OPENSTACK VS. GANETI

LANCE ALBERTSON

Director, Oregon State University Open Source Lab

http://osuosl.org

@ramereth @osuosl

Attribution-ShareAlike CC BY-SA ©2015-2017
SESSION OVERVIEW

- OpenStack quick overview
- Ganeti overview
- Ganeti walk-through
- Comparing OpenStack vs. Ganeti
ABOUT ME

• Lance Albertson
• Director, OSU Open Source Lab (OSUOSL)
  ▪ Provide infrastructure hosting for FOSS projects
  ▪ Linux Foundation, Python Software Foundation, Drupal, etc
• Ops guy
• Ganeti user since 2009
• OpenStack user since 2013
• http://osuosl.org
PRIVATE CLOUD CONSIDERATIONS

- Each organization has different requirements
- Some are small and simple
- Others are more complex and larger
- Cost is also a major factor (both in licenses and staff maintenance)
- High availability vs. Ephemeral resources
- Integration with current hardware
TYPICAL PRIVATE CLOUD SOLUTIONS

- VMWare
- OpenStack
- [insert favorite solution]
OPENSTACK OVERVIEW
OPENSTACK POPULARITY

- Designed to scale and function similar to AWS
- Many companies are investing in it
- The project itself is very large and is still maturing
- Includes a wide-array of features and optional components
OPENSTACK PROS

- Standard Cloud API
- Fast VM deployment and tear down
- Excellent for elastic computing needs
- Large community support
- Quickly growing and adding new features constantly
OPENSTACK CONS

- Extremely difficult to deploy and maintain
- Lots of moving parts to get it running
- Project is still maturing and unstable
- Not an ideal solution for small to medium sized organizations
- Requires several administrative machines to use effectively
- HA is possible but takes a lot of work
- Until recently, upgrades are very difficult
SO WHAT ABOUT GANETI?
FIRST, WHAT IS GANETI?

- Cluster management tool for virtual compute resources
- IaaS solution that uses either KVM or Xen hypervisors
- Provides fast and simple recovery from hardware failures
- N+1 redundancy built into Ganeti for VM resources
- Primarily CLI driven, sysadmin focused
- Supports live migration
- Cluster re-balancing
GANETI ARCHITECTURE OVERVIEW

- Comprised of several daemons running on each node
- One node is marked as a master which can be easily migrated
- Configuration is replicated to all master-capable nodes and uses flat files
PROJECT BACKGROUND

- Open Sourced in 2007 from an internal Google project
- Used widely internally at Google for back-office needs
- Active community, mailing list and IRC
- Started before libvirt/OpenStack existed
- New releases every 6-9 months
- Primarily written in Python / Haskell
- Recently migrated to Github (http://github.com/ganeti/ganeti)
GANETI GOALS

Low Entry Level

- Easy to install, manage and upgrade
- Architecture is fairly easy to understand

Enterprise Scale

- Manage 1 to 200 within a single cluster

Open Source Citizen

- Design and code discussions are open to the community
- Welcome third-party projects
ARCHITECTURE
Clusters are comprised of nodes, one of which is master
Nodes can be split up into logical groups
Instances (guests) run on nodes
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node</td>
<td>Virtualization host</td>
</tr>
<tr>
<td>Instance</td>
<td>Virtual Machine Guest</td>
</tr>
<tr>
<td>Cluster</td>
<td>Set of nodes, managed as a collective</td>
</tr>
<tr>
<td>Node Group</td>
<td>homogeneous set of nodes (i.e. rack of nodes)</td>
</tr>
<tr>
<td>Job</td>
<td>Ganeti operation</td>
</tr>
</tbody>
</table>
STORAGE IN GANETI

- Known as disk templates
- LVM, DRBD
- RBD (Ceph)
- File (both local and shared via NFS)
- External storage provider
  - Useful to interface with existing storage appliances
- Designed to be flexible
- Instances always run on primary node
- Uses secondary node for replication when using DRBD template
- Also works with RBD, ext and shared file templates
## LISTING NODES

```
root@node1:~# gnt-node list
<table>
<thead>
<tr>
<th>Node</th>
<th>DTotal</th>
<th>DFree</th>
<th>MTotal</th>
<th>MNode</th>
<th>MFree</th>
<th>Pinst</th>
<th>Sinst</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1.example.org</td>
<td>26.0G</td>
<td>25.5G</td>
<td>744M</td>
<td>186M</td>
<td>587M</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>node2.example.org</td>
<td>26.0G</td>
<td>25.5G</td>
<td>744M</td>
<td>116M</td>
<td>650M</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
```

