Bridges and Tunnels: A Drive Through OpenStack Networking

Mark McClain
mark@akanda.io
@gtwmm

Why Create Neutron?

- Rich Topologies
- Technology Agnostic
- Extensible
- Advance Services Support
  - Load Balancing, VPN, Firewall

Challenges in the Cloud

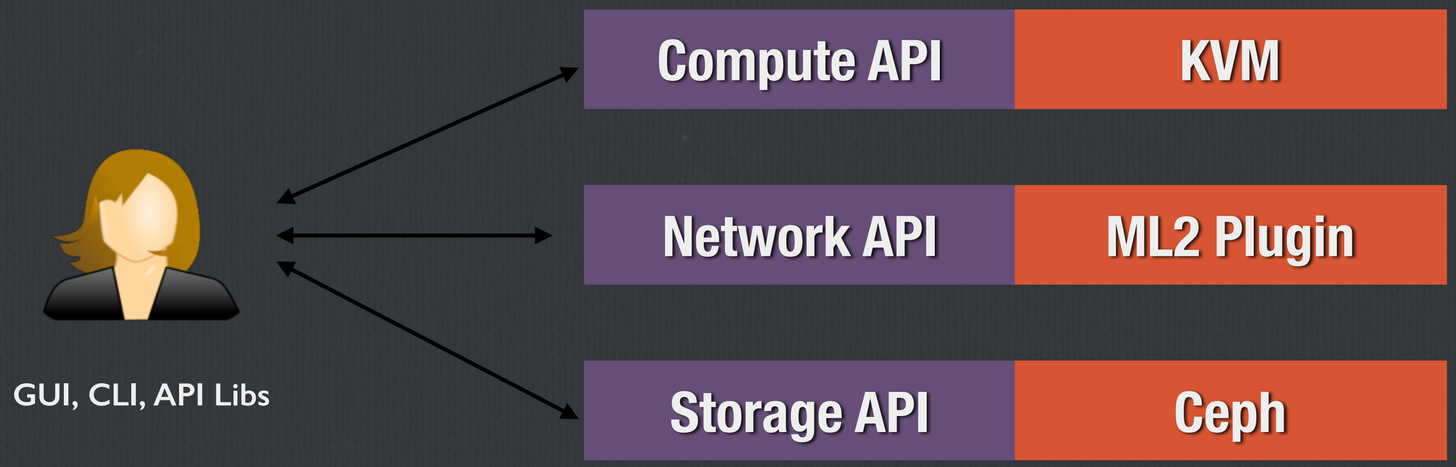
- High-density multi-tenancy
  - VLANs have trouble scaling
- On-demand provisioning
  - traditional solutions require manual configuration
- Need to place / move workloads
  - state tied (IP address) tied to location

Tackling these Challenges

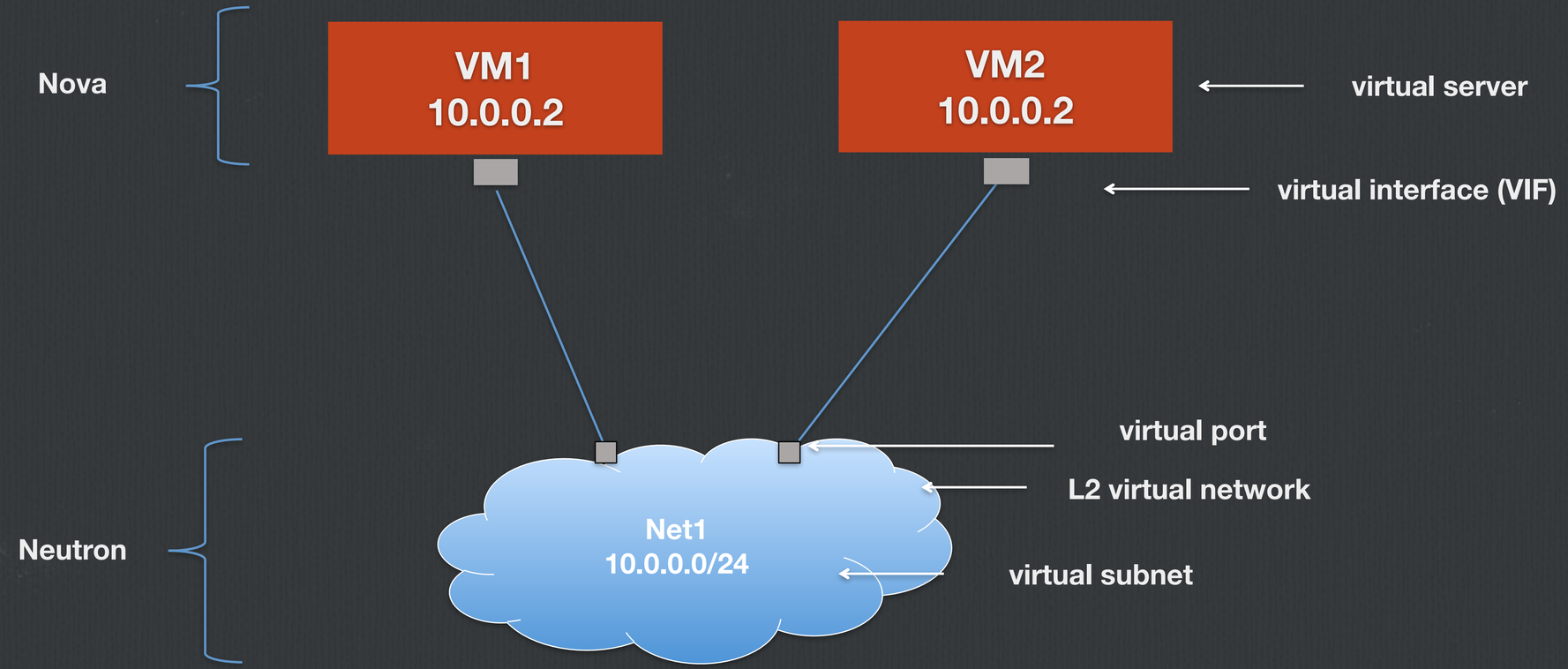
- Network virtualization
- Overlay tunneling
  - VXLAN, GRE, STT
- Software Defined Networking (SDN)
  - OpenFlow
- L2 Fabric Solution
- ???

