



Red Hat OpenStack Platform 13

Networking Guide

An advanced guide to Red Hat OpenStack Platform Networking

Red Hat OpenStack Platform 13 Networking Guide

An advanced guide to Red Hat OpenStack Platform Networking

OpenStack Team
rhos-docs@redhat.com

Legal Notice

Copyright © 2024 Red Hat, Inc.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux[®] is the registered trademark of Linus Torvalds in the United States and other countries.

Java[®] is a registered trademark of Oracle and/or its affiliates.

XFS[®] is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL[®] is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js[®] is an official trademark of Joyent. Red Hat is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack[®] Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

Abstract

A cookbook for common OpenStack Networking tasks.

Table of Contents

MAKING OPEN SOURCE MORE INCLUSIVE	7
CHAPTER 1. NETWORKING OVERVIEW	8
1.1. HOW NETWORKING WORKS	8
1.1.1. VLANs	8
1.2. CONNECTING TWO LANS TOGETHER	8
1.2.1. Firewalls	9
1.3. RED HAT OPENSTACK NETWORK FLOW MATRIX	9
1.4. WORKING WITH OPENSTACK NETWORKING (NEUTRON)	10
1.5. WORKING WITH CIDR FORMAT	10
CHAPTER 2. OPENSTACK NETWORKING CONCEPTS	11
2.1. INSTALLING OPENSTACK NETWORKING (NEUTRON)	11
2.2. OPENSTACK NETWORKING DIAGRAM	11
2.3. SECURITY GROUPS	12
2.4. OPEN VSWITCH	12
2.5. CHANGING THE OPENFLOW INTERFACE FOR OPEN VSWITCH	13
2.6. MODULAR LAYER 2 (ML2) NETWORKING	14
2.6.1. The reasoning behind ML2	14
2.6.2. ML2 network types	14
2.6.3. ML2 mechanism drivers	15
2.7. ML2 TYPE AND MECHANISM DRIVER COMPATIBILITY	15
2.8. LIMITS OF THE ML2/OVN MECHANISM DRIVER	15
2.9. LIMIT FOR NON-SECURE PORTS WITH ML2/OVN	16
2.10. CONFIGURING THE L2 POPULATION DRIVER	17
2.11. OPENSTACK NETWORKING SERVICES	18
2.11.1. L3 agent	18
2.11.2. DHCP agent	18
2.11.3. Open vSwitch agent	18
2.12. PROJECT AND PROVIDER NETWORKS	18
2.12.1. Project networks	19
2.12.2. Provider networks	19
2.13. LAYER 2 AND LAYER 3 NETWORKING	19
2.13.1. Use switching where possible	20
CHAPTER 3. COMMON ADMINISTRATIVE NETWORKING TASKS	22
3.1. CREATING A NETWORK	22
3.2. ADDING NETWORK ROUTING	24
3.3. DELETING A NETWORK	24
3.4. PURGING ALL RESOURCES AND DELETING A PROJECT	25
3.5. WORKING WITH SUBNETS	25
3.5.1. Creating a subnet	26
3.6. DELETING A SUBNET	27
3.7. ADDING A ROUTER	28
3.8. DELETING A ROUTER	28
3.9. TUNING KEEPALIVED TO AVOID VRRP PACKET LOSS	28
3.10. ADDING AN INTERFACE	29
3.11. DELETING AN INTERFACE	30
3.12. CREATING FLOATING IP POOLS	30
3.13. ASSIGNING A SPECIFIC FLOATING IP	31
3.14. CREATING AN ADVANCED NETWORK	32
3.15. ASSIGNING A RANDOM FLOATING IP	33

3.16. CREATING MULTIPLE FLOATING IP POOLS	35
3.17. BRIDGING THE PHYSICAL NETWORK	35
3.18. SPECIFYING THE NAME THAT DNS ASSIGNS TO PORTS	36
3.19. ASSIGNING DHCP ATTRIBUTES TO PORTS	39
3.20. LOADING KERNEL MODULES	41
CHAPTER 4. PLANNING IP ADDRESS USAGE	43
4.1. VLAN PLANNING	43
4.2. TYPES OF NETWORK TRAFFIC	43
4.3. IP ADDRESS CONSUMPTION	45
4.4. VIRTUAL NETWORKING	45
4.5. EXAMPLE NETWORK PLAN	45
CHAPTER 5. REVIEWING OPENSTACK NETWORKING ROUTER PORTS	47
5.1. VIEWING CURRENT PORT STATUS	47
CHAPTER 6. TROUBLESHOOTING PROVIDER NETWORKS	49
6.1. BASIC PING TESTING	49
6.2. TROUBLESHOOTING VLAN NETWORKS	51
6.2.1. Reviewing the VLAN configuration and log files	51
6.3. TROUBLESHOOTING FROM WITHIN PROJECT NETWORKS	52
6.3.1. Performing advanced ICMP testing within the namespace	53
CHAPTER 7. CONNECTING AN INSTANCE TO THE PHYSICAL NETWORK	55
7.1. OVERVIEW OF THE OPENSTACK NETWORKING TOPOLOGY	55
7.2. PLACEMENT OF OPENSTACK NETWORKING SERVICES	55
7.3. CONFIGURING FLAT PROVIDER NETWORKS	56
7.4. HOW DOES THE FLAT PROVIDER NETWORK PACKET FLOW WORK?	58
7.5. TROUBLESHOOTING INSTANCE-PHYSICAL NETWORK CONNECTIONS ON FLAT PROVIDER NETWORKS	62
7.6. CONFIGURING VLAN PROVIDER NETWORKS	64
7.7. HOW DOES THE VLAN PROVIDER NETWORK PACKET FLOW WORK?	66
7.8. TROUBLESHOOTING INSTANCE-PHYSICAL NETWORK CONNECTIONS ON VLAN PROVIDER NETWORKS	70
7.9. ENABLING MULTICAST SNOOPING FOR PROVIDER NETWORKS IN AN ML2/OVS DEPLOYMENT	71
7.10. ENABLING MULTICAST IN AN ML2/OVN DEPLOYMENT	73
7.11. ENABLING COMPUTE METADATA ACCESS	75
7.12. FLOATING IP ADDRESSES	75
CHAPTER 8. CONFIGURING PHYSICAL SWITCHES FOR OPENSTACK NETWORKING	76
8.1. PLANNING YOUR PHYSICAL NETWORK ENVIRONMENT	76
8.2. CONFIGURING A CISCO CATALYST SWITCH	76
8.2.1. About trunk ports	76
8.2.2. Configuring trunk ports for a Cisco Catalyst switch	77
8.2.3. About access ports	78
8.2.4. Configuring access ports for a Cisco Catalyst switch	78
8.2.5. About LACP port aggregation	79
8.2.6. Configuring LACP on the physical NIC	79
8.2.7. Configuring LACP for a Cisco Catalyst switch	79
8.2.8. About MTU settings	80
8.2.9. Configuring MTU settings for a Cisco Catalyst switch	81
8.2.10. About LLDP discovery	82
8.2.11. Configuring LLDP for a Cisco Catalyst switch	82
8.3. CONFIGURING A CISCO NEXUS SWITCH	82

8.3.1. About trunk ports	82
8.3.2. Configuring trunk ports for a Cisco Nexus switch	83
8.3.3. About access ports	83
8.3.4. Configuring access ports for a Cisco Nexus switch	83
8.3.5. About LACP port aggregation	83
8.3.6. Configuring LACP on the physical NIC	84
8.3.7. Configuring LACP for a Cisco Nexus switch	84
8.3.8. About MTU settings	85
8.3.9. Configuring MTU settings for a Cisco Nexus 7000 switch	85
8.3.10. About LLDP discovery	85
8.3.11. Configuring LLDP for a Cisco Nexus 7000 switch	85
8.4. CONFIGURING A CUMULUS LINUX SWITCH	86
8.4.1. About trunk ports	86
8.4.2. Configuring trunk ports for a Cumulus Linux switch	86
8.4.3. About access ports	86
8.4.4. Configuring access ports for a Cumulus Linux switch	86
8.4.5. About LACP port aggregation	87
8.4.6. About MTU settings	87
8.4.7. Configuring MTU settings for a Cumulus Linux switch	87
8.4.8. About LLDP discovery	88
8.4.9. Configuring LLDP for a Cumulus Linux switch	88
8.5. CONFIGURING A EXTREME EXOS SWITCH	88
8.5.1. About trunk ports	88
8.5.2. Configuring trunk ports on an Extreme Networks EXOS switch	88
8.5.3. About access ports	89
8.5.4. Configuring access ports for an Extreme Networks EXOS switch	89
8.5.5. About LACP port aggregation	89
8.5.6. Configuring LACP on the physical NIC	90
8.5.7. Configuring LACP on an Extreme Networks EXOS switch	90
8.5.8. About MTU settings	90
8.5.9. Configuring MTU settings on an Extreme Networks EXOS switch	91
8.5.10. About LLDP discovery	91
8.5.11. Configuring LLDP settings on an Extreme Networks EXOS switch	91
8.6. CONFIGURING A JUNIPER EX SERIES SWITCH	91
8.6.1. About trunk ports	91
8.6.2. Configuring trunk ports for a Juniper EX Series switch	92
8.6.3. About access ports	92
8.6.4. Configuring access ports for a Juniper EX Series switch	92
8.6.5. About LACP port aggregation	93
8.6.6. Configuring LACP on the physical NIC	93
8.6.7. Configuring LACP for a Juniper EX Series switch	93
8.6.8. About MTU settings	95
8.6.9. Configuring MTU settings for a Juniper EX Series switch	95
8.6.10. About LLDP discovery	96
8.6.11. Configuring LLDP for a Juniper EX Series switch	96
CHAPTER 9. CONFIGURING MAXIMUM TRANSMISSION UNIT (MTU) SETTINGS	97
9.1. MTU OVERVIEW	97
9.2. CONFIGURING MTU SETTINGS IN DIRECTOR	98
9.3. REVIEWING THE RESULTING MTU CALCULATION	98
CHAPTER 10. CONFIGURING QUALITY OF SERVICE (QOS) POLICIES	99
10.1. QOS POLICY SCOPE	99

10.2. CREATING AND APPLYING A QOS POLICY AND RULE	99
10.3. DSCP MARKING FOR EGRESS TRAFFIC	100
10.4. RBAC FOR QOS POLICIES	103
CHAPTER 11. CONFIGURING BRIDGE MAPPINGS	104
11.1. OVERVIEW OF BRIDGE MAPPINGS	104
11.2. TRAFFIC FLOW	104
11.3. CONFIGURING BRIDGE MAPPINGS	104
11.4. MAINTAINING BRIDGE MAPPINGS FOR OVS	105
11.4.1. Cleaning up OVS patch ports manually	105
11.4.2. Cleaning up OVS patch ports automatically	106
CHAPTER 12. VLAN-AWARE INSTANCES	108
12.1. VLAN TRUNKS AND VLAN TRANSPARENT NETWORKS	108
12.2. REVIEWING THE TRUNK PLUG-IN	108
12.3. CREATING A TRUNK CONNECTION	108
12.4. ADDING SUBPORTS TO THE TRUNK	110
12.5. CONFIGURING AN INSTANCE TO USE A TRUNK	111
12.6. CONFIGURING NETWORKING SERVICE RPC TIMEOUT	113
12.7. UNDERSTANDING TRUNK STATES	114
CHAPTER 13. CONFIGURING RBAC POLICIES	116
13.1. OVERVIEW OF RBAC POLICIES	116
13.2. CREATING RBAC POLICIES	116
13.3. REVIEWING RBAC POLICIES	117
13.4. DELETING RBAC POLICIES	117
13.5. GRANTING RBAC POLICY ACCESS FOR EXTERNAL NETWORKS	118
CHAPTER 14. CONFIGURING DISTRIBUTED VIRTUAL ROUTING (DVR)	119
14.1. UNDERSTANDING DISTRIBUTED VIRTUAL ROUTING (DVR)	119
14.1.1. Overview of Layer 3 routing	119
14.1.2. Routing flows	119
14.1.3. Centralized routing	120
14.2. DVR OVERVIEW	120
14.3. DVR KNOWN ISSUES AND CAVEATS	120
14.4. SUPPORTED ROUTING ARCHITECTURES	121
14.5. DEPLOYING DVR WITH ML2 OVS	122
14.6. MIGRATING CENTRALIZED ROUTERS TO DISTRIBUTED ROUTING	123
CHAPTER 15. CONFIGURE LOAD BALANCING-AS-A-SERVICE WITH THE NETWORKING LBAASV2 API	125
15.1. OVERVIEW OF LBAAS	125
15.2. OPENSTACK NETWORKING AND LBAAS TOPOLOGY	126
15.2.1. Support Status of LBaaS	126
15.3. CONFIGURING LBAAS	126
CHAPTER 16. PROJECT NETWORKING WITH IPV6	129
16.1. IPV6 SUBNET OPTIONS	129
16.2. CREATE AN IPV6 SUBNET USING STATEFUL DHCPV6	130
CHAPTER 17. MANAGING PROJECT QUOTAS	133
17.1. CONFIGURING PROJECT QUOTAS	133
17.2. L3 QUOTA OPTIONS	133
17.3. FIREWALL QUOTA OPTIONS	133
17.4. SECURITY GROUP QUOTA OPTIONS	133
17.5. MANAGEMENT QUOTA OPTIONS	134

CHAPTER 18. CONFIGURING FIREWALL-AS-A-SERVICE (FWAAS)	135
18.1. OVERVIEW OF FIREWALL-AS-A-SERVICE (FWAAS)	135
18.2. ENABLING FIREWALL-AS-A-SERVICE (FWAAS)	135
18.3. CONFIGURING FIREWALL-AS-A-SERVICE (FWAAS)	136
18.4. CREATING FIREWALLS	137
CHAPTER 19. CONFIGURING ALLOWED ADDRESS PAIRS	138
19.1. OVERVIEW OF ALLOWED ADDRESS PAIRS	138
19.2. CREATING A PORT AND ALLOWING ONE ADDRESS PAIR	138
19.3. ADDING ALLOWED ADDRESS PAIRS	139
CHAPTER 20. CONFIGURING LAYER 3 HIGH AVAILABILITY (HA)	141
20.1. RHOSP NETWORKING SERVICE WITHOUT HIGH AVAILABILITY (HA)	141
20.2. OVERVIEW OF LAYER 3 HIGH AVAILABILITY (HA)	141
20.3. LAYER 3 HIGH AVAILABILITY (HA) FAILOVER CONDITIONS	142
20.4. PROJECT CONSIDERATIONS FOR LAYER 3 HIGH AVAILABILITY (HA)	142
20.5. HIGH AVAILABILITY (HA) CHANGES TO THE RHOSP NETWORKING SERVICE	142
20.6. ENABLING LAYER 3 HIGH AVAILABILITY (HA) ON RHOSP NETWORKING SERVICE NODES	143
20.7. REVIEWING HIGH AVAILABILITY (HA) RHOSP NETWORKING SERVICE NODE CONFIGURATIONS	144
CHAPTER 21. IDENTIFYING VIRTUAL DEVICES WITH TAGS	146
21.1. OVERVIEW OF VIRTUAL DEVICE TAGGING	146
21.2. TAGGING VIRTUAL DEVICES	146

MAKING OPEN SOURCE MORE INCLUSIVE

Red Hat is committed to replacing problematic language in our code, documentation, and web properties. We are beginning with these four terms: master, slave, blacklist, and whitelist. Because of the enormity of this endeavor, these changes will be implemented gradually over several upcoming releases. For more details, see [our CTO Chris Wright's message](#).

CHAPTER 1. NETWORKING OVERVIEW

The OpenStack Networking service (codename *neutron*) is the software-defined networking component of Red Hat OpenStack Platform 13.

Network administrators can use software-defined networking (SDN) to manage network services through abstraction of lower-level functionality. While server workloads have been migrated into virtual environments, they are still just servers that look for a network connection to send and receive data. SDN meets this need by moving networking equipment (such as routers and switches) into the same virtualized space. If you are already familiar with basic networking concepts, then it is easy to consider that these physical networking concepts have now been virtualized, just like the servers that they connect.

