Considerations for Building a Private Cloud

Folsom Update

Ryan Richard, RHCA
OpenStack Architect - Private Cloud
ryan.richard@rackspace.com
@rackninja

April 2013

Why Folsom?

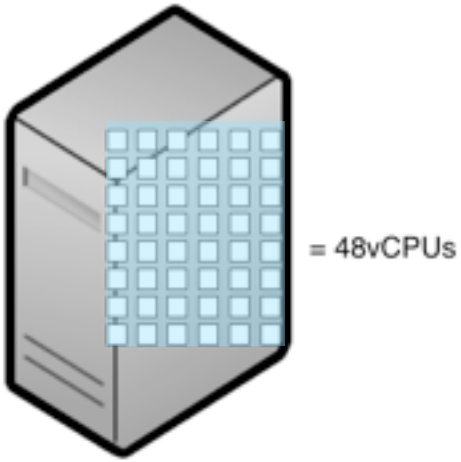
What is a Private Cloud?

- Elastic or Traditional Virtualization (hint: Pick One!)
- Multi-tenant (Multi-Application)
- Size
- Private API endpoints
- Limited inbound connectivity
- Customized for specific workloads

Build with the End in Mind

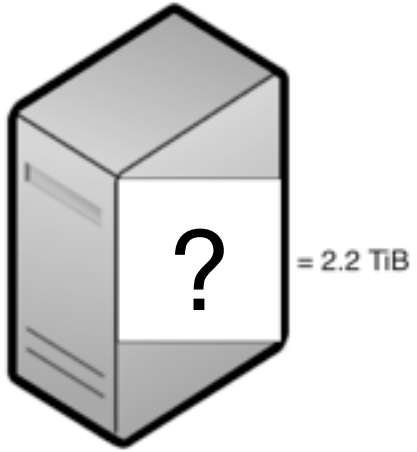
of instances on this physical host?

vCPU



Ram

Disk



Flavor: m1.tiny

Ram: 512MiB

vCPU: 1

Disk: 0 (size of image)

Total instances:

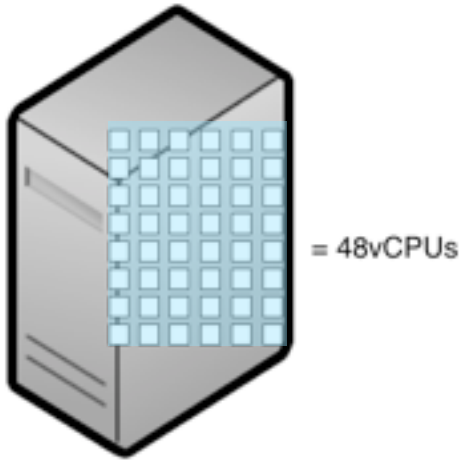
48

rackspace
the open cloud company

Build with the End in Mind

of instances on this physical host?

vCPU



Ram

Disk

Flavor: m1.medium

Ram: 4096MiB

vCPU: 2

Disk: 10GiB + 40GiB

Total instances:

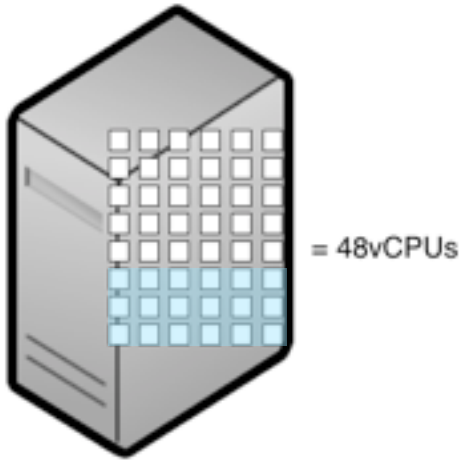
24

rackspace
the open cloud company

Build with the End in Mind

of instances on this physical host?

vCPU



Ram

Disk

Flavor: ram heavy

Ram: 64GiB

vCPU: 8

Disk: 100GiB

Total instances:

2

rackspace
the open cloud company

Capacity Thoughts

Don’t allow disk size of 0

Public Cloud: Limit flavors to image size

Private Cloud: Can add flavors for specific application profiles

Network Utilization

Capacity Thoughts

I can always add more machines....

But you can't change or modify the fixed network once instances are running

Capacity Thoughts

I can always add more machines....

