OpenStackでのダウンタイム・ゼロへの挑戦
OpenStack NeutronのHAとは

Rossella Sblendido
Software Engineer
SUSE
rsblendido@suse.com

Ryo Murakawa
Evangelist
Novell Japan
rmurakawa@suse.com

What is SUSE® OpenStack Cloud?

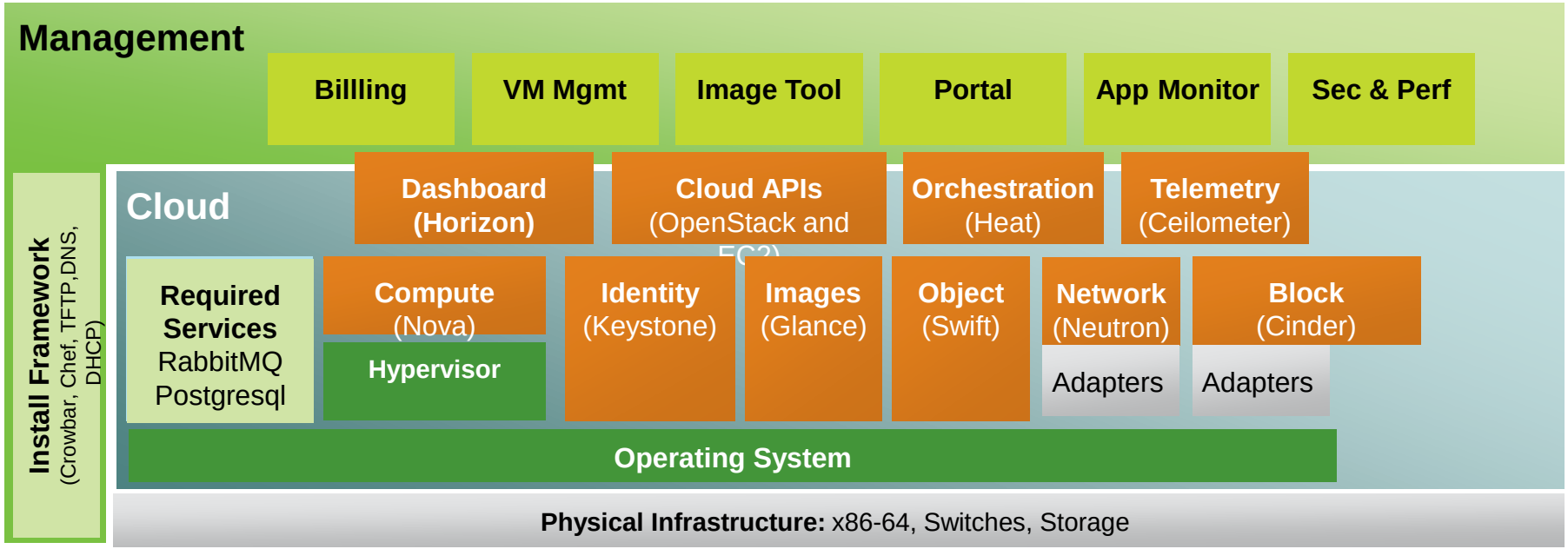
SUSE will change the name from SUSE Cloud to SUSE OpenStack Cloud as same as the release of next version OpenStack product.

- SUSE OpenStack Cloud 5

Before, we called our product SUSE Cloud.

Name	Based OpenStack
SUSE Cloud 2	Grizzly
SUSE Cloud 3	Havana
SUSE Cloud 4	Icehouse

SUSE® OpenStack Cloud



SUSE® OpenStack Cloud Feature

- Crowbar based Installation Framework

openstack™

CROWBAR

The cloud *un*boxer.

Deploying OpenStack Manually

>1400

Parameters

11

Components

2

Days

Deploying OpenSack with SUSE OpenStack Cloud

>1400
Parameters

11
Components

2
Hours

Multi Hyper Visor Support

Supported Hyper Visor

- ❑ KVM
- ❑ Xen
- ❑ vSphere
- ❑ Hyper-V

High Availability for Control Node

High Availability for Control Node

OpenStack Component	High Availability type
Data Base	Active/Standby
Messaging Service	Active/Standby
Keystone	Active/Active
Glance	Active/Active
Cinder	Active/Active
Neutron	Active/Active
Nova-Service	Active/Active
Horizon	Active/Active
Ceilometer	Active/Active

High Availability in Neutron

What's Neutron?