The Basics

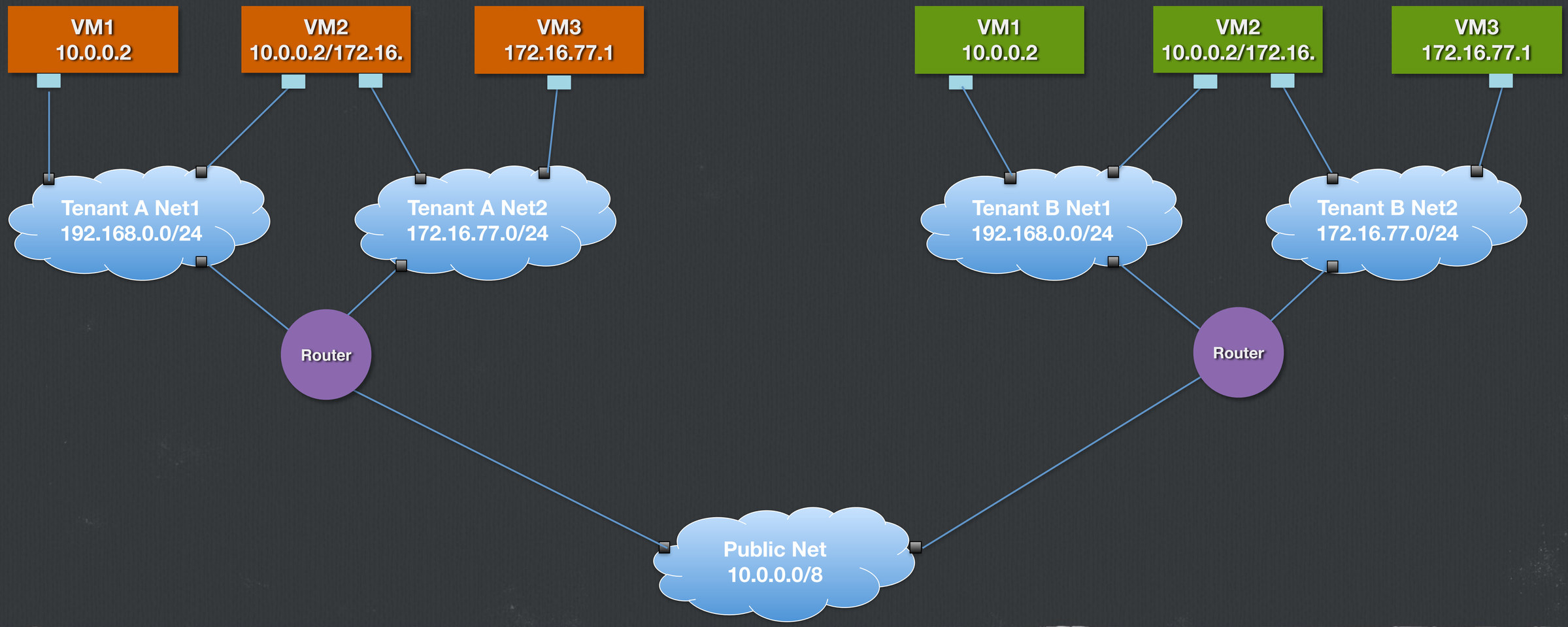
What does the user see?



Abstractions



Using the API…



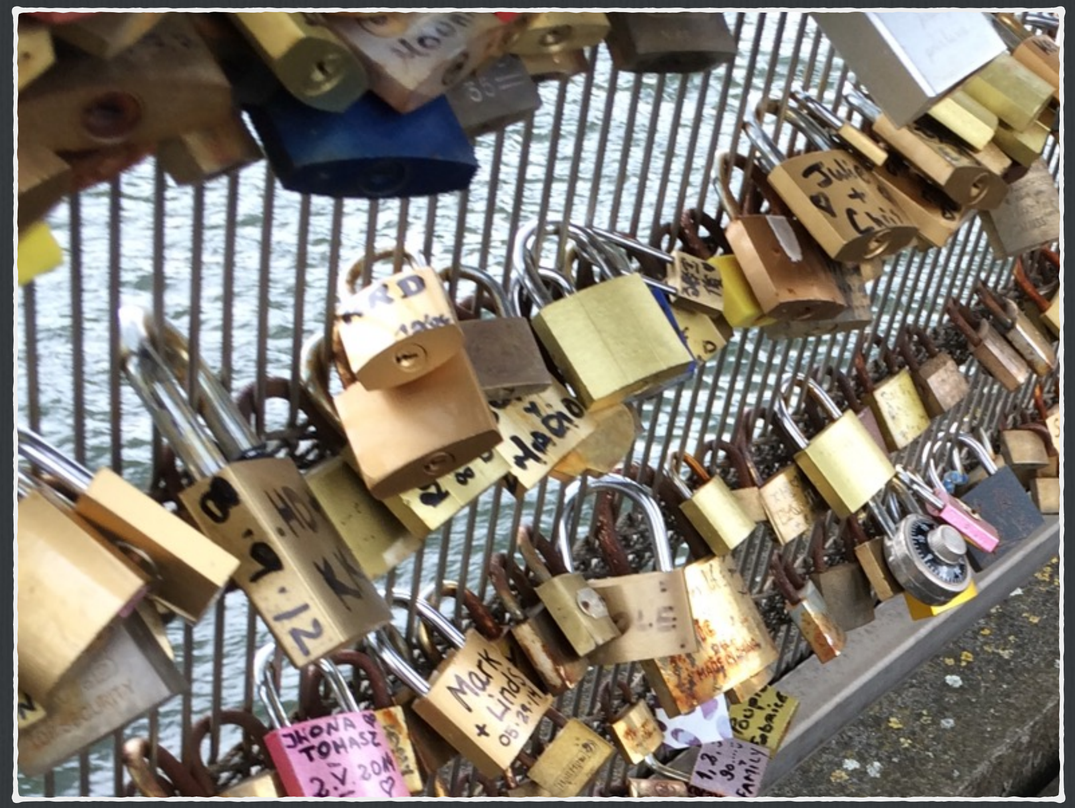
Design Goals

- [ ] Unified API
- [ ] Small Core
- [ ] Pluggable Open Architecture
- [ ] Extensible