But you can't change or modify the fixed network once instances are running

(resource / resource in smallest flavor) = maximum # of instances per machine

Fixed Network - Double or quadruple number of expected instances

Build with the End in Mind

nova-network

2 networks (3 w/ floating IPs)

1) Host Network

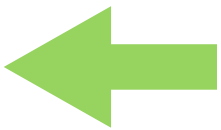
2) Fixed Network

3) Floating network

Build with the End in Mind

nova-network

2 networks (3 w/ floating IPs)

1) Host Network 

2) Fixed Network

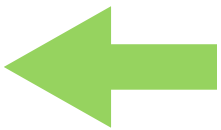
3) Floating network

Build with the End in Mind

nova-network

2 networks (3 w/ floating IPs)

1) Host Network



Easy to add physical nodes and/or networks

2) Fixed Network

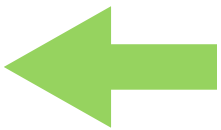
3) Floating network

Build with the End in Mind

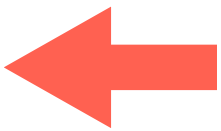
nova-network

2 networks (3 w/ floating IPs)

1) Host Network



2) Fixed Network



3) Floating network

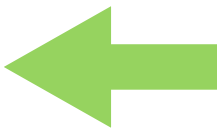
Easy to add physical nodes and/or networks

Build with the End in Mind

nova-network

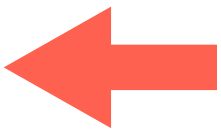
2 networks (3 w/ floating IPs)

1) Host Network



Easy to add physical nodes and/or networks

2) Fixed Network



Don't try to change the fixed network once in production

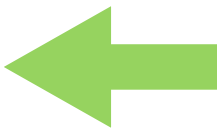
3) Floating network

Build with the End in Mind

nova-network

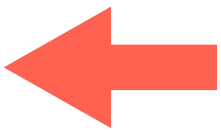
2 networks (3 w/ floating IPs)

1) Host Network



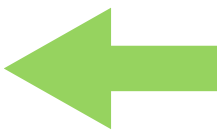
Easy to add physical nodes and/or networks

2) Fixed Network



Don't try to change the fixed network once in production

3) Floating network

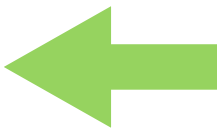


Build with the End in Mind

nova-network

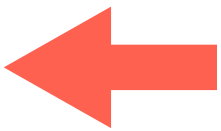
2 networks (3 w/ floating IPs)

1) Host Network



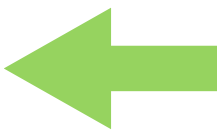
Easy to add physical nodes and/or networks

2) Fixed Network



Don't try to change the fixed network once in production

3) Floating network



Easy to add additional floating networks

Images and Storage

Images

Drivers: virtio vhost_net xen ...
Format: raw qcow2 vdi ...
Container: bare ami bundle ...
Cloud tech: cloud-init none
Partitioning: fixed dynamic

Building images talk tomorrow @1:50 in room C123

Images and Storage

- Glance
  - File backed (local)
  - Alternatives: Swift, CloudFiles, NFS (locally mounted)
  - Snapshots?
  - qcow2 format: “snapshot_image_format=qcow2“

Images and Storage

Glance Performance

- Network Throughput: 1Gbps = 125MBps (expect ~112MB/s)
- RAID 5 - large sequential read/writes
- Disk Bandwidth > IOPS
- Improve cache - Reduce # of images

Images and Storage

Glance Performance

- Network Throughput: 1Gbps = 125MBps (expect ~112MB/s)
- RAID 5 - large sequential read/writes
- Disk Bandwidth > IOPS
- Improve cache - Reduce # of images

Image Size	Not Cached	Cached
1.4GB	20secs	1sec
16.4GB	2min 21secs	1sec

Images and Storage

Glance Performance

- Network Throughput: 1Gbps = 125MBps (expect ~112MB/s)
- RAID 5 - large sequential read/writes
- Disk Bandwidth > IOPS
- Improve cache - Reduce # of images

Image Size	Not Cached	Cached
1.4GB	20secs	1sec
16.4GB	2min 21secs	1sec

*times from "creating image" to "qemu-img create"