1.1. HOW NETWORKING WORKS

The term networking refers to the act of moving information from one computer to another. At the most basic level, this is performed by running a cable between two machines, each with network interface cards (NICs) installed. In the OSI networking model, the cable represents layer 1.

Now, if you want more than two computers to get involved in the conversation, you would need to scale out this configuration by adding a device called a switch. Enterprise switches have multiple Ethernet ports where you can connect additional machines. A network of multiple machines is called a Local Area Network (LAN).

Because they increase complexity, switches represent another layer of the OSI model, layer two. Each NIC has a unique MAC address number assigned to the hardware, and this number enables machines connected to the same switch to find each other. The switch maintains a list of which MAC addresses are plugged into which ports, so that when one computer attempts to send data to another, the switch knows where they are both situated, and adjusts entries in the CAM (Content Addressable Memory), which monitors of MAC-address-to-port mappings.

1.1.1. VLANs

You can use VLANs to segment network traffic for computers running on the same switch. This means that you can logically divide your switch by configuring the ports to be members of different networks – they are basically mini-LANs that you can use to separate traffic for security reasons.

For example, if your switch has 24 ports in total, you can assign ports 1-6 to VLAN200, and ports 7-18 to VLAN201. As a result, computers connected to VLAN200 are completely separate from those on VLAN201; they cannot communicate directly, and if they wanted to, the traffic must pass through a router as if they were two separate physical switches. Firewalls can also be useful for governing which VLANs can communicate with each other.

1.2. CONNECTING TWO LANS TOGETHER

If you have two LANs running on two separate switches, and you want them to share information with each other. You have two options for configuring this communication:

- **Use 802.1Q VLAN tagging to configure a single VLAN that spans across both physical switches:**

You must connect one end of a network cable to a port on one switch, connect the other end to a port on the other switch, and then configure these ports as 802.1Q tagged ports (sometimes known as trunk ports). These two switches act as one big logical switch, and the connected computers can find each other.

The downside to this option is scalability. You can only daisy-chain a limited number of switches until overhead becomes an issue.

- **Obtain a router and use cables to connect it to each switch:**

The router is aware of the networks configured on both switches. Each end of the cable plugged into the switch receives an IP address, known as the default gateway for that network. A default gateway defines the destination where traffic is sent when it is clear that the destination machine is not on the same LAN as the source machine. By establishing a default gateway, each computer can send traffic to other computers without knowing specific information about the destination. Each computer sends traffic to the default gateway, and the router determines which destination computer receives the traffic. Routing works on layer 3 of the OSI model, and is where the familiar concepts like IP addresses and subnets operate.

1.2.1. Firewalls

Firewalls can filter traffic across multiple OSI layers, including layer 7 (for inspecting actual content). Firewalls are often situated in the same network segments as routers, where they govern the traffic moving between all the networks. Firewalls refer to a predefined set of rules that prescribe which traffic can enter a network. These rules can become very granular, for example:

"Servers on **VLAN200** may only communicate with computers on **VLAN201**, and only on a Thursday afternoon, and only if they are sending encrypted web traffic (HTTPS) in one direction".

To help enforce these rules, some firewalls also perform Deep Packet Inspection (DPI) at layers 5-7, whereby they examine the contents of packets to ensure that the packets are legitimate. Hackers can exfiltrate data by having the traffic masquerade as something it is not. DPI is one of the means that you can use to mitigate that threat.

1.3. RED HAT OPENSTACK NETWORK FLOW MATRIX

The network flow matrix is a comma separated values (CSV) file that describes flows to and from OpenStack services.

NOTE: The network flow matrix describes common traffic flows. It does not describe every possible flow. Some flows that are not described in this matrix might be critical to operation. For instance, if you block all traffic and then selectively open only the flows described here, you might unintentionally block a necessary flow. That could cause issues that are difficult to troubleshoot.

The matrix describes flows in the following columns.

Service

The OpenStack service.

Protocol

Transmission protocol.

Dest. Port

Destination port.

Source Object

Source of data.

Dest. Object

Destination of data.

Source/Dest Pairs

Valid source and destination pairs.

Dest. Network

Destination network.

ServiceNetMap Parent

Determines the network type used for each service.

Traffic Description

Notes about the traffic flow.

Download the network flow matrix file from the following location:

[Red Hat OpenStack Network Flows](#).

1.4. WORKING WITH OPENSTACK NETWORKING (NEUTRON)

These same networking concepts apply in OpenStack, where they are known as Software-defined networking (SDN). The OpenStack Networking (neutron) component provides the API for virtual networking capabilities, and includes switches, routers, and firewalls. The virtual network infrastructure allows your instances to communicate with each other and also externally using the physical network. The Open vSwitch bridge allocates virtual ports to instances, and can span across the network infrastructure to the physical network for incoming and outgoing traffic.

1.5. WORKING WITH CIDR FORMAT

IP addresses are generally first allocated in blocks of subnets. For example, the IP address range **192.168.100.0 - 192.168.100.255** with a subnet mask of **255.255.255.0** allows for **254** IP addresses (the first and last addresses are reserved).

These subnets can be represented in a number of ways:

- **Common usage:**

Subnet addresses are traditionally displayed using the network address accompanied by the subnet mask:

- Network Address: *192.168.100.0*
- Subnet mask: *255.255.255.0*

- **CIDR format:**

The subnet mask is shortened into its total number of active bits.

For example, in **192.168.100.0/24**, **/24** is a shortened representation of **255.255.255.0**, and is a total of the number of flipped bits when converted to binary.

Also, CIDR format can be used in **ifcfg-xxx** scripts instead of the **NETMASK** value:

```
#NETMASK=255.255.255.0  
PREFIX=24
```

CHAPTER 2. OPENSTACK NETWORKING CONCEPTS

OpenStack Networking has system services to manage core services such as routing, DHCP, and metadata. Together, these services are included in the concept of the Controller node, which is a conceptual role assigned to a physical server.

A physical server is typically assigned the role of *Network node* and dedicated to the task of managing Layer 3 routing for network traffic to and from instances. In OpenStack Networking, you can have multiple physical hosts performing this role, allowing for redundant service in the event of hardware failure. For more information, see the chapter on *Layer 3 High Availability*.



NOTE

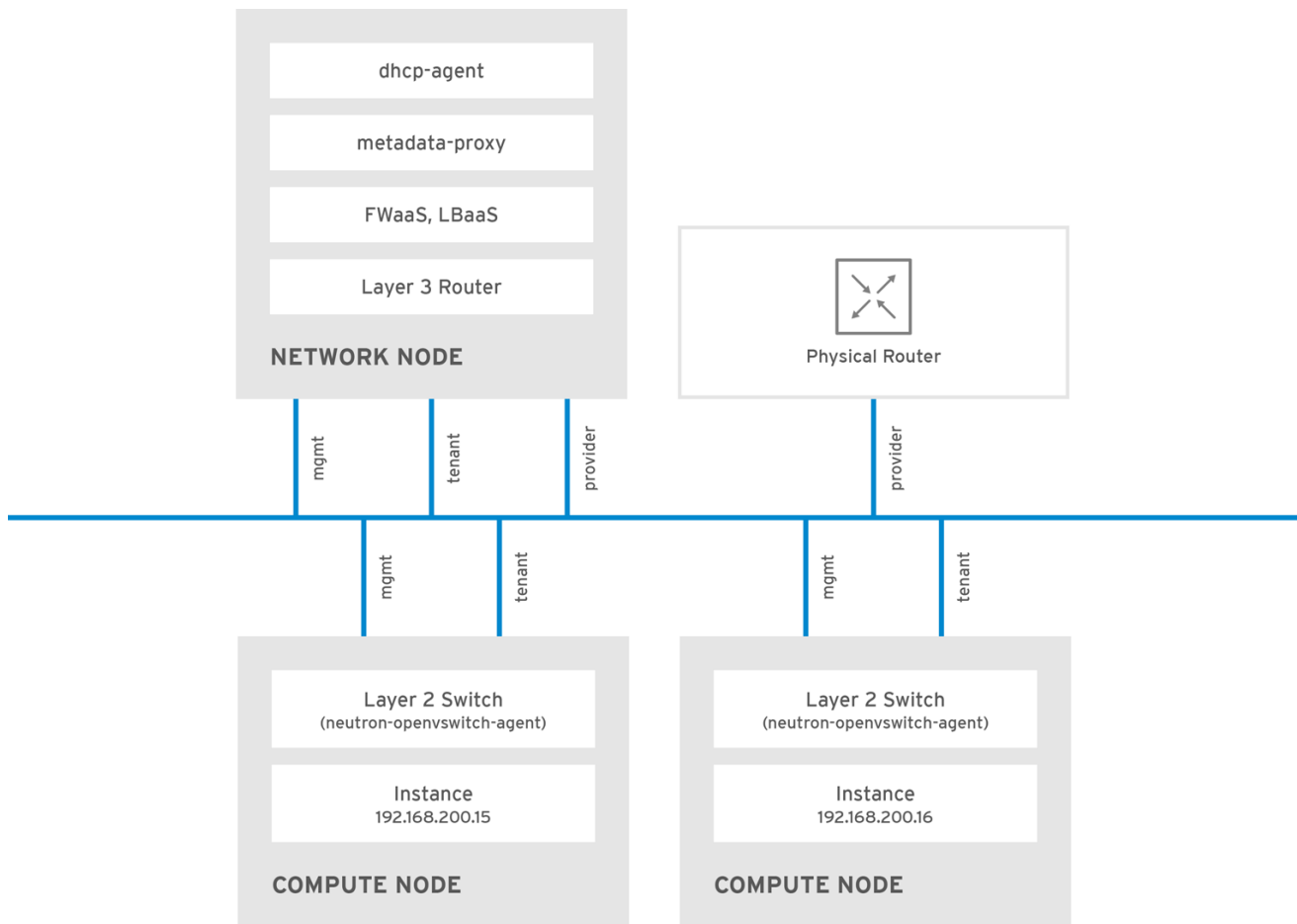
Red Hat OpenStack Platform 11 added support for composable roles, allowing you to separate network services into a custom role. However, for simplicity, this guide assumes that a deployment uses the default controller role.

2.1. INSTALLING OPENSTACK NETWORKING (NEUTRON)

The OpenStack Networking component is installed as part of a Red Hat OpenStack Platform director deployment. For more information about director deployment, see [Director Installation and Usage](#).

2.2. OPENSTACK NETWORKING DIAGRAM

This diagram depicts a sample OpenStack Networking deployment, with a dedicated OpenStack Networking node performing layer 3 routing and DHCP, and running the advanced service load balancing as a Service (LBaaS). Two Compute nodes run the Open vSwitch (openvswitch-agent) and have two physical network cards each, one for project traffic, and another for management connectivity. The OpenStack Networking node has a third network card specifically for provider traffic:



OPENSTACK_450456_0617

2.3. SECURITY GROUPS

Security groups and rules filter the type and direction of network traffic that neutron ports send and receive. This provides an additional layer of security to complement any firewall rules present on the compute instance. The security group is a container object with one or more security rules. A single security group can manage traffic to multiple compute instances.

Ports created for floating IP addresses, OpenStack Networking LBaaS VIPs, and instances are associated with a security group. If you do not specify a security group, then the port is associated with the default security group. By default, this group drops all inbound traffic and allows all outbound traffic. However, traffic flows between instances that are members of the default security group, because the group has a remote group ID that points to itself.

To change the filtering behavior of the default security group, you can add security rules to the group, or create entirely new security groups.

2.4. OPEN VSWITCH

Open vSwitch (OVS) is a software-defined networking (SDN) virtual switch similar to the Linux software bridge. OVS provides switching services to virtualized networks with support for industry standard, OpenFlow, and sFlow. OVS can also integrate with physical switches using layer 2 features, such as STP, LACP, and 802.1Q VLAN tagging. Open vSwitch version 1.11.0-1.el6 or later also supports tunneling with VXLAN and GRE.

For more information about network interface bonds, see the [Network Interface Bonding chapter of the Advanced Overcloud Customization guide](#).

**NOTE**

To mitigate the risk of network loops in OVS, only a single interface or a single bond can be a member of a given bridge. If you require multiple bonds or interfaces, you can configure multiple bridges.

**IMPORTANT**

Using single root I/O virtualization (SR-IOV) on bonded interfaces is not supported.

2.5. CHANGING THE OPENFLOW INTERFACE FOR OPEN VSWITCH

In Red Hat OpenStack Platform 13, the Networking service (neutron) uses Python 2.7 which does not work well with the **python-ryu** library that Open vSwitch depends on for managing OpenFlow rules.

If you experience timeouts when the neutron Open vSwitch (OVS) agent connects to OVS, then you must change the value for the OpenFlow interface and OVS database options.

Prerequisites

- You are using Open vSwitch in RHOSP 13.

Procedure

1. On the undercloud host, logged in as the stack user, create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-ovs-environment.yaml
```

TIP

The Orchestration service (heat) uses a set of plans called templates to install and configure your environment. You can customize aspects of the overcloud with a custom environment file, which is a special type of template that provides customization for your heat templates.

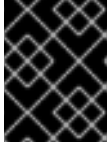
2. In the YAML environment file under **parameter_defaults**, add the following Puppet variables:

```
parameter_defaults:
  ExtraConfig:
    neutron::agents::ml2::ovs::of_interface: ovs-ofctl
    neutron::agents::ml2::ovs::ovsdb_interface: vsctl
    ...
```

**IMPORTANT**

Ensure that you add a whitespace character between the single colon (:) and the value.

3. Run the **openstack overcloud deploy** command and include the core heat templates, environment files, and this new custom environment file.



IMPORTANT

The order of the environment files is important as the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-ovs-
environment.yaml
```

Additional resources

- [Puppet: Customizing Hieradata for Individual Nodes](#) in the *Advanced Overcloud Customization* guide
- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

2.6. MODULAR LAYER 2 (ML2) NETWORKING

ML2 is the OpenStack Networking core plug-in introduced in the OpenStack Havana release. Superseding the previous model of monolithic plug-ins, the ML2 modular design enables the concurrent operation of mixed network technologies. The monolithic Open vSwitch and Linux Bridge plug-ins have been deprecated and removed; their functionality is now implemented by ML2 mechanism drivers.



NOTE

ML2 is the default OpenStack Networking plug-in, with OVN configured as the default mechanism driver.

2.6.1. The reasoning behind ML2

Previously, OpenStack Networking deployments could use only the plug-in selected at implementation time. For example, a deployment running the Open vSwitch (OVS) plug-in was required to use the OVS plug-in exclusively. The monolithic plug-in did not support the simultaneous use of another plug-in such as linuxbridge. This limitation made it difficult to meet the needs of environments with heterogeneous requirements.

2.6.2. ML2 network types

Multiple network segment types can be operated concurrently. In addition, these network segments can interconnect using ML2 support for multi-segmented networks. Ports are automatically bound to the segment with connectivity; it is not necessary to bind ports to a specific segment. Depending on the mechanism driver, ML2 supports the following network segment types:

- flat
- GRE
- local

- VLAN
- VXLAN
- Geneve

Enable Type drivers in the ML2 section of the `ml2_conf.ini` file. For example:

```
[ml2]
type_drivers = local,flat,vlan,gre,vxlan,geneve
```

2.6.3. ML2 mechanism drivers

Plug-ins are implemented as mechanisms with a common code base. This approach enables code reuse and eliminates much of the complexity around code maintenance and testing.

The default mechanism driver is OVN. You enable mechanism drivers using the Orchestration service (heat) parameter, `NeutronMechanismDrivers`. Here is an example from a heat custom environment file:

```
parameter_defaults:
...
NeutronMechanismDrivers: ansible,ovn,baremetal
...
```

The order in which you specify the mechanism drivers matters. In the earlier example, if you want to bind a port using the baremetal mechanism driver, then you must specify `baremetal` before `ansible`. Otherwise, the ansible driver will bind the port, because it precedes `baremetal` in the list of values for `NeutronMechanismDrivers`.

Additional resources

- [Neutron](#) in *Component, Plug-In, and Driver Support in Red Hat OpenStack Platform*
- [Environment Files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

2.7. ML2 TYPE AND MECHANISM DRIVER COMPATIBILITY

Mechanism Driver	Type Driver				
	flat	gre	vlan	vxlan	geneve
ovn	yes	no	yes	no	yes
openvswitch	yes	yes	yes	yes	no

2.8. LIMITS OF THE ML2/OVN MECHANISM DRIVER

The following table describes features that Red Hat does not yet support with ML2/OVN. Red Hat plans to support each of these features in a future Red Hat OpenStack Platform release.

