

When One Cloud is Not Enough

**An Overview of Sites, Regions, Edges,
Distributed Clouds, and More**

Who are we?

Curtis Collicutt



OpenStack Architect

OpenStack Operators
Telecom/NFV Working Group
Chair

<https://wiki.openstack.org/wiki/Ops-telecom-nfv>

INTERdynamix

Chaoyi Huang



OpenStack Tricircle PTL:

<https://wiki.openstack.org/wiki/Tricircle>

OPNFV Multisite PTL:

<https://wiki.opnfv.org/display/multisite/Multisite>



Adrien Lebre



OpenStack Fog/Edge/
Massively Distributed Clouds
WG Chair

https://wiki.openstack.org/wiki/Fog_Edge_Massively_Distributed_Clouds

Discovery Initiative Chair
<http://beyondtheclouds.github.io>



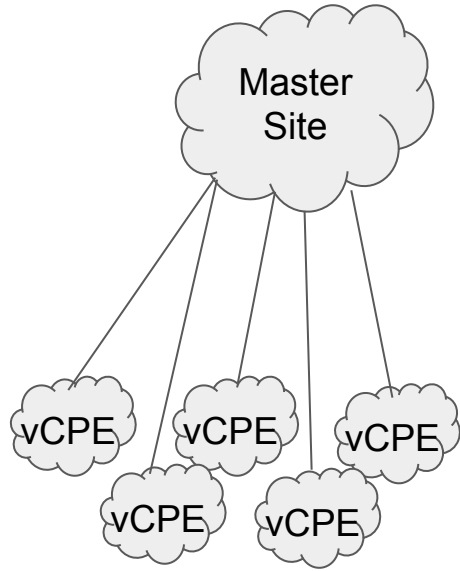
Agenda

- Context
 - NFV working group
 - Fog/Edge/Massively Distributed Clouds working group
 - Common interests/challenges: multi-site/multi-location deployments.
- (Exhaustive) overview of available solutions and building blocks
 - Try to clarify the eco-system
- Focus On OpenNFV Side (MultiSite use-case, TriCircle and Kingbird)
- Focus on the Fog/Edge/Massively Distributed Clouds WG
- Conclusion and takeaway message

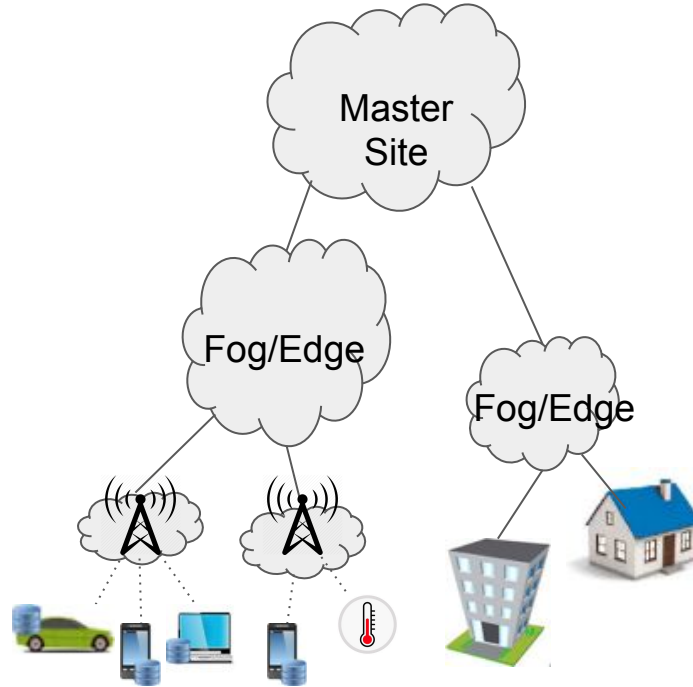
Federated clouds
MANO
Fog/Edge/Massively Distributed WG
vCPE
OpenBaton
CloudLets
NFV
Failure Domains
Distributed clouds
Availability Zones
Hyper-distributed clouds
“Cascading” OpenStack
ONAP
Multi-site
5G
Multi-region
LDAP
Cells V1
Segregations
Cells V2
CloudRAN
Edge clouds
Latency aware applications
NeCTAR
Internet of Things
Public Clouds
Mobile Computing
CERN
Fog Computing
Micro/Nano DCs

A lot of words ...

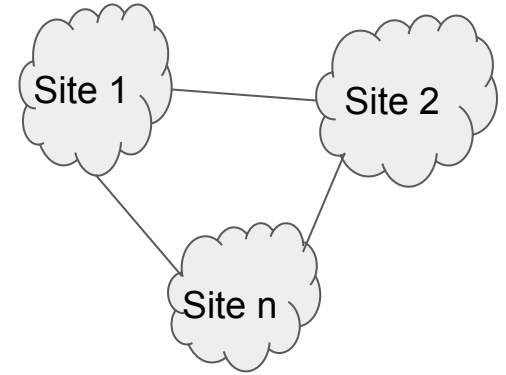
NFV, Fog, Distributed Clouds Use-cases



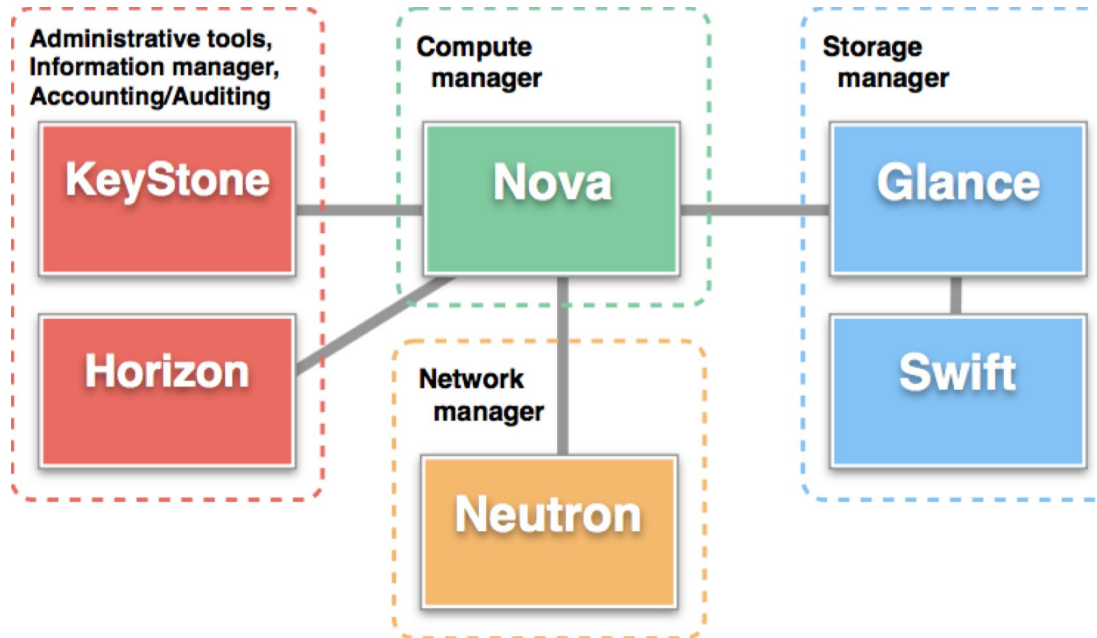
NFV/SDN
(large scale but provisioning
frequency rather low)



