
CERN


CERN

CERN DataCenter/Cloud Network

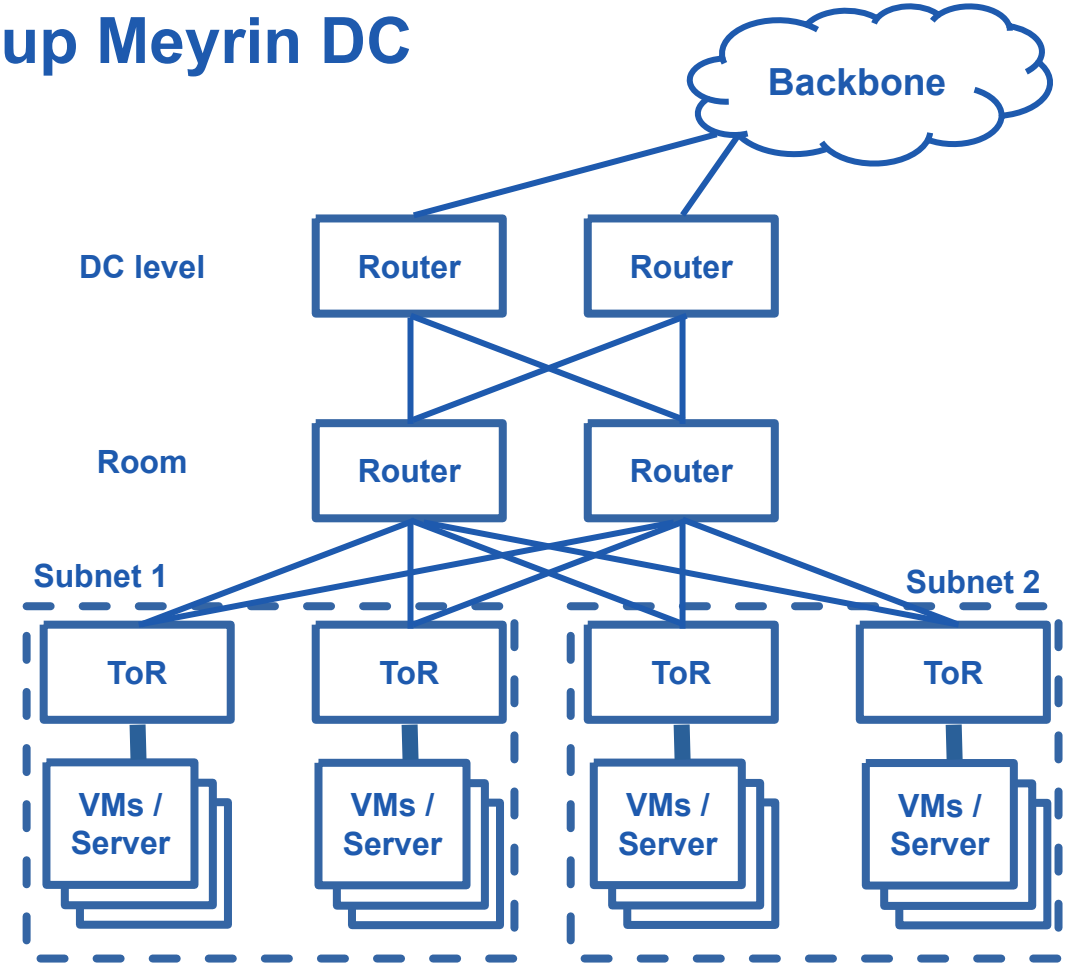
Overview

- DataCenter Network
- Cloud Network
- Feature set
- QA

DataCenter/CERN Network

Context – Current network setup Meyrin DC

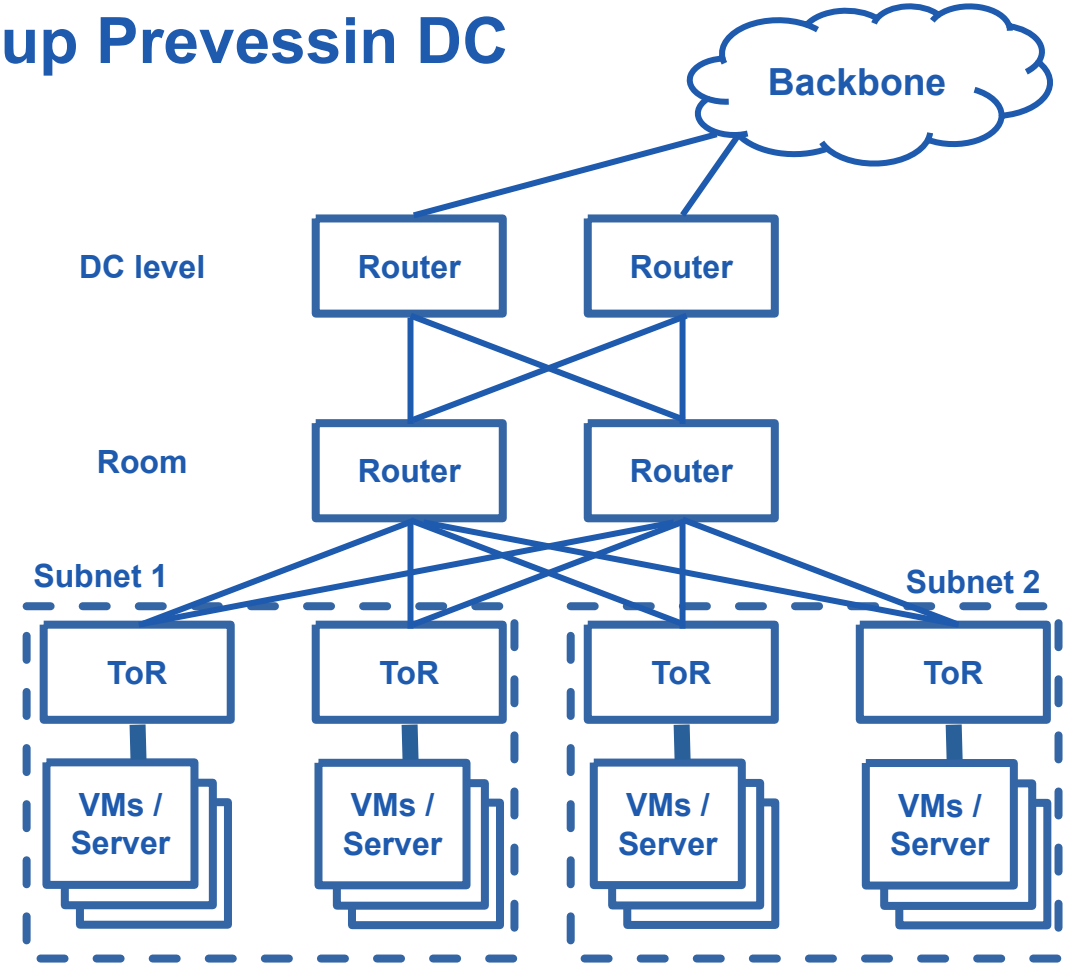
- Servers connected to
  - ToR or EoR (per density)
- Multiple routed L2 domains