- Neutron is an OpenStack project to provide “networking as a service” between interface devices (e.g., vNICs) managed by other Openstack services (e.g., nova)
- provides a powerful API to define the network connectivity

Neutron abstractions

- Network: L2 broadcast domain
- Subnet: a block of v4 or v6 IP addresses and associated configuration state.
- Port: a connection point for attaching a single device, such as the NIC of a virtual server, to a virtual network. Also describes the associated network configuration, such as the MAC and IP addresses to be used on that port.
- Router: interconnects networks

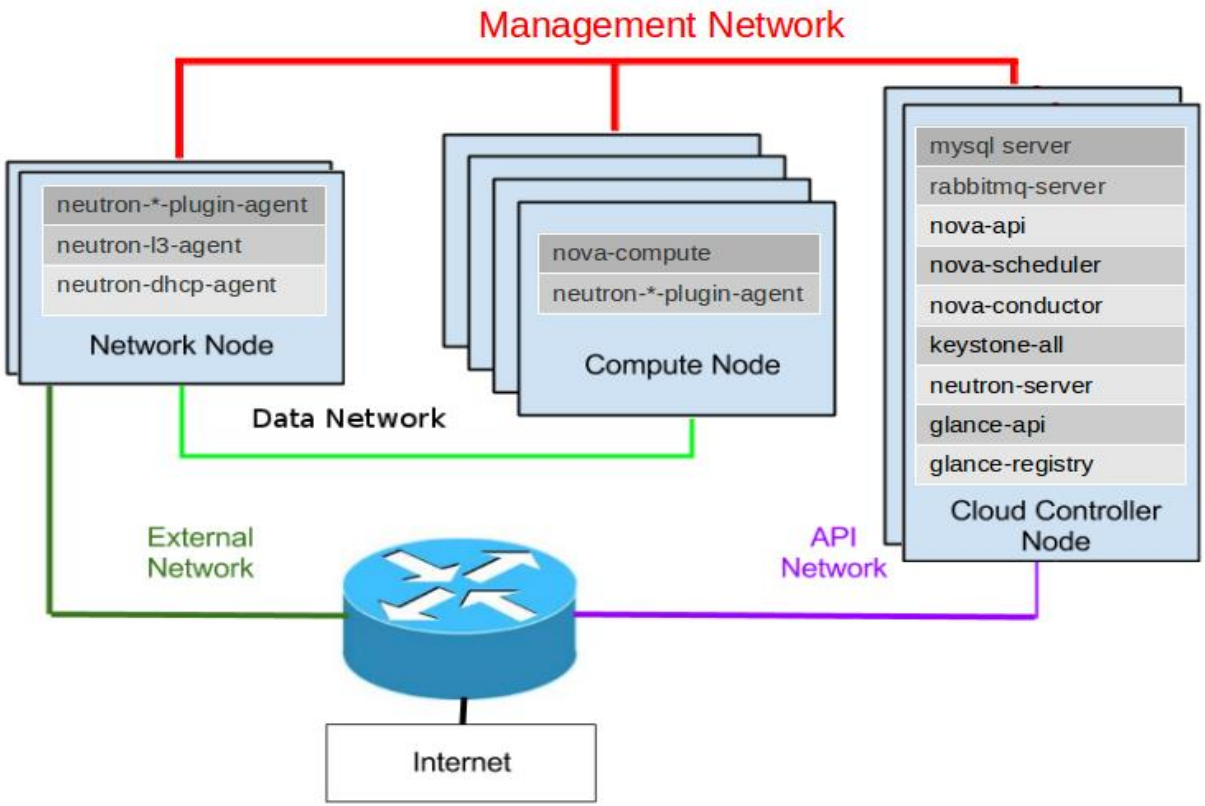
Modular architecture

- Plugin: custom back-end implementation of the Networking API
- Neutron-server: exposes the API
- Several agents (L2, L3, DHCP, Metadata, etc)

Plugin

- Monolithic plugin (direct control of core resources)
- ML2
  - Modular, delegates calls to proper drivers
  - Two kind of drives:
    - Type drivers (support specific network type)
    - Mechanism drivers (ensure the information established by the TypeDriver is properly applied)
  - 2 default implementations: OpenVSwitch and LinuxBridge