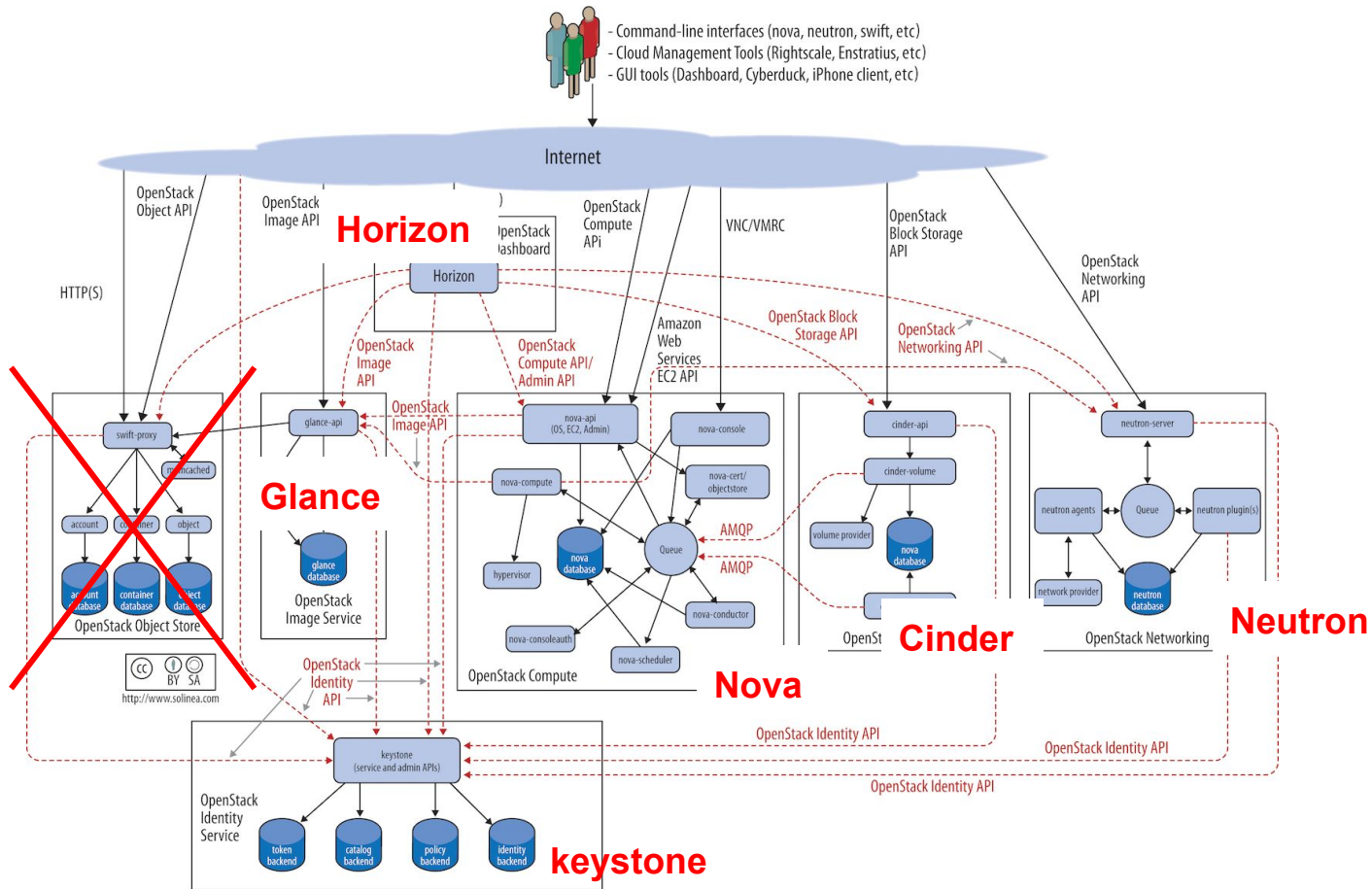
Fog/Edge Computing
(Latency aware / smart* / IoT Apps)



Academic cloud federation
(NeCTAR /EGI Federated Cloud)



How deploying OpenStack main services to operate such infrastructures?



How deploying OpenStack main services to operate such infrastructures?

Potential solutions for the three use-cases

- OpenStack “Core” Designs
 - Service nodes are deployed within a master site and compute nodes are deployed remotely
 - Segregation tools
 - Host aggregates/availability zones (not discussed in this presentation)
 - Nova Cells CERN and NECTAR approach (Cells V1 based)
 - Multi region - Shared Keystone
 - Completely separate clouds (no coordination/synchronization between each cloud)
 - Authentication/Authorization/Identity Solutions
 - Keystone Federation - Example: EGI cloud compute, Barcelona presentation
 - Shared authentication backend - Such as LDAP or Active Directory, oauth, x509, kerberos, etc (not discussed in depth)
- External from the OpenStack ecosystem
 - Kingbird
 - ETSI MANO layer (not discussed in this presentation, <https://www.onap.org/>)
 - Brokering (fogbow, See Barcelona presentation)

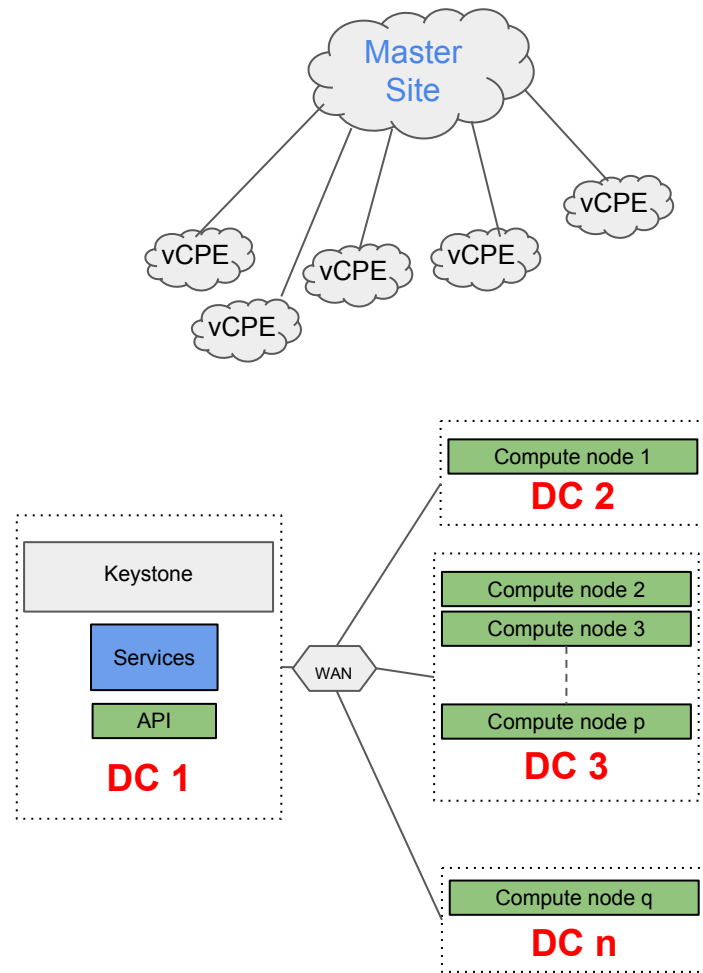
Definitions

- **Data Center:** one facility that delivers compute/storage/network resources in one location (nano/micro DCs: a few servers / mega DCs: thousands of servers)
- **WAN/DC Interconnect:** A WAN link connecting data centers
- **Multi-site:** More than one geographical location interconnected through WAN links or dedicated lines.
- **Fog/Edge Computing:** a multi-site like infrastructure but resources are deployed through micro/nano DCs deployed at the edge of the backbone (a Fog/Edge infrastructure is composed of a significant numbers of sites)
- **Services:** all services (nova, neutron, glance, cinder... but keystone)
 - **Services' nodes:** nodes that are hosting services
 - **Services' DB:** all DBs related to the services
 - **Services' API:** API related to the services
- **Compute nodes:** nodes that are hosting VMs.



OpenStack WANWide

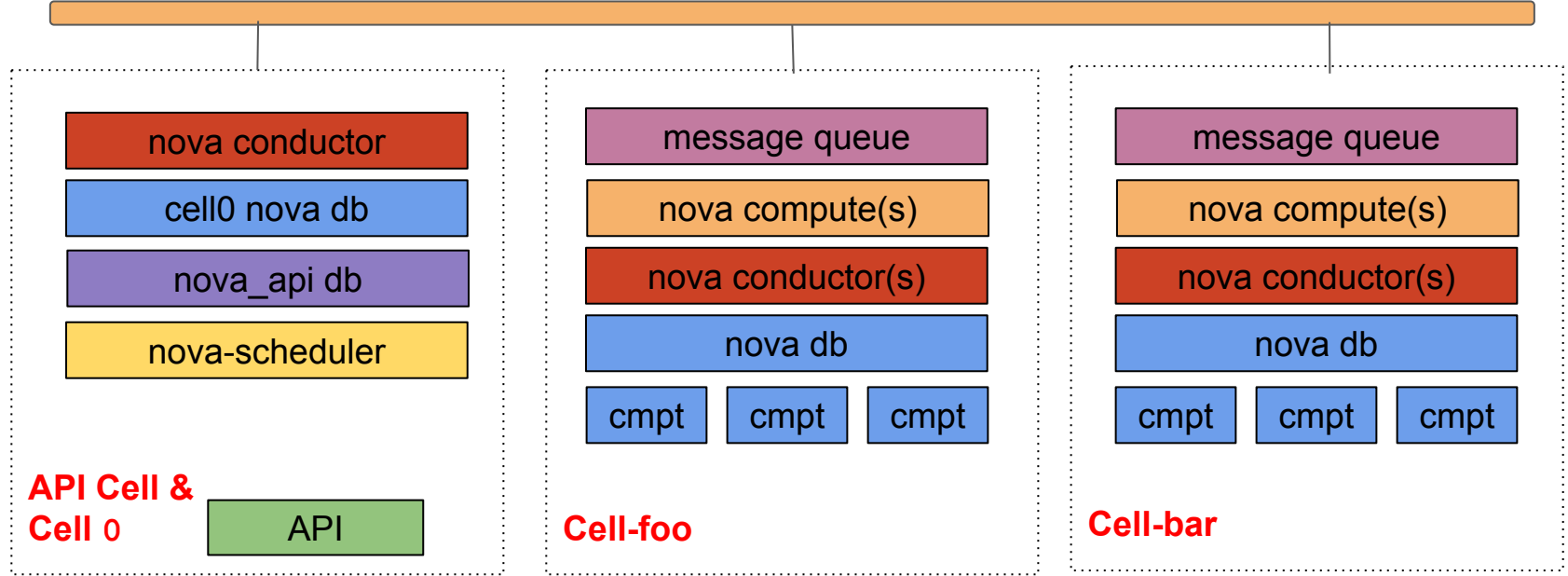
- A Single OpenStack to operate remote compute resources deployed at the edge
 - All control services are deployed into the master.
 - The RabbitMQ bus is deployed across all locations (i.e., through each server composing the infrastructure)
- Possibility to define availability zone/host aggregate
- Pros: “simple”
- Cons:
 - Network impact (latency/bandwidth)?
 - Neutron L3 (need to use DVR?)
 - Cinder (remote attached volume WANwide)?
 - High availability?
 - Security management for RPC message and port?
 - Scalability?