In addition, this release of the Red Hat OpenStack Platform (RHOSP) does not provide a supported migration from the ML2/OVS mechanism driver to the ML2/OVN mechanism driver. This RHOSP release does not support the OpenStack community migration strategy. Migration support is planned for a future RHOSP release.

Feature	Notes	Track this Feature
Fragmentation / Jumbo Frames	OVN does not yet support sending ICMP "fragmentation needed" packets. Larger ICMP/UDP packets that require fragmentation do not work with ML2/OVN as they would with the ML2/OVS driver implementation. TCP traffic is handled by maximum segment sized (MSS) clamping.	https://bugzilla.redhat.com/show_bug.cgi?id=1547074 (ovn-network) https://bugzilla.redhat.com/show_bug.cgi?id=1702331 (Core ovn)
Port Forwarding	OVN does not support port forwarding.	https://bugzilla.redhat.com/show_bug.cgi?id=1654608 https://blueprints.launchpad.net/neutron/+spec/port-forwarding
Security Groups Logging API	ML2/OVN does not provide a log file that logs security group events such as an instance trying to execute restricted operations or access restricted ports in remote servers.	https://bugzilla.redhat.com/show_bug.cgi?id=1619266
Multicast	When using ML2/OVN as the integration bridge, multicast traffic is treated as broadcast traffic. The integration bridge operates in FLOW mode, so IGMP snooping is not available. To support this, core OVN must support IGMP snooping.	https://bugzilla.redhat.com/show_bug.cgi?id=1672278
SR-IOV	Presently, SR-IOV only works with the neutron DHCP agent deployed.	https://bugzilla.redhat.com/show_bug.cgi?id=1666684
Provisioning Baremetal Machines with OVN DHCP	The built-in DHCP server on OVN presently can not provision baremetal nodes. It cannot serve DHCP for the provisioning networks. Chainbooting iPXE requires tagging (--dhcp-match in dnsmasq), which is not supported in the OVN DHCP server.	https://bugzilla.redhat.com/show_bug.cgi?id=1622154
OVS_DPDK	OVS_DPDK is presently not supported with OVN.	

2.9. LIMIT FOR NON-SECURE PORTS WITH ML2/OVN

Ports might become unreachable if you disable the port security plug-in extension in Red Hat Open Stack Platform (RHOSP) deployments with the default ML2/OVN mechanism driver and a large number of ports.

In some large ML2/OVN RHOSP deployments, a flow chain limit inside ML2/OVN can drop ARP requests that are targeted to ports where the security plug-in is disabled.

There is no documented maximum limit for the actual number of logical switch ports that ML2/OVN can support, but the limit approximates 4,000 ports.

Attributes that contribute to the approximated limit are the number of resubmits in the OpenFlow pipeline that ML2/OVN generates, and changes to the overall logical topology.

2.10. CONFIGURING THE L2 POPULATION DRIVER

The L2 Population driver enables broadcast, multicast, and unicast traffic to scale out on large overlay networks. By default, Open vSwitch GRE and VXLAN replicate broadcasts to every agent, including those that do not host the destination network. This design requires the acceptance of significant network and processing overhead. The alternative design introduced by the L2 Population driver implements a partial mesh for ARP resolution and MAC learning traffic; it also creates tunnels for a particular network only between the nodes that host the network. This traffic is sent only to the necessary agent by encapsulating it as a targeted unicast.

To enable the L2 Population driver, complete the following steps:

1. Enable the L2 population driver by adding it to the list of mechanism drivers. You also must enable at least one tunneling driver enabled; either GRE, VXLAN, or both. Add the appropriate configuration options to the `ml2_conf.ini` file:

```
[ml2]
type_drivers = local,flat,vlan,gre,vxlan
mechanism_drivers = openvswitch,linuxbridge,l2population
```



NOTE

Neutron's Linux Bridge ML2 driver and agent were deprecated in Red Hat OpenStack Platform 11. The Open vSwitch (OVS) plugin OpenStack Platform director default, and is recommended by Red Hat for general usage.

2. Enable L2 population in the `openvswitch_agent.ini` file. Enable it on each node that contains the L2 agent:

```
[agent]
l2_population = True
```



NOTE

To install ARP reply flows, configure the **arp_responder** flag:

```
[agent]
l2_population = True
arp_responder = True
```

2.11. OPENSTACK NETWORKING SERVICES

By default, Red Hat OpenStack Platform includes components that integrate with the ML2 and Open vSwitch plugin to provide networking functionality in your deployment:

2.11.1. L3 agent

The L3 agent is part of the **openstack-neutron** package. Use network namespaces to provide each project with its own isolated layer 3 routers, which direct traffic and provide gateway services for the layer 2 networks. The L3 agent assists with managing these routers. The nodes that host the L3 agent must not have a manually-configured IP address on a network interface that is connected to an external network. Instead there must be a range of IP addresses from the external network that are available for use by OpenStack Networking. Neutron assigns these IP addresses to the routers that provide the link between the internal and external networks. The IP range that you select must be large enough to provide a unique IP address for each router in the deployment as well as each floating IP.

2.11.2. DHCP agent

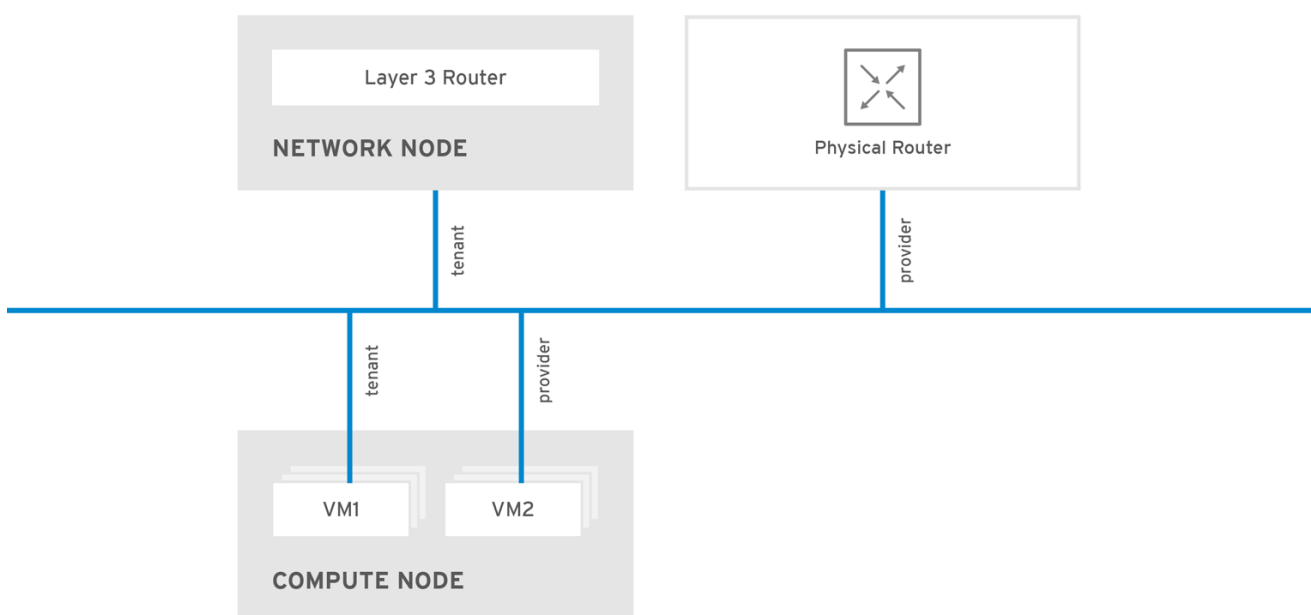
The OpenStack Networking DHCP agent manages the network namespaces that are spawned for each project subnet to act as DHCP server. Each namespace runs a dnsmasq process that can allocate IP addresses to virtual machines on the network. If the agent is enabled and running when a subnet is created then by default that subnet has DHCP enabled.

2.11.3. Open vSwitch agent

The Open vSwitch (OVS) neutron plug-in uses its own agent, which runs on each node and manages the OVS bridges. The ML2 plugin integrates with a dedicated agent to manage L2 networks. By default, Red Hat OpenStack Platform uses **ovs-agent**, which builds overlay networks using OVS bridges.

2.12. PROJECT AND PROVIDER NETWORKS

The following diagram presents an overview of the project and provider network types, and illustrates how they interact within the overall OpenStack Networking topology:



OPENSTACK_450456_0617

2.12.1. Project networks

Users create project networks for connectivity within projects. Project networks are fully isolated by default and are not shared with other projects. OpenStack Networking supports a range of project network types:

- **Flat** - All instances reside on the same network, which can also be shared with the hosts. No VLAN tagging or other network segregation occurs.
- **VLAN** - OpenStack Networking allows users to create multiple provider or project networks using VLAN IDs (802.1Q tagged) that correspond to VLANs present in the physical network. This allows instances to communicate with each other across the environment. They can also communicate with dedicated servers, firewalls, load balancers and other network infrastructure on the same layer 2 VLAN.
- **VXLAN and GRE tunnels**- VXLAN and GRE use network overlays to support private communication between instances. An OpenStack Networking router is required to enable traffic to traverse outside of the GRE or VXLAN project network. A router is also required to connect directly-connected project networks with external networks, including the Internet; the router provides the ability to connect to instances directly from an external network using floating IP addresses. VXLAN and GRE type drivers are compatible with the ML2/OVS mechanism driver.
- **GENEVE tunnels** - GENEVE recognizes and accommodates changing capabilities and needs of different devices in network virtualization. It provides a framework for tunneling rather than being prescriptive about the entire system. Geneve defines the content of the metadata flexibly that is added during encapsulation and tries to adapt to various virtualization scenarios. It uses UDP as its transport protocol and is dynamic in size using extensible option headers. Geneve supports unicast, multicast, and broadcast. The GENEVE type driver is compatible with the ML2/OVN mechanism driver.

Additional resources

- [Configuring Quality of Service \(QoS\) policies](#)

2.12.2. Provider networks

The OpenStack administrator creates provider networks. Provider networks map directly to an existing physical network in the data center. Useful network types in this category include flat (untagged) and VLAN (802.1Q tagged). You can also share provider networks among projects as part of the network creation process.

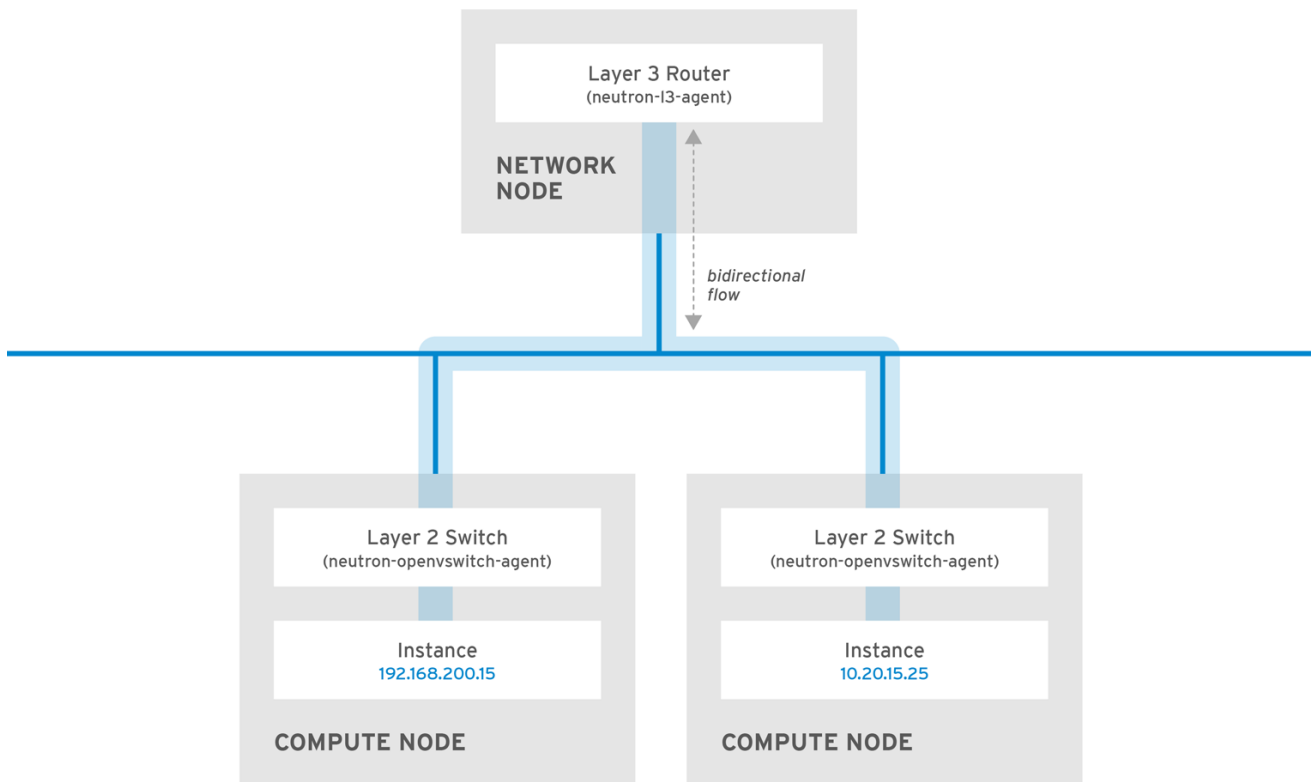
Additional resources

- [Configuring flat provider networks](#)
- [Configuring VLAN provider networks](#)

2.13. LAYER 2 AND LAYER 3 NETWORKING

When designing your virtual network, anticipate where the majority of traffic is going to be sent. Network traffic moves faster within the same logical network, rather than between multiple logical networks. This is because traffic between logical networks (using different subnets) must pass through a router, resulting in additional latency.

Consider the diagram below which has network traffic flowing between instances on separate VLANs:



OPENSTACK_450456_0617



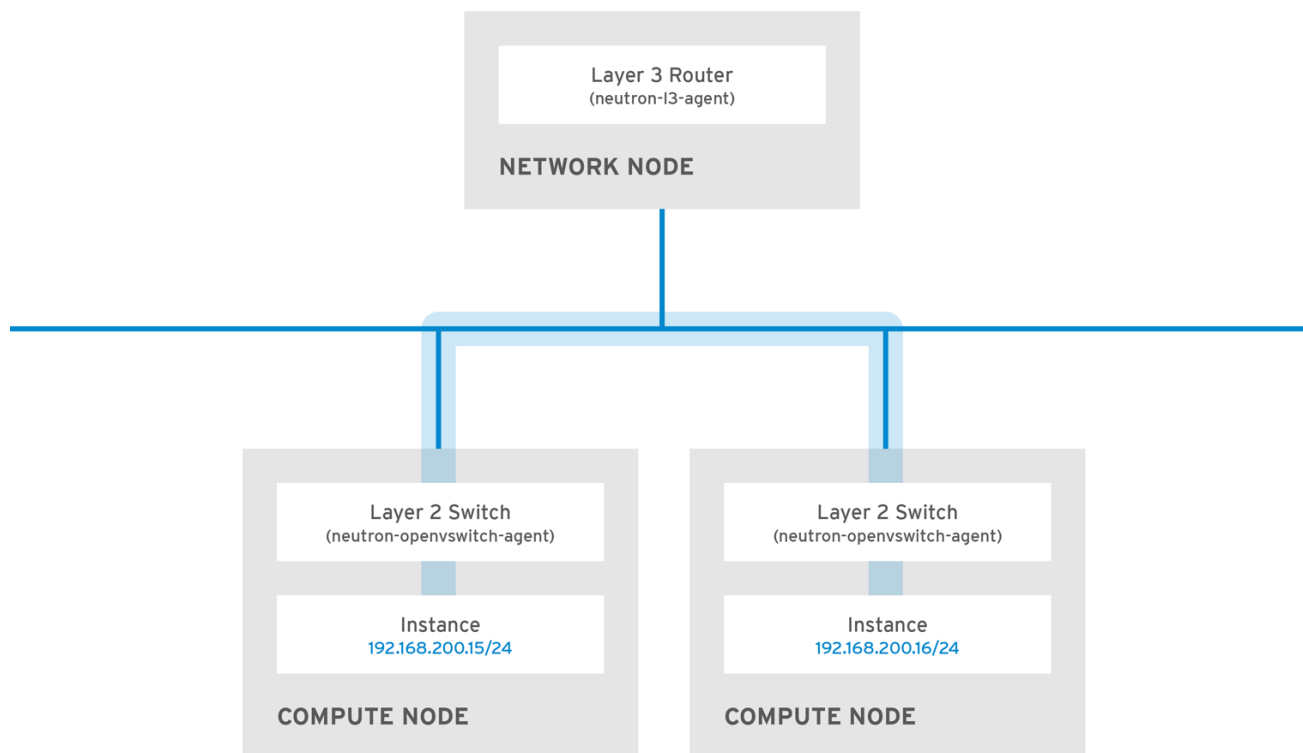
NOTE

Even a high performance hardware router adds latency to this configuration.