- Spine/Leaf Routers
- BGP+some remaining OSPF in Spine
- Full Dual-Stack IPv4 / IPv6
- Mix of private and public IPs



Context – Current network setup Prevessin DC

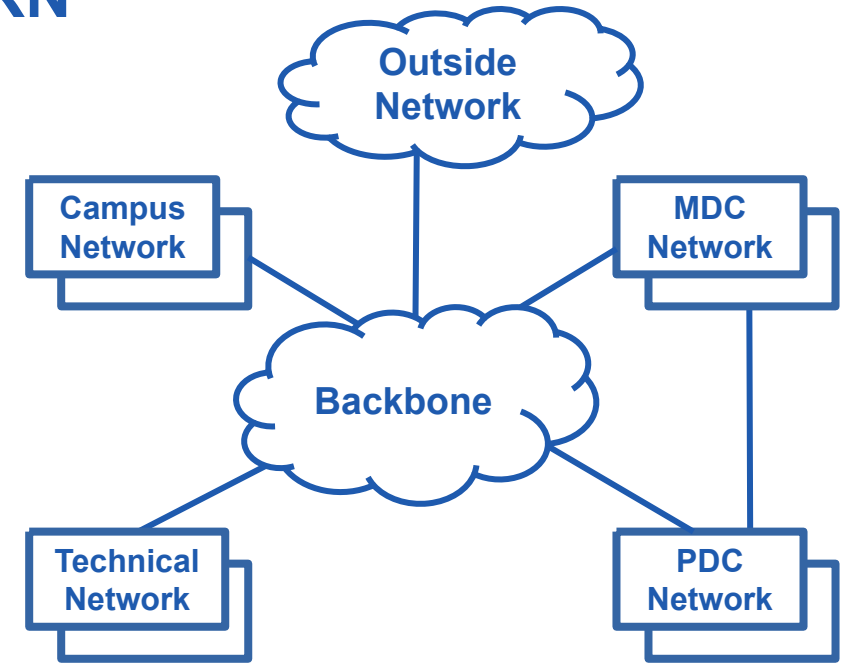
- Servers connected to
  - ToR
- Multiple routed L2 domains
- Spine/Leaf Routers
- BGP only
- Full Dual-Stack IPv4 / IPv6
- Mix of private and public IPs