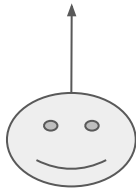
OpenStack Nova Cells

- A strategy for scaling Nova
- Each cell will have a message queue and database for nova
 - not the `nova_api` db but `nova` db
- Separate failure domains for groups of compute nodes (eg. rack)
- For Ocata+ default is to have a v2 cell in every deployment
 - V2 has created the requirement for the `nova_api` database - contains global information.
 - A central DB (SearchLight) is being integrated to support cross cells query.
- Optional grouping mechanism
- May help for testing, ie. can add a cell, test it, then provide it to users

Nova Cells*



API Endpoint



* Borrowed from <https://www.openstack.org/videos/video/nova-cells-v2-whats-going-on>

OpenStack Nova Cells - Pros and Cons

- Pros

- Reduce size of failure domains
- Help to scale Nova
- Quite powerful when used in conjunction with other scaling methods/technology

- Cons

- V2 is still relatively new, not complete, but is being worked on heavily. Because it's default now, will get a lot of use.
- Historically (V1) not as many users, operationally required unique in-house expertise. Typically used by large multi-rack deployments. V2 may change this.
- May be best suited with particular network models (such as routed provider networks*)
- Only for Nova

- Notes

- Used by large scale clouds such as CERN and NECTAR and others

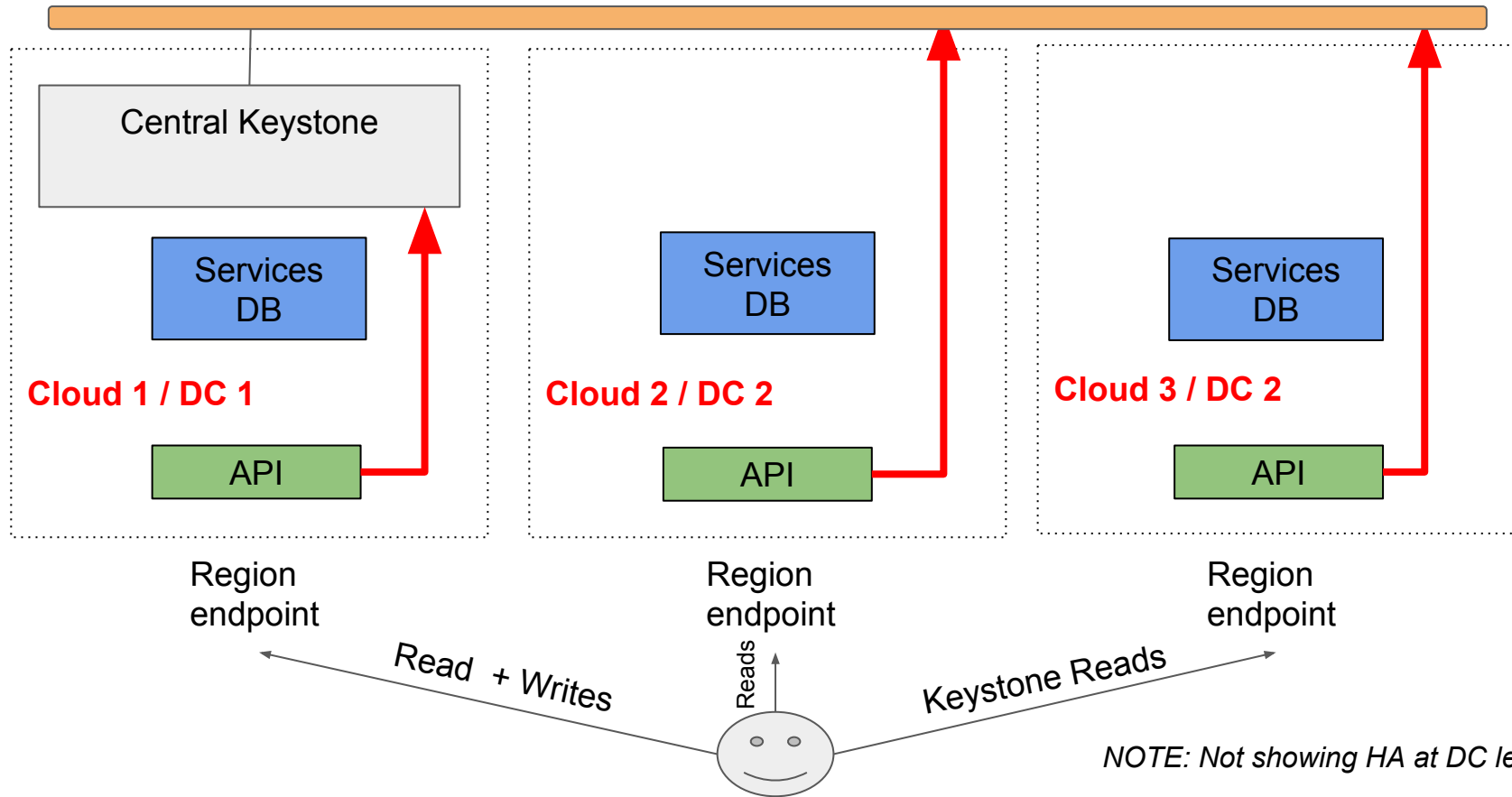
* <https://docs.openstack.org/newton/networking-guide/config-routed-networks.html>

OpenStack Regions

- Multiple complete OpenStack deployments (except Keystone) - each cloud has endpoints/control plane for everything.
 - Usually one cloud in each datacenter.
- A centralized or shared Keystone DB will be available to each cloud
 - Horizon and Glance databases could potentially be shared as well
- Some limitations -- Keystone authentication and authorization could be shared, but typically not quotas, ssh keys, images, etc. Only authentication and authorization.

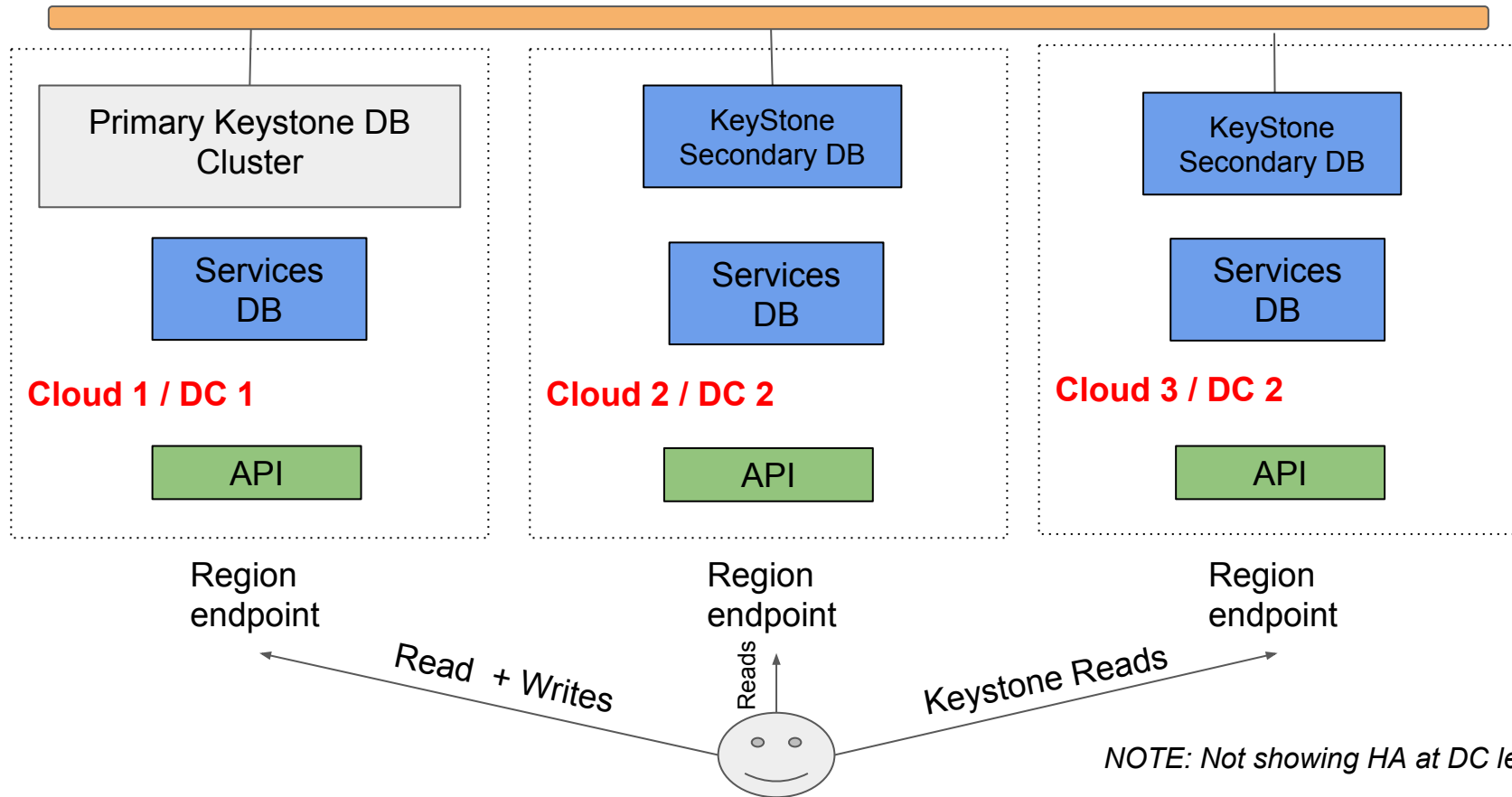
Relatively low latency, Cross DC
secure link

