tbody>
</table>

*Agile Infrastructure: an updated overview of IaaS at CERN*
In progress… Ceilometer

Unique point of contact to acquire all counters across all OpenStack components

• CERN delivers resources in form of virtual machines and via traditional batch and Grid computing
  – Accounting should cover both use cases
• Interesting metrics for us: What is the resource usage of experiment A during December? What is the resource usage of user B last year?

• Use of ceilometer as single source of truth for accounting data
• Batch data is put in the ceilometer database for accounting purpose

• The grow rate of the mongodb database is about 2GB/day
• Hyper-V accounting is expected during Havana release
In progress… Monitoring

- Infrastructure monitoring
  - Cloud hardware and services
  - User’s cloud ecosystem
  - Performance of cloud resources
- No OpenStack project
  - Service probing, operator dashboard,…
- Integration with General Notification Infrastructure (GNI) developed by AI Monitoring group
- Windows compute nodes
  - Integrate windows nova nodes in CERN windows monitoring: Microsoft SCOM
  - Development of SCOM Management Pack for OpenStack Nova Compute.
  - Deliver SCOM messages to GNI
Windows deployment

- Deployment of Hyper-V Nova Compute nodes
- Why emphasize Hyper-V?
  - Migration of current virtualization infrastructure based on SCVMM to OpenStack. ~1000 Windows instances
  - Deployment of new Windows machines in a fully Microsoft supported stack
- Aim for deploying Windows compute nodes with Puppet modules
- Working closely with the Hyper-V OpenStack team
  - Puppet modules, cloud-init, rdp console design,…
- In progress: Monitoring, Ceilometer agent…

Agile Infrastructure: an updated overview of IaaS at CERN - 16
Service Implementation

- **Preparation for production**
  - First steps towards high availability deployment

- **OpenStack services running behind DNS load balancer:**
  - Two servers
  - Keystone, Network, Scheduler…
  - Working on glance and console

- **Database provided by CERN Database Services**
  - Database on Demand
  - MySQL
  - External database for: keystone, nova and glance
Current status

• Working on Folsom code base from the EPEL repository
  – Excellent experience with the Fedora Cloud SIG team
• Pre-production facility with around 200 Hypervisors
• Number of VMs growing every day!
  – ~360 instances
  – 164 projects (personal & shared)
• Maintain the project linked to upstream releases
• Deployment delay:
  – Time for testing
  – Checking new features suitability for us
  – CERN patches
• We have been delivering a series of pre-production services…
  – Adding functionality…
  – Scaling out…
  – Integrating into IT infrastructure…
Outlook (I)

• Going to **production** for managed servers
  – End of May, 2013
  – Based on Grizzly release
  – ~2000 Hypervisors (KVM and Hyper-V)
  – Load Balancing

• **Grizzly** **new features**
  – Preview release of cells functionality
  – No-DB-Compute
  – Keystone Domains
  – Bare metal provisioning
  – Better Hyper-V support, resizing, cinder, quantum…
  – Database archiving, admin APIs, UI enhancements,…
Outlook (II)

• Non-managed machines, missing requirements:
  – Shared storage for volume service and live-migration
• Investigating alternatives for block storage:
  – NetApp
  – Ceph
• Special attention to cells
  – Exploit cells functionality for scalability
  – Multi-site operations with new data centre
• And more:
  – Load balancing as a service
  – Heat for Orchestration
  – …
More information…

http://cern.ch/openstack

…@ HEPiX Spring 2013…

• Agile Infrastructure Monitoring. Pedro Manuel RODRIGUES DE SOUSA ANDRADE

• High Availability Load Balancing in the Agile Infrastructure (CERN). Vaggelis ATLIDAKIS

• Experiences running a production Puppet infrastructure at CERN. Ben JONES

• Ceph: A scalable, organic option for Storage-as-a-Service at CERN. Arne WIEBALCK
Backup Slides
IaaS Community Analysis

Figure 8 - Active Community Population

- Total
- Last Quarter
- Last Month

Figure 5 - Accumulated Community Population

OpenStack | OpenNebula | Eucalyptus | CloudStack

http://www.qyjohn.net/?p=3120
• Following industry reference structure
  – Infrastructure, Platform and Applications
• Start first with Infrastructure-as-a-Service at scale
  – Just provide Amazon functionality
• Value add increases up the chain so we need a vibrant ecosystem for PaaS and SaaS