=> Similar to MDC



Context – Current network setup CERN

- Multiple domains
- Routers distributed on different sites
- Different Firewalls + ACL
- All networks directly routable internally
  - Certain private networks



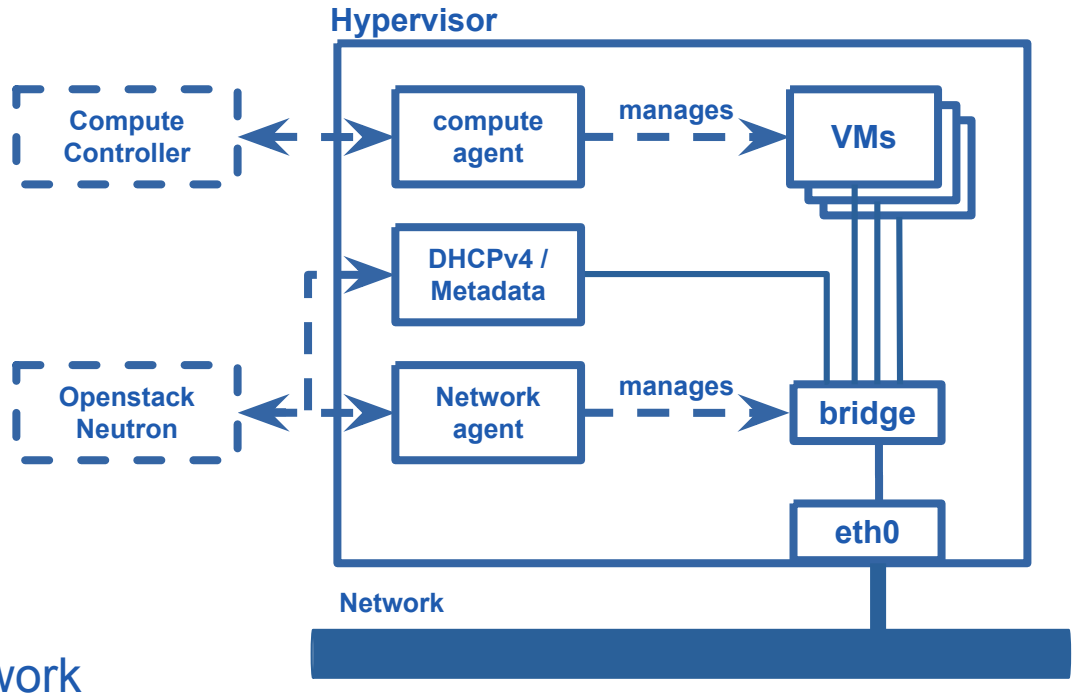
Cloud Network

Cloud Network - Physical Servers for users

- One IPMI network port
- One operating system port, one IP
  - Hypervisors have additional subnet on the same port for VMs
- IPs stay (mostly) during the full time
- Network not managed by Cloud team, using site wide DHCP
  - +PXE for setup with OpenStack Ironic

Cloud Network - Current offering for VMs (Meyrin)

- VMs connect over LinuxBridge
- Separated Subnets / Segmented
  - hidden from user
- Mantra: “Everything in same network”
  - => no E/W isolation