Common Features

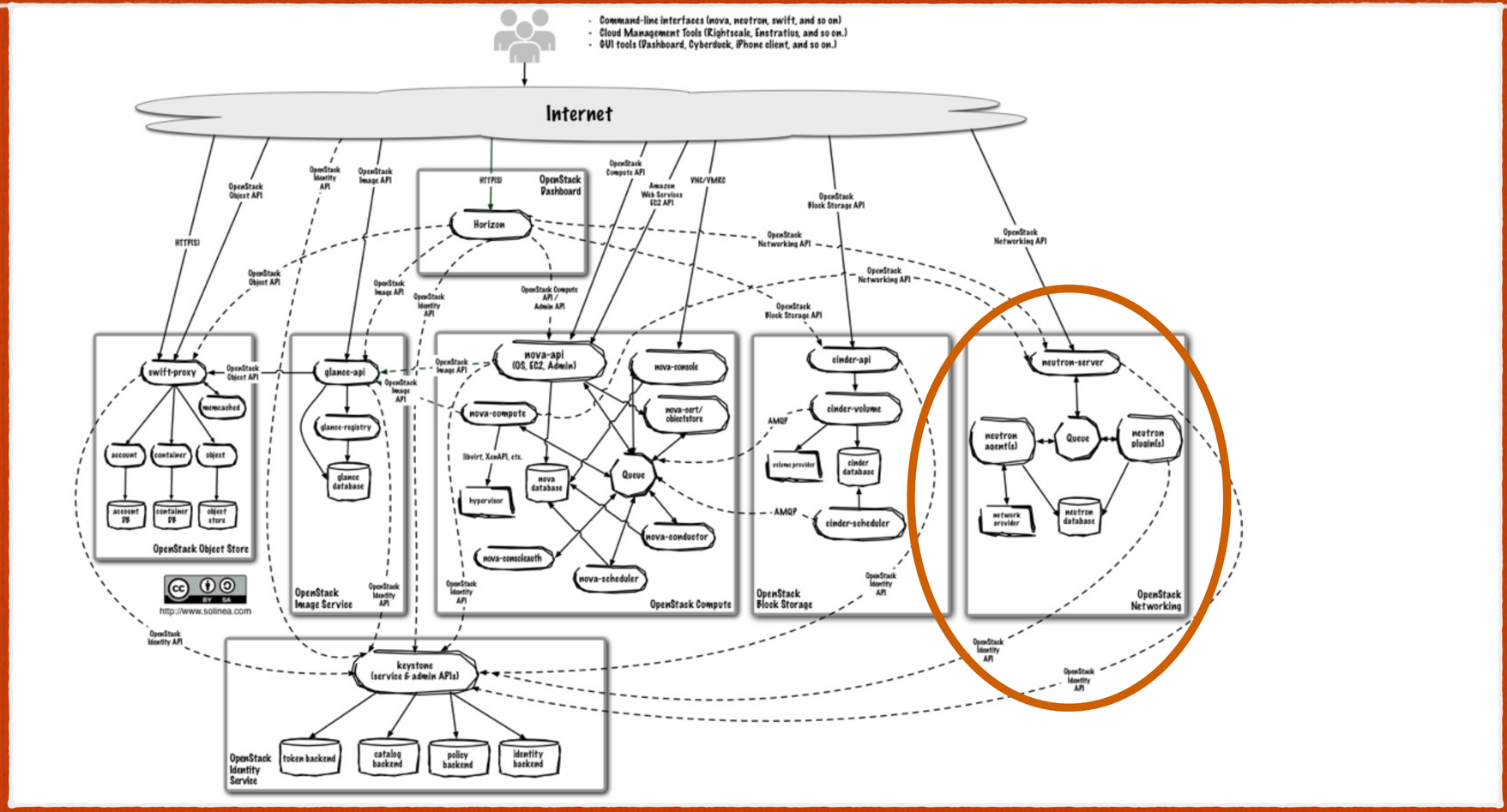
- Support for Overlapping IPs
  - Tenant A: 192.168.0.0/24
  - Tenant B: 192.168.0.0/24
- Configuration
  - DHCP/Metadata
- Floating IPs