2.13.1. Use switching where possible

Because switching occurs at a lower level of the network (layer 2) it can function faster than the routing that occurs at layer 3. Design as few hops as possible between systems that communicate frequently.

For example, the following diagram depicts a switched network that spans two physical nodes, allowing the two instances to communicate directly without using a router for navigation first. Note that the instances now share the same subnet, to indicate that they are on the same logical network:



OPENSTACK_450456_0617

To allow instances on separate nodes to communicate as if they are on the same logical network, use an encapsulation tunnel such as VXLAN or GRE. Red Hat recommends adjusting the MTU size from end-to-end to accommodate the additional bits required for the tunnel header, otherwise network performance can be negatively impacted as a result of fragmentation. For more information, see *Configure MTU Settings*.

You can further improve the performance of VXLAN tunneling by using supported hardware that features VXLAN offload capabilities.

Additional resources

- <https://access.redhat.com/articles/1390483>

CHAPTER 3. COMMON ADMINISTRATIVE NETWORKING TASKS

OpenStack Networking (neutron) is the software-defined networking component of Red Hat OpenStack Platform. The virtual network infrastructure enables connectivity between instances and the physical external network.

This section contains information about common administration tasks, such as adding and removing subnets and routers to suit your Red Hat OpenStack Platform deployment.

3.1. CREATING A NETWORK

Create a network so that your instances can communicate with each other and receive IP addresses using DHCP. You can also integrate a network with external networks in your Red Hat OpenStack Platform deployment, or elsewhere, such as the physical network. This integration allows your instances to communicate with outside systems. For more information, see *Bridge the physical network*.

When creating networks, it is important to know that networks can host multiple subnets. This is useful if you intend to host distinctly different systems in the same network, and prefer a measure of isolation between them. For example, you can designate that only webserver traffic is present on one subnet, while database traffic traverses another. Subnets are isolated from each other, and any instance that wants to communicate with another subnet must have their traffic directed by a router. Consider placing systems that require a high volume of traffic amongst themselves in the same subnet, so that they do not require routing, and can avoid the subsequent latency and load.

1. In the dashboard, select **Project > Network > Networks**
2. Click **+Create Network** and specify the following values:

Field	Description
Network Name	Descriptive name, based on the role that the network will perform. If you are integrating the network with an external VLAN, consider appending the VLAN ID number to the name. For example, webservers_122 , if you are hosting HTTP web servers in this subnet, and your VLAN tag is 122 . Or you might use internal-only if you intend to keep the network traffic private, and not integrate the network with an external network.
Admin State	Controls whether the network is immediately available. Use this field to create the network in a Down state, where it is logically present but inactive. This is useful if you do not intend to enter the network into production immediately.

3. Click the **Next** button, and specify the following values in the **Subnet** tab:

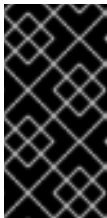
Field	Description
-------	-------------

Field	Description
Create Subnet	Determines whether to create a subnet. For example, you might not want to create a subnet if you intend to keep this network as a placeholder without network connectivity.
Subnet Name	Enter a descriptive name for the subnet.
Network Address	Enter the address in CIDR format, which contains the IP address range and subnet mask in one value. To determine the address, calculate the number of bits masked in the subnet mask and append that value to the IP address range. For example, the subnet mask 255.255.255.0 has 24 masked bits. To use this mask with the IPv4 address range 192.168.122.0, specify the address 192.168.122.0/24.
IP Version	Specifies the internet protocol version, where valid types are IPv4 or IPv6. The IP address range in the Network Address field must match whichever version you select.
Gateway IP	IP address of the router interface for your default gateway. This address is the next hop for routing any traffic destined for an external location, and must be within the range that you specify in the Network Address field. For example, if your CIDR network address is 192.168.122.0/24, then your default gateway is likely to be 192.168.122.1.
Disable Gateway	Disables forwarding and isolates the subnet.

4. Click **Next** to specify **DHCP** options:

- **Enable DHCP** - Enables DHCP services for this subnet. You can use DHCP to automate the distribution of IP settings to your instances.
- **IPv6 Address** - Configuration Modes. If you create an IPv6 network, you must specify how to allocate IPv6 addresses and additional information:
 - **No Options Specified** - Select this option if you want to set IP addresses manually, or if you use a non OpenStack-aware method for address allocation.
 - **SLAAC (Stateless Address Autoconfiguration)** - Instances generate IPv6 addresses based on Router Advertisement (RA) messages sent from the OpenStack Networking router. Use this configuration to create an OpenStack Networking subnet with `ra_mode` set to `slaac` and `address_mode` set to `slaac`.

- **DHCPv6 stateful** - Instances receive IPv6 addresses as well as additional options (for example, DNS) from the OpenStack Networking DHCPv6 service. Use this configuration to create a subnet with `ra_mode` set to `dhcpv6-stateful` and `address_mode` set to `dhcpv6-stateful`.
- **DHCPv6 stateless** - Instances generate IPv6 addresses based on Router Advertisement (RA) messages sent from the OpenStack Networking router. Additional options (for example, DNS) are allocated from the OpenStack Networking DHCPv6 service. Use this configuration to create a subnet with `ra_mode` set to `dhcpv6-stateless` and `address_mode` set to `dhcpv6-stateless`.
- **Allocation Pools** - Range of IP addresses that you want DHCP to assign. For example, the value `192.168.22.100,192.168.22.100` considers all up addresses in that range as available for allocation.
- **DNS Name Servers** - IP addresses of the DNS servers available on the network. DHCP distributes these addresses to the instances for name resolution.



IMPORTANT

For strategic services such as DNS, it is a best practice not to host them on your cloud. For example, if your cloud hosts DNS and your cloud becomes inoperable, DNS is unavailable and the cloud components cannot do lookups on each other.

- **Host Routes** - Static host routes. First, specify the destination network in CIDR format, followed by the next hop that you want to use for routing (for example, `192.168.23.0/24, 10.1.31.1`). Provide this value if you need to distribute static routes to instances.
5. Click **Create**.
You can view the complete network in the **Networks** tab. You can also click **Edit** to change any options as needed. When you create instances, you can configure them now to use its subnet, and they receive any specified DHCP options.

3.2. ADDING NETWORK ROUTING

To allow traffic to be routed to and from your new network, you must add its subnet as an interface to an existing virtual router:

1. In the dashboard, select **Project > Network > Routers**
2. Select your virtual router name in the **Routers** list, and click **Add Interface**.
In the Subnet list, select the name of your new subnet. You can optionally specify an IP address for the interface in this field.
3. Click **Add Interface**.
Instances on your network can now communicate with systems outside the subnet.

3.3. DELETING A NETWORK

There are occasions where it becomes necessary to delete a network that was previously created, perhaps as housekeeping or as part of a decommissioning process. You must first remove or detach any interfaces where the network is still in use, before you can successfully delete a network.

To delete a network in your project, together with any dependent interfaces, complete the following steps:

1. In the dashboard, select **Project > Network > Networks**

Remove all router interfaces associated with the target network subnets.

To remove an interface, find the ID number of the network that you want to delete by clicking on your target network in the **Networks** list, and looking at the ID field. All the subnets associated with the network share this value in the **Network ID** field.

2. Navigate to **Project > Network > Routers**, click the name of your virtual router in the **Routers** list, and locate the interface attached to the subnet that you want to delete. You can distinguish this subnet from the other subnets by the IP address that served as the gateway IP. You can further validate the distinction by ensuring that the network ID of the interface matches the ID that you noted in the previous step.
3. Click the **Delete Interface** button for the interface that you want to delete.
4. Select **Project > Network > Networks** and click the name of your network.
5. Click the **Delete Subnet** button for the subnet that you want to delete.



NOTE

If you are still unable to remove the subnet at this point, ensure it is not already being used by any instances.

6. Select **Project > Network > Networks** and select the network you would like to delete.
7. Click **Delete Networks**.

3.4. PURGING ALL RESOURCES AND DELETING A PROJECT

Use the **openstack project purge** command to delete all resources that belong to a particular project as well as deleting the project, too.

For example, to purge the resources of the **test-project** project, and then delete the project, run the following commands:

```
# openstack project list
+-----+-----+
| ID                | Name      |
+-----+-----+
| 02e501908c5b438dbc73536c10c9aac0 | test-project |
| 519e6344f82e4c079c8e2eabb690023b | services    |
| 80bf5732752a41128e612fe615c886c6 | demo       |
| 98a2f53c20ce4d50a40dac4a38016c69 | admin      |
+-----+-----+

# openstack project purge --project 02e501908c5b438dbc73536c10c9aac0
```

3.5. WORKING WITH SUBNETS

Use subnets to grant network connectivity to instances. Each instance is assigned to a subnet as part of the instance creation process, therefore it's important to consider proper placement of instances to best accommodate their connectivity requirements.

You can create subnets only in pre-existing networks. Remember that project networks in OpenStack Networking can host multiple subnets. This is useful if you intend to host distinctly different systems in the same network, and prefer a measure of isolation between them.

For example, you can designate that only webserver traffic is present on one subnet, while database traffic traverse another.

Subnets are isolated from each other, and any instance that wants to communicate with another subnet must have their traffic directed by a router. Therefore, you can lessen network latency and load by grouping systems in the same subnet that require a high volume of traffic between each other.

3.5.1. Creating a subnet

To create a subnet, follow these steps:

1. In the dashboard, select **Project > Network > Networks** and click the name of your network in the **Networks** view.
2. Click **Create Subnet**, and specify the following values:

Field	Description
Subnet Name	Descriptive subnet name.
Network Address	Address in CIDR format, which contains the IP address range and subnet mask in one value. To determine the CIDR address, calculate the number of bits masked in the subnet mask and append that value to the IP address range. For example, the subnet mask 255.255.255.0 has 24 masked bits. To use this mask with the IPv4 address range 192.168.122.0, specify the address 192.168.122.0/24.
IP Version	Internet protocol version, where valid types are IPv4 or IPv6. The IP address range in the Network Address field must match whichever protocol version you select.
Gateway IP	IP address of the router interface for your default gateway. This address is the next hop for routing any traffic destined for an external location, and must be within the range that you specify in the Network Address field. For example, if your CIDR network address is 192.168.122.0/24, then your default gateway is likely to be 192.168.122.1.
Disable Gateway	Disables forwarding and isolates the subnet.

3. Click **Next** to specify **DHCP** options:

- **Enable DHCP** - Enables DHCP services for this subnet. You can use DHCP to automate the distribution of IP settings to your instances.
- **IPv6 Address** - Configuration Modes. If you create an IPv6 network, you must specify how to allocate IPv6 addresses and additional information:
 - **No Options Specified** - Select this option if you want to set IP addresses manually, or if you use a non OpenStack-aware method for address allocation.
 - **SLAAC (Stateless Address Autoconfiguration)** - Instances generate IPv6 addresses based on Router Advertisement (RA) messages sent from the OpenStack Networking router. Use this configuration to create an OpenStack Networking subnet with `ra_mode` set to `slaac` and `address_mode` set to `slaac`.
 - **DHCPv6 stateful** - Instances receive IPv6 addresses as well as additional options (for example, DNS) from the OpenStack Networking DHCPv6 service. Use this configuration to create a subnet with `ra_mode` set to `dhcpv6-stateful` and `address_mode` set to `dhcpv6-stateful`.
 - **DHCPv6 stateless** - Instances generate IPv6 addresses based on Router Advertisement (RA) messages sent from the OpenStack Networking router. Additional options (for example, DNS) are allocated from the OpenStack Networking DHCPv6 service. Use this configuration to create a subnet with `ra_mode` set to `dhcpv6-stateless` and `address_mode` set to `dhcpv6-stateless`.
- **Allocation Pools** - Range of IP addresses that you want DHCP to assign. For example, the value `192.168.22.100,192.168.22.100` considers all up addresses in that range as available for allocation.
- **DNS Name Servers** - IP addresses of the DNS servers available on the network. DHCP distributes these addresses to the instances for name resolution.
- **Host Routes** - Static host routes. First, specify the destination network in CIDR format, followed by the next hop that you want to use for routing (for example, `192.168.23.0/24, 10.1.31.1`). Provide this value if you need to distribute static routes to instances.

4. Click **Create**.

You can view the subnet in the **Subnets** list. You can also click **Edit** to change any options as needed. When you create instances, you can configure them now to use its subnet, and they receive any specified DHCP options.

3.6. DELETING A SUBNET

You can delete a subnet if it is no longer in use. However, if any instances are still configured to use the subnet, the deletion attempt fails and the dashboard displays an error message.

Complete the following steps to delete a specific subnet in a network:

1. In the dashboard, select **Project > Network > Networks**
2. Click the name of your network.
3. Select the target subnet, and click **Delete Subnets**.

3.7. ADDING A ROUTER

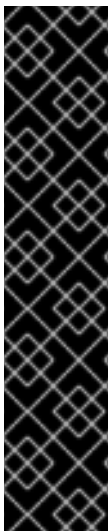
OpenStack Networking provides routing services using an SDN-based virtual router. Routers are a requirement for your instances to communicate with external subnets, including those in the physical network. Routers and subnets connect using interfaces, with each subnet requiring its own interface to the router.

The default gateway of a router defines the next hop for any traffic received by the router. Its network is typically configured to route traffic to the external physical network using a virtual bridge.

To create a router, complete the following steps:

1. In the dashboard, select **Project > Network > Routers** and click **Create Router**.
2. Enter a descriptive name for the new router, and click **Create router**.
3. Click **Set Gateway** next to the entry for the new router in the **Routers** list.
4. In the **External Network** list, specify the network that you want to receive traffic destined for an external location.
5. Click **Set Gateway**.

After you add a router, you must configure any subnets you have created to send traffic using this router. You do this by creating interfaces between the subnet and the router.



IMPORTANT

The default routes for subnets must not be overwritten. When the default route for a subnet is removed, the L3 agent automatically removes the corresponding route in the router namespace too, and network traffic cannot flow to and from the associated subnet. If the existing router namespace route has been removed, to fix this problem, perform these steps:

1. Disassociate all floating IPs on the subnet.
2. Detach the router from the subnet.
3. Re-attach the router to the subnet.
4. Re-attach all floating IPs.

3.8. DELETING A ROUTER

You can delete a router if it has no connected interfaces.

To remove its interfaces and delete a router, complete the following steps:

1. In the dashboard, select **Project > Network > Routers** and click the name of the router that you want to delete.
2. Select the interfaces of type **Internal Interface**, and click **Delete Interfaces**.
3. From the Routers list, select the target router and click **Delete Routers**.

3.9. TUNING KEEPALIVED TO AVOID VRRP PACKET LOSS

If the number of highly available (HA) routers on a single host is high, when an HA router fail over occurs, the Virtual Router Redundancy Protocol (VRRP) messages might overflow the IRQ queues. This overflow stops Open vSwitch (OVS) from responding and forwarding those VRRP messages.

To avoid VRRP packet overload, you must increase the VRRP advertisement interval using the **ha_vrrp_advert_int** option in Networking service (neutron) agent configuration file, **l3_agent.conf**.

