Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules

OpenStack Networking : Neutron

11èmes Journées Informatique IN2P3/IRFU

CCIN2P3

Octobre 2018

cnrs

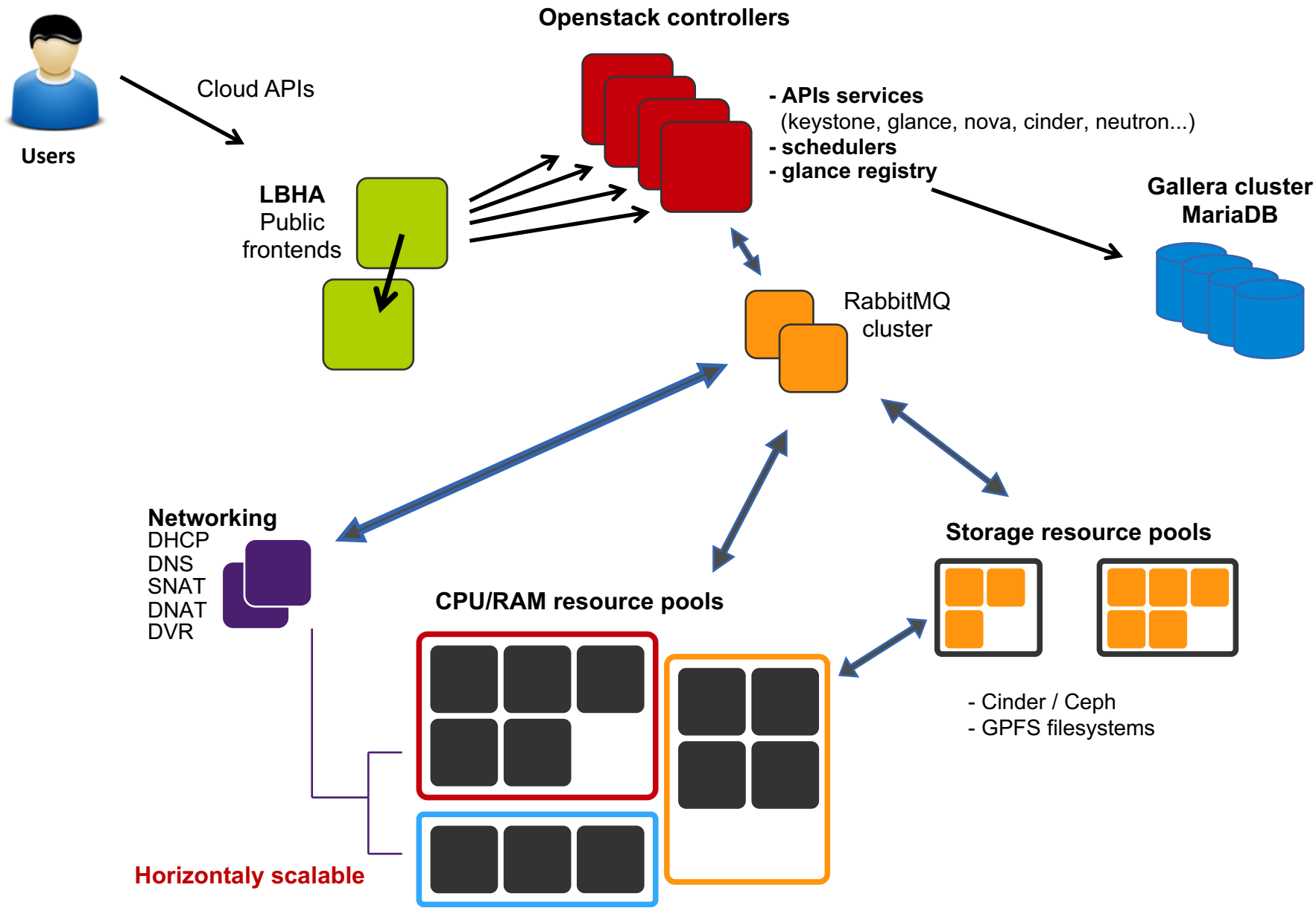
irfu

Cloud OpenStack au CC-IN2P3 : Introduction

- Plusieurs cas d'usage
  - R&D, Services, Hébergé, Computing.
- Différents niveaux d'exigences
  - Disponibilité, performance, stockage, accessibilité réseau..
- Clusters et Plateformes déployés
  - (*) Production, Pre-production, Testing.
- Déploiement et gestion des configurations
  - Foreman, Puppet.
- Environnement système et logiciel
  - CentOS 7, paquets RPM RedHat RDO.