- User perspective:
  - 1 VM, 1 port, all in same public network

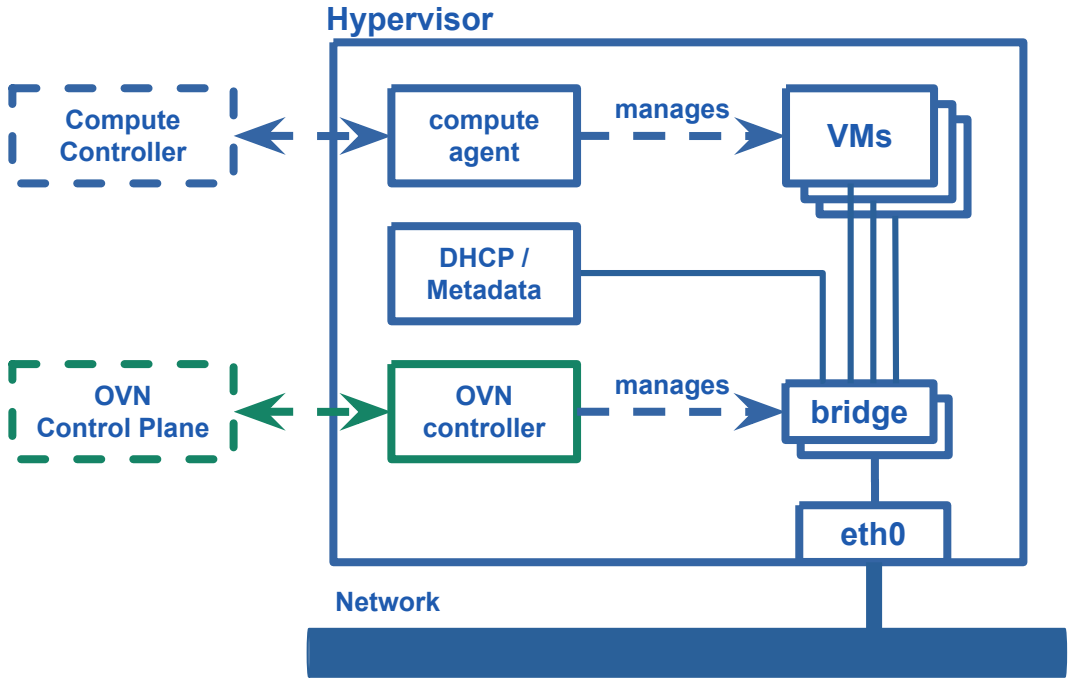
New requirements / Old limitations

- Add ability to migrate VMs between hypervisors in different subnets
- New hypervisors can easily host over 100 VMs
  - some performance issues seen in current setup with higher VM count
- LinuxBridge implementation upstream marked experimental / unsupported
- Different teams ask for
  - Private Networks
  - Security Groups (Firewall rules on HV level)
  - Floatings IPs

Cloud Network Setup (New/PDC)

- Open Virtual Network (OVN)
- Open VSwitch (OVS) per Hypervisor
- Central OVN Network DB for configuration
  - Flow rules in local OVS bridge
- DHCP, Metadata for VMs in Hypervisor

- Public networks leaves eth0 directly
- Private networks tunneled to target host



User Facing Features

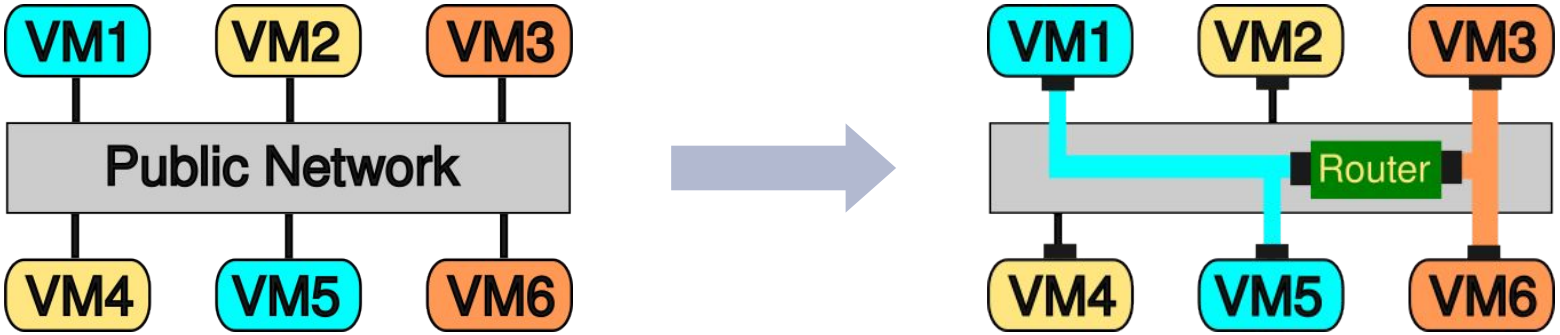
Different network types in PDC

- **Public/Provider** network (from start)
  - Keep existing functionality to simplify on-boarding
  - one subnet per 16 servers/2 racks (approx 1000 IPv4 + IPv6)