Procedure

1. Log in as the stack user on the overcloud Networking controller host and open **l3_agent.conf** in a text editor.

Example

```
vi /var/lib/config-data/neutron/etc/neutron/l3_agent.conf
```

2. Locate the **ha_vrrp_advert_int** option, and increase the VRRP advert interval. (The default is 2 seconds.)

Example

```
[DEFAULT]
ha_vrrp_advert_int = 7
ha_vrrp_garp_master_repeat = 5
ha_vrrp_garp_master_delay = 5
```

There are also options for gratuitous ARP messages that you can set:

ha_vrrp_garp_master_repeat

The number of gratuitous ARP messages to send at one time after the transition to the master state. (The default is 5 messages.)

ha_vrrp_garp_master_delay

The delay for second set of gratuitous ARP messages after the lower priority advert is received in the master state. (The default is 5 seconds.)

TIP

The next time that the **openstack overcloud deploy** command is run, any manual changes like this one to **l3_agent.conf** are overwritten. You can write a post-deployment script that re-applies your changes after future overcloud updates. If you incorporate the puppet hieradata role name described in the following document in your script, the script will run only on the overcloud node roles that must have your **l3_agent.conf** changes re-applied. For more information, see [Post-Configuration: Customizing All Overcloud Roles](#) in the *Advanced Overcloud Customization* guide.

Additional resources

- [2.1.2 Data Forwarding Rules, Subsection 2](#) in *RFC 4541*

3.10. ADDING AN INTERFACE

You can use interfaces to interconnect routers with subnets so that routers can direct any traffic that instances send to destinations outside of their intermediate subnet.

To add a router interface and connect the new interface to a subnet, complete these steps:



NOTE

This procedure uses the Network Topology feature. Using this feature, you can see a graphical representation of all your virtual routers and networks while you perform network management tasks.

1. In the dashboard, select **Project > Network > Network Topology**
2. Locate the router that you want to manage, hover your mouse over it, and click **Add Interface**.
3. Specify the Subnet that you want to connect to the router.
You can also specify an IP address. The address is useful for testing and troubleshooting purposes, since a successful ping to this interface indicates that the traffic is routing as expected.
4. Click **Add interface**.
The Network Topology diagram automatically updates to reflect the new interface connection between the router and subnet.

3.11. DELETING AN INTERFACE

You can remove an interface to a subnet if you no longer require the router to direct traffic for the subnet.

To delete an interface, complete the following steps:

1. In the dashboard, select **Project > Network > Routers**
2. Click the name of the router that hosts the interface that you want to delete.
3. Select the interface type (**Internal Interface**), and click **Delete Interfaces**

3.12. CREATING FLOATING IP POOLS

You can use floating IP addresses to direct ingress network traffic to your OpenStack instances. First, you must define a pool of validly routable external IP addresses, which you can then assign to instances dynamically. OpenStack Networking routes all incoming traffic destined for that floating IP to the instance that you associate with the floating IP.



NOTE

OpenStack Networking allocates floating IP addresses to all projects (tenants) from the same IP ranges in CIDR format. As a result, all projects can consume floating IPs from every floating IP subnet. You can manage this behavior using quotas for specific projects. For example, you can set the default to **10** for **ProjectA** and **ProjectB**, while setting the quota for **ProjectC** to **0**.

Procedure

- When you create an external subnet, you can also define the floating IP allocation pool.

```
$ openstack subnet create --no-dhcp --allocation-pool
start=IP_ADDRESS,end=IP_ADDRESS --gateway IP_ADDRESS --network
SUBNET_RANGE NETWORK_NAME
```

If the subnet hosts only floating IP addresses, consider disabling DHCP allocation with the **--no-dhcp** option in the **openstack subnet create** command.

Example

```
$ openstack subnet create --no-dhcp --allocation_pool
start=192.168.100.20,end=192.168.100.100 --gateway 192.168.100.1 --network
192.168.100.0/24 public
```

Verification steps

- You can verify that the pool is configured properly by assigning a random floating IP to an instance. (See the later link that follows.)

Additional resources

- [subnet create](#) in the *Command Line Interface Reference*
- [Assigning a random floating IP](#)

3.13. ASSIGNING A SPECIFIC FLOATING IP

You can assign a specific floating IP address to a VM instance.

Procedure

- Allocate a floating IP address to an instance by using the **openstack server add floating ip** command.

Example

```
$ openstack server add floating ip prod-serv1 192.0.2.200
```

Validation steps

- Confirm that your floating IP is associated with your instance by using the **openstack server show** command.

Example

```
$ openstack server show prod-serv1
```

Sample output

```
+-----+-----+
| Field          | Value                |
+-----+-----+
| OS-DCF:diskConfig | MANUAL               |
```

```

| OS-EXT-AZ:availability_zone | nova |
| OS-EXT-STS:power_state | Running |
| OS-EXT-STS:task_state | None |
| OS-EXT-STS:vm_state | active |
| OS-SRV-USG:launched_at | 2021-08-11T14:45:37.000000 |
| OS-SRV-USG:terminated_at | None |
| accessIPv4 | |
| accessIPv6 | |
| addresses | public=198.51.100.56,192.0.2.200 |
| config_drive | |
| created | 2021-08-11T14:44:54Z |
| flavor | review-ephemeral |
| | (8130dd45-78f6-44dc-8173-4d6426b8e520) |
| hostId | 2308c8d8f60ed5394b1525122fb5bf8ea55c78b8 |
| | 0ec6157eca4488c9 |
| id | aef3ca09-887d-4d20-872d-1d1b49081958 |
| image | rhel8 |
| | (20724bfe-93a9-4341-a5a3-78b37b3a5dfb) |
| key_name | example-keypair |
| name | prod-serv1 |
| progress | 0 |
| project_id | bd7a8c4a19424cf09a82627566b434fa |
| properties | |
| security_groups | name='default' |
| status | ACTIVE |
| updated | 2021-08-11T14:45:37Z |
| user_id | 4b7e19a0d723310fd92911eb2fe59743a3a5cd32 |
| | 45f76ffced91096196f646b5 |
| volumes_attached | |
+-----+

```

Additional resources

- [server add floating ip](#) in the *Command Line Interface Reference*
- [server show](#) in the *Command Line Interface Reference*
- [Assigning a random floating IP](#)

3.14. CREATING AN ADVANCED NETWORK

Advanced network options are available for administrators, when creating a network from the **Admin** view. Use these options to specify projects and to define the network type that you want to use.

Procedure

1. In the dashboard, select **Admin > Networks > Create Network > Project**
2. Select the project that you want to host the new network with the **Project** drop-down list.
3. Review the options in **Provider Network Type**:
 - **Local** - Traffic remains on the local Compute host and is effectively isolated from any external networks.

- **Flat** - Traffic remains on a single network and can also be shared with the host. No VLAN tagging or other network segregation takes place.
 - **VLAN** - Create a network using a VLAN ID that corresponds to a VLAN present in the physical network. This option allows instances to communicate with systems on the same layer 2 VLAN.
 - **GRE** - Use a network overlay that spans multiple nodes for private communication between instances. Traffic egressing the overlay must be routed.
 - **VXLAN** - Similar to GRE, and uses a network overlay to span multiple nodes for private communication between instances. Traffic egressing the overlay must be routed.
4. Click **Create Network**.
Review the **Project Network Topology** to validate that the network has been successfully created.

Additional resources

- [Assigning a specific floating IP](#)
- [Assigning a random floating IP](#)

3.15. ASSIGNING A RANDOM FLOATING IP

You can dynamically allocate floating IP addresses to VM instances from a pool of external IP addresses.

Prerequisites

- A pool of routable external IP addresses.
For more information, see [Section 3.12, "Creating floating IP pools"](#).

Procedure

1. Enter the following command to allocate a floating IP address from the pool. In this example, the network is named **public**.

Example

```
$ openstack floating ip create public
```

Sample output

In the following example, the newly allocated floating IP is **192.0.2.200**. You can assign it to an instance.

```
+-----+-----+
| Field      | Value                               |
+-----+-----+
| fixed_ip_address | None                                 |
| floating_ip_address | 192.0.2.200                         |
| floating_network_id | f0dcc603-f693-4258-a940-0a31fd4b80d9 |
| id          | 6352284c-c5df-4792-b168-e6f6348e2620 |
| port_id     | None                                 |
```

```
| router_id      | None          |
| status        | ACTIVE       |
+-----+-----+
```

2. Enter the following command to locate your instance:

```
$ openstack server list
```

Sample output

```
+-----+-----+-----+-----+-----+-----+
| ID      | Name      | Status | Networks | Image | Flavor |
+-----+-----+-----+-----+-----+-----+
| aef3ca09-88 | prod-serv1 | ACTIVE | public=198. | rhel8 | review- |
| 7d-4d20-872 |           |       | 51.100.56  |       | ephemeral |
| d-1d1b49081 |           |       |           |       |           |
| 958        |           |       |           |       |           |
|           |           |       |           |       |           |
+-----+-----+-----+-----+-----+-----+
```

3. Associate the instance name or ID with the floating IP.

Example

```
$ openstack server add floating ip prod-serv1 192.0.2.200
```

Validation steps

- Enter the following command to confirm that your floating IP is associated with your instance.

Example

```
$ openstack server show prod-serv1
```

Sample output

```
+-----+-----+-----+-----+-----+-----+
| Field          | Value          |
+-----+-----+-----+-----+-----+-----+
| OS-DCF:diskConfig | MANUAL        |
| OS-EXT-AZ:availability_zone | nova          |
| OS-EXT-STS:power_state | Running      |
| OS-EXT-STS:task_state | None         |
| OS-EXT-STS:vm_state | active       |
| OS-SRV-USG:launched_at | 2021-08-11T14:45:37.000000 |
| OS-SRV-USG:terminated_at | None         |
| accessIPv4      |               |
| accessIPv6      |               |
| addresses       | public=198.51.100.56,192.0.2.200 |
| config_drive    |               |
| created         | 2021-08-11T14:44:54Z |
| flavor          | review-ephemeral |
```

```

| (8130dd45-78f6-44dc-8173-4d6426b8e520) |
| hostId | 2308c8d8f60ed5394b1525122fb5bf8ea55c78b8 |
| | 0ec6157eca4488c9 |
| id | aef3ca09-887d-4d20-872d-1d1b49081958 |
| image | rhel8 |
| | (20724bfe-93a9-4341-a5a3-78b37b3a5dfb) |
| key_name | example-keypair |
| name | prod-serv1 |
| progress | 0 |
| project_id | bd7a8c4a19424cf09a82627566b434fa |
| properties | |
| security_groups | name='default' |
| status | ACTIVE |
| updated | 2021-08-11T14:45:37Z |
| user_id | 4b7e19a0d723310fd92911eb2fe59743a3a5cd32 |
| | 45f76ffced91096196f646b5 |
| volumes_attached | |
+-----+-----+

```

Additional resources

- [floating ip create](#) in the *Command Line Interface Reference*
- [server add floating ip](#) in the *Command Line Interface Reference*
- [server show](#) in the *Command Line Interface Reference*
- [Creating floating IP pools](#)

3.16. CREATING MULTIPLE FLOATING IP POOLS

OpenStack Networking supports one floating IP pool for each L3 agent. Therefore, you must scale your L3 agents to create additional floating IP pools.

Procedure

- Make sure that in `/var/lib/config-data/neutron/etc/neutron/neutron.conf` the property **handle_internal_only_routers** is set to **True** for only one L3 agent in your environment. This option configures the L3 agent to manage only non-external routers.

Additional resources

- [Creating floating IP pools](#)
- [Assigning a random floating IP](#)

3.17. BRIDGING THE PHYSICAL NETWORK

Bridge your virtual network to the physical network to enable connectivity to and from virtual instances.

In this procedure, the example physical interface, **eth0**, is mapped to the bridge, **br-ex**; the virtual bridge acts as the intermediary between the physical network and any virtual networks.

As a result, all traffic traversing **eth0** uses the configured Open vSwitch to reach instances.

To map a physical NIC to the virtual Open vSwitch bridge, complete the following steps:

Procedure

1. Open `/etc/sysconfig/network-scripts/ifcfg-eth0` in a text editor, and update the following parameters with values appropriate for the network at your site:

- IPADDR
- NETMASK GATEWAY
- DNS1 (name server)
Here is an example:

```
# vi /etc/sysconfig/network-scripts/ifcfg-eth0
DEVICE=eth0
TYPE=OVSPort
DEVICETYPE=ovs
OVS_BRIDGE=br-ex
ONBOOT=yes
```

2. Open `/etc/sysconfig/network-scripts/ifcfg-br-ex` in a text editor and update the virtual bridge parameters with the IP address values that were previously allocated to eth0:

```
# vi /etc/sysconfig/network-scripts/ifcfg-br-ex
DEVICE=br-ex
DEVICETYPE=ovs
TYPE=OVSBridge
BOOTPROTO=static
IPADDR=192.168.120.10
NETMASK=255.255.255.0
GATEWAY=192.168.120.1
DNS1=192.168.120.1
ONBOOT=yes
```

You can now assign floating IP addresses to instances and make them available to the physical network.

Additional resources

- [Configuring bridge mappings](#)

3.18. SPECIFYING THE NAME THAT DNS ASSIGNS TO PORTS

You can specify the name assigned to ports by the internal DNS when you enable the Red Hat OpenStack Platform (RHOSP) Networking service (neutron) `dns_domain` for ports extension (`dns_domain_ports`).

You enable the `dns_domain` for ports extension by declaring the RHOSP Orchestration (heat) `NeutronPluginExtensions` parameter in a YAML-formatted environment file. Using a corresponding parameter, `NeutronDnsDomain`, you specify your domain name, which overrides the default value, `openstacklocal`. After redeploying your overcloud, you can use the OpenStack Client port commands, `port set` or `port create`, with `--dns-name` to assign a port name.

Also, when the `dns_domain` for ports extension is enabled, the Compute service automatically populates the `dns_name` attribute with the `hostname` attribute of the instance during the boot of VM instances. At the end of the boot process, `dnsmasq` recognizes the allocated ports by their instance hostname.

Procedure

1. Log in to the undercloud as the stack user, and source the `stackrc` file to enable the director command line tools.

Example

```
$ source ~/stackrc
```

2. Create a custom YAML environment file (`my-neutron-environment.yaml`).



NOTE

Values inside parentheses are sample values that are used in the example commands in this procedure. Substitute these sample values with ones that are appropriate for your site.

Example

```
$ vi /home/stack/templates/my-neutron-environment.yaml
```

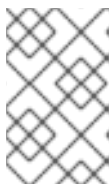
TIP

The undercloud includes a set of Orchestration service templates that form the plan for your overcloud creation. You can customize aspects of the overcloud with environment files, which are YAML-formatted files that override parameters and resources in the core Orchestration service template collection. You can include as many environment files as necessary.

3. In the environment file, add a `parameter_defaults` section. Under this section, add the `dns_domain` for ports extension, `dns_domain_ports`.

Example

```
parameter_defaults:
  NeutronPluginExtensions: "qos,port_security,dns_domain_ports"
```



NOTE

If you set `dns_domain_ports`, ensure that the deployment does not also use `dns_domain`, the DNS Integration extension. These extensions are incompatible, and both extensions cannot be defined simultaneously.

4. Also in the `parameter_defaults` section, add your domain name (`example.com`) using the `NeutronDnsDomain` parameter.

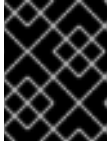
Example

```
parameter_defaults:
```

```
  NeutronPluginExtensions: "qos,port_security,dns_domain_ports"
```

```
  NeutronDnsDomain: "example.com"
```

5. Run the **openstack overcloud deploy** command and include the core Orchestration templates, environment files, and this new environment file.



