Enabling cloud for e-Science with OpenNebula

cloud-devroom@Fosdem’13

Zeeshan Ali Shah
System Administrator, PDC-HPC
KTH, Stockholm
zashah@pdc.kth.se

disclaimer: opinions expressed in this talk are solely those of the presenter and do not reflect of KTH or any other institute.
Agenda

- eScience users !! customer segmentation
- Projects
  - NEON
  - Venus-C
  - SNIC Cloud
- Challenges
- Federation
  - EGI Fed cloud project
- QA
eScience users

Initially we focused on bioinformatics

Currently running on HPC machines

That was an issue

Think in HPC, but an elastic way

That was a challenge

Not all, but some need bursty peaks (longtail users)

That was a core proposition
NEON Project

- Northern europe cloud initiative, 2009-2010
- Eucalyptus 2.0
- Federated with other centre
Venus-C Project

- Finished May 2012
- OpenNebula with CDMI, OVF
SNIC cloud Project

Currently running

Started with Public cloud Amazon AWS.

Why?

Prebuilt images, App stacks, Trainings, Workshops ..

Planning to connect with Private Cloud (with new challenges)
Challenges @ Private Cloud

- Non-Technical (changing hearts n minds)
- Technical (Using minds to solve issues)
Non-Technical

- Myth of security
- Living in HPC world.
- It is hard to digest elasticity, self-provisioning, on demand and other cloud benefits.
- But Private cloud has to live with other computer/network admins ... or may be configure by same sysadmin
more cloud = jobless sysadmins!
is this true?

Same Role (Driver)
Different styles
(Choose yours!)
Technical

- Network Latency
- Infiniband
- Multiple NICs

Storage

- Image repository
- EBS style (for application data)
- CDMI
Technical (more)

- Public IPs
  - For Cluster deployment limit to 1 master and X private ips
  - We want more ...
  - (hybrid cloud)
- Sharing of images (but be careful about private data)
  - Hooking mechanisms in Open Nebula
Technical (more)

Usage

- UI based on Sunstone
- command line with OCA client
- For Programmers, use SDK (Java n Ruby)
agile developer vs static administrator
SDK

Think compute and storage as an object (Create/Operate/Destroy whenever you want)

VirtualMachine vm = new VirtualMachine
......
Spark sp = new Spark(4); # spark with 1 master 4 slaves
......
if load increase 90% launch more #(just an e.g.)
...
PlainDebian pd = new PlainDebian();
....
#Expand as you want :)

NoDep on sysadmin
Technical (more)

Security

VLANs

Bridge firewalls

Network auditing and other traditional tools

Q: What if legitimate user install legal software but left its mgt panel opened, e.g. tomcat
EGI Cloud Federation

https://wiki.egi.eu/wiki/Fedcloud-tf
Areas

VM Management
- OCCI 1.1 proxy for multiple IaaS

Data Management
- CDMI proxy for multiple IaaS
- OVF

Information Systems
- Extended GLUE2 schema
- LDAP server

Accounting
- Cloud Usage Record (UR) schema
- UR server
- UR client for each IaaS

Monitoring
- Nagios with cloud probes

Federated AAI
- X509 certificates
- Support for Virtual Organisations (Vos)

Image catalogue
- StratusLab marketplace
Consolidation

Federation
- OCCI/CDMI deployment

Information System
- GLUE2 extension
- upload from the RPs

Accounting
- OGF UR extension
- upload clients

Monitoring
- Metrics
- Availability -> performance

Notification
- Review available implementations.

Federated AAI
- RP account integration
- VOMS?

Image management
- Multiple storage model
EGI Cloud Federation

Members
- 60 individuals.
- 23 institutions.
- 13 countries.

Technologies
- 7 OpenNebula.
- 3 StratusLab.
- 3 OpenStack.
- 1 Okeanos.
- 1 WNoDeS.

Stakeholders
- 15 Resource Providers.
- 7 Technology Providers.
- 5 User Communities.
- 3 Liaisons.
Components

- Federation Test bed
- Resource Providers: OpenNebula, OpenStack, Okeanos, WNoDeS...
- VM Mngmt: OCCI 1.1
- VM sharing: Marketplace
- Information: GLUE 2.0+, LDAP/BDII
- AAI: X509 [SSO, OpenID, Shib]
- Resource Providers: OpenNebula, OpenStack, Okeanos, WNoDeS...
- Data: CDMI 1.0 [OVF 1.1]
- Accounting: OGF UR+ [StAR]
- Monitoring: Nagios
- Message Bus

Clients
Brokers
Portals
...
Issues @ Hands

- Workflow (like SWF of Amazon AWS)
- Orchestration service, Scale up/down (like cloudwatch of AWS)
- Auditing VM from inside! should we?
- Bare metal IaaS to apps. (Galaxy, Mapreduce, Spark ...)

- But platform security is even harder!
Thanks
QA
@zeeshanaisalishah
zashah@pdc.kth.se