- **Private/Tenant** networks (Q2-Q3/2024)
  - Overlay network with OVN
  - Geneve tunnel between hypervisors

Private/Tenant Networks

- Isolated tunneled network creatable by user (Geneve tunnels)
- (Virtual) Routers to connect to other private or public networks
- For now limited to VMs in the setup



Security Groups

- Firewall for VMs on hypervisor level
- Allow certain groups of servers to talk to each other
- Whitelist approach
- Break of current mantra: “Everything can communicate with everything”
- Experience to be gained for large-scale deployment
  - Performance
  - User feedback

Load-Balancing as a Service

- Based on Openstack Octavia
- managed HaProxy in a VM
- Guaranteed fixed (public) IP address
- configurable via API
- centralized monitoring

Beyond / Plans

- Short term:
  - ➢ validating functionality with users
  - ➢ scalability test and gain confidence
- Migrate existing setup (old DC) to OVN (~15000 VMs, ~1700 hypervisors)
  - ➢ Migration path not straightforward
- Better integration with routers (e.g. BGP, EVPN)
  - ➢ Active/Active LBaaS
  - ➢ Floating IPs
  - ➢ Even greater flexibility to move VMs

Upstream Network Setups (Quick Summary)

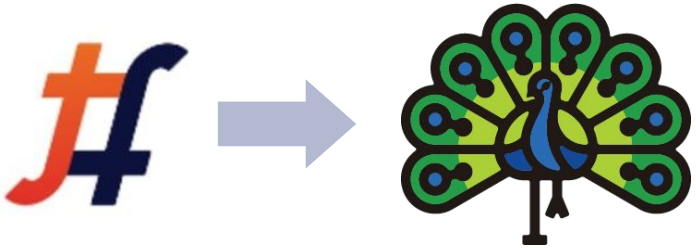
- Network Service: Openstack Neutron
- Support for multiple vendors:
  - LinuxBridge
    - ➢ currently in use for other DC, marked experimental now
  - OpenVSwitch (OVS)
    - ➢ widely used
  - Open Virtual Network (OVN)
    - ➢ more flexible, widely used and recommended upstream
  - hardware vendor specific drivers
    - ➢ potential vendor lock-in, typically only for hardware switches/routers