Central Keystone, No Replication



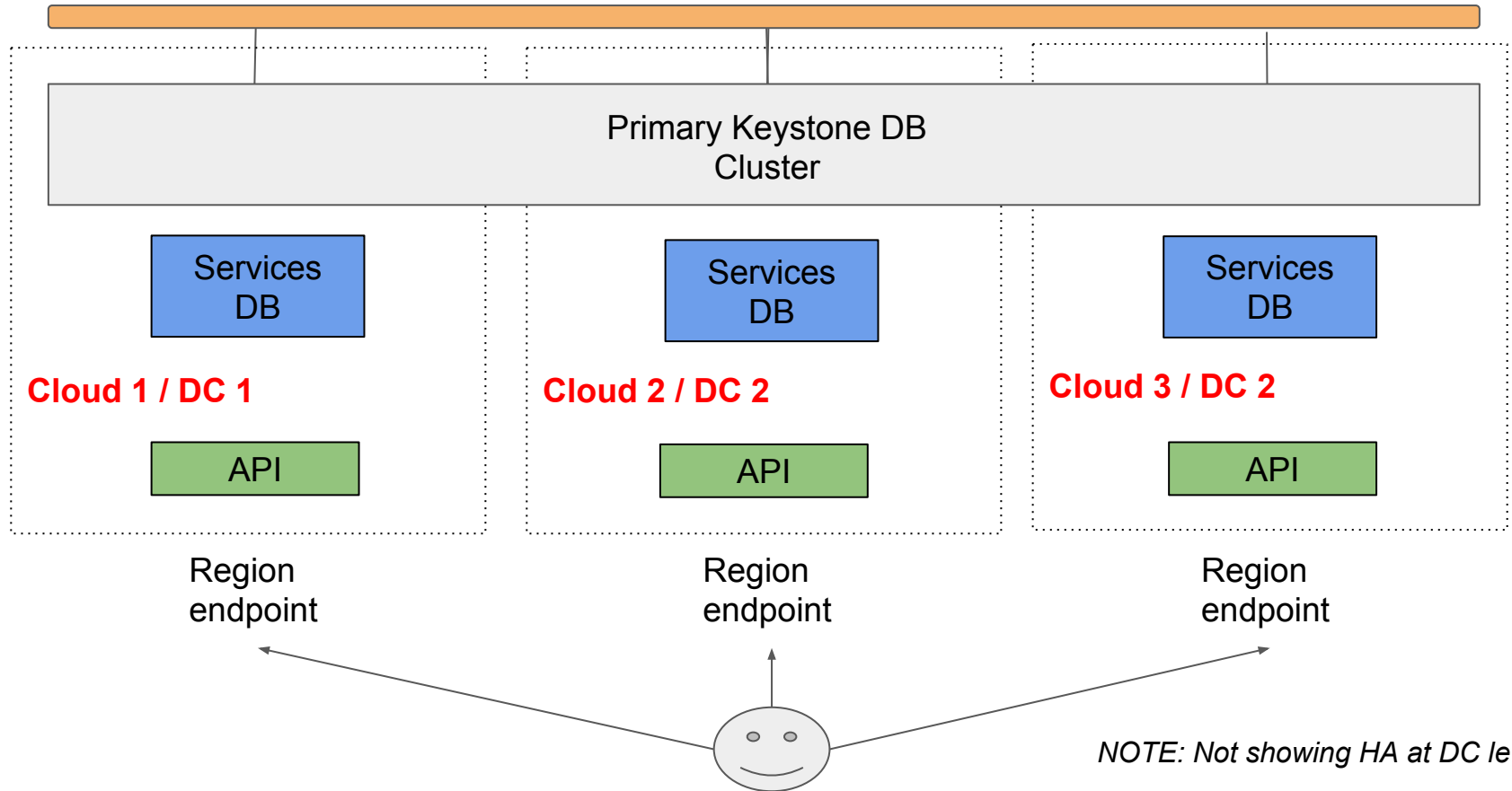
Relatively low latency, Cross DC
secure link

Primary Keystone, Async Replication to Other Regions



Relatively low latency, Cross DC
secure link

Shared Synchronous Keystone



OpenStack Regions - Pros and Cons

- Pros

- Shared authentication and authorization across multiple clouds
- Looks like what public clouds typically look like (regions, endpoints, etc)

- Cons

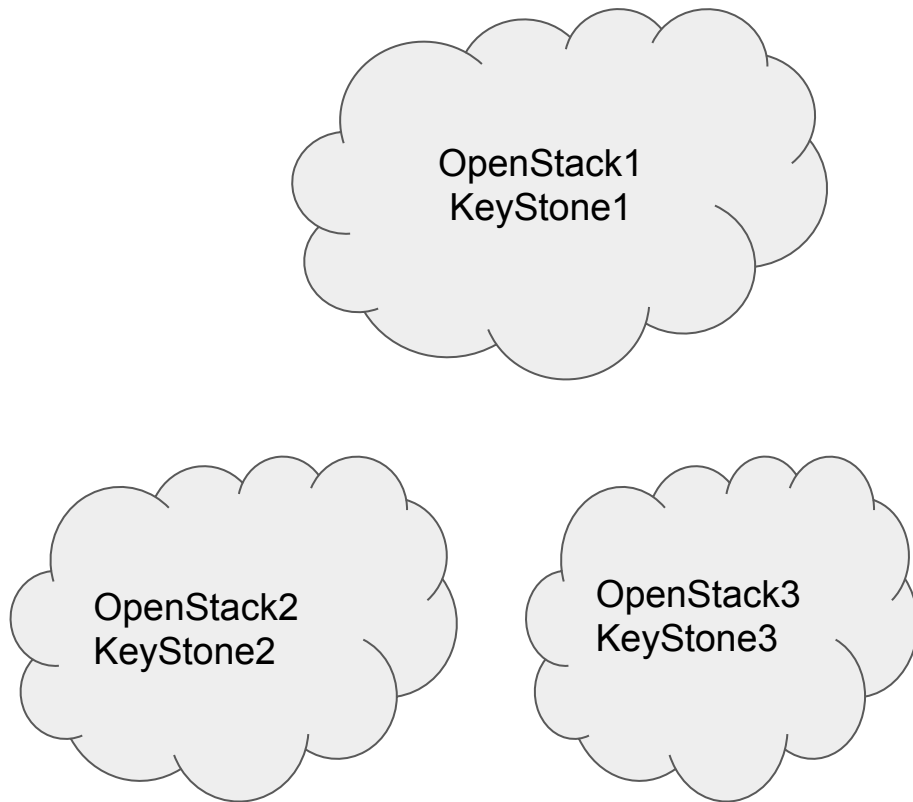
- Decisions to be made on architecture (central, async, sync...)
- Requires some kind of secure, usually private, link for database cluster to communicate over
- May be limited to the number of database cluster members you are willing to have: 2,3 clouds seems ok, but 10, 20...probably not? Using secondary model may help.
- Typically only manages authentication and authorization over multiple clouds. Does not help with things like quotas, networking, etc. Sometimes that is enough, sometimes it's not.
- Can make upgrading clouds more difficult
- Operational complexity

OpenStack Multi-Clouds

Each OpenStack is a separate cloud, even for KeyStone. Re-login when switching clouds.