(*) *Agile : Development, Testing, Staging and Production Environments*

Cloud OpenStack au CC-IN2P3 : Architecture



Neutron : Introduction

- Neutron est un projet OpenSource qui gère le réseau dans le Cloud OpenStack
- Fournit du NaaS (Network as a Service)
- Basé sur SDN (Software Defined Networking)
- Permet l'orchestration du réseau physique et virtuel
- Nécessite une base de données
  - Contient les réseaux, sous-réseaux, ports réseaux, groupes de sécurité..
- Nécessite un service de messaging queue
  - AMQP (Advanced Message Queuing Protocol).
  - Permet de recevoir les ordres et de les transmettre aux agents et à tous les services dont il a besoin.

NEUTRON

an OpenStack Community Project

Neutron : Historique

- Réseau OpenStack
  - A l'origine géré par « nova-network » (composant Nova).
  - Utilise LinuxBridge.
- Projet Quantum
  - Initié au début de la release Essex.
- Release Folsom (sept. 2012)
  - Quantum supporté et intégré dans OpenStack.
- Release Havana (oct. 2013)
  - Quantum devient Neutron.
- Release Newton (oct. 2016)
  - nova-network : Deprecated.
- Release Ocata (fev. 2017)
  - (*) nova-network : Non supporté ?

(*) *nova-network was deprecated in the OpenStack Newton release. In Ocata and future releases, you can start nova-network only with a cells v1 configuration. This is not a recommended configuration for deployment.*

Neutron : Généralités

- Permet de créer des réseaux autonomes par tenant, des switchs virtuels, routeurs virtuels, etc.
- Permet d'utiliser un contrôleur externe
  - Dragonflow (mode distribué), OpenDaylight , Contrail, Nuage.
  - VMWare (NSX), Cisco (ACI, VTS), etc.
- Offre la possibilité d'utiliser des plugins et leur délègue certaines actions
- Core plugins : IP et connectivité niveau 2 des VMs
  - Open vSwitch, Linux bridge.
- Service plugins : Services à la demande (en option)
  - NFV (Network Functions Virtualization)
    - FWaaS (Firewall as a Service).
    - VPNaaS (VPN as a Service) : Tunnels IPSec.
    - LBaaS (Load-Balancer as a service).

SDN - NFV

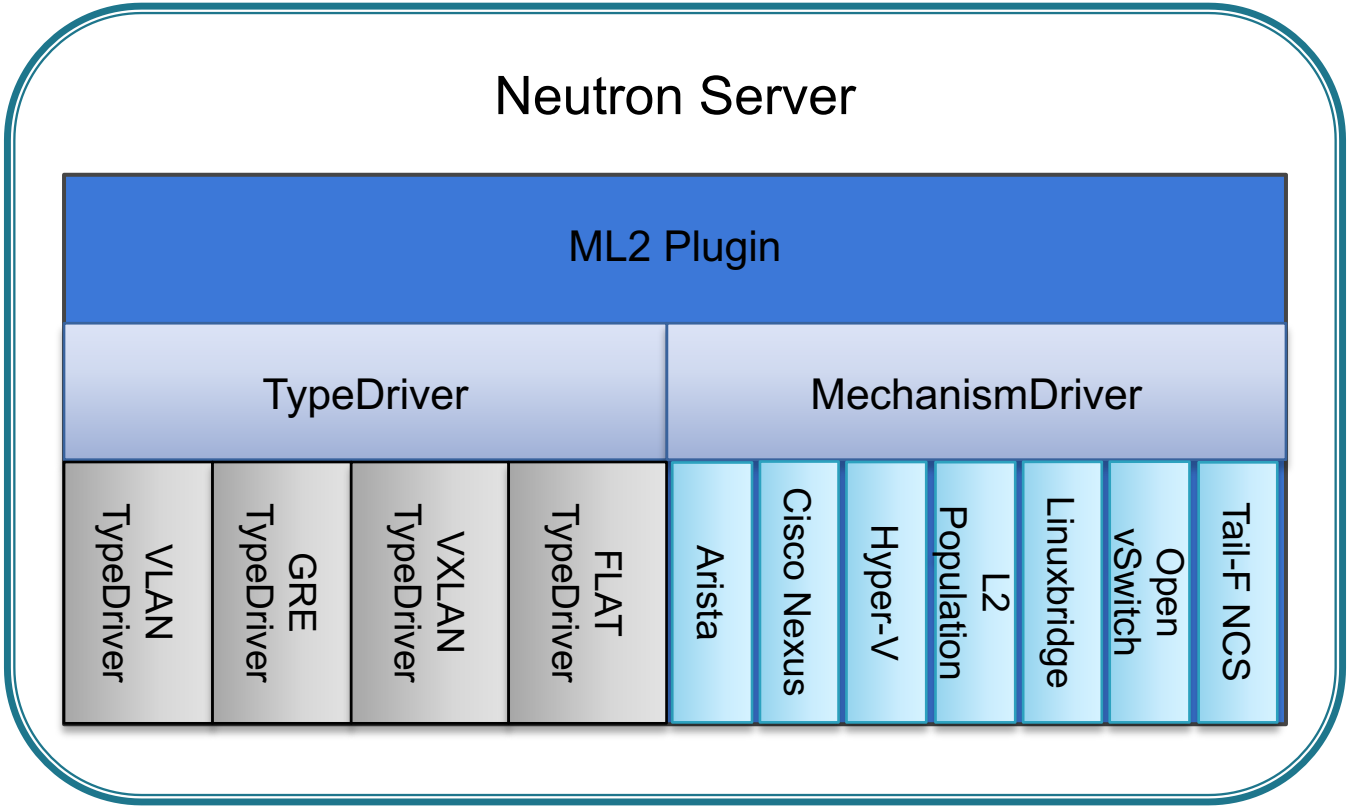
- (*) Le SDN est donc plus globalement reconnu aujourd'hui comme une architecture permettant d'ouvrir le réseau aux applications. Cela intègre les deux volets suivants :
  - permettre aux applications de programmer le réseau afin d'en accélérer le déploiement.
  - permettre au réseau de mieux identifier les applications transportées pour mieux les gérer (qualité de service, sécurité, ingénierie de trafic...).