Different kind of nodes



L2 Agent

- Runs on hypervisor
- Configure the local vswitch
- Communicates with the server over RPC
- Wires new devices
- Security Group Rules

L3 Agent

- Provides L3/NAT
- Runs network node
- Uses namespaces
- External network access

DHCP agent

- Provides DHCP services
- Communication over RPC
- Isolation through namespaces
- dnsmasq

Metadata Agent

- Proxies Metadata requests to Nova
- Routed Networks
  - Process embedded in router

- Non-routed Networks
  - Static routes redirect traffic running in the DHCP namespace

High Availability

What HA means

- Minimize two things:
- System downtime — occurs when a user-facing service is unavailable beyond a specified maximum amount of time
- Data loss — accidental deletion or destruction of data

Stateless vs. Stateful services

- A stateless service is one that provides a response after your request, and then requires no further attention
- A stateful service is one where subsequent requests to the service depend on the results of the first request

Active/passive Active/active

- Active/passive (one instance is only receiving notification but not processing requests):

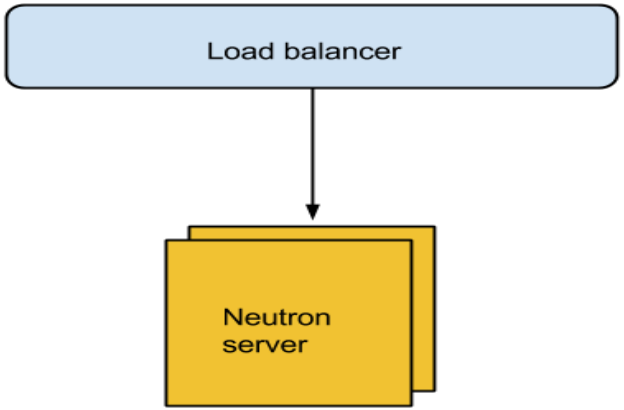
- Active/active: there's a backup but both the main and redundant systems run concurrently

How to make Neutron HA

How to make Neutron HA

- L2 agent runs on every compute node, no need of HA
- Neutron server, DHCP and L3 agent need HA