IMPORTANT

The order of the environment files is important because the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-neutron-
environment.yaml
```

Verification steps

1. Log in to the overcloud, and create a new port (**new_port**) on a network (**public**). Assign a DNS name (**my_port**) to the port.

Example

```
$ source ~/overcloudrc
$ openstack port create --network public --dns-name my_port new_port
```

2. Display the details for your port (**new_port**).

Example

```
$ openstack port show -c dns_assignment -c dns_domain -c dns_name -c name new_port
```

Output

```
+-----+-----+
| Field          | Value                                     |
+-----+-----+
| dns_assignment | fqdn='my_port.example.com',           |
|                 | hostname='my_port',                   |
|                 | ip_address='10.65.176.113'            |
| dns_domain     | example.com                             |
| dns_name       | my_port                                 |
| name           | new_port                               |
+-----+-----+
```

Under **dns_assignment**, the fully qualified domain name (**fqdn**) value for the port contains a concatenation of the DNS name (**my_port**) and the domain name (**example.com**) that you set earlier with **NeutronDnsDomain**.

3. Create a new VM instance (**my_vm**) using the port (**new_port**) that you just created.

Example

```
$ openstack server create --image rhel --flavor m1.small --port new_port my_vm
```

4. Display the details for your port (**new_port**).

Example

```
$ openstack port show -c dns_assignment -c dns_domain -c dns_name -c name new_port
```

Output

```
+-----+-----+
| Field          | Value                                     |
+-----+-----+
| dns_assignment | fqdn='my_vm.example.com',               |
|                 | hostname='my_vm',                       |
|                 | ip_address='10.65.176.113'              |
| dns_domain     | example.com                             |
| dns_name       | my_vm                                   |
| name           | new_port                                |
+-----+-----+
```

Note that the Compute service changes the **dns_name** attribute from its original value (**my_port**) to the name of the instance with which the port is associated (**my_vm**).

Additional resources

- [Environment files](#) in the *Advanced Openstack Customization* guide
- [Including Environment Files in Openstack Creation](#) in the *Advanced Openstack Customization* guide
- [port](#) in the *Command Line Interface Reference*
- [server create](#) in the *Command Line Interface Reference*

3.19. ASSIGNING DHCP ATTRIBUTES TO PORTS

You can use Red Hat Openstack Platform (RHOSP) Networking service (neutron) extensions to add networking functions. You can use the extra DHCP option extension (**extra_dhcp_opt**) to configure ports of DHCP clients with DHCP attributes. For example, you can add a PXE boot option such as **ftpp-server**, **server-ip-address**, or **bootfile-name** to a DHCP client port.

The value of the **extra_dhcp_opt** attribute is an array of DHCP option objects, where each object contains an **opt_name** and an **opt_value**. IPv4 is the default version, but you can change this to IPv6 by including a third option, **ip-version=6**.

When a VM instance starts, the RHOSP Networking service supplies port information to the instance using DHCP protocol. If you add DHCP information to a port already connected to a running instance, the instance only uses the new DHCP port information when the instance is restarted.

Some of the more common DHCP port attributes are: **bootfile-name**, **dns-server**, **domain-name**, **mtu**, **server-ip-address**, and **tftp-server**. For the complete set of acceptable values for **opt_name**, refer to the DHCP specification.

Prerequisites

- You must have RHOSP administrator privileges.

Procedure

1. Log in to the undercloud host as the **stack** user.
2. Source the undercloud credentials file:

```
$ source ~/stackrc
```

3. Create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-octavia-environment.yaml
```

4. Your environment file must contain the keywords **parameter_defaults**. Under these keywords, add the extra DHCP option extension, **extra_dhcp_opt**.

Example

```
parameter_defaults:  
  NeutronPluginExtensions: "qos,port_security,extra_dhcp_opt"
```

5. Run the deployment command and include the core heat templates, environment files, and this new custom environment file.

```
The order of the environment files is important because the parameters and resources  
defined in subsequent environment files take precedence.
```

Example

```
$ openstack overcloud deploy --templates \  
-e <your_environment_files> \  
-e /usr/share/openstack-tripleo-heat-templates/environments/services/octavia.yaml \  
-e /home/stack/templates/my-octavia-environment.yaml
```

Verification

1. Source your credentials file.

Example

```
$ source ~/overcloudrc
```

2. Create a new port (**new_port**) on a network (**public**). Assign a valid attribute from the DHCP specification to the new port.

Example

```
$ openstack port create --extra-dhcp-option name=domain-name,value=test.domain --extra-dhcp-option name=ntp-server,value=192.0.2.123 --network public new_port
```

3. Display the details for your port (**new_port**).

Example

```
$ openstack port show new_port -c extra_dhcp_opts
```

Sample output

```
+-----+-----+
| Field      | Value                                     |
+-----+-----+
| extra_dhcp_opts | ip_version='4', opt_name='domain-name', opt_value='test.domain' |
|                | ip_version='4', opt_name='ntp-server', opt_value='192.0.2.123' |
+-----+-----+
```

Additional resources

- [OVN supported DHCP options](#)
- [Dynamic Host Configuration Protocol \(DHCP\) and Bootstrap Protocol \(BOOTP\) Parameters](#)
- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide
- [port create](#) in the *Command Line Interface Reference*
- [port show](#) in the *Command Line Interface Reference*

3.20. LOADING KERNEL MODULES

Some features in Red Hat OpenStack Platform (RHOSP) require certain kernel modules to be loaded. For example, the OVS firewall driver requires you to load the **nf_conntrack_proto_gre** kernel module to support GRE tunneling between two VM instances.

By using a special Orchestration service (heat) parameter, **ExtraKernelModules**, you can ensure that heat stores configuration information about the required kernel modules needed for features like GRE tunneling. Later, during normal module management, these required kernel modules are loaded.

Procedure

1. On the undercloud host, logged in as the stack user, create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-modules-environment.yaml
```

TIP

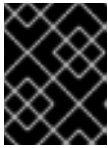
Heat uses a set of plans called *templates* to install and configure your environment. You can customize aspects of the overcloud with a *custom environment file*, which is a special type of template that provides customization for your heat templates.

- In the YAML environment file under **parameter_defaults**, set **ExtraKernelModules** to the name of the module that you want to load.

Example

```
ComputeParameters:
  ExtraKernelModules:
    nf_conntrack_proto_gre: {}
ControllerParameters:
  ExtraKernelModules:
    nf_conntrack_proto_gre: {}
```

- Run the **openstack overcloud deploy** command and include the core heat templates, environment files, and this new custom environment file.



IMPORTANT

The order of the environment files is important as the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-modules-
environment.yaml
```

Verification steps

- If heat has properly loaded the module, you should see output when you run the **lsmod** command on the Compute node:

Example

```
sudo lsmod | grep nf_conntrack_proto_gre
```

Additional resources

- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

CHAPTER 4. PLANNING IP ADDRESS USAGE

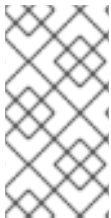
An OpenStack deployment can consume a larger number of IP addresses than might be expected. This section contains information about correctly anticipating the quantity of addresses that you require, and where the addresses are used in your environment.

4.1. VLAN PLANNING

When you plan your Red Hat OpenStack Platform deployment, you start with a number of subnets, from which you allocate individual IP addresses. When you use multiple subnets you can segregate traffic between systems into VLANs.

For example, it is ideal that your management or API traffic is not on the same network as systems that serve web traffic. Traffic between VLANs travels through a router where you can implement firewalls to govern traffic flow.

You must plan your VLANs as part of your overall plan that includes traffic isolation, high availability, and IP address utilization for the various types of virtual networking resources in your deployment.



NOTE

The maximum number of VLANs in a single network, or in one OVS agent for a network node, is 4094. In situations where you require more than the maximum number of VLANs, you can create several provider networks (VXLAN networks) and several network nodes, one per network. Each node can contain up to 4094 private networks.

4.2. TYPES OF NETWORK TRAFFIC

You can allocate separate VLANs for the different types of network traffic that you want to host. For example, you can have separate VLANs for each of these types of networks. Only the External network must be routable to the external physical network. In this release, director provides DHCP services.



NOTE

You do not require all of the isolated VLANs in this section for every OpenStack deployment.. For example, if your cloud users do not create ad hoc virtual networks on demand, then you may not require a project network. If you want each VM to connect directly to the same switch as any other physical system, connect your Compute nodes directly to a provider network and configure your instances to use that provider network directly.

- **Provisioning network** - This VLAN is dedicated to deploying new nodes using director over PXE boot. OpenStack Orchestration (heat) installs OpenStack onto the overcloud bare metal servers. These servers attach to the physical network to receive the platform installation image from the undercloud infrastructure.
- **Internal API network** - The OpenStack services use the Internal API network for communication, including API communication, RPC messages, and database communication. In addition, this network is used for operational messages between controller nodes. When planning your IP address allocation, note that each API service requires its own IP address. Specifically, you must plan IP addresses for each of the following services:
 - vip-msg (ampq)

- vip-keystone-int
- vip-glance-int
- vip-cinder-int
- vip-nova-int
- vip-neutron-int
- vip-horizon-int
- vip-heat-int
- vip-ceilometer-int
- vip-swift-int
- vip-keystone-pub
- vip-glance-pub
- vip-cinder-pub
- vip-nova-pub
- vip-neutron-pub
- vip-horizon-pub
- vip-heat-pub
- vip-ceilometer-pub
- vip-swift-pub



NOTE

When using High Availability, Pacemaker moves VIP addresses between the physical nodes.

- **Storage** - Block Storage, NFS, iSCSI, and other storage services. Isolate this network to separate physical Ethernet links for performance reasons.
- **Storage Management** - OpenStack Object Storage (swift) uses this network to synchronise data objects between participating replica nodes. The proxy service acts as the intermediary interface between user requests and the underlying storage layer. The proxy receives incoming requests and locates the necessary replica to retrieve the requested data. Services that use a Ceph back end connect over the Storage Management network, since they do not interact with Ceph directly but rather use the front end service. Note that the RBD driver is an exception; this traffic connects directly to Ceph.
- **Project networks** - Neutron provides each project with their own networks using either VLAN segregation (where each project network is a network VLAN), or tunneling using VXLAN or GRE. Network traffic is isolated within each project network. Each project network has an IP subnet associated with it, and multiple project networks may use the same addresses.

- **External** - The External network hosts the public API endpoints and connections to the Dashboard (horizon). You can also use this network for SNAT. In a production deployment, it is common to use a separate network for floating IP addresses and NAT.
- **Provider networks** - Use provider networks to attach instances to existing network infrastructure. You can use provider networks to map directly to an existing physical network in the data center, using flat networking or VLAN tags. This allows an instance to share the same layer-2 network as a system external to the OpenStack Networking infrastructure.

4.3. IP ADDRESS CONSUMPTION

The following systems consume IP addresses from your allocated range:

- **Physical nodes** - Each physical NIC requires one IP address. It is common practice to dedicate physical NICs to specific functions. For example, allocate management and NFS traffic to distinct physical NICs, sometimes with multiple NICs connecting across to different switches for redundancy purposes.
- **Virtual IPs (VIPs) for High Availability** - Plan to allocate between one and three VIPs for each network that controller nodes share.

4.4. VIRTUAL NETWORKING

The following virtual resources consume IP addresses in OpenStack Networking. These resources are considered local to the cloud infrastructure, and do not need to be reachable by systems in the external physical network:

- **Project networks** - Each project network requires a subnet that it can use to allocate IP addresses to instances.
- **Virtual routers** - Each router interface plugging into a subnet requires one IP address. If you want to use DHCP, each router interface requires two IP addresses.
- **Instances** - Each instance requires an address from the project subnet that hosts the instance. If you require ingress traffic, you must allocate a floating IP address to the instance from the designated external network.
- **Management traffic** - Includes OpenStack Services and API traffic. All services share a small number of VIPs. API, RPC and database services communicate on the internal API VIP.

4.5. EXAMPLE NETWORK PLAN

This example shows a number of networks that accommodate multiple subnets, with each subnet being assigned a range of IP addresses:

Table 4.1. Example subnet plan

Subnet name	Address range	Number of addresses	Subnet Mask
Provisioning network	192.168.100.1 - 192.168.100.250	250	255.255.255.0
Internal API network	172.16.1.10 - 172.16.1.250	241	255.255.255.0

Subnet name	Address range	Number of addresses	Subnet Mask
Storage	172.16.2.10 - 172.16.2.250	241	255.255.255.0
Storage Management	172.16.3.10 - 172.16.3.250	241	255.255.255.0
Tenant network (GRE/VXLAN)	172.16.4.10 - 172.16.4.250	241	255.255.255.0
External network (incl. floating IPs)	10.1.2.10 - 10.1.3.222	469	255.255.254.0
Provider network (infrastructure)	10.10.3.10 - 10.10.3.250	241	255.255.252.0

CHAPTER 5. REVIEWING OPENSTACK NETWORKING ROUTER PORTS

Virtual routers in OpenStack Networking use ports to interconnect with subnets. You can review the state of these ports to determine whether they connect as expected.

5.1. VIEWING CURRENT PORT STATUS

Complete the following steps to lists all of the ports that attach to a particular router and to retrieve the current state of a port (DOWN or ACTIVE):

1. To view all the ports that attach to the router named `r1`, run the following command:

```
# openstack port list --router r1
```

Example result:

```
+-----+-----+-----+-----+
| id                | name | mac_address    | fixed_ips
|
+-----+-----+-----+-----+
| b58d26f0-cc03-43c1-ab23-ccdb1018252a |    | fa:16:3e:94:a7:df | {"subnet_id": "a592fdbababd-48e0-96e8-2dd9117614d3", "ip_address": "192.168.200.1"} |
| c45e998d-98a1-4b23-bb41-5d24797a12a4 |    | fa:16:3e:ee:6a:f7 | {"subnet_id": "43f8f625-c773-4f18-a691-fd4ebfb3be54", "ip_address": "172.24.4.225"} |
+-----+-----+-----+-----+
```

2. To view the details of each port, run the following command. Include the port ID of the port that you want to view. The result includes the port status, indicated in the following example as having an **ACTIVE** state:

```
# openstack port show b58d26f0-cc03-43c1-ab23-ccdb1018252a
```

Example result:

```
+-----+-----+-----+-----+
| Field            | Value
+-----+-----+-----+-----+
| admin_state_up  | True
| allowed_address_pairs |
| binding:host_id | node.example.com
| binding:profile | {}
| binding:vif_details | {"port_filter": true, "ovs_hybrid_plug": true}
| binding:vif_type | ovs
| binding:vnic_type | normal
| device_id       | 49c6ebdc-0e62-49ad-a9ca-58cea464472f
| device_owner    | network:router_interface
| extra_dhcp_opts |
```

```
| fixed_ips      | {"subnet_id": "a592fdbd-babd-48e0-96e8-2dd9117614d3", "ip_address":  
"192.168.200.1"} |  
| id            | b58d26f0-cc03-43c1-ab23-ccdb1018252a |  
| mac_address   | fa:16:3e:94:a7:df |  
| name          | |  
| network_id    | 63c24160-47ac-4140-903d-8f9a670b0ca4 |  
| security_groups | |  
| status        | ACTIVE |  
| tenant_id     | d588d1112e0f496fb6cac22f9be45d49 |  
+-----+-----+  
+
```

3. Perform step 2 for each port to retrieve its status.

CHAPTER 6. TROUBLESHOOTING PROVIDER NETWORKS

A deployment of virtual routers and switches, also known as software-defined networking (SDN), may seem to introduce complexity. However, the diagnostic process of troubleshooting network connectivity in OpenStack Networking is similar to the diagnostic process for physical networks. If you use VLANs, you can consider the virtual infrastructure as a trunked extension of the physical network, rather than a wholly separate environment.

6.1. BASIC PING TESTING

The **ping** command is a useful tool for analyzing network connectivity problems. The results serve as a basic indicator of network connectivity, but might not entirely exclude all connectivity issues, such as a firewall blocking the actual application traffic. The ping command sends traffic to specific destinations, and then reports back whether the attempts were successful.



NOTE

The ping command is an ICMP operation. To use **ping**, you must allow ICMP traffic to traverse any intermediary firewalls.

Ping tests are most useful when run from the machine experiencing network issues, so it may be necessary to connect to the command line via the VNC management console if the machine seems to be completely offline.