- (*) NFV est une approche consistant à réaliser certaines fonctions réseau, traditionnellement effectuées sur du matériel dédié, sur des serveurs x86.

(*) *Le SDN pour les nuls, Jérôme Durand (Cisco Systems) JRES 2015 - Montpellier*

Neutron : Plugin ML2

- Neutron ne supporte qu'un seul plugin par déploiement
- Le plugin ML2 (Modular Layer 2)
  - Permet d'intégrer plusieurs mécaniques réseau dans un même cluster.
  - Fait l'interface entre un plugin réseau et neutron.

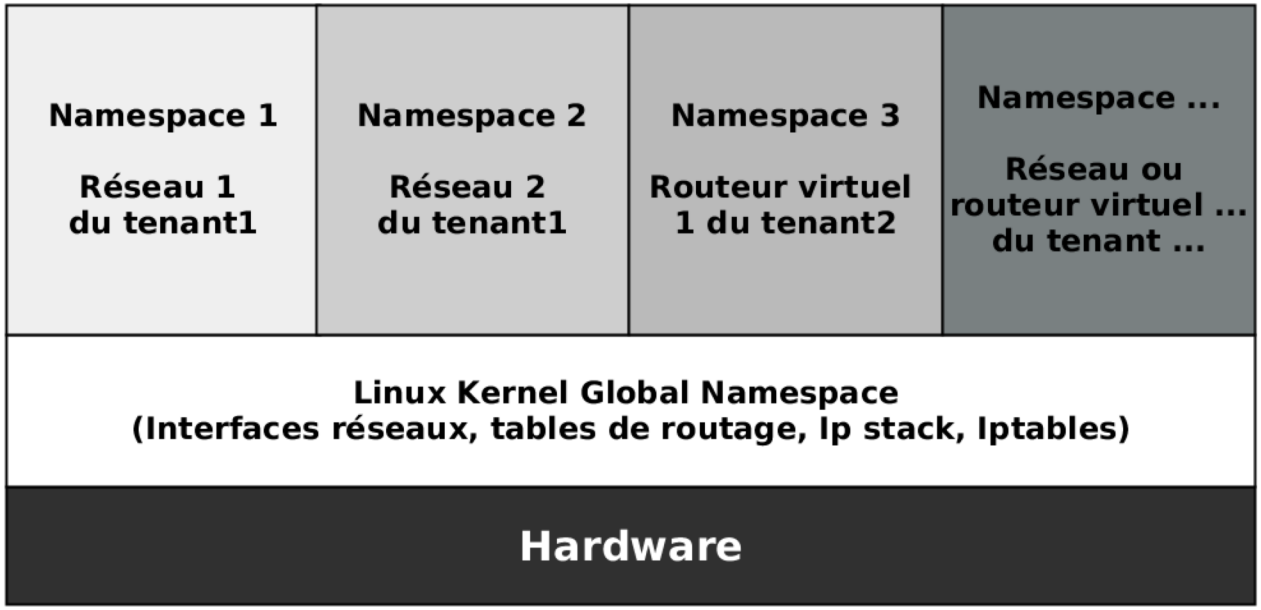


Neutron : Les composants

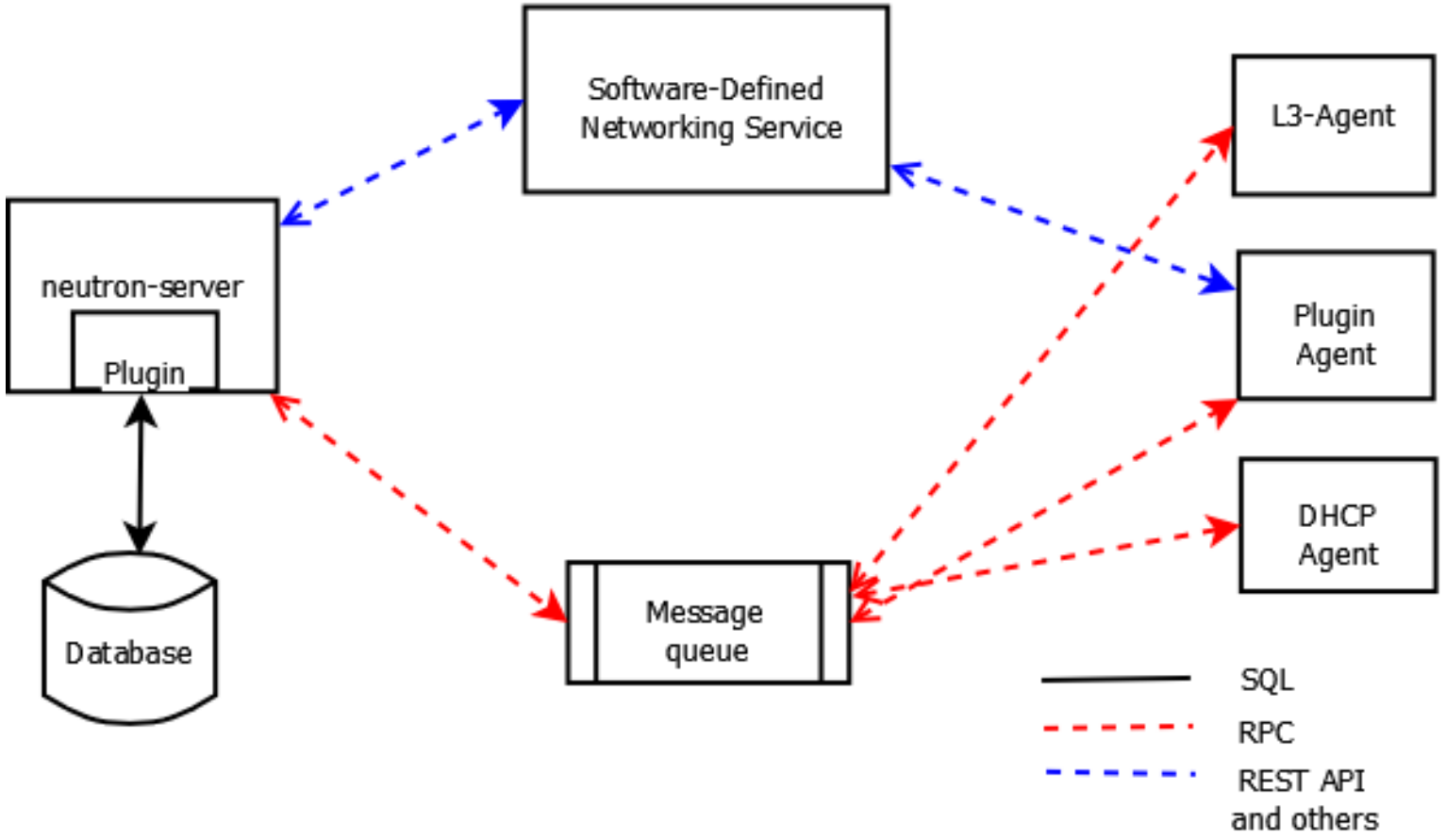
- neutron-server
  - Service d'API (contrôleurs OpenStack).
- Network node
  - Nœud dédié au routage, filtrage, adressage (L3, DHCP..).
  - Fournit un proxy pour le service de métadonnées.
- neutron-dhcp-agent
  - Configure et démarre un processus dnsmasq.
  - Crée des namespaces avec un préfix « qdhcp- ».
  - Besoin d'un agent L2 sur la même machine.
- neutron-openvswitch-agent
  - Configure et alloue le réseau au niveau L2.
- neutron-l3-agent
  - Configure le routage est-ouest et nord-sud, NAT, Floating Ips.
  - Crée des namespaces avec un préfix « qrouter- ».
- neutron-metadata-agent
  - Passerelle entre les instances et le serveur de métadonnées.