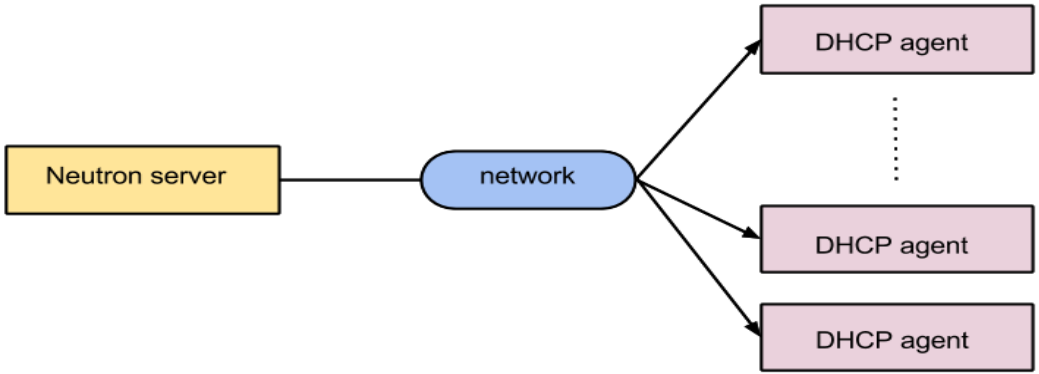
HA Neutron server



- It's stateless

HA DHCP agent

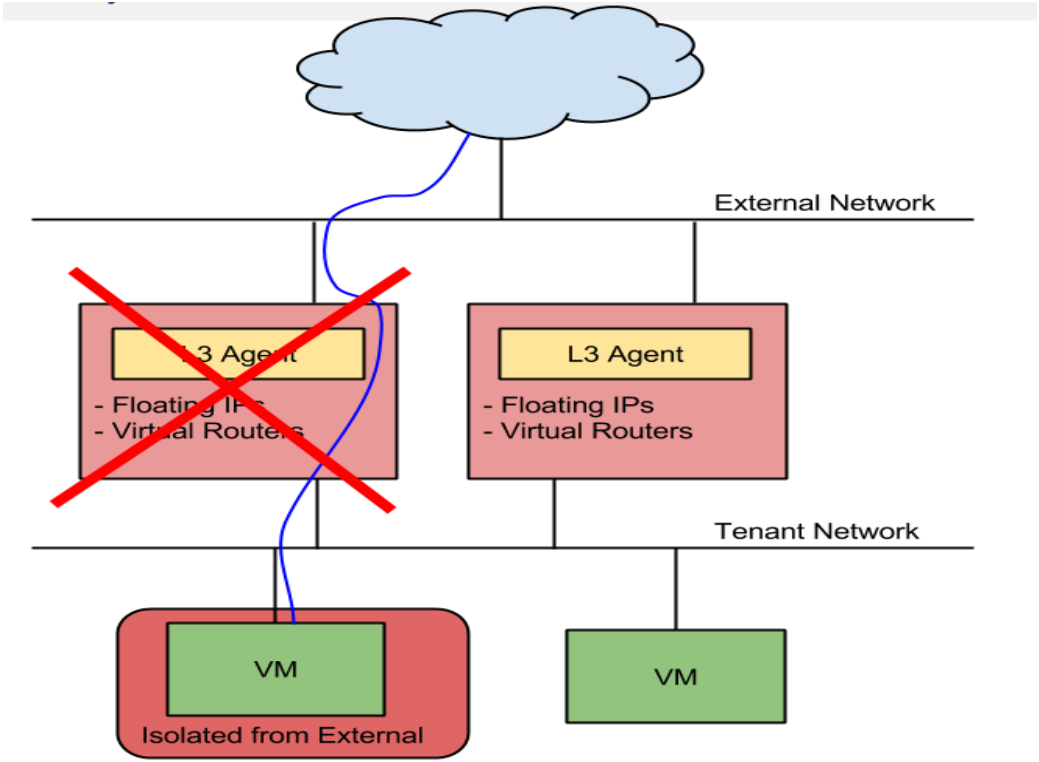
- OpenStack Networking service has a scheduler that lets you run multiple agents across nodes.
- The DHCP agent can be natively highly available.



HA L3 agent

- It is scalable thanks to the scheduler that allows distribution of virtual routers across multiple nodes.
- But no native HA, routers need to be migrated if a L3 agent fails

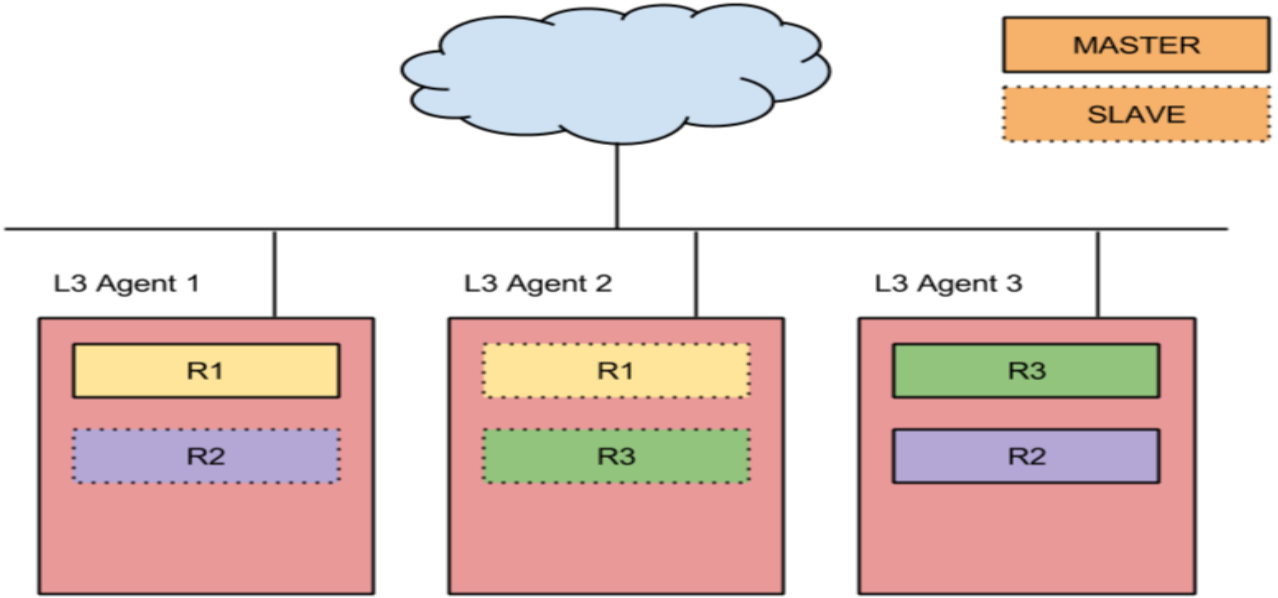
L3 agent fails...