Security Groups



- Support Overlapping IPs
- Ingress/Egress Rules
- IPv6
- VMs with multiple VIFs

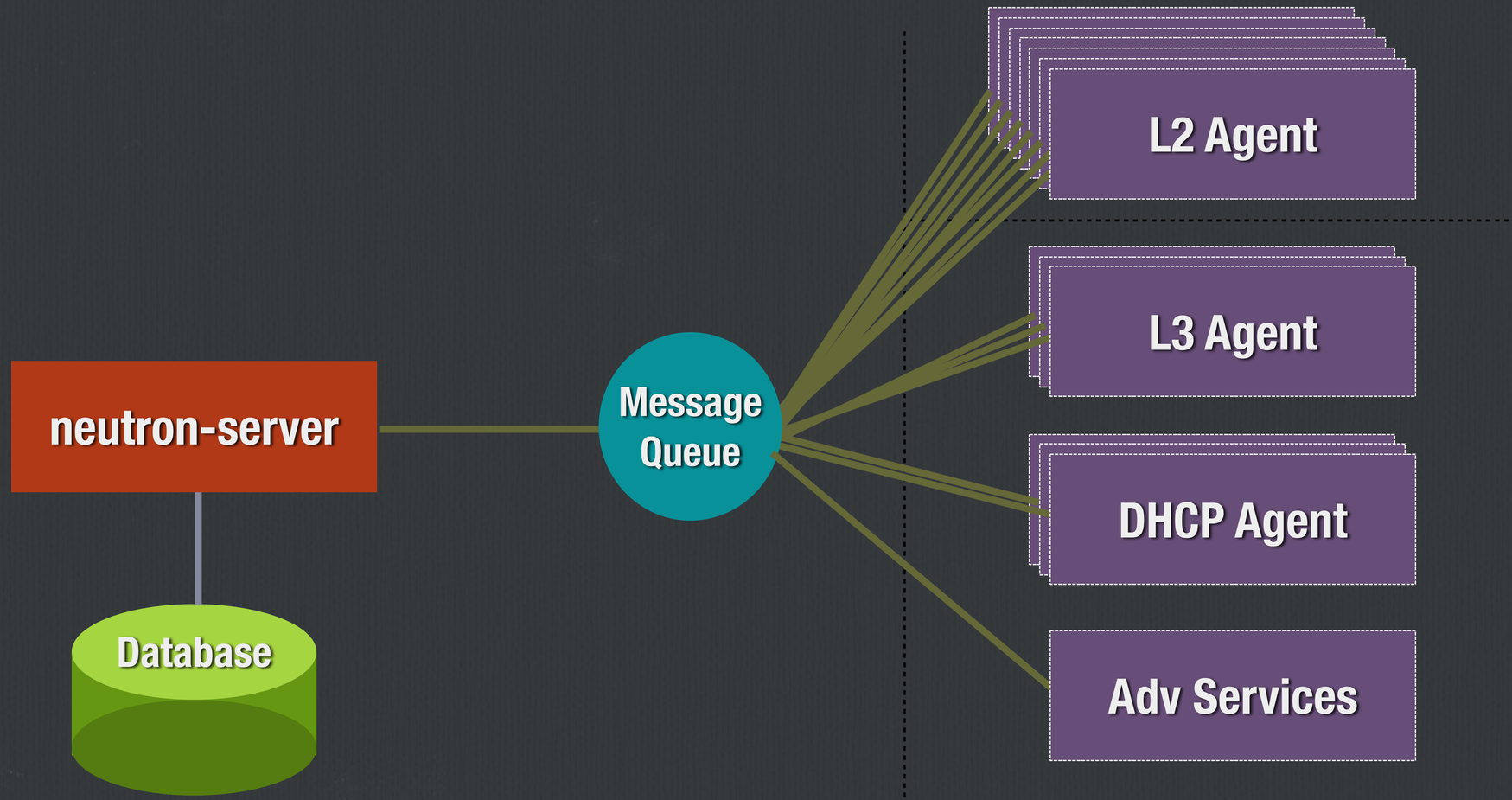
Architecture



OpenStack

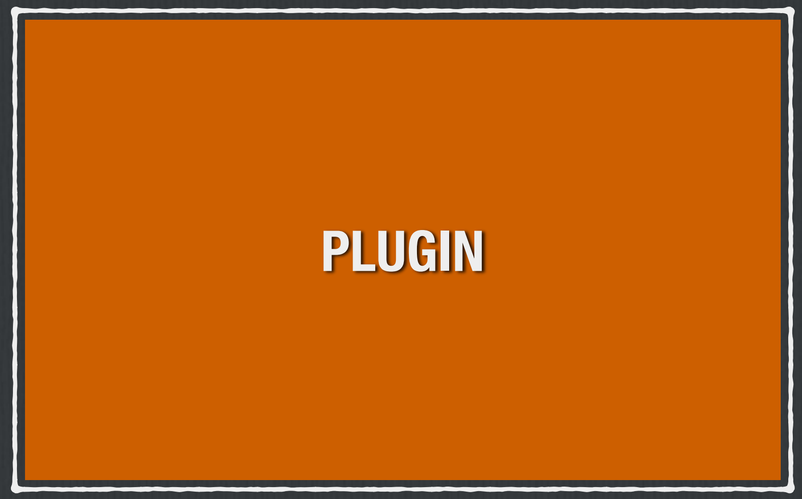
The Operator View

Basic Deployment



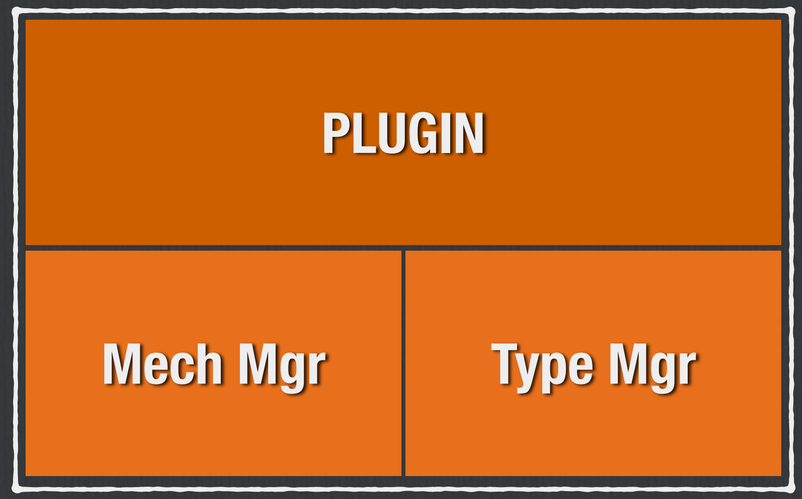
Monolithic Plugin

- Full implementation of core resources
- Two types:
  - Proxy
  - Direct control



ML2: Modular Layer 2 Plugin

- Full V2 Plugin Implementation
- Delegates calls to proper L2 drivers
- Two kinds of drivers
  - Type Driver
  - Mechanism Driver



Plugin Extensions

- Add logical resources to the REST API
- Discovered by server at startup
  - REST: /v2.0/extensions
- Common Extensions
  - Binding, DHCP, L3, Provider, Quota, Security Group
- Other Extensions
  - Allowed Addresses, Extra Routes, Metering

L2 Agent

- Runs on hypervisor
- Communicates with server via RPC
- Watch and notify when devices added/removed
- Wires new devices
    - Proper network segment
    - Security Group Rules