Pros: Fully isolated failure domain

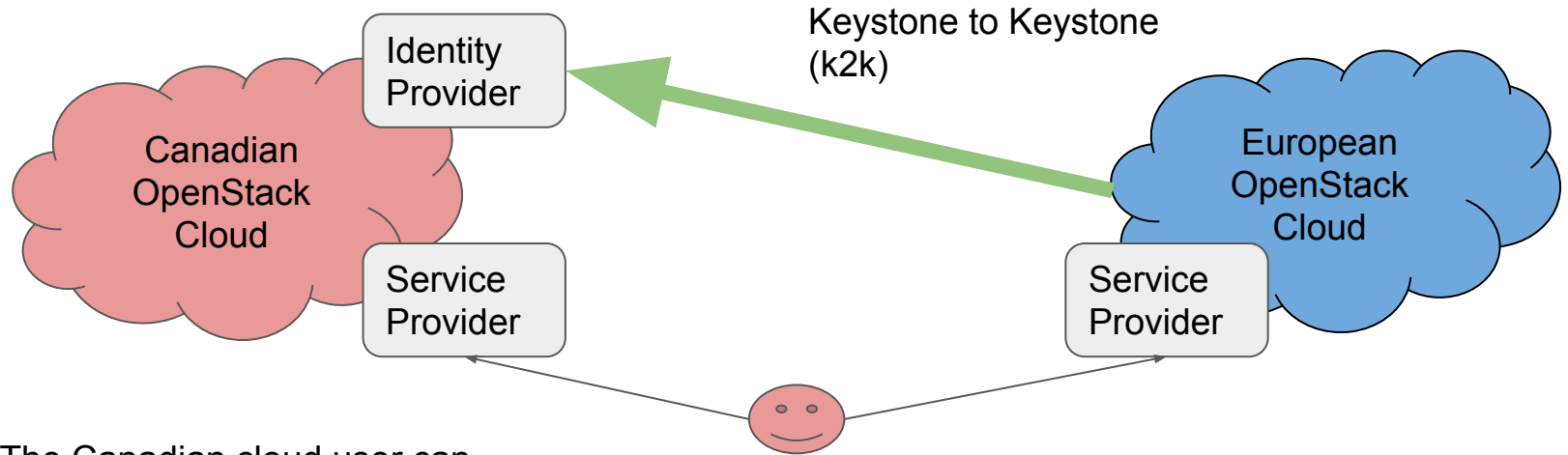
Cons: No global resource view and synchronization, has to be operated one by one at the application level, lack of tenant's network automation



OpenStack Federation

- The canonical use is to allow users of another, completely separate, organization's OpenStack cloud to utilize resources in your cloud (and potentially vice-versa)
 - Often a desirable feature academic or non-profit organizations who want to be able to share resources
 - Also could be useful for public clouds, for example one in Canada and one in Europe
 - Can also provide a centralized identity
- Essentially allows establishing trusts between service providers and identity providers
- Also can be used to help provide “single sign on” (SSO)

OpenStack Federation Example



The Canadian cloud user can access resources in the European cloud because the European clouds trusts the Canadian cloud as an IdP.

OpenStack Federation - Pros and Cons

- Pros

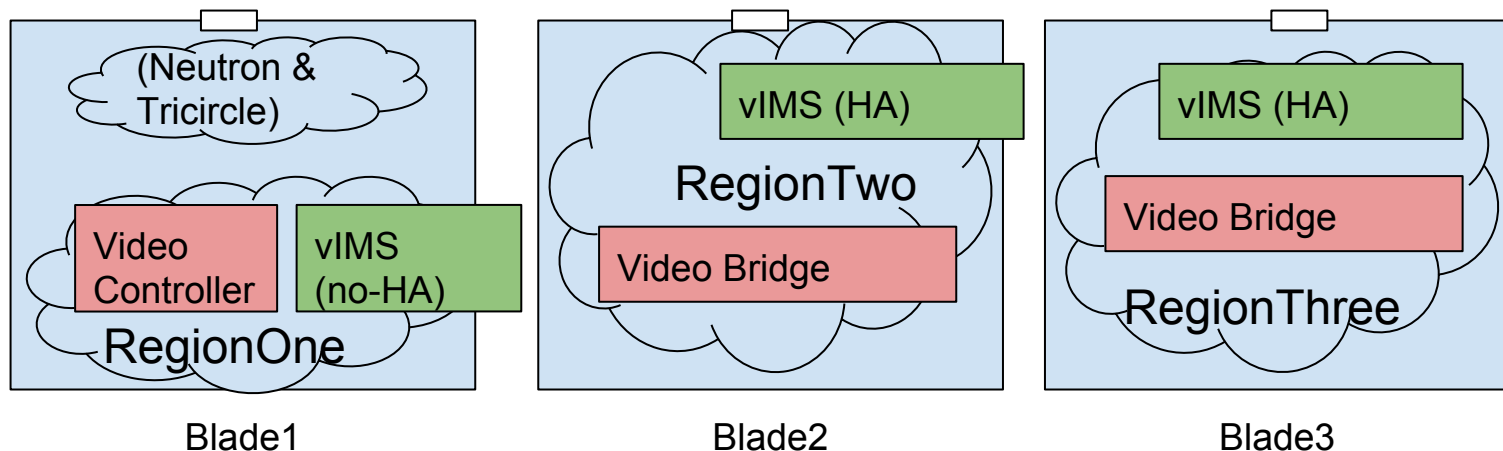
- Shared authentication without the need for a shared Keystone database
- Remove need to synchronize identity across multiple clouds (if not using shared keystone)
- Could be used as centralized identity for a single organization as well

- Cons

- Can involve complex configuration, requires a good understanding of how federation works, as well as the chosen authentication method
- Mapping to groups for authorization also requires careful consideration
- Still work to do, such as enabling roles from group membership*
- May be additional work in cleaning up resources on user removal

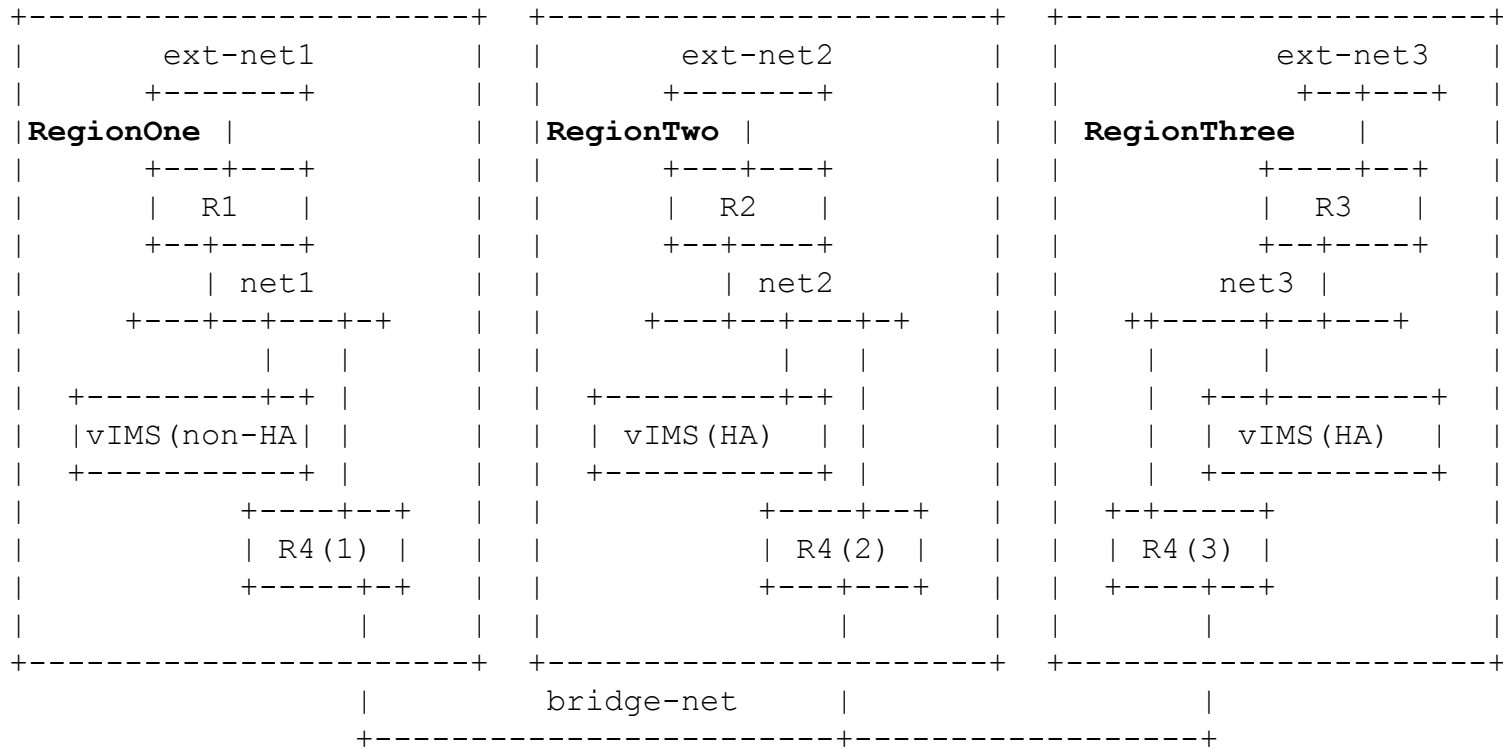
* <http://markmail.org/message/up52vwlgno3ca3tv>