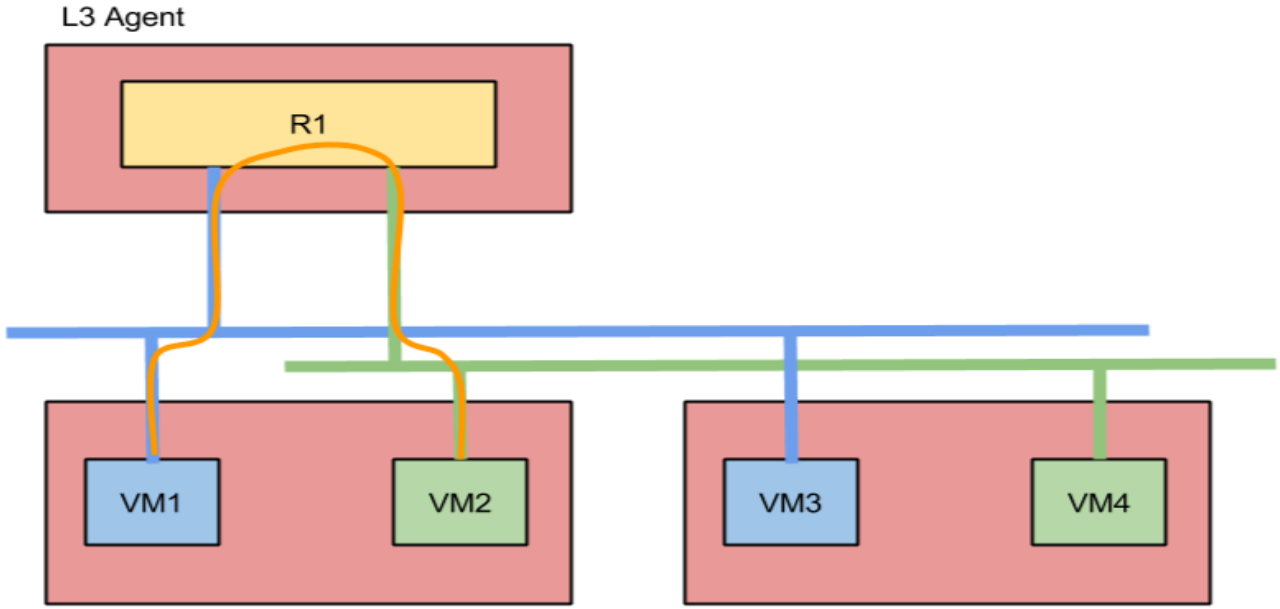
HA upstream solutions

- VRRP (Virtual Router Redundancy Protocol)
- DVR (Distributed Virtual Router)