For example, the following ping test command validates multiple layers of network infrastructure in order to succeed; name resolution, IP routing, and network switching must all function correctly:

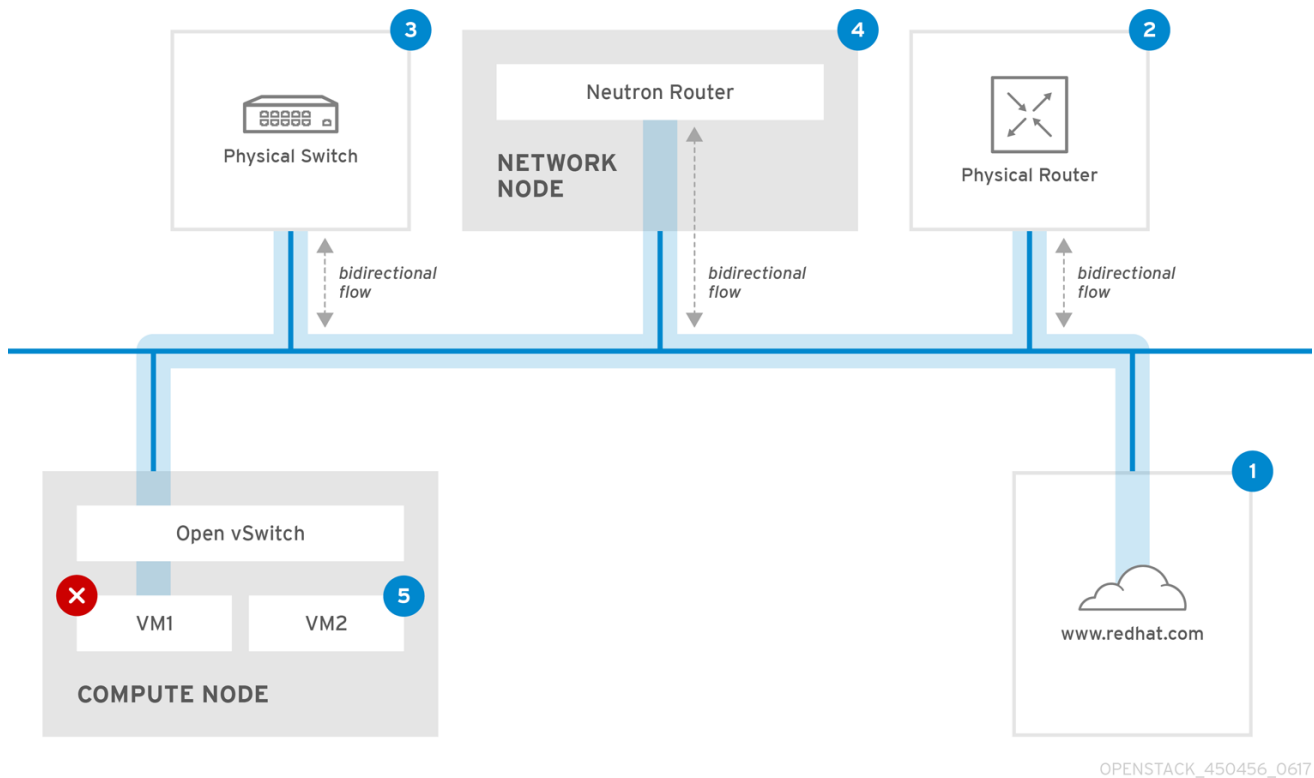
```
$ ping www.redhat.com

PING e1890.b.akamaiedge.net (125.56.247.214) 56(84) bytes of data:
64 bytes from a125-56.247-214.deploy.akamaitechnologies.com (125.56.247.214): icmp_seq=1
ttl=54 time=13.4 ms
64 bytes from a125-56.247-214.deploy.akamaitechnologies.com (125.56.247.214): icmp_seq=2
ttl=54 time=13.5 ms
64 bytes from a125-56.247-214.deploy.akamaitechnologies.com (125.56.247.214): icmp_seq=3
ttl=54 time=13.4 ms
^C
```

You can terminate the ping command with Ctrl-c, after which a summary of the results is presented. Zero percent packet loss indicates that the connection was stable and did not time out.

```
--- e1890.b.akamaiedge.net ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2003ms
rtt min/avg/max/mdev = 13.461/13.498/13.541/0.100 ms
```

The results of a ping test can be very revealing, depending on which destination you test. For example, in the following diagram VM1 is experiencing some form of connectivity issue. The possible destinations are numbered in blue, and the conclusions drawn from a successful or failed result are presented:



1. The internet - a common first step is to send a ping test to an internet location, such as www.redhat.com.

- *Success*: This test indicates that all the various network points in between the machine and the Internet are functioning correctly. This includes the virtual and physical network infrastructure.
- *Failure*: There are various ways in which a ping test to a distant internet location can fail. If other machines on your network are able to successfully ping the internet, that proves the internet connection is working, and the issue is likely within the configuration of the local machine.

2. Physical router - This is the router interface that the network administrator designates to direct traffic onward to external destinations.

- *Success*: Ping tests to the physical router can determine whether the local network and underlying switches are functioning. These packets do not traverse the router, so they do not prove whether there is a routing issue present on the default gateway.
- *Failure*: This indicates that the problem lies between VM1 and the default gateway. The router/switches might be down, or you may be using an incorrect default gateway. Compare the configuration with that on another server that you know is functioning correctly. Try pinging another server on the local network.

3. Neutron router - This is the virtual SDN (Software-defined Networking) router that Red Hat OpenStack Platform uses to direct the traffic of virtual machines.

- *Success*: Firewall is allowing ICMP traffic, the Networking node is online.
- *Failure*: Confirm whether ICMP traffic is permitted in the security group of the instance. Check that the Networking node is online, confirm that all the required services are running, and review the L3 agent log (`/var/log/neutron/l3-agent.log`).

4. Physical switch - The physical switch manages traffic between nodes on the same physical network.

- *Success*: Traffic sent by a VM to the physical switch must pass through the virtual network infrastructure, indicating that this segment is functioning correctly.
- *Failure*: Check that the physical switch port is configured to trunk the required VLANs.

5. **VM2** - Attempt to ping a VM on the same subnet, on the same Compute node.

- *Success*: The NIC driver and basic IP configuration on VM1 are functional.
- *Failure*: Validate the network configuration on VM1. Or, firewall on VM2 might simply be blocking ping traffic. In addition, verify the virtual switching configuration and review the Open vSwitch (or Linux Bridge) log files.

6.2. TROUBLESHOOTING VLAN NETWORKS

OpenStack Networking can trunk VLAN networks through to the SDN switches. Support for VLAN-tagged provider networks means that virtual instances can integrate with server subnets in the physical network.

To troubleshoot connectivity to a VLAN Provider network, complete these steps:

1. Ping the gateway with **ping <gateway-IP-address>**.

Consider this example, in which a network is created with these commands:

```
# openstack network create --provider-network-type vlan --provider-physical-network phy-
eno1 --provider-segment 120 provider
# openstack subnet create --no-dhcp --allocation-pool
start=192.168.120.1,end=192.168.120.153 --gateway 192.168.120.254 --network provider
public_subnet
```

In this example, the gateway IP address is **192.168.120.254**.

```
$ ping 192.168.120.254
```

2. If the ping fails, do the following:
 - a. Confirm that you have network flow for the associated VLAN. It is possible that the VLAN ID has not been set. In this example, OpenStack Networking is configured to trunk VLAN 120 to the provider network. (See **--provider:segmentation_id=120** in the example in step 1.)
 - b. Confirm the VLAN flow on the bridge interface using the command, **ovs-ofctl dump-flows <bridge-name>**.

In this example the bridge is named **br-ex**:

```
# ovs-ofctl dump-flows br-ex

NXST_FLOW reply (xid=0x4):
  cookie=0x0, duration=987.521s, table=0, n_packets=67897, n_bytes=14065247,
  idle_age=0, priority=1 actions=NORMAL
  cookie=0x0, duration=986.979s, table=0, n_packets=8, n_bytes=648, idle_age=977,
  priority=2,in_port=12 actions=drop
```

6.2.1. Reviewing the VLAN configuration and log files

- **OpenStack Networking (neutron) agents** – Use the **openstack network agent list** command to verify that all agents are up and registered with the correct names:

```
(overcloud)[stack@undercloud~]$ openstack network agent list
+-----+-----+-----+-----+-----+
| id                | agent_type   | host                | alive | admin_state_up |
+-----+-----+-----+-----+-----+
| a08397a8-6600-437d-9013-b2c5b3730c0c | Metadata agent | rhelosp.example.com | :- ) | True            |
| a5153cd2-5881-4fc8-b0ad-be0c97734e6a | L3 agent      | rhelosp.example.com | :- ) | True            |
| b54f0be7-c555-43da-ad19-5593a075ddf0 | DHCP agent    | rhelosp.example.com | :- ) | True            |
| d2be3cb0-4010-4458-b459-c5eb0d4d354b | Open vSwitch agent | rhelosp.example.com | :- ) | True            |
+-----+-----+-----+-----+-----+
```

- Review `/var/log/containers/neutron/openvswitch-agent.log` – this log should provide confirmation that the creation process used the **ovs-ofctl** command to configure VLAN trunking.
- Validate **external_network_bridge** in the `/etc/neutron/l3_agent.ini` file. If there is a hardcoded value in the **external_network_bridge** parameter, you cannot use a provider network with the L3-agent, and you cannot create the necessary flows. The **external_network_bridge** value must be in the format `external_network_bridge = ""``.
- Check the **network_vlan_ranges** value in the `/etc/neutron/plugin.ini` file. For provider networks, do not specify the numeric VLAN ID. Specify IDs only when using VLAN isolated project networks.
- Validate the OVS agent configuration file bridge mappings, confirm that the bridge mapped to **phy-eno1** exists and is properly connected to **eno1**.

6.3. TROUBLESHOOTING FROM WITHIN PROJECT NETWORKS

In OpenStack Networking, all project traffic is contained within network namespaces so that projects can configure networks without interfering with each other. For example, network namespaces allow different projects to have the same subnet range of 192.168.1.1/24 without interference between them.

To begin troubleshooting a project network, first determine which network namespace contains the network:

1. List all the project networks using the **openstack network list** command:

```
# (overcloud)[stack@osp13-undercloud ~]$ openstack network list
+-----+-----+-----+-----+-----+
| id                | name         | subnets            |
+-----+-----+-----+-----+-----+
| 9cb32fe0-d7fb-432c-b116-f483c6497b08 | web-servers | 453d6769-fcde-4796-a205-66ee01680bba 192.168.212.0/24 |
| a0cc8cdd-575f-4788-a3e3-5df8c6d0dd81 | private     | c1e58160-707f-44a7-bf94-8694f29e74d3 10.0.0.0/24 |
| baadd774-87e9-4e97-a055-326bb422b29b | private     | 340c58e1-7fe7-4cf2-96a7-96a0a4ff3231 192.168.200.0/24 |
+-----+-----+-----+-----+-----+
```



```
| 24ba3a36-5645-4f46-be47-f6af2a7d8af2 | public | 35f3d2cb-6e4b-4527-a932-
952a395c4bb3 172.24.4.224/28 |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
```

In this example, examine the **web-servers** network. Make a note of the id value in the web-server row (**9cb32fe0-d7fb-432c-b116-f483c6497b08**). This value is appended to the network namespace, which helps you identify the namespace in the next step.

- List all the network namespaces using the **ip netns list** command:

```
# ip netns list
qdhcp-9cb32fe0-d7fb-432c-b116-f483c6497b08
qrouter-31680a1c-9b3e-4906-bd69-cb39ed5faa01
qrouter-62ed467e-abae-4ab4-87f4-13a9937fbd6b
qdhcp-a0cc8cdd-575f-4788-a3e3-5df8c6d0dd81
qrouter-e9281608-52a6-4576-86a6-92955df46f56
```

The output contains a namespace that matches the **web-servers** network id. In this example the namespace is **qdhcp-9cb32fe0-d7fb-432c-b116-f483c6497b08**.

- Examine the configuration of the **web-servers** network by running commands within the namespace, prefixing the troubleshooting commands with **ip netns exec <namespace>**:

```
# ip netns exec qrouter-62ed467e-abae-4ab4-87f4-13a9937fbd6b route -n

Kernel IP routing table
Destination Gateway Genmask Flags Metric Ref Use Iface
0.0.0.0 172.24.4.225 0.0.0.0 UG 0 0 0 qg-8d128f89-87
172.24.4.224 0.0.0.0 255.255.255.240 U 0 0 0 qg-8d128f89-87
192.168.200.0 0.0.0.0 255.255.255.0 U 0 0 0 qr-8efd6357-96
```

6.3.1. Performing advanced ICMP testing within the namespace

- Capture ICMP traffic using the **tcpdump** command:

```
# ip netns exec qrouter-62ed467e-abae-4ab4-87f4-13a9937fbd6b tcpdump -qnntpi any icmp
```

- In a separate command line window, perform a ping test to an external network:

```
# ip netns exec qrouter-62ed467e-abae-4ab4-87f4-13a9937fbd6b ping www.redhat.com
```

- In the terminal running the **tcpdump** session, observe detailed results of the ping test.

```
tcpdump: listening on any, link-type LINUX_SLL (Linux cooked), capture size 65535 bytes
IP (tos 0xc0, ttl 64, id 55447, offset 0, flags [none], proto ICMP (1), length 88)
 172.24.4.228 > 172.24.4.228: ICMP host 192.168.200.20 unreachable, length 68
IP (tos 0x0, ttl 64, id 22976, offset 0, flags [DF], proto UDP (17), length 60)
 172.24.4.228.40278 > 192.168.200.21: [bad udp cksum 0xfa7b -> 0xe235!] UDP, length 32
```

**NOTE**

When you perform a tcpdump analysis of traffic, you might observe the responding packets heading to the router interface rather than the instance. This is expected behavior, as the qrouter performs DNAT on the return packets.

CHAPTER 7. CONNECTING AN INSTANCE TO THE PHYSICAL NETWORK

This chapter contains information about using provider networks to connect instances directly to an external network.

7.1. OVERVIEW OF THE OPENSTACK NETWORKING TOPOLOGY

OpenStack Networking (neutron) has two categories of services distributed across a number of node types.

- **Neutron server** - This service runs the OpenStack Networking API server, which provides the API for end-users and services to interact with OpenStack Networking. This server also integrates with the underlying database to store and retrieve project network, router, and loadbalancer details, among others.
- **Neutron agents** - These are the services that perform the network functions for OpenStack Networking:
 - **neutron-dhcp-agent** - manages DHCP IP addressing for project private networks.
 - **neutron-l3-agent** - performs layer 3 routing between project private networks, the external network, and others.
- **Compute node** - This node hosts the hypervisor that runs the virtual machines, also known as instances. A Compute node must be wired directly to the network in order to provide external connectivity for instances. This node is typically where the I2 agents run, such as **neutron-openvswitch-agent**.

Additional resources

- [Placement of OpenStack Networking services](#)

7.2. PLACEMENT OF OPENSTACK NETWORKING SERVICES

The OpenStack Networking services can either run together on the same physical server, or on separate dedicated servers, which are named according to their roles:

- *Controller node* - The server that runs API service.
- *Network node* - The server that runs the OpenStack Networking agents.
- *Compute node* - The hypervisor server that hosts the instances.

The steps in this chapter apply to an environment that contains these three node types. If your deployment has both the Controller and Network node roles on the same physical node, then you must perform the steps from both sections on that server. This also applies for a High Availability (HA) environment, where all three nodes might be running the Controller node and Network node services with HA. As a result, you must complete the steps in sections applicable to Controller and Network nodes on all three nodes.

Additional resources

- [Overview of the OpenStack Networking topology](#)

7.3. CONFIGURING FLAT PROVIDER NETWORKS

You can use flat provider networks to connect instances directly to the external network. This is useful if you have multiple physical networks and separate physical interfaces, and intend to connect each Compute and Network node to those external networks.

Prerequisites

- You have multiple physical networks.
This example uses physical networks called **physnet1**, and **physnet2**, respectively.
- You have separate physical interfaces.
This example uses separate physical interfaces, **eth0** and **eth1**, respectively.