## Listing OS images

```
root@node1:~# gnt-os list
<table>
<thead>
<tr>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>image+cirros</td>
</tr>
<tr>
<td>image+default</td>
</tr>
</tbody>
</table>
```
CREATING A NEW VM

```bash
root@node1:~# gnt-instance add -n node1 -o image+cirros -t plain -s 1G \  
    --no-start instance1
Thu Jun  7 06:05:58 2015 * disk 0, vg ganeti, name 780af428-3942-4fa9-8307-1323de416519.disk0
Thu Jun  7 06:05:58 2015 * creating instance disks...
Thu Jun  7 06:05:58 2015 adding instance instance1.example.org to cluster config
Thu Jun  7 06:05:58 2015  - INFO: Waiting for instance instance1.example.org to sync disks.
Thu Jun  7 06:05:58 2015  - INFO: Instance instance1.example.org's disks are in sync.
Thu Jun  7 06:05:58 2015 * running the instance OS create scripts...

root@node1:~# gnt-instance list
Instance              Hypervisor OS           Primary_node      Status     Memory
instance1.example.org kvm        image+cirros node1.example.org ADMIN_down      -
```
root@node1:~# gnt-instance info instance1
Instance name: instance1.example.org
UUID: bb87da5b-05f9-4dd6-9bc9-48592c1e091f
Serial number: 1
Creation time: 2015-06-07 06:05:58
Modification time: 2015-06-07 06:05:58
State: configured to be down, actual state is down
   Nodes:
      - primary: node1.example.org
      - secondaries:
Operating system: image+cirros
Allocated network port: 11000
Hypervisor: kvm
   - console connection: vnc to node1.example.org:11000 (display 5100)
...
Hardware:
   - VCPUs: 1
   - memory: 128MiB
   - NICs:
      - nic/0: MAC: aa:00:00:dd:ac:db, IP: None, mode: bridged, link: br0
Disk template: plain
Disks:
   - disk/0: lvm, size 1.0G
CONVERTING DISK TEMPLATE

root@node1:~# gnt-instance shutdown instance1
Waiting for job 11 for instance1.example.org ...

root@node1:~# gnt-instance modify -t drbd -n node2 instance1
Thu Jun  7 06:09:07 2015 Converting template to drbd
Thu Jun  7 06:09:08 2015 Creating additional volumes...
Thu Jun  7 06:09:08 2015 Renaming original volumes...
Thu Jun  7 06:09:08 2015 Initializing DRBD devices...
Thu Jun  7 06:09:09 2015  - INFO: Waiting for instance instance1.example.org to sync disks.
Thu Jun  7 06:09:11 2015  - INFO:  - device disk/0:  5.10% done, 20s remaining (estimated)
Thu Jun  7 06:09:31 2015  - INFO:  - device disk/0: 86.00% done, 3s remaining (estimated)
Thu Jun  7 06:09:34 2015  - INFO:  - device disk/0: 98.10% done, 0s remaining (estimated)
Thu Jun  7 06:09:34 2015  - INFO: Instance instance1.example.org's disks are in sync.
Modified instance instance1
  - disk_template -> drbd
Please don't forget that most parameters take effect only at the next start of the instance.
LIVE MIGRATION

root@node1:~# gnt-instance start instance1
Waiting for job 14 for instance1.example.org ...

root@node1:~# gnt-instance migrate -f instance1
Thu Jun  7 06:10:38 2015 Migrating instance instance1.example.org
Thu Jun  7 06:10:38 2015 * checking disk consistency between source and target
Thu Jun  7 06:10:38 2015 * switching node node1.example.org to secondary mode
Thu Jun  7 06:10:38 2015 * changing into standalone mode
Thu Jun  7 06:10:38 2015 * changing disks into dual-master mode
Thu Jun  7 06:10:39 2015 * wait until resync is done
Thu Jun  7 06:10:39 2015 * preparing node1.example.org to accept the instance
Thu Jun  7 06:10:39 2015 * migrating instance to node1.example.org
Thu Jun  7 06:10:44 2015 * switching node node2.example.org to secondary mode
Thu Jun  7 06:10:44 2015 * wait until resync is done
Thu Jun  7 06:10:44 2015 * changing into standalone mode
Thu Jun  7 06:10:45 2015 * changing disks into single-master mode
Thu Jun  7 06:10:46 2015 * wait until resync is done
Thu Jun  7 06:10:46 2015 * done
CLUSTER REBALANCING