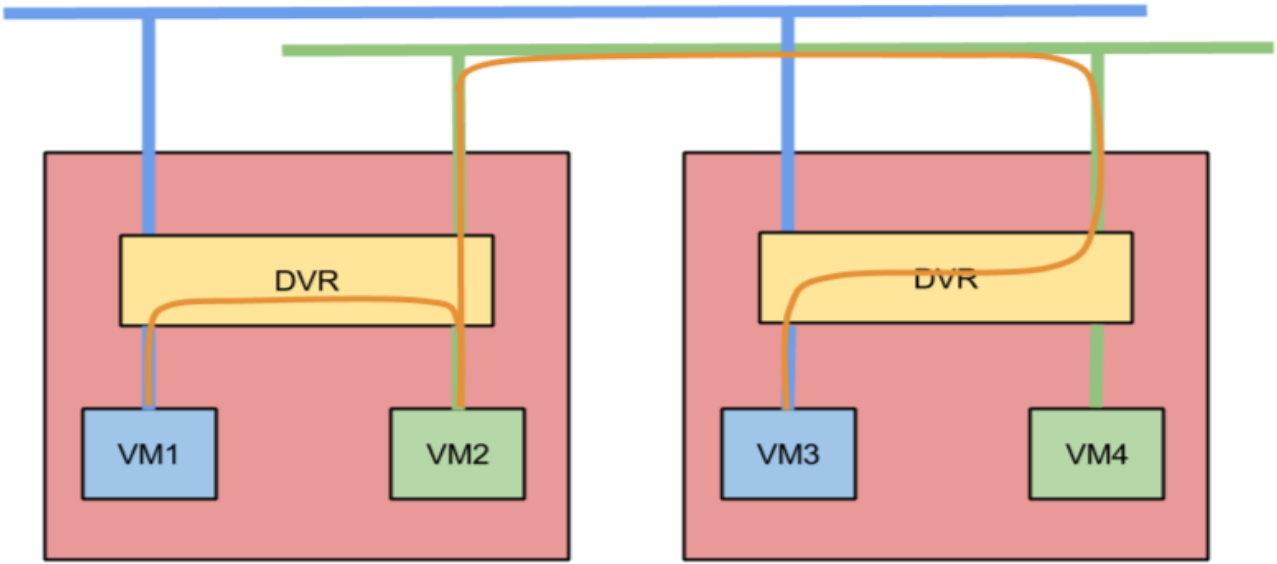
VRRP



Without DVR



With DVR



Real world example

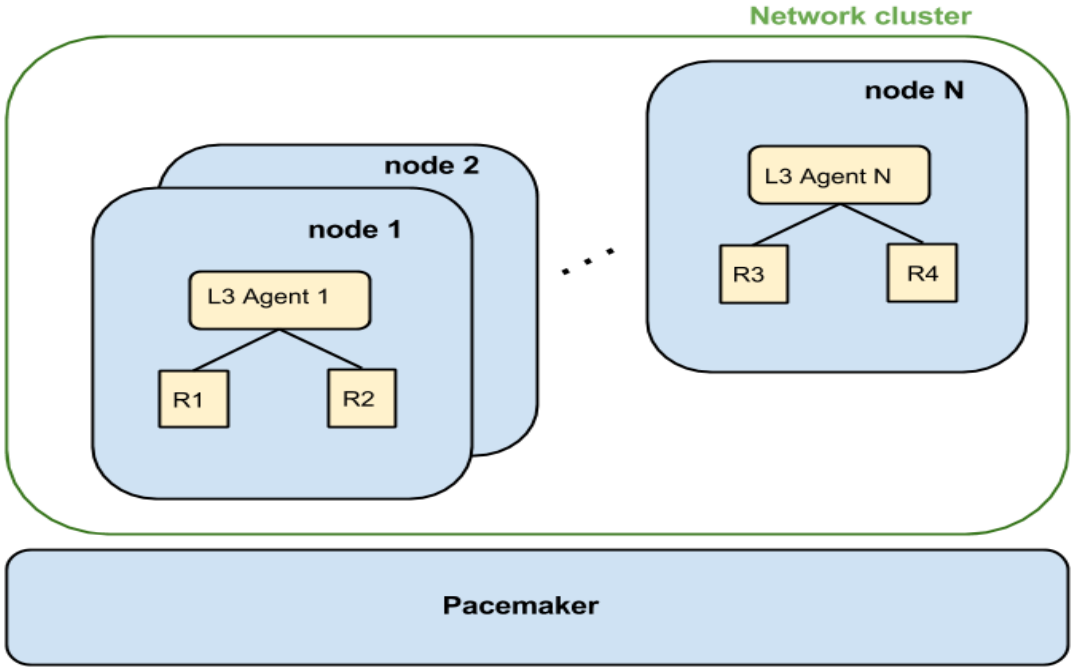
Neutron configuration - Real world scenario

- SUSE Cloud 4 (Icehouse) and tests for Cloud 5 (Juno)
- ML2, OVS with VLAN or LinuxBridge (multi-hypervisor support)
- HA!