OPNFV Multisite - VNF High availability across multi-region OpenStack clouds



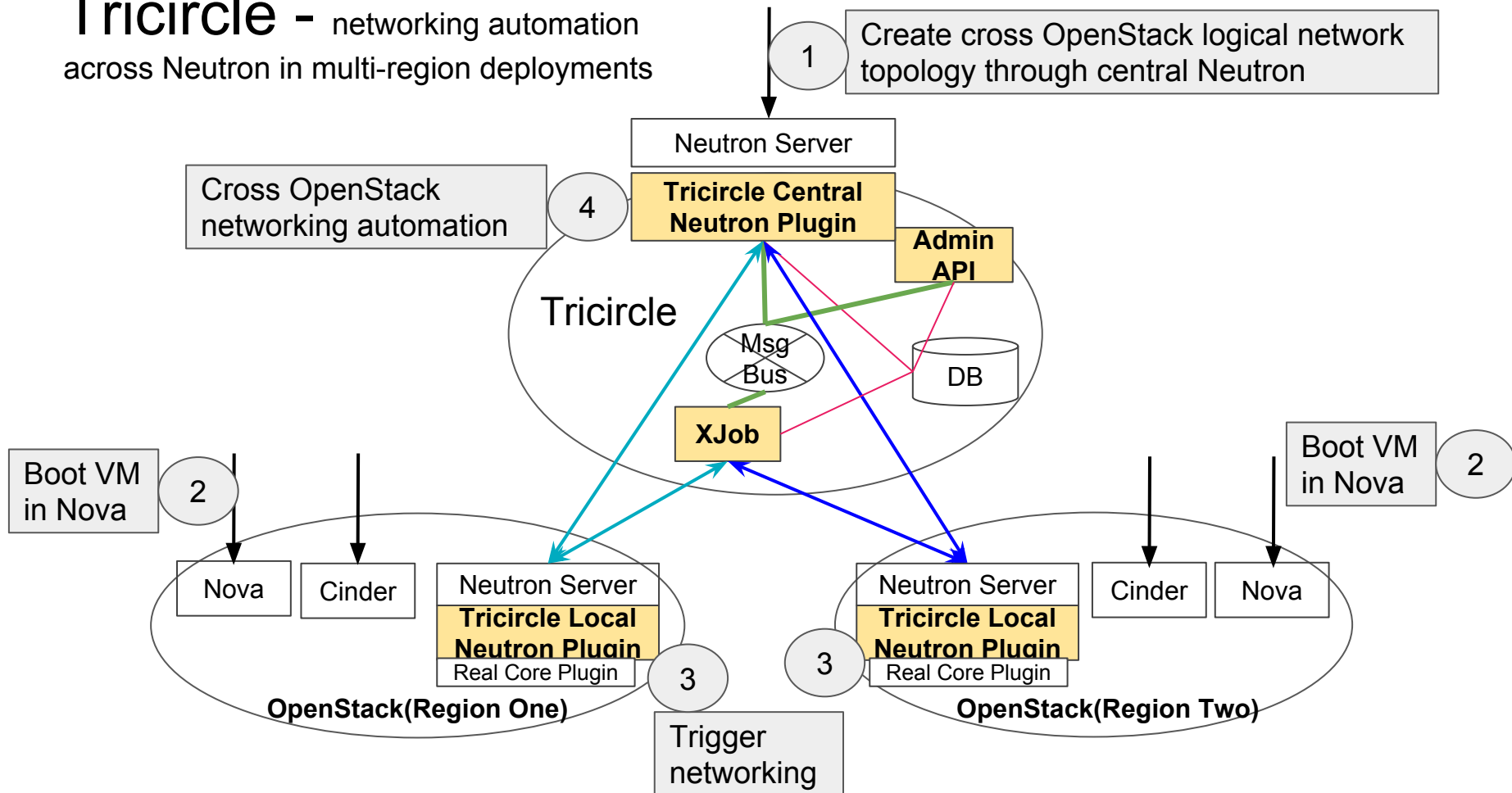
- Demo in OPNFV Beijing Summit, Jun. 2017.
- Multi-Region deployment, three OpenStack regions, running **Newton** version.
- Shared KeyStone, KeyStone is deployed in RegionOne.
- **Tricircle and Central Neutron** are deployed as one VM in RegionOne, it's **Pike** version, and run as Central Region. Cross version running together works.
- **Environment ready, now APP on-boarding**

vIMS overlay networking in multi-region deployment



The networking automation will be done through Tricircle

Tricircle - networking automation across Neutron in multi-region deployments



Tricircle

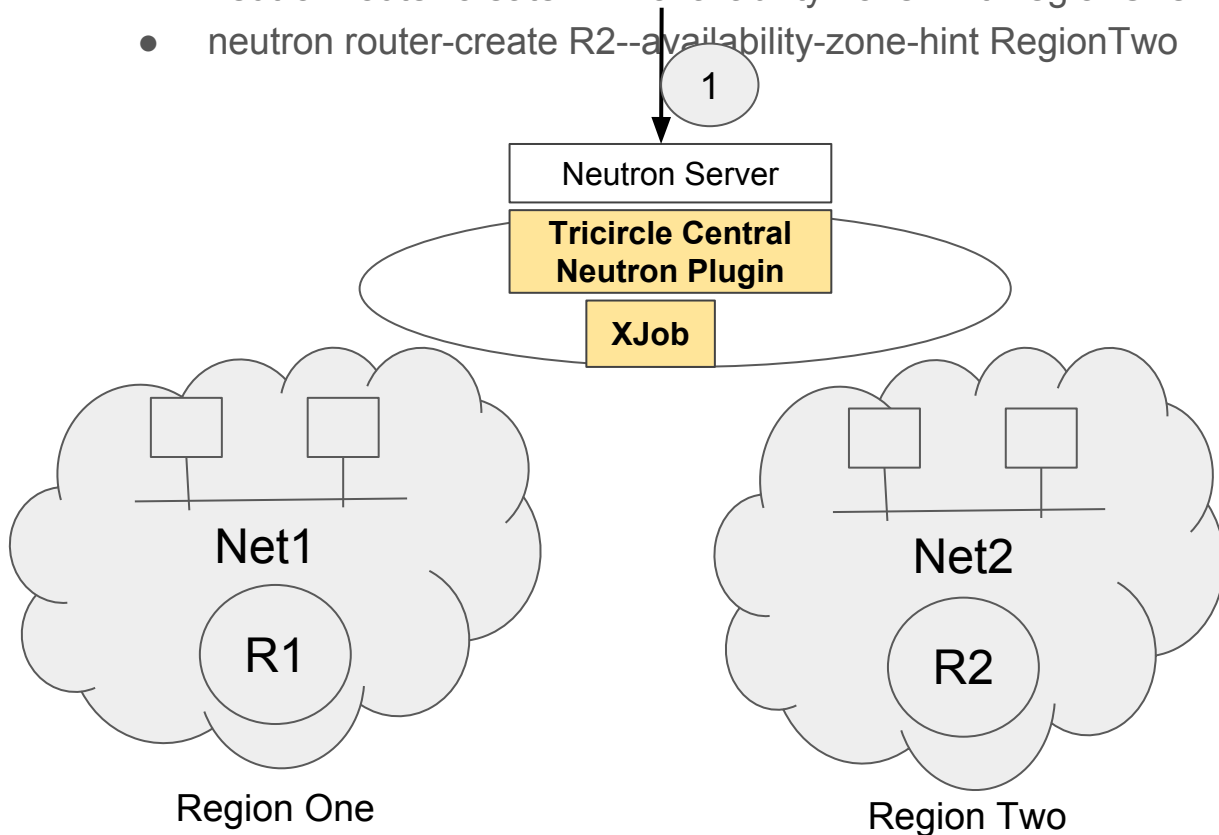
Local network

- A network will only reside in one OpenStack cloud.
- Network type could be VLAN, VxLAN, Flat

Local router

- A router will only reside in one OpenStack cloud

- neutron net-create Net1 --availability-zone-hint RegionOne
- neutron net-create Net1 --availability-zone-hint RegionTwo
- neutron router-create R1 --availability-zone-hint RegionOne
- neutron router-create R2 --availability-zone-hint RegionTwo