OVS L2 Agent

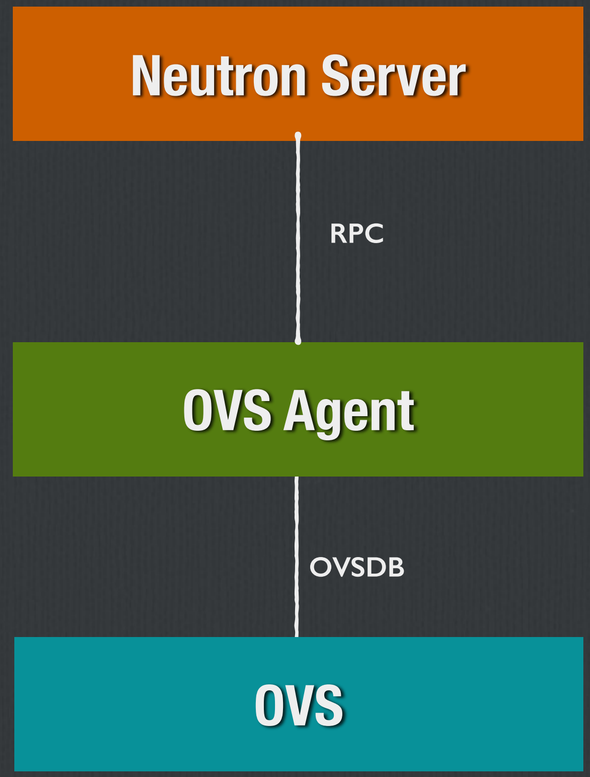
- Open vSwitch

  Open Source Virtual Switch

  http://openvswitch.org

- Tenant Isolation

  VLAN, GRE, VXLAN



Isolation

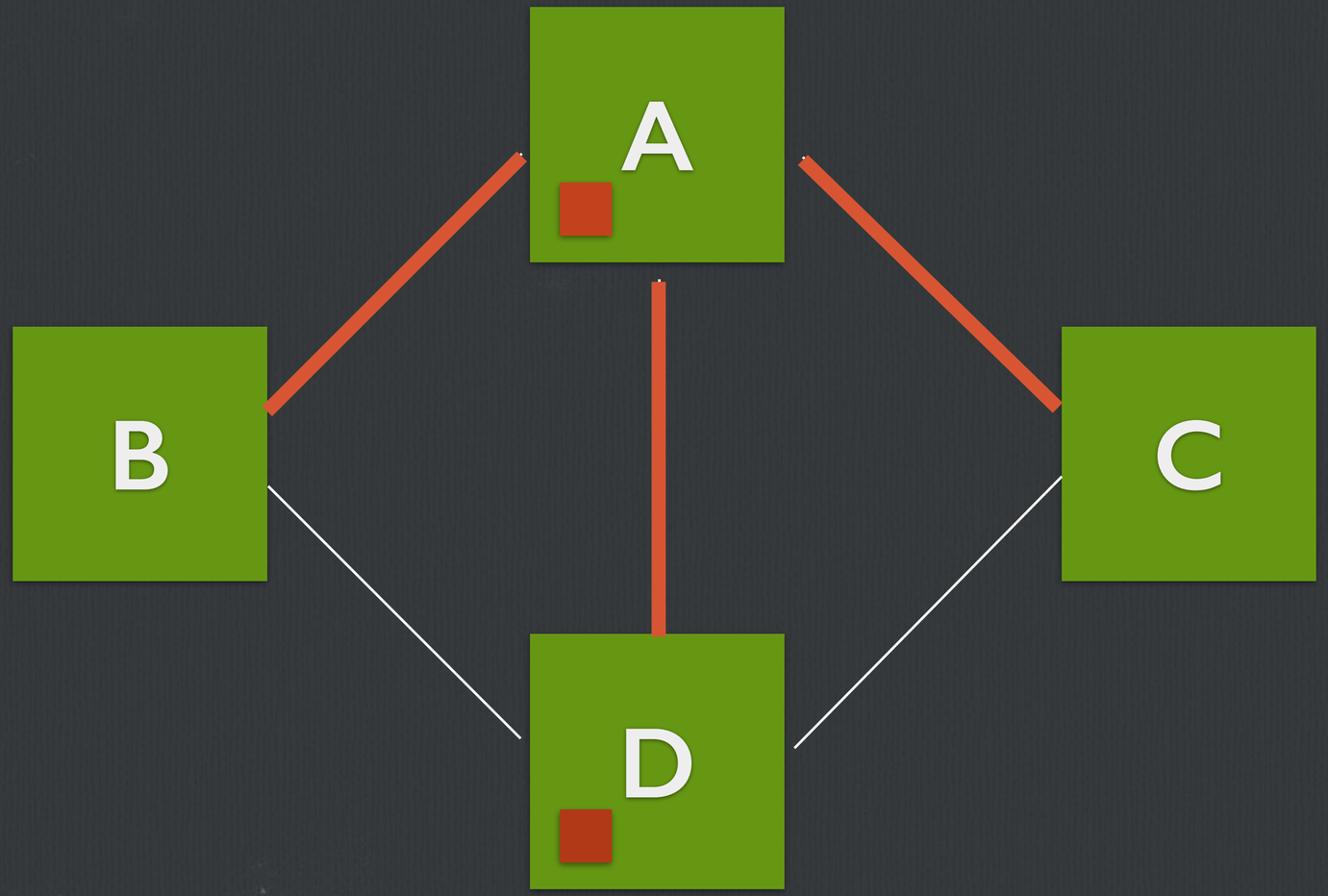
VLAN

- 802.1Q
- limited
- underlay must support

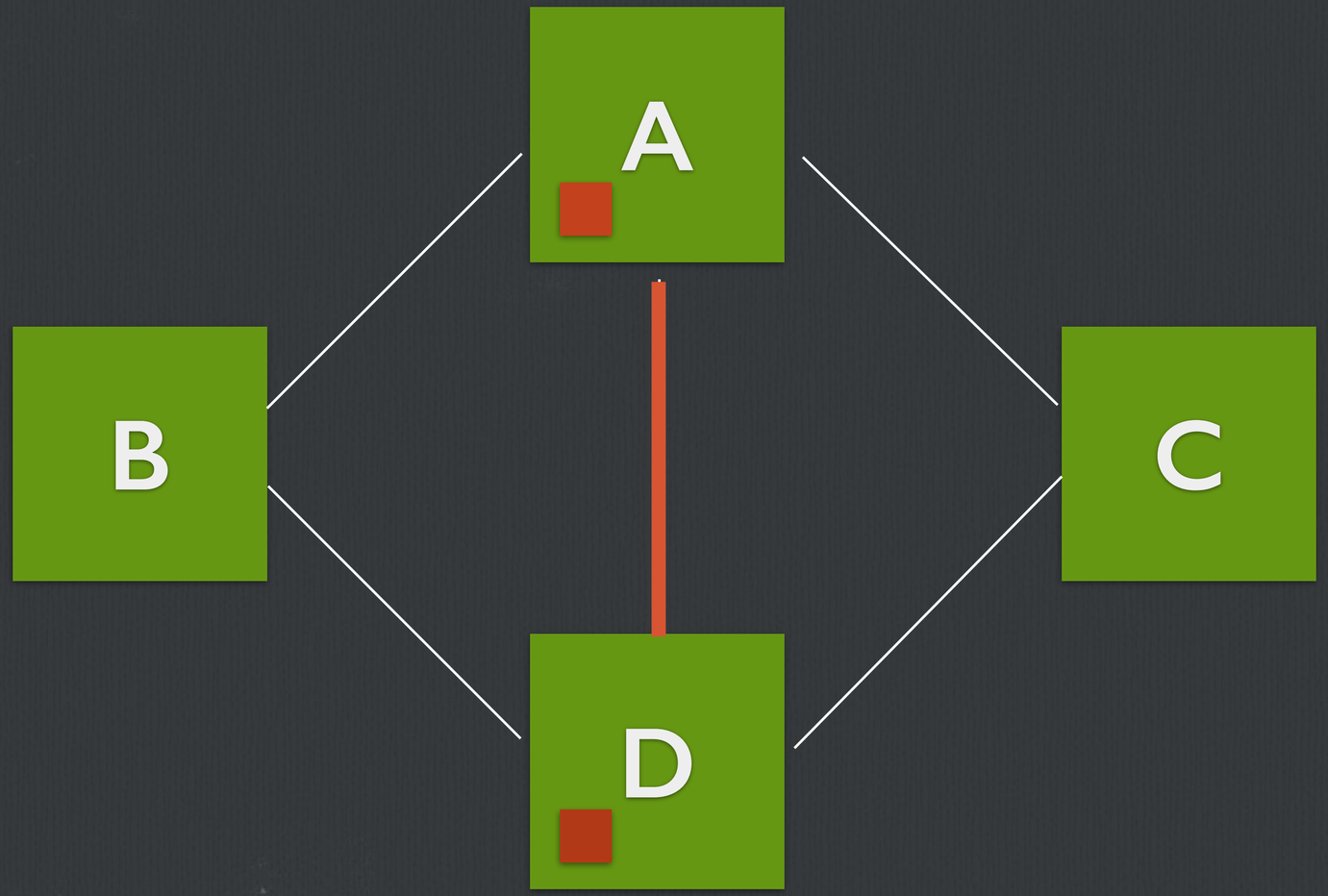
GRE/VXLAN

- L2 encapsulated in L3
- routable
- overlay independence

Tunneling



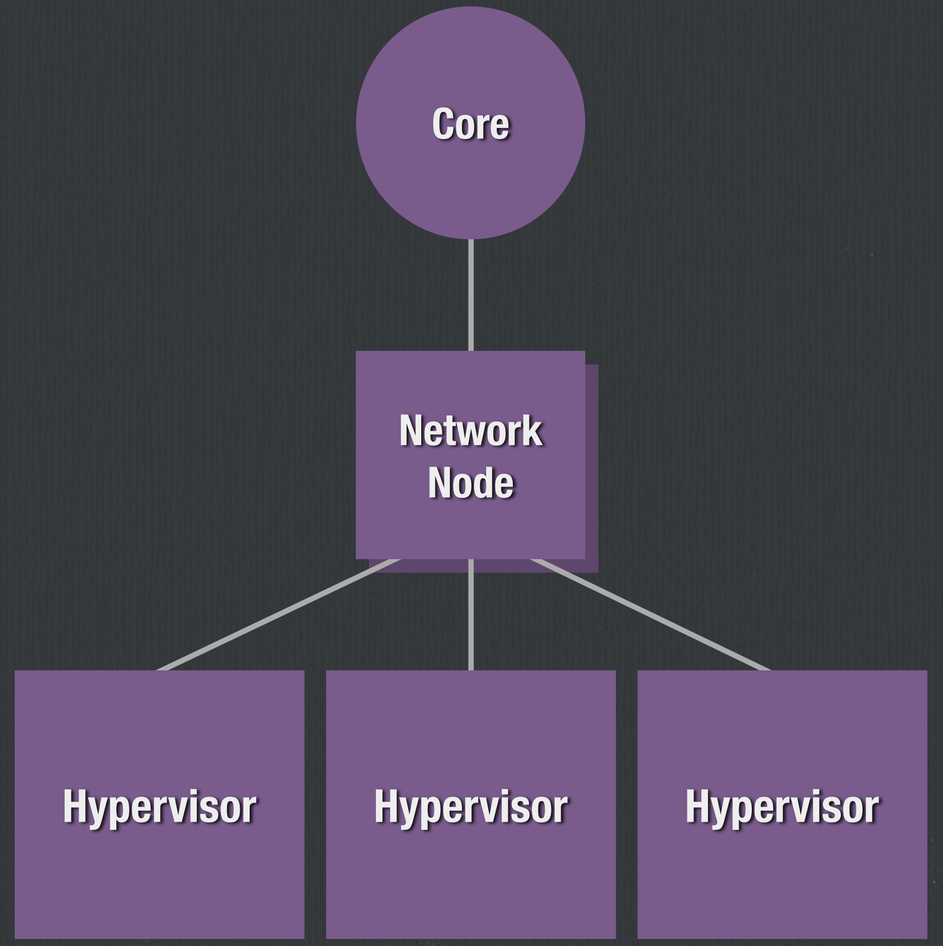
… with L2 Population



L3 Agents

L3 Agent

- Run on Network Node
- Uses Namespaces
- Metadata Agent (if enabled)