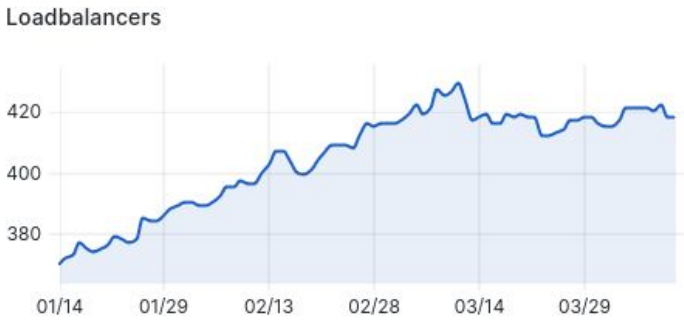
Q&A / Discussion

Backup Slides

Load-Balancing as a Service

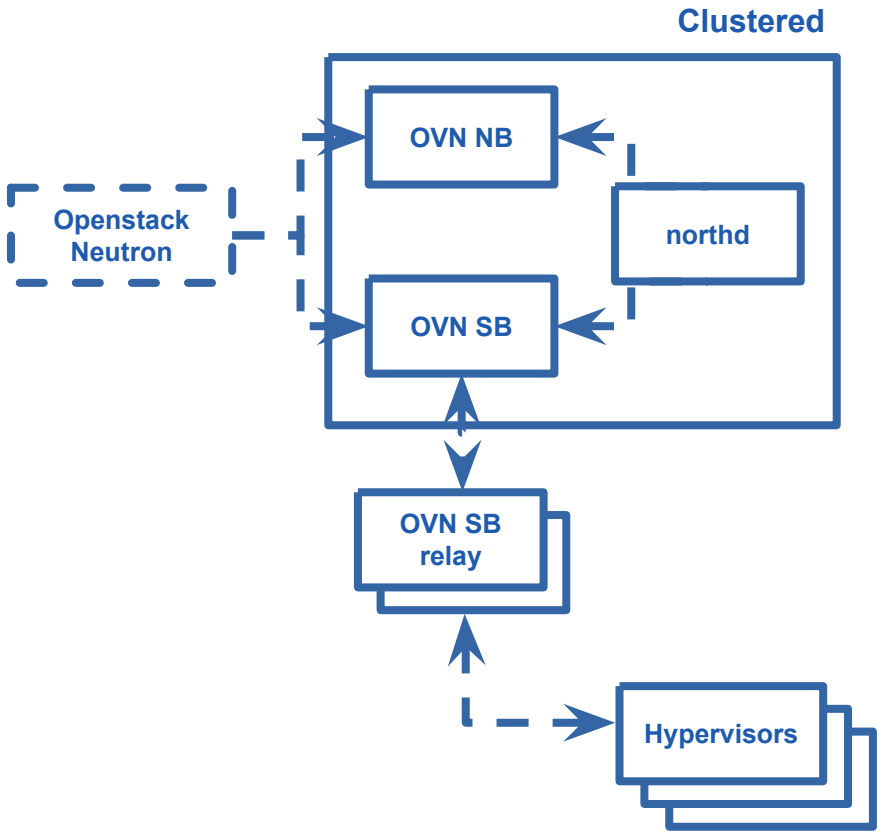
- Migrated out of TungstenFabric Q4/2023
  - Now based on OpenStack Octavia
- HaProxy in a VM
- Guaranteed fixed (public) IP address
- managed via API
- Users
  - Kubernetes Clusters, OpenShift
  - Windows Terminal Servers, Registry, …
  - Project with ATS
- Plan: Active/Active LBs in 2024