Tricricle

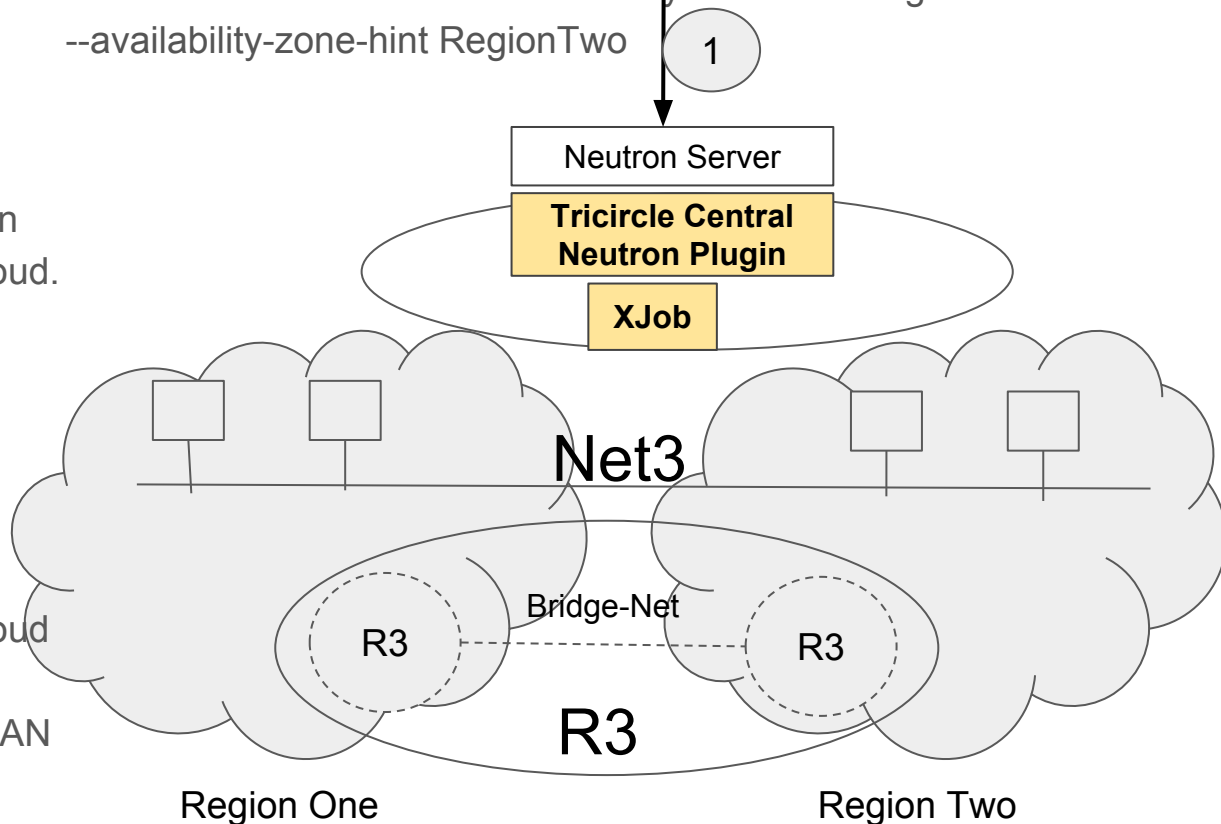
Cross OpenStack L2 network

- A network can be presented in more than one OpenStack cloud.
- Network type could be VLAN, VxLAN, Flat

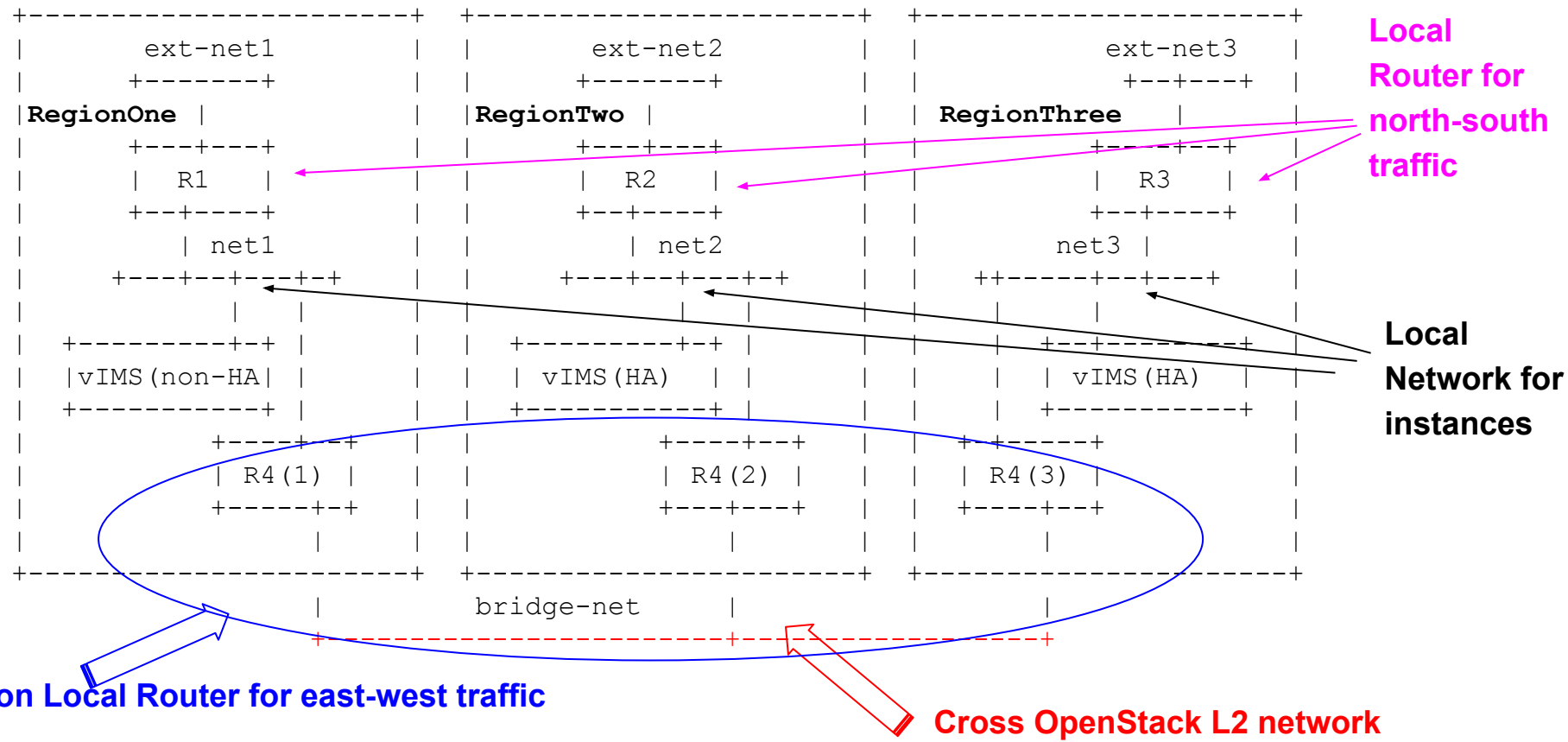
Non local router

- A router can be presented in more than one OpenStack cloud
- Inter-connected via bridge network, could be VLAN, VxLAN or Flat.

- neutron net-create Net3 --availability-zone-hint RegionOne
--availability-zone-hint RegionTwo
- neutron router-create R1 --availability-zone-hint RegionOne
--availability-zone-hint RegionTwo