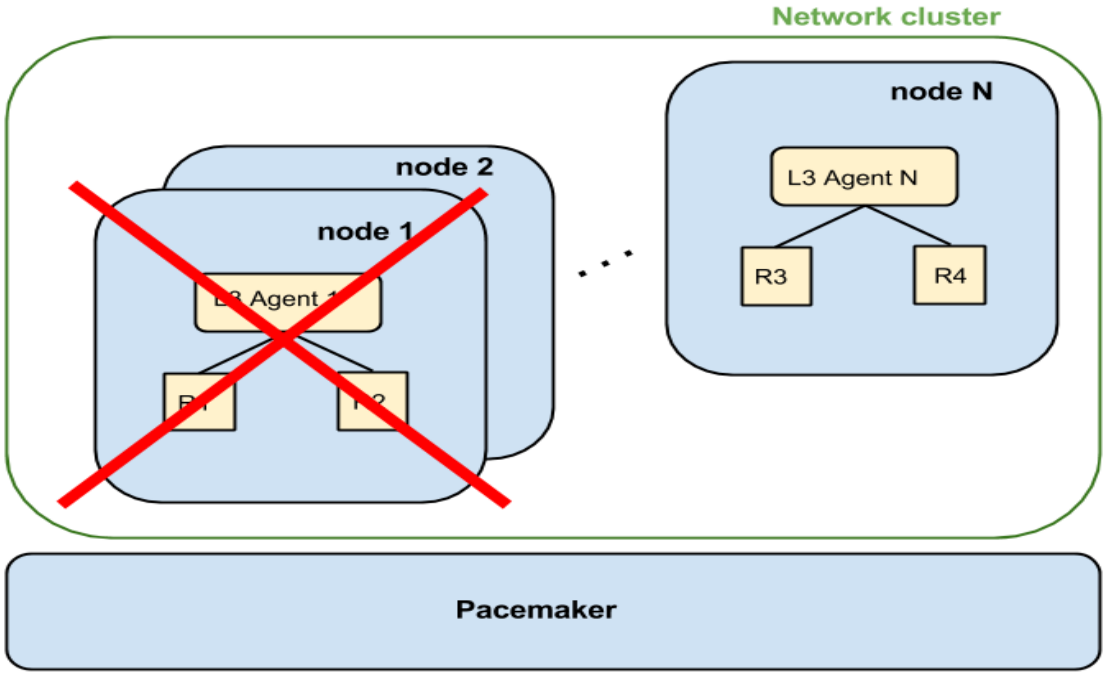
L3 Agent HA

- Active/active
- Migration tools that monitors the agents
  - Migrates routers from dead agents to alive ones
  - Replicates networks to all DHCP agents running

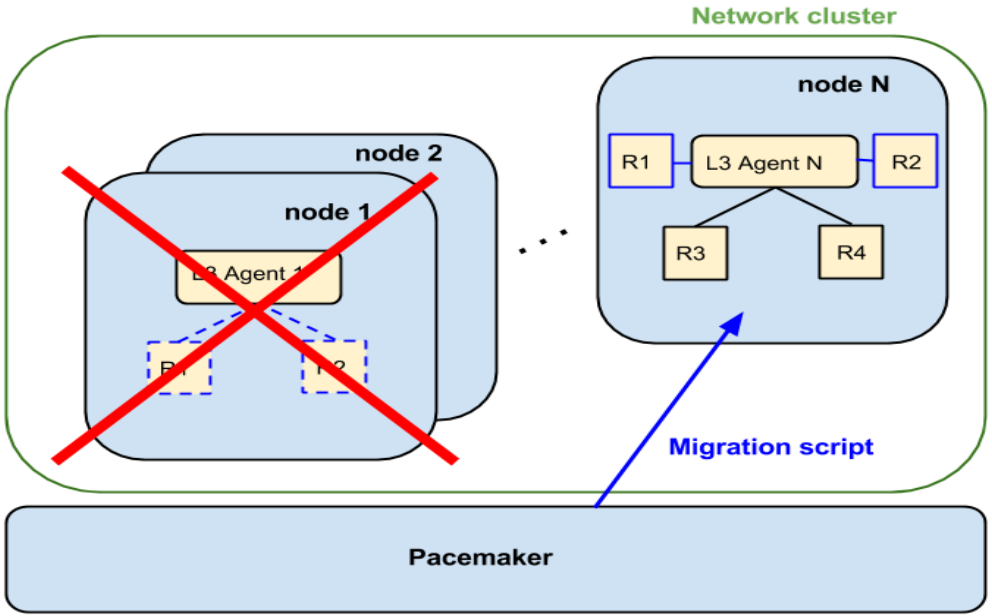
Network cluster



Failure...


Network cluster
node N
L3 Agent N
R3
R4
node 2
node 1
Pacemaker

Failover


Network cluster
node N
R1
L3 Agent N
R2
R3
R4
node 2
node 1
Migration script
Pacemaker

If you have questions...come to our booth!

Thank you.


SUSE®
We adapt. You succeed.