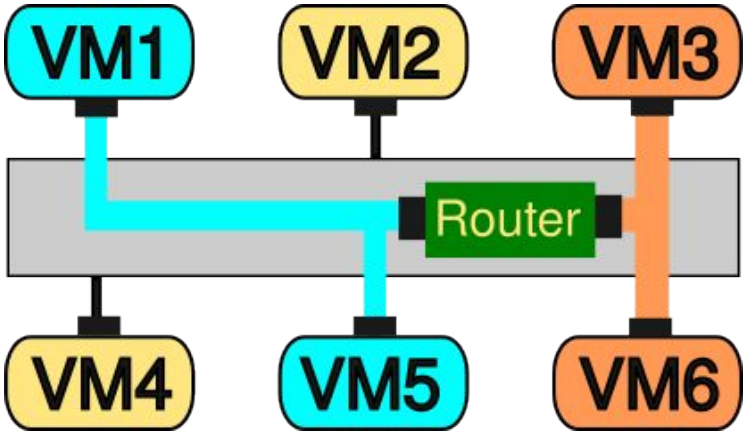
OVN Components

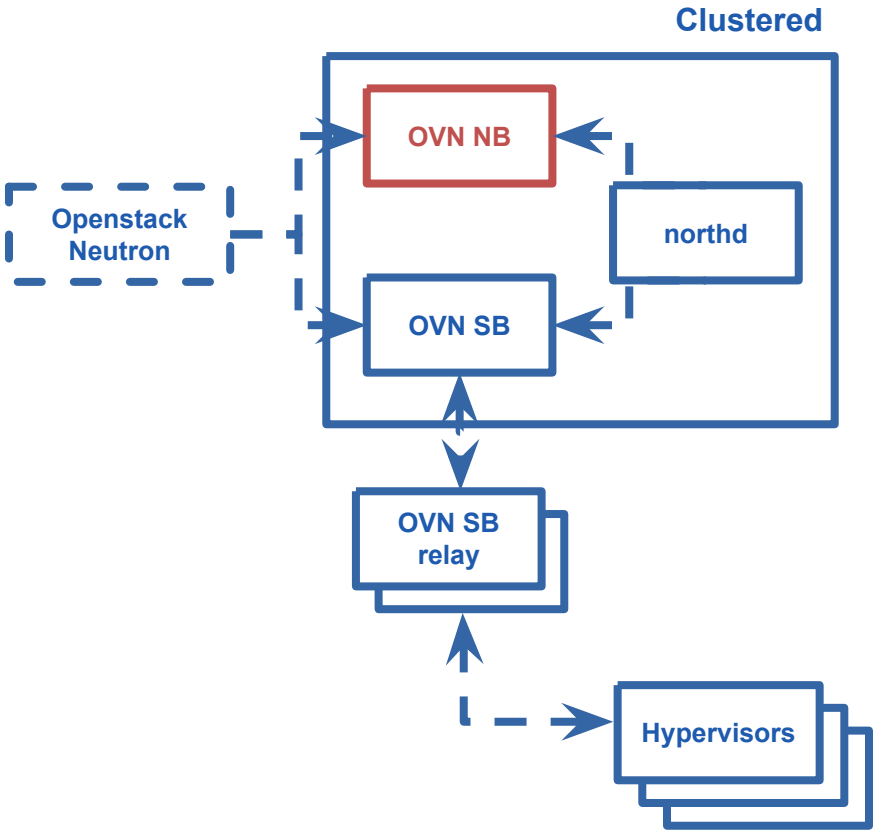
- OVN Northbound (NB) DB
  - ➢ “Port”, “Router”, “Switch”
- OVN Southbound (SB) DB
  - ➢ Hypervisor, Flow rules
- OVN northd
  - ➢ Translation between NB and SB
- OVN SB Relay
  - ➢ Relay/Cache for scaling



OVN Northbound

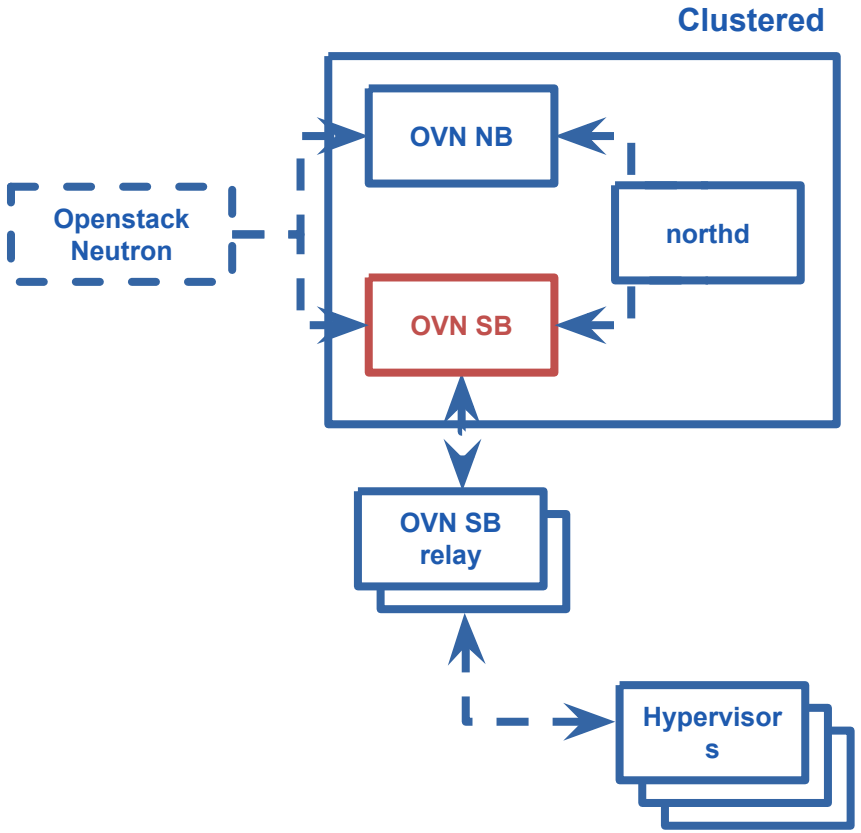
- Stores global abstract network view
- OVN NB: “Port”, “Router”, “Switch”





OVN Southbound

- OVN SB: Hypervisor, Flow rules
- Example FlowRule routed packet:
  - packet from port A
  - verify IP/MAC
  - check TTL
  - modify SRC IP
  - check firewall
  - send out port Z



Hypervisor

- OVN Controller + OpenVSwitch (OVS)
- Central OVN Network DB for configuration
  - Flow rules in local OVS bridge
- DHCP, Metadata for VMs in Hypervisor

- Public networks leaves eth0 directly
- Private networks tunneled to target host