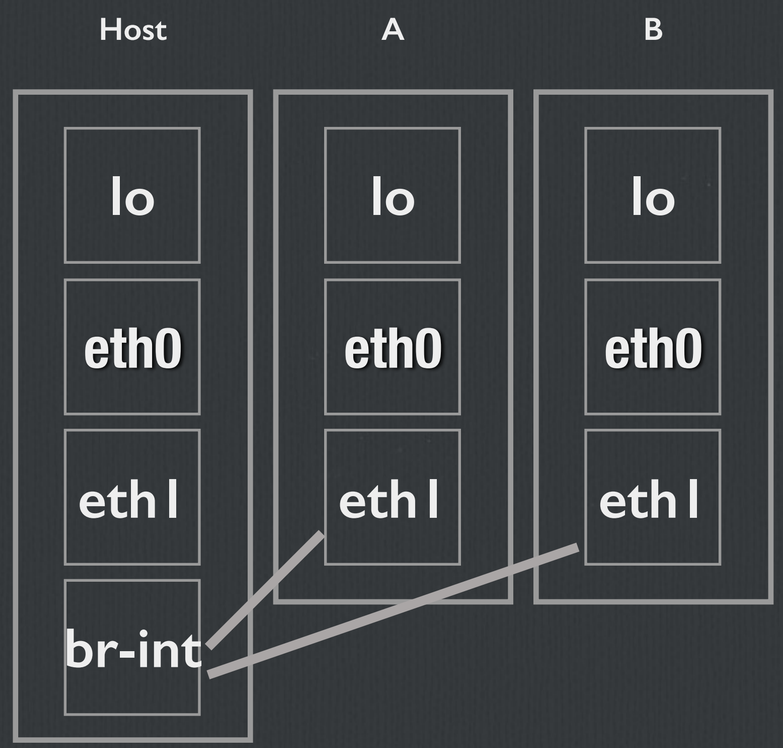
L3 Agent How it's implemented

- Manages Collection of Network Namespaces
  - Isolated IP Stacks
  - Forwarding Enabled

```
net.ipv4.ip_forward=1

net.ipv6.conf.all.forwarding=1
```

  - Static Routing
  - Metadata Proxy

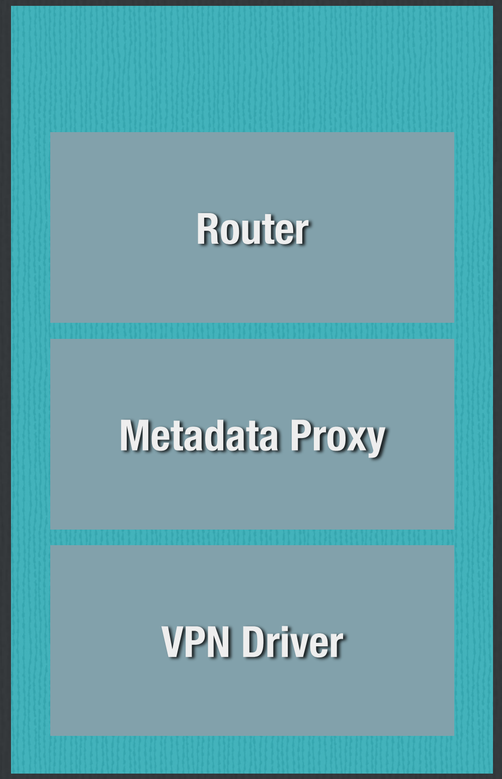


Load Balancer as a Service

- Service Plugin
  - Driver based
- Agent w/Driver
  - Agent communicates over RPC
  - Open Source requires namespaces
  - Others interact with other systems

VPN as a Service

- Service Plugin
  - Driver based
- Agent w/Driver
- Communicates over RPC
- Openswan



What's New in Juno

IPv6

IPv6: Basics

Router Advertisement Support

IPAM Algorithms:

- SLAAC
- Sequential

RA secured with security groups

IPv6: SLAAC

RA Autoconfiguration

IPv6 address generated from EUI-64 address

No DHCP

IPv6: DHCPv6 Stateless

Same as SLAAC

- IP Address from EUI-64 address

DHCP enables clients to review extra options

IPv6: DHCPv6 Stateful

Most similar to existing v4 support

Backed by dnsmasq and radvd

IPv6: Dual vs Single Stack

Dual Stack

- Applications have both v4/v6 access
- Support by latest long term support releases