Neutron : Les namespaces

- Les « namespaces » du noyau Linux
  - Permet l'isolation réseau.
  - Possibilité d'avoir des plages réseaux identiques dans plusieurs tenants.

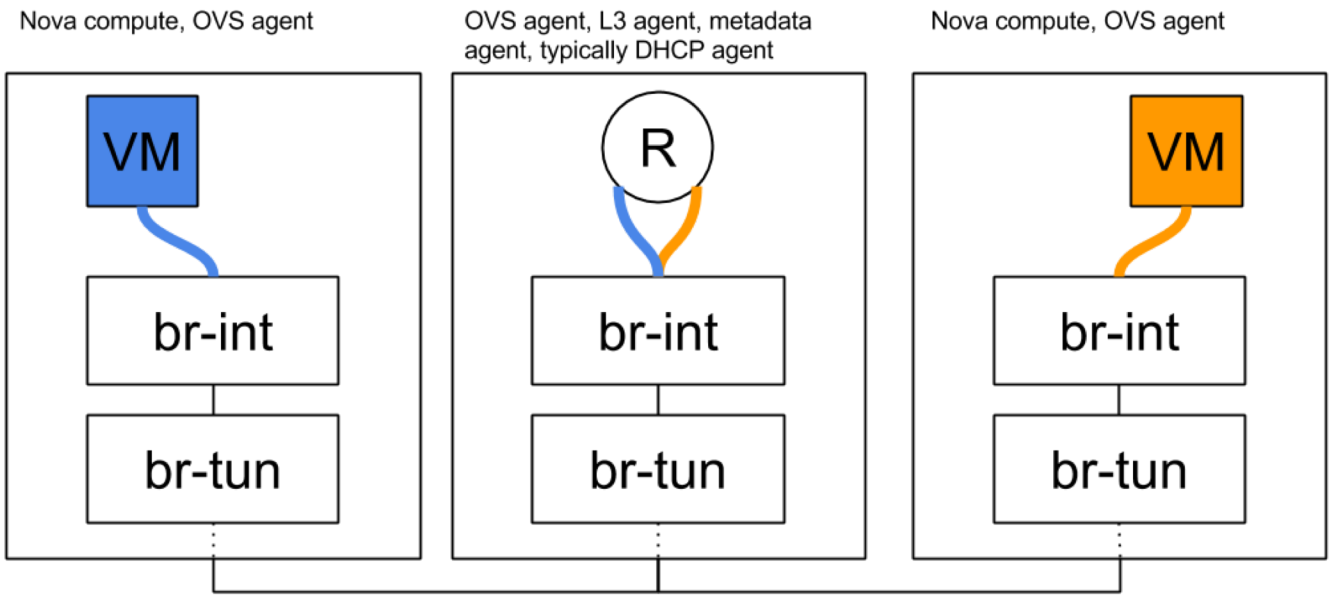


Neutron : Architecture