- Ability to rebalance CPU, memory and storage across the nodes
- Useful when adding or removing nodes
- Initiated using the `hbal` command

```
$ hbal -L -q
Loaded 8 nodes, 130 instances
Group size 8 nodes, 130 instances
Selected node group: default
Cluster is not N+1 happy, continuing but no guarantee that the cluster will end N+1 happy.
Initial score: 59.58075308
Trying to minimize the CV...
  1. instance1.osuosl.org  gprod3:gprod8 => gprod3:gprod4  58.23987138 a=r:gprod4
  2. instance2.osuosl.org  gprod3:gprod8 => gprod3:gprod4  56.95979668 a=r:gprod4
  3. instance3.osuosl.org  gprod3:gprod8 => gprod3:gprod4  55.75769557 a=r:gprod4
  4. instance4.osuosl.org  gprod2:gprod8 => gprod2:gprod4  54.65118990 a=r:gprod4
Cluster score improved from 59.58075308 to 54.65118990
```
MISC

- Network Management
- Storage integration
- Linux-HA Support
- Auto Repair VMs
- Hotplug (NIC or Disks)
- Open vSwitch support
- Hsqueeze (automatically drain and power down nodes)
COMPARING OPENSTACK TO GANETI
COMMON USE CASES FOR GANETI

- Cheap, stable and reliable virtual compute resources
- Hosting web sites and other misc services in a private
- Useful for hosting “pet” virtual machines
- Need a highly reliable IaaS
- Small to medium size organizations with few sysadmins
GANETI PROS

- Architecture is fairly easy to deploy and understand
- Requires a minimal staff to maintain and upgrade
- Scales well for small/medium organization needs
- Highly customizable backend
- Built-in redundancy
- Upgrades are easy and quick
- It just works!
GANETI CONS

- No GUI frontend by default (third party projects exist)
- API isn’t very cloud compatible
- API not intended to be open to general users of the platform
- Management becomes slower the larger the cluster gets (although, its improving)
- Deploying VMs can be slower than compared to OpenStack
- VM image support is lacking but helped with ganeti-instance-image
- Doesn’t scale as well as OpenStack
GANETI + SYNNEFO = OPENSTACK-ISH

Synnefo is a complete open source IaaS cloud stack written in Python that provides Compute, Network, Image, Volume and Object Storage services

https://www.synnefo.org

- Manages multiple Ganeti clusters
- Provides API/Accounting/Quota/Block/Object storage
- Written in Python by GRNET
- Transforms Ganeti into an OpenStack/AWS-like platform
HOW THE OSL IS USING GANETI

- Hosting all of the production VMs which power our infrastructure
- Project specific VM(s)
- Mix of shared web infrastructure
  - Load balancers, web frontends, backend services
- OpenStack controller node :)  
- Whenever we want to host something that needs to have high reliability
HOW THE OSL IS USING OPENSTACK

- Two clusters:
  - OSL-internal x86
  - Public POWER8 based
- OSL-internal
  - Chef cookbook integration testing
  - Developer staging/development VMs
  - Multi-node testing
- POWER8 (soon POWER9)
  - FOSS project ppc64/ppc64le porting efforts
  - POWER8 software testing
OPENSTACK / GANETI SIDE-BY-SIDE

OpenStack:

- Pro: Great for quickly creating test vms for whatever needs you have
- Con: Extremely complicated to setup and maintain

Ganeti:

- Pro: Extremely fault tolerant and stable VM hosting and easy to use/maintain
- Con: Doesn’t scale well for cloud-specific needs
OUR PLANS WITH BOTH PLATFORMS

- Open up our x86 OpenStack cluster to our hosted projects soon(TM)
- Continue using Ganeti along-side OpenStack for production services
- Continue supporting both platforms long term
FINAL SUMMARY

- Both fill a specific niche in the ecosystem
- OpenStack will eventually mature and become more stable
- Give Ganeti a look, might be what you’re looking for if OpenStack is too complicated
- Make sure you experiment with both and fully understand their maintenance needs
- And if you’re able to, running both gives you the best of both worlds!
QUESTIONS?

- Lance Albertson
- lance@osuosl.org
- @ramereth @osuosl

References:

- http://osuosl.org
- http://www.ganeti.org/
- https://supermarket.chef.io/cookbooks/ganeti
- https://github.com/osuosl/ganeti-instance-image

Attribution-ShareAlike CC BY-SA ©2015-2017