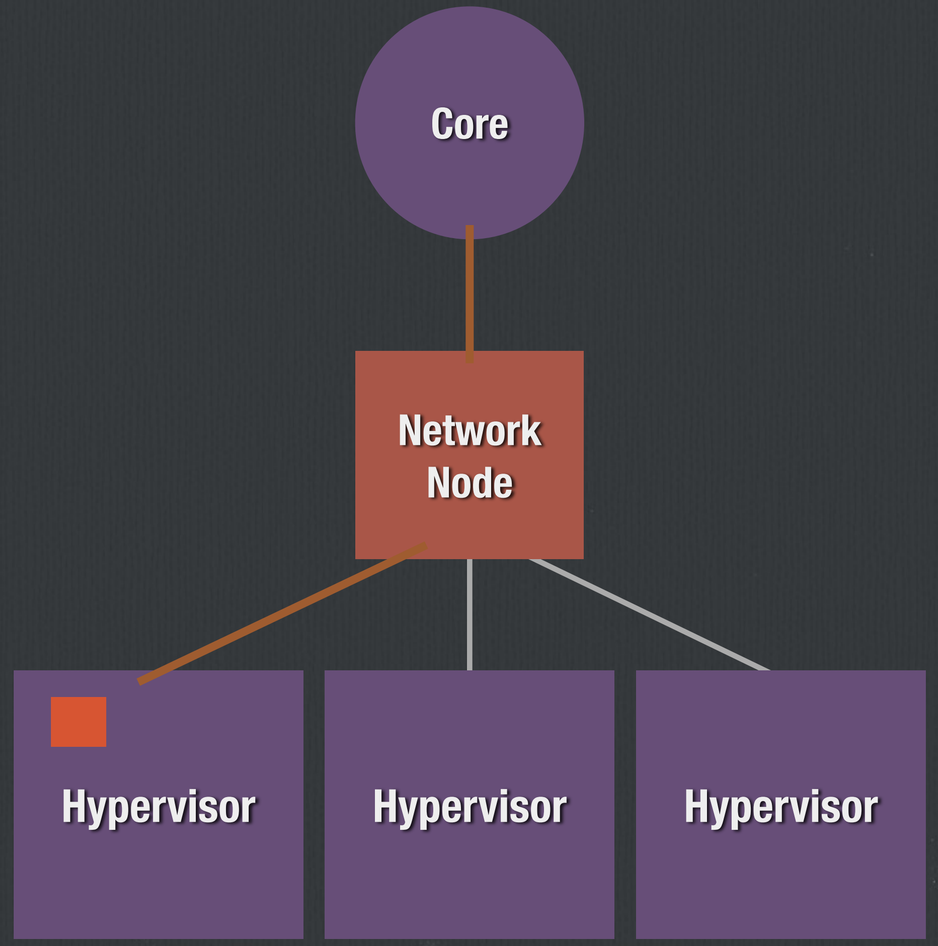
Single Stack v6

- Metadata service does not work
- Config drive required*

Distributed Virtual Routing

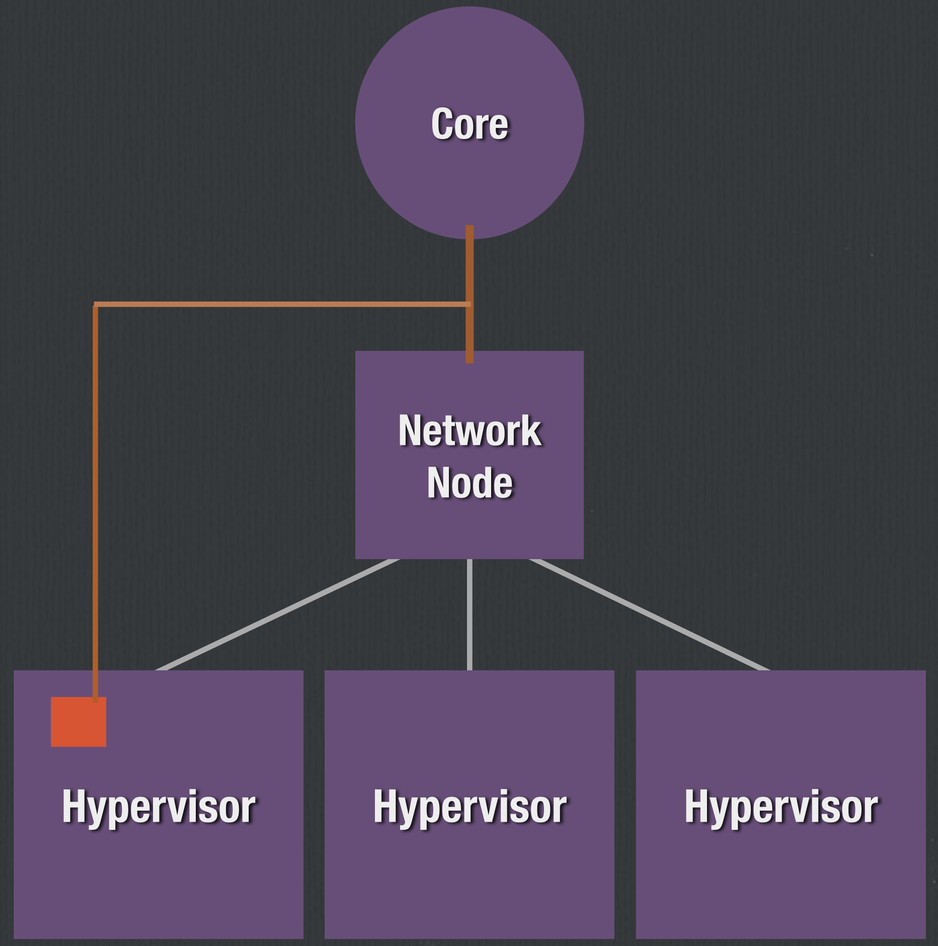
DVR: Before

- L3 Service run on Network Node
- Uses Namespaces
- Metadata Agent (if enabled)



DVR: After

- L3 Service run on Network Node
- Uses Namespaces
- Metadata Agent (if enabled)



DVR: How it works

1) Operator deploys DVR L3 Agent

    Agent runs on each Hypervisor

2) Associate floating IP with instance

3) Profit!!!

