An Introduction to OpenStack and its use of KVM

Daniel P. Berrangé <berrange@redhat.com>
About me

- Contributor to multiple virt projects
- Libvirts Developer / Architect 8 years
- OpenStack contributor 1 year
- Nova Core Team Reviewer
- Focused on Nova libvirt + KVM integration
Talk Structure

• Introduction to OpenStack services
• Compute service architecture
• Guest boot sequence
• Recent developments
What is OpenStack?

- Public or private cloud
- Self-service user API and dashboard
- Apache 2.0 licensed
- Broad community contribution
What is in OpenStack?

- Compute (Nova)
- Network (Neutron)
- Image storage (Glance)
- Block storage (Cinder)
- Object storage (Swift)*
- Identity (Keystone)*
- Metering (Ceilometer)*
- Orchestration (Heat)*
- Dashboard (Horizon)*

* not discussed in this presentation
What is Nova?

- Execution of compute workloads
- Technology agnostic
  - Virtual machine or container virt
- Virtualization agnostic
  - Libvirt (KVM, QEMU, Xen, LXC), XenAPI, Hyper-V, VMware ESX, PowerVM, Docker, Bare-metal
What is Glance?

- Write-once, read-many storage of images
- Image copied on use by Nova
- Format agnostic
  - eg raw, qcow2, etc
- Metadata properties
  - eg specify virtual hardware preferences
What is Cinder?

• Persistent block storage
• Multiple storage backends
  – eg LVM, RBD, Gluster, Sheepdog, ...+ more...
• Exposed to compute host via iSCSI
• Optional direct access by compute
  – Gluster
• Pre-requisite for live migration
What is Neutron?

- Network infrastructure management
- Concepts
  - Networks
  - Routers
  - Subnets
  - Ports
- Multiple technologies
  - OpenVSwitch, Linux Bridge, Vendor plugins
Nova architecture

• Concepts
  - Instances
  - Flavours
  - Virt drivers
  - Security group

• Dual APIs
  - OpenStack REST
  - EC2 compatible REST
Nova architecture (simple)
Nova architecture (scaling)
Nova cells

- Partition cloud infrastructure
  - Resilience within a data center
  - Scale out across data centers
  - Technology variation (eg KVM vs Hyper-V)
Nova cells architecture (Part 1)
Nova cells architecture (Part 2)
Nova cells architecture (Part 3)
Nova scheduler

- Places instances on compute hosts
- Pluggable filtering rules
  - CPU model / architecture
  - Virtualization type
  - PCI device availability
  - CPU, RAM, Disk usage
  - Trusted boot (TXT)
  - +more...
Nova conductor

- Mediates database access
- No database access from compute hosts
  - Compute hosts relatively untrusted / high risk
  - Scalability bottleneck for database
- Compute hosts issues RPC calls
- Conductor updates database state
Nova graphics proxy

- No direct compute access for VNC/SPICE
- Nova VNC/SPICE websockets proxy
- HTML5 VNC/SPICE browser clients
- Obtain auth token via REST API
- Pass to websockets proxy to authenticate
- Data proxied between compute & proxy
Instance boot step 1

1

Identity
Dashboard
Orchestration
Metering
Object Storage
Image Service
Block Storage
Networking
Compute
Instance boot step 2

1. Identity
2. Dashboard
Orchestration
Metering

Object Storage
Image Service
Block Storage
Networking

Compute
Instance boot step 3
Instance boot step 4

1. Identity
2. Dashboard
3. Orchestration
4. Metering

Object Storage
Image Service
Block Storage
Networking
Compute
Instance boot step 5
Instance boot step 6

1. Identity
2. Dashboard
3. Orchestration
4. Metering
5. Object Storage
6. Image Service
7. Block Storage
8. Networking
9. Compute
Nova KVM config part 1

- CPU
  - Named model or host model or host passthrough
- NIC model
- Disk bus type
- PCI device assignment
- Serial console x2 (1x log, 1x interactive)
- Disk devices
Nova KVM config part 2

- SMBios info
- CPU pinning
- VNC or SPICE
- QEMU + SPICE agents
- Clock (PIT, RTC) parameters
- Scheduler, disk, network tunables
New in Havana

• Released Thursday 17\textsuperscript{th} Oct
• Notable features
  – Block storage backend migration
  – Store images in RBD
  – Gluster native boot
  – QEMU guest agent assisted snapshots
  – PCI device assignment
Coming in Icehouse

• Target 17\textsuperscript{th} Oct + 6 months
• Planning summit in Hong Kong Nov 4\textsuperscript{th}-8\textsuperscript{th}
• Notable blueprints
  – VM ensembles
  – VM migration with storage
  – Live snapshots (disk + RAM)
  – Host reservation (user request entire host)