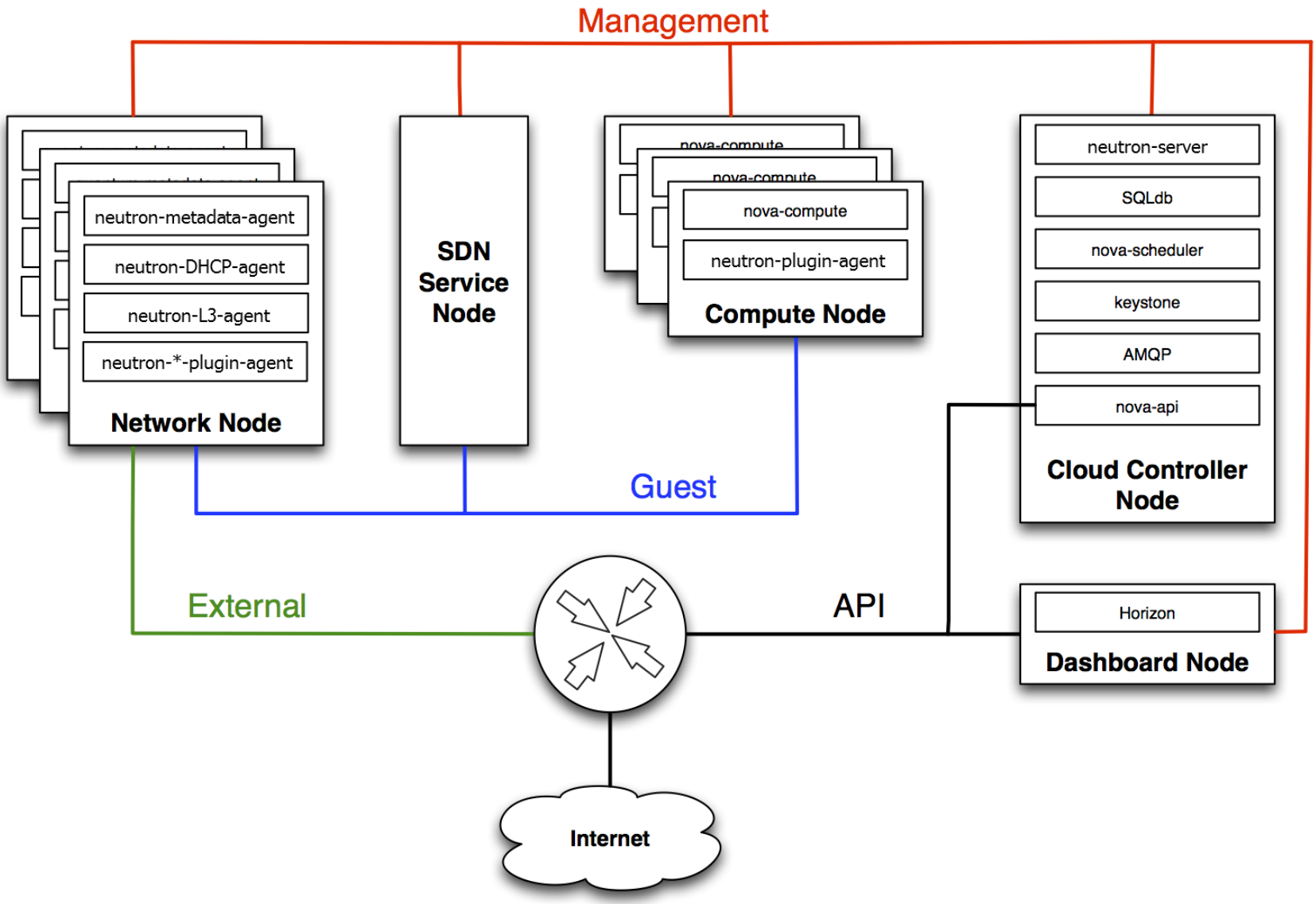


Neutron : Legacy Routing

- Tout le trafic réseau passe par le network node
  - Routage, Floating IP, SNAT.
  - Métadonnées.
- Problèmes
  - Performance, Passage à l'échelle, SPoF.