DVR: How it works

1) Operator deploys DVR L3 Agent

    Agent runs on each Hypervisor

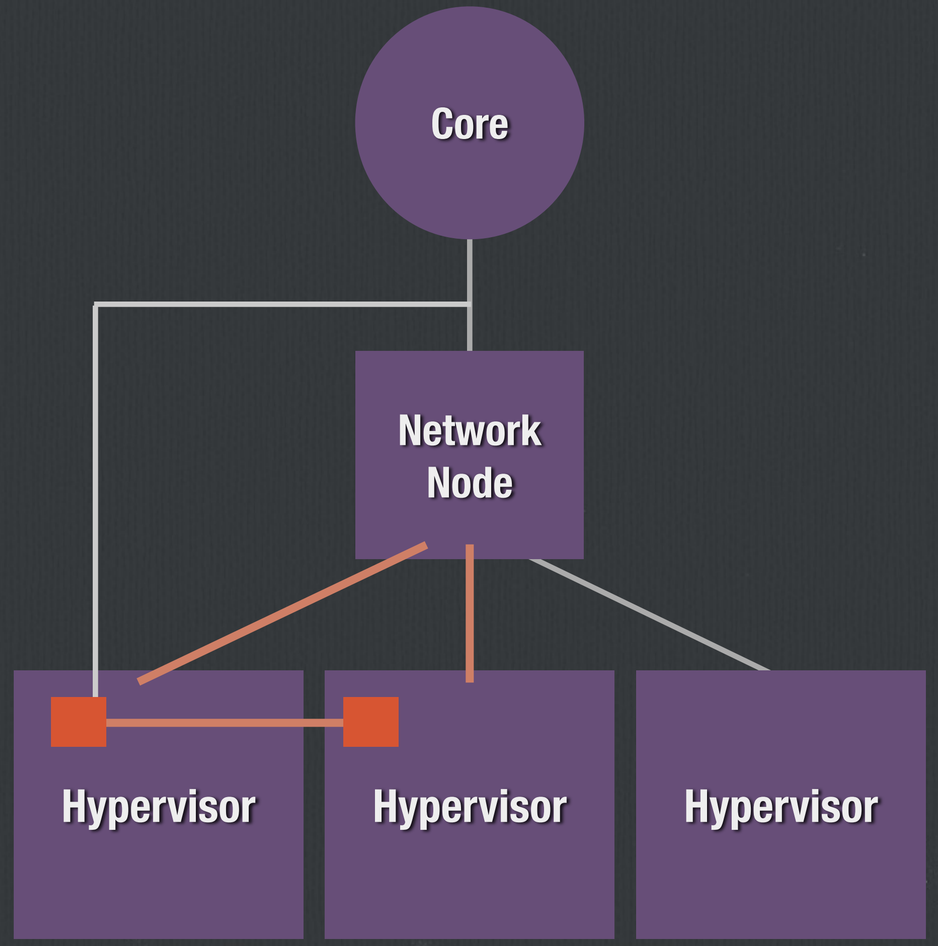
2) Associate floating IP with instance

~~3) Profit!!!~~

3) All N/S instance traffic is NAT’d directly from hypervisor

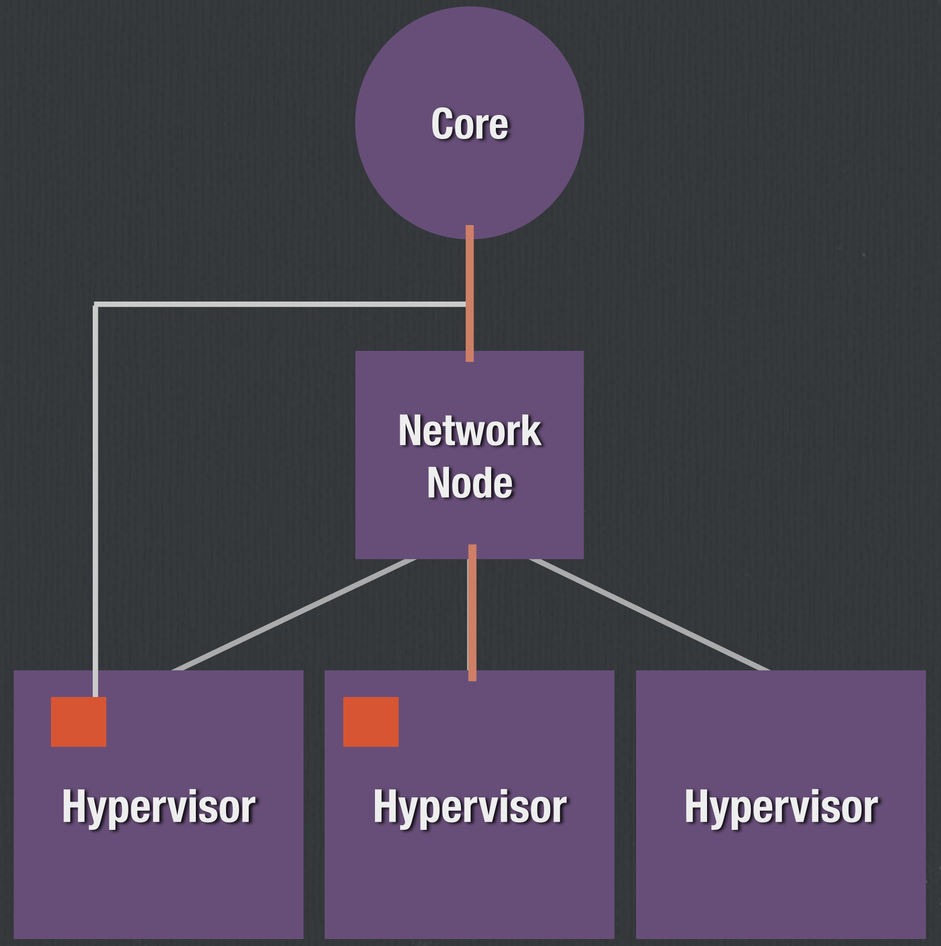
DVR: East/West

- L3 Service run on Network Node
- Uses Namespaces
- Metadata Agent (if enabled)



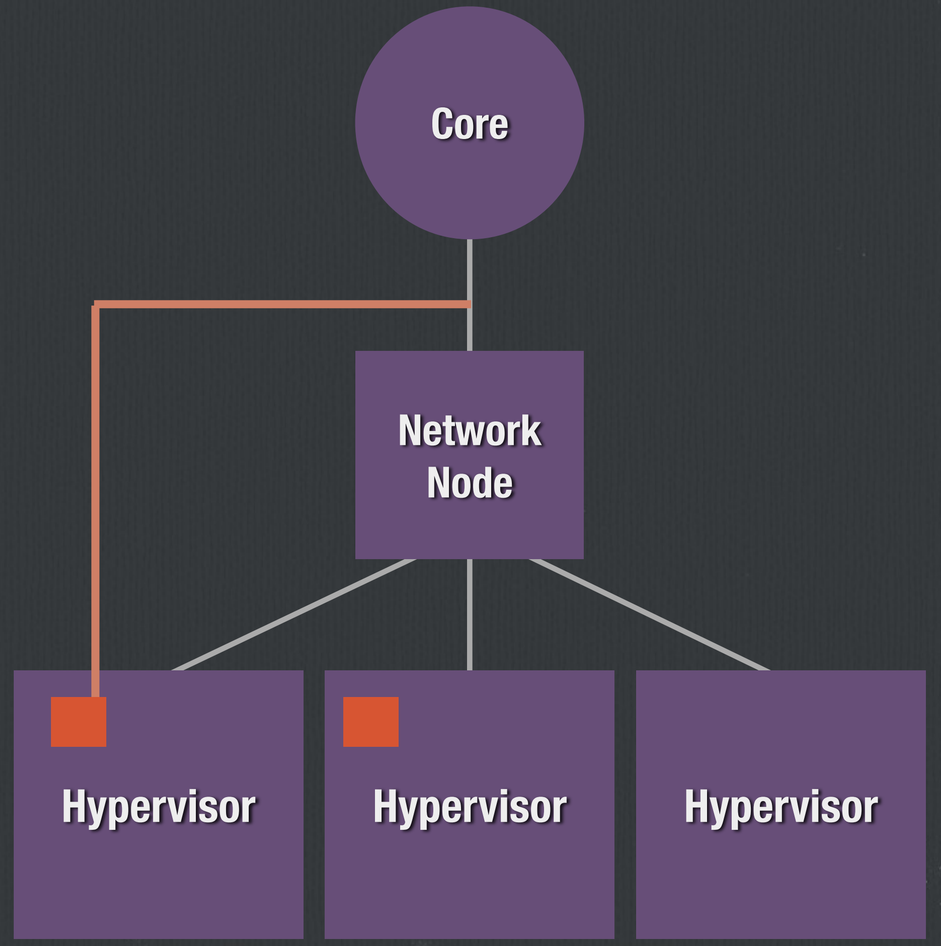
DVR: North/South SNAT w/o Floating IP

- L3 Service run on Network Node
- Uses Namespaces
- Metadata Agent (if enabled)



DVR: North/South SNAT w/ Floating IP

- L3 Service run on Network Node
- Uses Namespaces
- Metadata Agent (if enabled)



Other Improvements

- Security Groups
  - Now uses IPsets
- L3 Agent HA
  - via Namespace Pairs

Looking Ahead to Kilo

- Paying Down Technical Debt
- IPv6
  - Prefix delegation
  - Metadata Service
- IPAM
- Facilitate Dynamic Routing
- Enabling NFV Applications

Summary

Unified API

Small Core

Pluggable Open Architecture

Multiple Vendor Support

Extensible

More Information

- Cloud Administrator Guide
  - http://docs.openstack.org/admin-guide-cloud/content/ch_networking.html
- OpenStack Network v2.0 API
  - http://developer.openstack.org/api-ref-networking-v2.html

Thank You