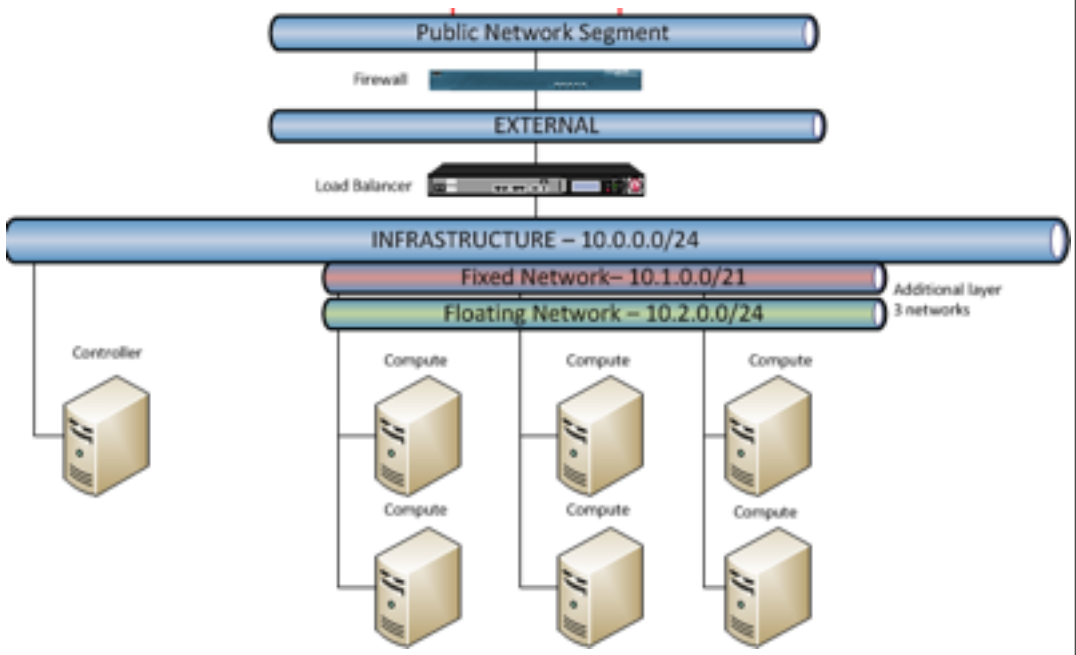
Images and Storage

- Storage
  - 4 focus points: Glance, Compute (local), Cinder, Swift
  - Glance - Space, Sequential read/write
  - Compute - random IO
  - Cinder - Performance, Density
  - Swift - JBOD, density

Architecture Examples and Thoughts

1 - 20 physical servers

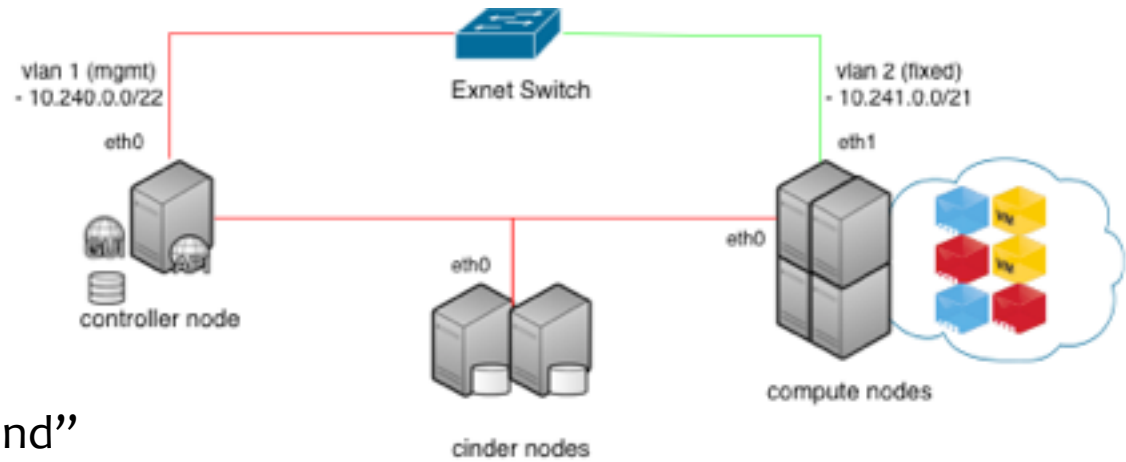
- Single controller
- Single API
- Single network (1Gbps or 2Gbps)



Architecture Examples and Thoughts

20-100 servers

- HA controllers and load balanced APIs
- Swift/CloudFiles for Glance
- Availability Zones
- Consider “frontend” and “backend” networks
- Compute Node metrics - Dedicated Machine(s)