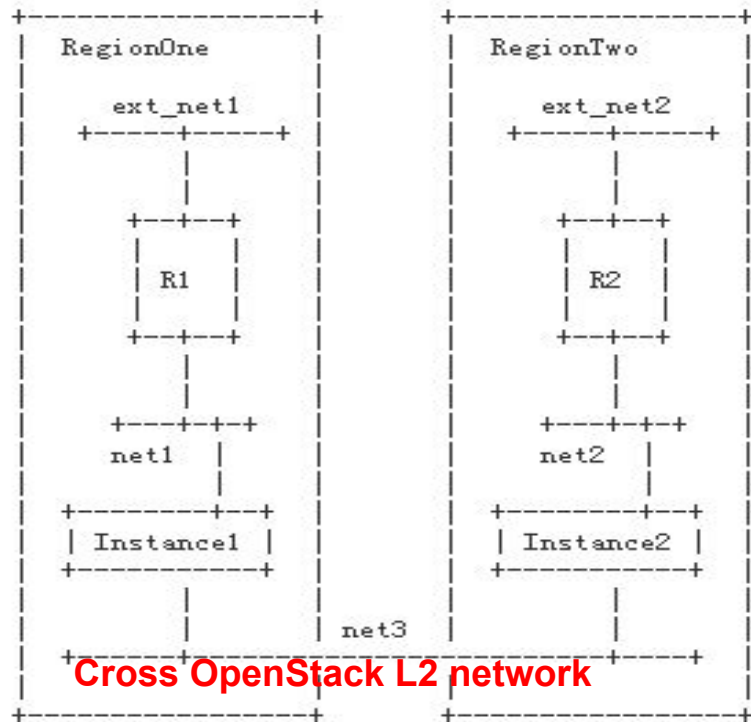


vIMS overlay networking in multi-region deployment

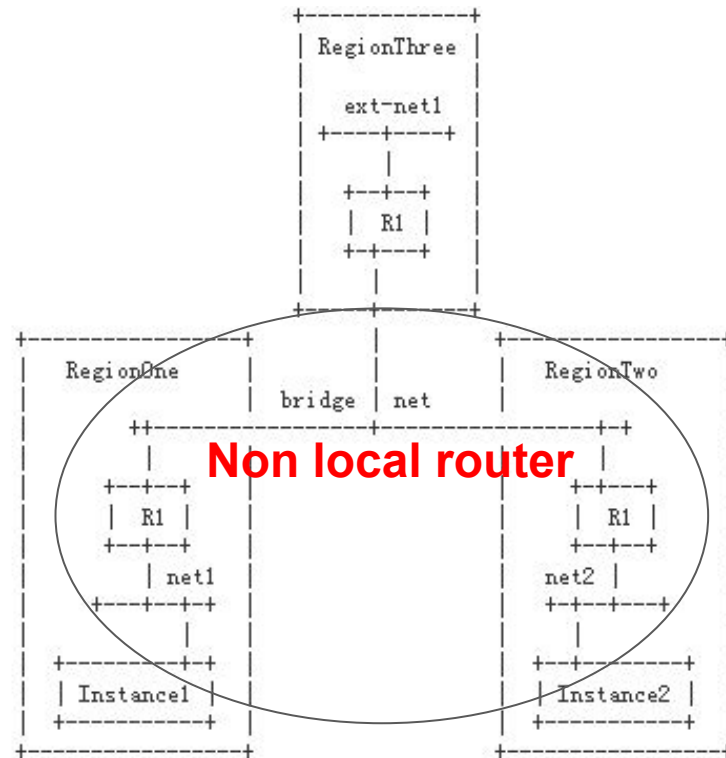


Locality North-South Traffic with East-West L3 networking enabled

Tricircle- other networking topology



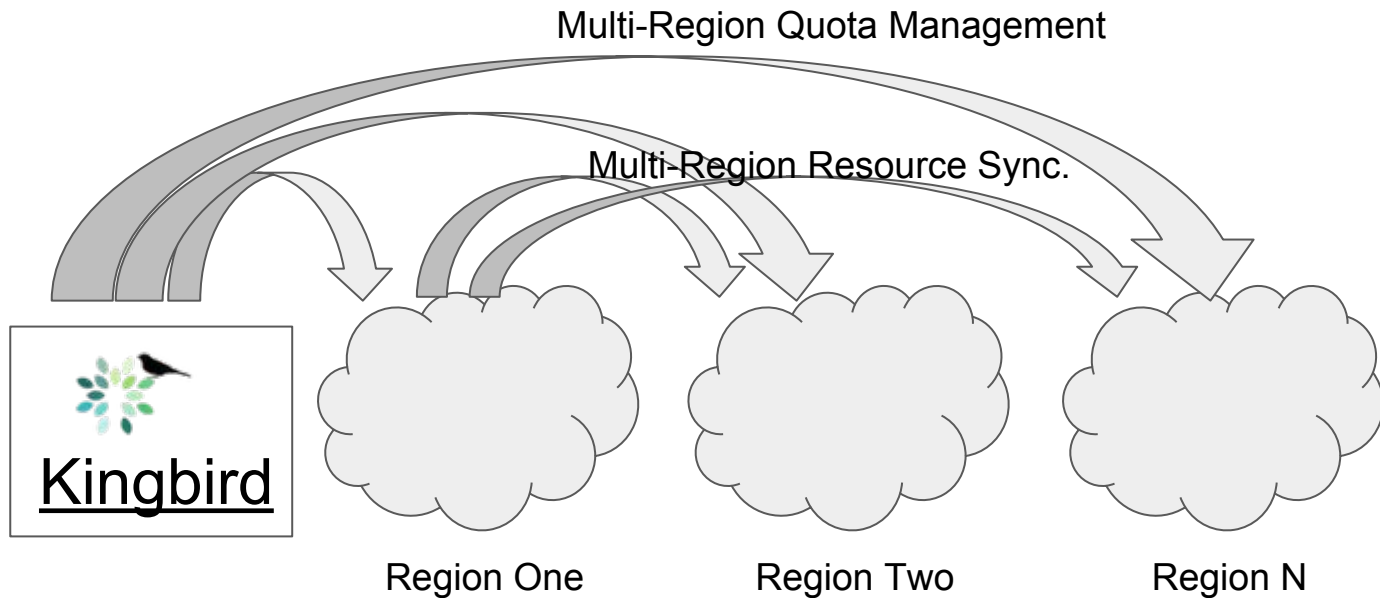
Instances plugged into cross
OpenStack L2 network



Centralized North-South Traffic with
East-West L3 networking enabled

Join us Tricircle on-boarding session on
Tuesday 4:40pm-5:25pm Hynes - MR 101

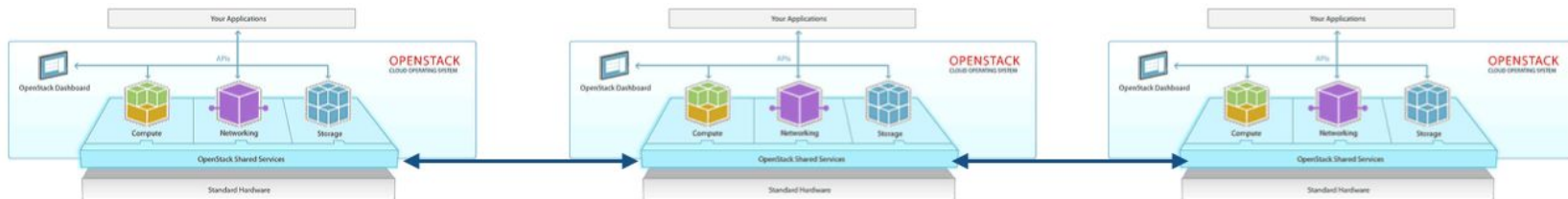
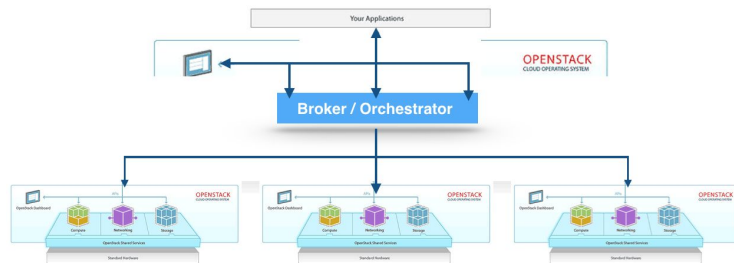
KingBird




Kingbird provides features like centralized quota management, synchronisation of ssh keys, images, flavors etc. across multi-regions, otherwise you have to do that region by region.

Fog/Edge/Massively Distributed Clouds WG

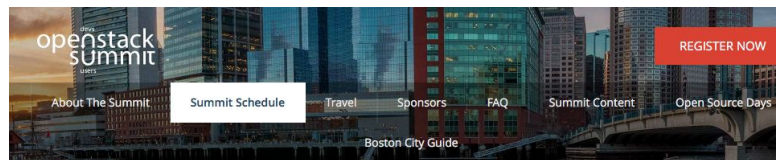
- Investigate how OpenStack can address Fog/Edge Computing use-cases
- Top/Down vs Bottom/Up approaches
 - Top/Down: add a substrate to pilot independent OpenStack Instances
 - **Bottom/Up: Investigate whether/how OpenStack core services can be cooperative by default using Self-* and P2P mechanisms**



Fog/Edge/Massively Distributed Clouds WG - cont.

- Ongoing action: study pros/cons and performance aspects of major deployment scenarios (scalability, network limitations, split brain...Glance, Cinder...)
- Experimental eNvironment for OpenStack 
 - Enos deploys OpenStack and targets reproducible experiments. It allows easy:
 - deployment of the system
 - customization of the system
 - benchmarking of the system
 - visualization of various metrics

See more on Wed. 2.40 - 3:20 am - Level 2, Ballroom A
Toward Fog/Edge and NFV deployments: Evaluating OpenStack
WanWide?




Event Details

[<< Go back](#)

Toward Fog, Edge, and NFV Deployments: Evaluating OpenStack WANwide

Telecom / NFV Operations

[Schedule](#) [Watch Later](#)

 Wednesday, May 10, 2:40pm-3:20pm

 [Hynes Convention Center - Level 3 - Ballroom A](#)

 Will be recorded

 Level: Intermediate

 Tags: [Public Clouds](#) [Telecom operator Architect Upstream Scientific](#)

In this talk, we present results of a study focusing on the impact of Wide Area Network communications on real OpenStack

Conclusion and Takeaways

- OpenStack has fairly rich, and complex, authentication options
- Some OpenStack components (eg. Nova) have improving segregation methods
 - Cells V2 in combination with things like Neutron “routed provider networks” start to look very powerful for some deployment models
- OpenStack Regions are useful for authentication/authorization, but probably only for a handful of regions, depending on your risk profile
 - NFV deployments may be 10,20,100,3000+ clouds for a single organization.
 - Traditional OpenStack regions will not work at that scale.
 - Perform large scale experiments in terms of regions (how many?) Massively Distributed WG/Performance Team.
 - Shared authentication is doable, but we are missing some other requirements.
 - How can we upgrade the code between distinct versions?

Conclusion and Takeaways Cont'd

- Centralized identity using federation may be more useful than currently realized
- Distributed clouds is an important use-case (NFV, Fog/Edge, Federation) but documentation has been (temporarily?) removed !
- Having a full HA control plane directly next to each set of compute nodes is not always going to be feasible (especially for vCPE or Edge use-case where remote resources -- aka servers -- should be limited)
- No matter what solution, we still have problems with things like keys, quotas...
 - Fog/Edge/Massively Distributed WG aims to identify such issues and propose countermeasures.
- MANO layer solutions are looking to “own” this space, is that what we as an ecosystem/community want to see happen?
- Ongoing effort: OpenStack Operators Telecom/NFV Team documentation project

Join us during our face to face session on Thursday
11:00 - 11:40 am - Hynes, MR202

When One Cloud is Not Enough

**An Overview of Sites, Regions, Edges,
Distributed Clouds, and More**

Thanks

Photo: Sy Clark
<https://flic.kr/p/2o6Djr>