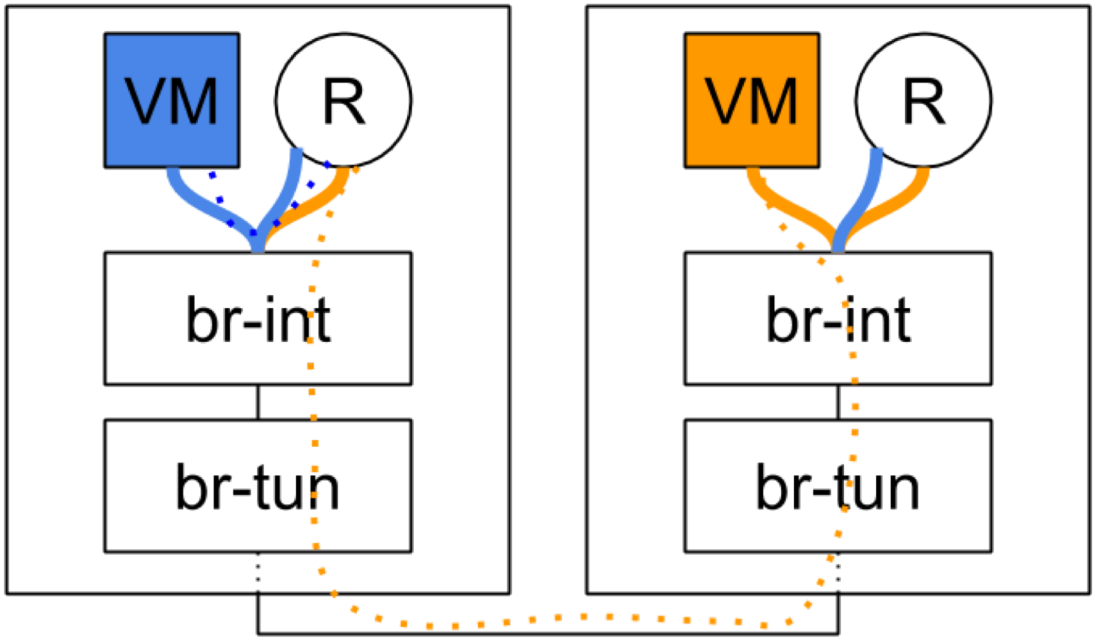


Neutron : Exemple de connectivité réseau

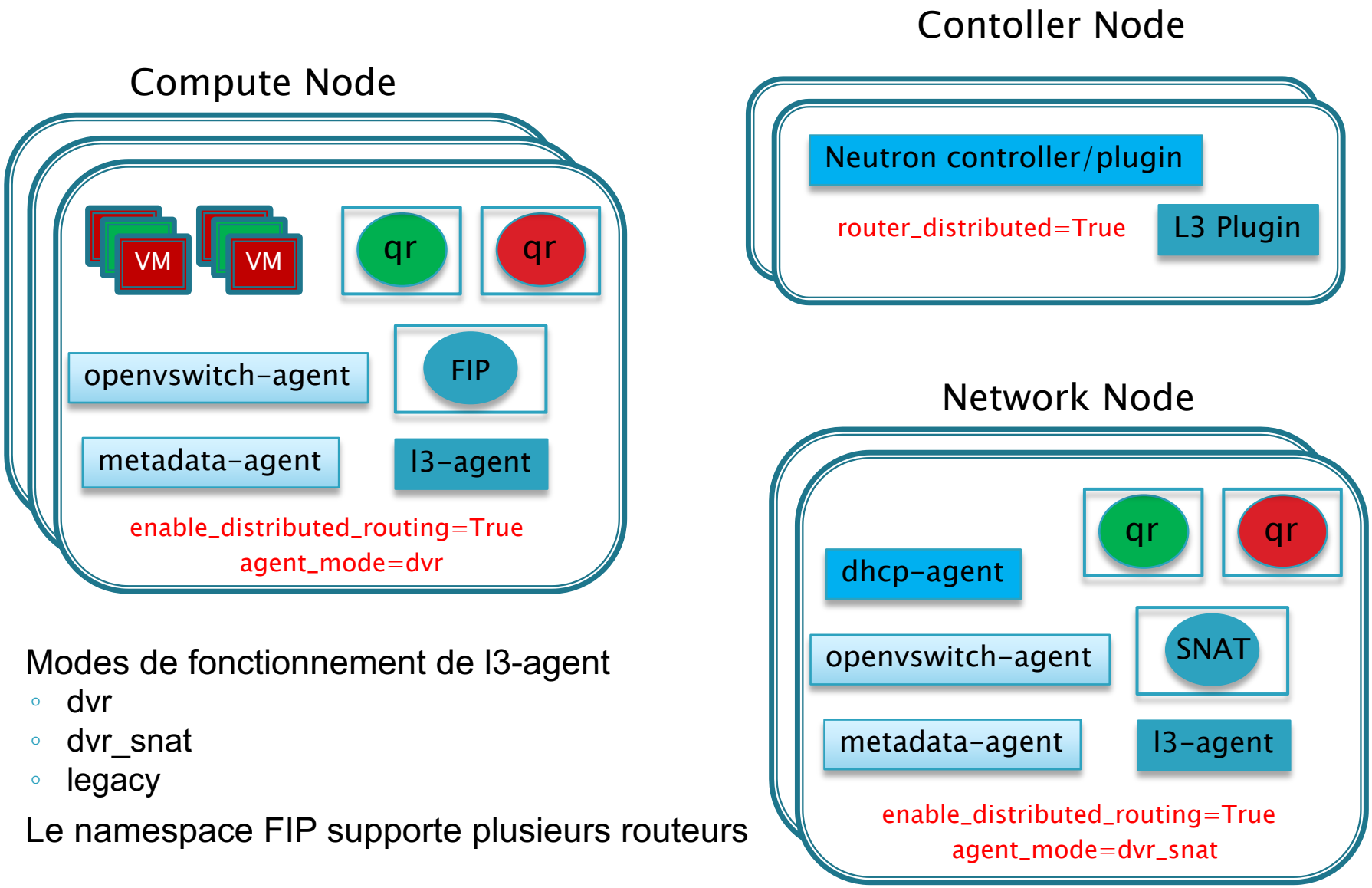


Neutron : DVR (Distributed Virtual Routing)

- Le trafic réseau distribué sur les hyperviseurs
  - Routage, Floating IP.
  - Métadonnées.
- Avantages
  - Performance, Passage à l'échelle.
- Limitation
  - SNAT centralisé.