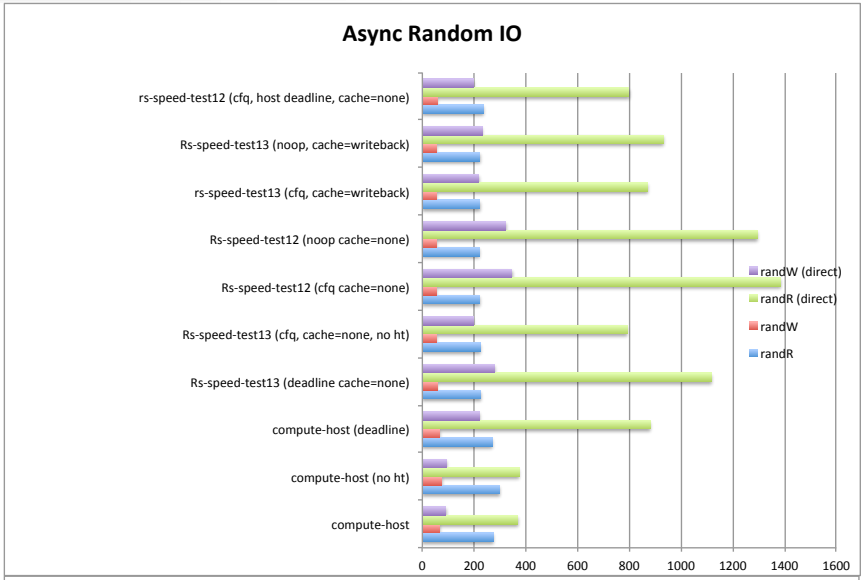


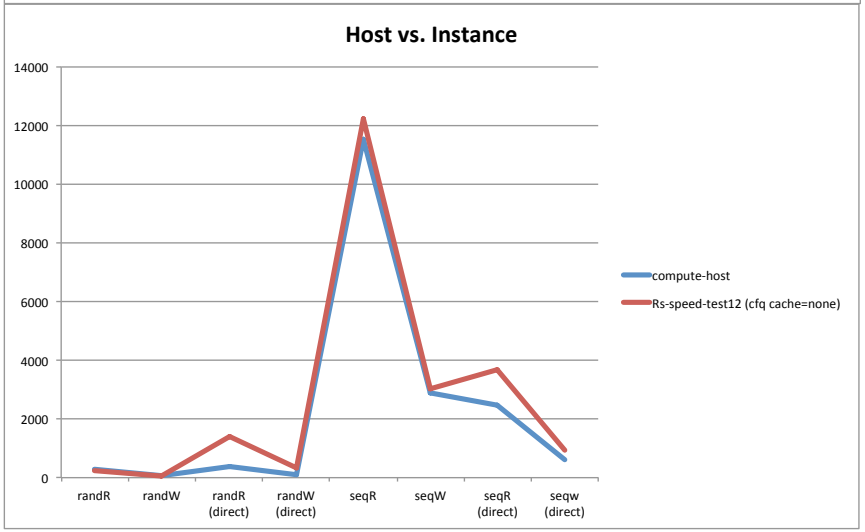
Performance Considerations and Bottlenecks

- IO
  - Random IO
  - Reduce IO per instance
  - Block Storage!
  - Review Hypervisor best practices

Performance Considerations and Bottlenecks

- IO
  - Random IO
  - Reduce IO per instance
  - Block Storage!
  - Review Hypervisor best practices





rackspace
the open cloud company

Lessons Learned

- Lessons learned
  - Floating IPs MUST be associated with the “public_interface”
  - Each piece of OpenStack has its own architecture
  - Folsom is stable
  - Migration (live, block) works but scenarios exist where it doesn’t
  - OpenStack is still changing often, keep up to date with current state of the projects
  - Don’t do Heterogeneous Nodes

Operational Updates

- nova hypervisor* calls
- Image types in Glance
- policy.json
- Coming in Grizzly:
  - Cells
  - Quantum
  - Better LDAP/AD support

But....

- But this is a design summit also
  - Open to discussions/thoughts/ questions

THANK YOU

Rackspace is hiring

www.rackertalent.com

RACKSPACE® HOSTING | 5000 WALZEM ROAD | SAN ANTONIO, TX 78218
US SALES: 1-800-961-2888 | **US SUPPORT:** 1-800-961-4454 | WWW.**RACKSPACE**.COM