Neutron : DVR (Distributed Virtual Routing)



- Modes de fonctionnement de l3-agent
  - dvr
  - dvr_snat
  - legacy
- Le namespace FIP supporte plusieurs routeurs

- Création d'un réseau

```
$ openstack network create --description "Test" --project test --provider-network-type vxlan test
```

- Création d'un sous réseau

```
$ openstack subnet create test-subnet \
--description "Subnet pour les tests" \
--project test \
--network test \
--subnet-range 10.0.1.0/24
```

- Ajout du sous réseau au routeur virtuel

```
$ openstack router add subnet router01 test-subnet
```

Boot d'une VM : Interaction avec nova

- Extraits de logs nova et neutron

```
neutron.plugins.ml2.drivers.openvswitch.agent.ovs_neutron_agent ... Port 6c1af981-7c40-4085-9184-7507fad7faec updated.

ovs-vsctl[28800]: ... iface-id=6c1af981-7c40-4085-9184-7507fad7faec ... attached-mac=fa:16:3e:18:e5:16 external-ids:vm-uuid=2b5226e9-f772-4070-a471-9532064be4f4

nova-compute[6169]: { ... "os_instance":"2b5226e9-f772-4070-a471-9532064be4f4", "os_message":"Instance spawned successfully." }
```

Neutron : DVR et namespaces

- Liste des réseau

```
$ openstack network list -f value -c ID -c Name
8358e0b1-cc39-490a-a86e-a141094b5c95 floating
aaf4298b-0663-4ea6-8b78-e075660925de ccin2p3
b147d054-3d19-4d9f-b1d7-7ef295fdb59e test
```

Réseau partagé

Réseaux de tenant

- Network node

```
$ ip netns
qrouter-f7db668a-7a58-434c-9241-93b983feed9c
snat-f7db668a-7a58-434c-9241-93b983feed9c
qdhcp-aaf4298b-0663-4ea6-8b78-e075660925de
qdhcp-b147d054-3d19-4d9f-b1d7-7ef295fdb59e
```

- Compute node

```
$ ip netns
fip-8358e0b1-cc39-490a-a86e-a141094b5c95
qrouter-f7db668a-7a58-434c-9241-93b983feed9c
```

Neutron : Usage des namespaces

```
$ ip netns exec qdhcp-b147d054-3d19-4d9f-b1d7-7ef295fdb59e ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN qlen 1
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
83: tap60f92ee8-bf: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state
UNKNOWN qlen 1000
    link/ether fa:16:3e:82:cc:3b brd ff:ff:ff:ff:ff:ff
    inet 172.17.39.3/24 brd 172.17.39.255 scope global tap60f92ee8-bf
       valid_lft forever preferred_lft forever
    inet 169.254.169.254/16 brd 169.254.255.255 scope global tap60f92ee8-bf
       valid_lft forever preferred_lft forever
    inet6 fe80::f816:3eff:fe82:cc3b/64 scope link
       valid_lft forever preferred_lft forever
```

Neutron : Usage des namespaces

```
$ ip netns exec qdhcp-b147d054-3d19-4d9f-b1d7-7ef295fdb59e route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         172.17.39.1     0.0.0.0         UG    0      0        0 tap60f92ee8-bf
169.254.0.0     0.0.0.0         255.255.0.0     U     0      0        0 tap60f92ee8-bf
172.17.39.0     0.0.0.0         255.255.255.0   U     0      0        0 tap60f92ee8-bf
```

```
$ ip netns exec qdhcp-b147d054-3d19-4d9f-b1d7-7ef295fdb59e ping -c3 172.17.39.1
PING 172.17.39.1 (172.17.39.1) 56(84) bytes of data.
64 bytes from 172.17.39.1: icmp_seq=1 ttl=64 time=0.074 ms
64 bytes from 172.17.39.1: icmp_seq=2 ttl=64 time=0.057 ms
64 bytes from 172.17.39.1: icmp_seq=3 ttl=64 time=0.057 ms

--- 172.17.39.1 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2000ms
rtt min/avg/max/mdev = 0.057/0.062/0.074/0.012 ms
```

Questions ?