Procedure

1. On the undercloud host, logged in as the stack user, create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-modules-environment.yaml
```

TIP

The Red Hat OpenStack Platform Orchestration service (heat) uses a set of plans called *templates* to install and configure your environment. You can customize aspects of the overcloud with a *custom environment file*, which is a special type of template that provides customization for your orchestration templates.

2. In the YAML environment file under **parameter_defaults**, use the **NeutronBridgeMappings** to specify which OVS bridges are used for accessing external networks.

Example

```
parameter_defaults:
  NeutronBridgeMappings: 'physnet1:br-net1,physnet2:br-net2'
```

3. In the custom NIC configuration template for the Controller and Compute nodes, configure the bridges with interfaces attached.

Example

```
...
- type: ovs_bridge
  name: br-net1
  mtu: 1500
  use_dhcp: false
  members:
  - type: interface
    name: eth0
    mtu: 1500
    use_dhcp: false
    primary: true
```

```

- type: ovs_bridge
  name: br-net2
  mtu: 1500
  use_dhcp: false
  members:
- type: interface
  name: eth1
  mtu: 1500
  use_dhcp: false
  primary: true
...

```

4. Run the **openstack overcloud deploy** command and include the templates and the environment files, including this modified custom NIC template and the new environment file.



IMPORTANT

The order of the environment files is important because the parameters and resources defined in subsequent environment files take precedence.

Example

```

$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-neutron-
environment.yaml

```

Verification steps

1. Create an external network (**public1**) as a flat network and associate it with the configured physical network (**physnet1**).
Configure it as a shared network (using **--share**) to let other users create VM instances that connect to the external network directly.

Example

```

# openstack network create --share --provider-network-type flat --provider-physical-network
physnet1 --external public01

```

2. Create a subnet (**public_subnet**) using the **openstack subnet create** command.

Example

```

# openstack subnet create --no-dhcp --allocation-pool
start=192.168.100.20,end=192.168.100.100 --gateway 192.168.100.1 --network public01
public_subnet

```

3. Create a VM instance and connect it directly to the newly-created external network.

Example

```

$ openstack server create --image rhel --flavor my_flavor --network public01 my_instance

```

Additional resources

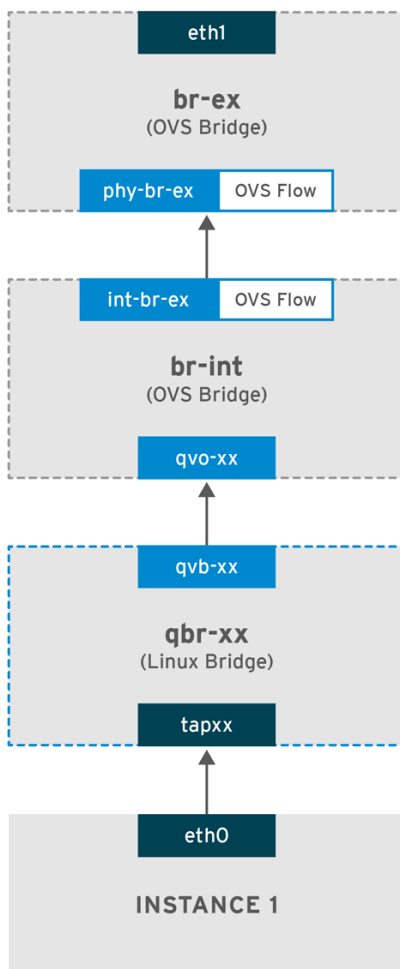
- [Custom network interface templates](#) in the *Advanced Overcloud Customization* guide
- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide
- [network create](#) in the *Command Line Interface Reference*
- [subnet create](#) in the *Command Line Interface Reference*
- [server create](#) in the *Command Line Interface Reference*

7.4. HOW DOES THE FLAT PROVIDER NETWORK PACKET FLOW WORK?

This section describes in detail how traffic flows to and from an instance with flat provider network configuration.

The flow of outgoing traffic in a flat provider network

The following diagram describes the packet flow for traffic leaving an instance and arriving directly at an external network. After you configure the **br-ex** external bridge, add the physical interface to the bridge, and spawn an instance to a Compute node, the resulting configuration of interfaces and bridges resembles the configuration in the following diagram (if using the **iptables_hybrid** firewall driver):



1. Packets leave the **eth0** interface of the instance and arrive at the linux bridge **qbr-xx**.
2. Bridge **qbr-xx** is connected to **br-int** using veth pair **qvb-xx <-> qvo-xxx**. This is because the bridge is used to apply the inbound/outbound firewall rules defined by the security group.
3. Interface **qvb-xx** is connected to the **qbr-xx** linux bridge, and **qvoxx** is connected to the **br-int** Open vSwitch (OVS) bridge.

An example configuration of `qbr-xx` Linux bridge:

```
# brctl show
qbr269d4d73-e7 8000.061943266ebb no qvb269d4d73-e7
tap269d4d73-e7
```

The configuration of **qvo-xx** on **br-int**:

```
# ovs-vsctl show
Bridge br-int
  fail_mode: secure
  Interface "qvof63599ba-8f"
  Port "qvo269d4d73-e7"
    tag: 5
    Interface "qvo269d4d73-e7"
```



NOTE

Port **qvo-xx** is tagged with the internal VLAN tag associated with the flat provider network. In this example, the VLAN tag is **5**. When the packet reaches **qvo-xx**, the VLAN tag is appended to the packet header.

The packet is then moved to the **br-ex** OVS bridge using the patch-peer **int-br-ex <-> phy-br-ex**.

Example configuration of the patch-peer on **br-int**:

```
# ovs-vsctl show
Bridge br-int
  fail_mode: secure
  Port int-br-ex
    Interface int-br-ex
      type: patch
      options: {peer=phy-br-ex}
```

Example configuration of the patch-peer on **br-ex**:

```
Bridge br-ex
  Port phy-br-ex
    Interface phy-br-ex
      type: patch
      options: {peer=int-br-ex}
  Port br-ex
    Interface br-ex
      type: internal
```

When this packet reaches **phy-br-ex** on **br-ex**, an OVS flow inside **br-ex** strips the VLAN tag (5) and forwards it to the physical interface.

In the following example, the output shows the port number of **phy-br-ex** as **2**.

```
# ovs-ofctl show br-ex
OFPT_FEATURES_REPLY (xid=0x2): dpid:00003440b5c90dc6
n_tables:254, n_buffers:256
capabilities: FLOW_STATS TABLE_STATS PORT_STATS QUEUE_STATS ARP_MATCH_IP
actions: OUTPUT SET_VLAN_VID SET_VLAN_PCP STRIP_VLAN SET_DL_SRC SET_DL_DST
SET_NW_SRC SET_NW_DST SET_NW_TOS SET_TP_SRC SET_TP_DST ENQUEUE

2(phy-br-ex): addr:ba:b5:7b:ae:5c:a2
  config:  0
  state:   0
  speed: 0 Mbps now, 0 Mbps max
```

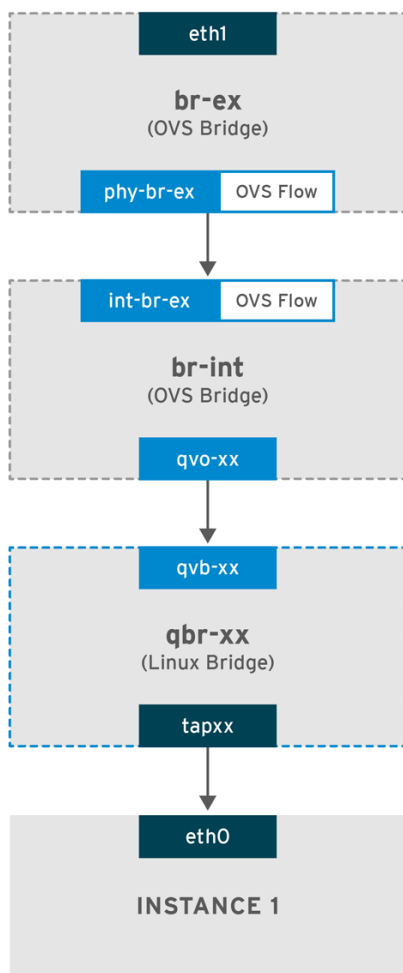
The following output shows any packet that arrives on **phy-br-ex (in_port=2)** with a VLAN tag of **5 (dl_vlan=5)**. In addition, an OVS flow in **br-ex** strips the VLAN tag and forwards the packet to the physical interface.

```
# ovs-ofctl dump-flows br-ex
NXST_FLOW reply (xid=0x4):
  cookie=0x0, duration=4703.491s, table=0, n_packets=3620, n_bytes=333744, idle_age=0, priority=1
  actions=NORMAL
  cookie=0x0, duration=3890.038s, table=0, n_packets=13, n_bytes=1714, idle_age=3764,
  priority=4,in_port=2,dl_vlan=5 actions=strip_vlan,NORMAL
  cookie=0x0, duration=4702.644s, table=0, n_packets=10650, n_bytes=447632, idle_age=0,
  priority=2,in_port=2 actions=drop
```

If the physical interface is another VLAN-tagged interface, then the physical interface adds the tag to the packet.

The flow of incoming traffic in a flat provider network

This section contains information about the flow of incoming traffic from the external network until it arrives at the interface of the instance.



OPENSTACK_450456_0617

1. Incoming traffic arrives at **eth1** on the physical node.
2. The packet passes to the **br-ex** bridge.
3. The packet moves to **br-int** via the patch-peer **phy-br-ex <--> int-br-ex**.

In the following example, **int-br-ex** uses port number **15**. See the entry containing **15(int-br-ex)**:

```

# ovs-ofctl show br-int
OFPT_FEATURES_REPLY (xid=0x2): dpid:00004e67212f644d
n_tables:254, n_buffers:256
capabilities: FLOW_STATS TABLE_STATS PORT_STATS QUEUE_STATS ARP_MATCH_IP
actions: OUTPUT SET_VLAN_VID SET_VLAN_PCP STRIP_VLAN SET_DL_SRC SET_DL_DST
SET_NW_SRC SET_NW_DST SET_NW_TOS SET_TP_SRC SET_TP_DST ENQUEUE
15(int-br-ex): addr:12:4e:44:a9:50:f4
  config: 0
  state: 0
  speed: 0 Mbps now, 0 Mbps max
  
```

Observing the traffic flow on br-int

1. When the packet arrives at **int-br-ex**, an OVS flow rule within the **br-int** bridge amends the packet to add the internal VLAN tag **5**. See the entry for **actions=mod_vlan_vid:5**:

```

# ovs-ofctl dump-flows br-int
NXST_FLOW reply (xid=0x4):
  
```

```

cookie=0x0, duration=5351.536s, table=0, n_packets=12118, n_bytes=510456, idle_age=0,
priority=1 actions=NORMAL
cookie=0x0, duration=4537.553s, table=0, n_packets=3489, n_bytes=321696, idle_age=0,
priority=3,in_port=15,vlan_tci=0x0000 actions=mod_vlan_vid:5,NORMAL
cookie=0x0, duration=5350.365s, table=0, n_packets=628, n_bytes=57892, idle_age=4538,
priority=2,in_port=15 actions=drop
cookie=0x0, duration=5351.432s, table=23, n_packets=0, n_bytes=0, idle_age=5351, priority=0
actions=drop

```

1. The second rule manages packets that arrive on int-br-ex (in_port=15) with no VLAN tag (vlan_tci=0x0000): This rule adds VLAN tag 5 to the packet (**actions=mod_vlan_vid:5,NORMAL**) and forwards it to **qvxxx**.
2. **qvxxx** accepts the packet and forwards it to **qvbxx**, after stripping away the VLAN tag.
3. The packet then reaches the instance.



NOTE

VLAN tag 5 is an example VLAN that was used on a test Compute node with a flat provider network; this value was assigned automatically by **neutron-openvswitch-agent**. This value may be different for your own flat provider network, and can differ for the same network on two separate Compute nodes.

Additional resources

- [Troubleshooting instance-physical network connections on flat provider networks](#)

7.5. TROUBLESHOOTING INSTANCE-PHYSICAL NETWORK CONNECTIONS ON FLAT PROVIDER NETWORKS

The output provided in "How does the flat provider network packet flow work?" provides sufficient debugging information for troubleshooting a flat provider network, should anything go wrong. The following steps contain further information about the troubleshooting process.

Procedure

1. Review the bridge_mappings:

Verify that the physical network name you use (for example, **physnet1**) is consistent with the contents of the **bridge_mapping** configuration as shown in this example:

```

# grep bridge_mapping /etc/neutron/plugins/ml2/openvswitch_agent.ini
bridge_mappings = physnet1:br-ex

# openstack network show provider-flat
...
| provider:physical_network | physnet1
...

```

2. Review the network configuration:

Confirm that the network is created as **external**, and uses the **flat** type:

```
# openstack network show provider-flat
```

```

...
| provider:network_type | flat
| router:external      | True
...

```

3. Review the patch-peer:

Run the **ovs-vsctl show** command, and verify that **br-int** and **br-ex** are connected using a patch-peer **int-br-ex <--> phy-br-ex**.

```

# ovs-vsctl show
Bridge br-int
  fail_mode: secure
  Port int-br-ex
    Interface int-br-ex
      type: patch
      options: {peer=phy-br-ex}

```

Example configuration of the patch-peer on **br-ex**:

```

Bridge br-ex
  Port phy-br-ex
    Interface phy-br-ex
      type: patch
      options: {peer=int-br-ex}
  Port br-ex
    Interface br-ex
      type: internal

```

This connection is created when you restart the **neutron-openvswitch-agent** service, if **bridge_mapping** is correctly configured in **/etc/neutron/plugins/ml2/openvswitch_agent.ini**. Re-check the **bridge_mapping** setting if the connection is not created after you restart the service.

4. Review the network flows:

Run **ovs-ofctl dump-flows br-ex** and **ovs-ofctl dump-flows br-int** and review whether the flows strip the internal VLAN IDs for outgoing packets, and add VLAN IDs for incoming packets. This flow is first added when you spawn an instance to this network on a specific Compute node.

- If this flow is not created after spawning the instance, verify that the network is created as **flat**, is **external**, and that the **physical_network** name is correct. In addition, review the **bridge_mapping** settings.
- Finally, review the **ifcfg-br-ex** and **ifcfg-ethx** configuration. Ensure that **ethX** is added as a port within **br-ex**, and that **ifcfg-br-ex** and **ifcfg-ethx** have an **UP** flag in the output of **ip a**.

The following output shows **eth1** is a port in **br-ex**:

```

Bridge br-ex
  Port phy-br-ex
    Interface phy-br-ex
      type: patch
      options: {peer=int-br-ex}
  Port "eth1"
    Interface "eth1"

```

The following example demonstrates that **eth1** is configured as an OVS port, and that the kernel knows to transfer all packets from the interface, and send them to the OVS bridge **br-ex**. This can be observed in the entry: **master ovs-system**.

```
# ip a
5: eth1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq master ovs-system state
UP qlen 1000
```

Additional resources

- [How does the flat provider network packet flow work?](#)
- [Configuring bridge mappings](#)

7.6. CONFIGURING VLAN PROVIDER NETWORKS

When you connect multiple VLAN-tagged interfaces on a single NIC to multiple provider networks, these new VLAN provider networks can connect VM instances directly to external networks.

Prerequisites

- You have a physical network, with a range of VLANs.
This example uses a physical network called **physnet1**, with a range of VLANs, **171-172**.
- Your Network nodes and Compute nodes are connected to a physical network using a physical interface.
This example uses Network nodes and Compute nodes that are connected to a physical network, **physnet1**, using a physical interface, **eth1**.
- The switch ports that these interfaces connect to must be configured to trunk the required VLAN ranges.

Procedure

1. On the undercloud host, logged in as the stack user, create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-modules-environment.yaml
```

TIP

The Red Hat OpenStack Platform Orchestration service (heat) uses a set of plans called *templates* to install and configure your environment. You can customize aspects of the overcloud with a *custom environment file*, which is a special type of template that provides customization for your orchestration templates.

2. In the YAML environment file under **parameter_defaults**, use **NeutronTypeDrivers** to specify your network type drivers.

Example

```
parameter_defaults:
  NeutronTypeDrivers: vxlan,flat,vlan
```

3. Configure the **NeutronNetworkVLANRanges** setting to reflect the physical network and VLAN ranges in use:

Example

```
parameter_defaults:
  NeutronTypeDrivers: 'vxlan,flat,vlan'
  NeutronNetworkVLANRanges: 'physnet1:171,172'
```

4. Create an external network bridge (*br-ex*), and associate a port (*eth1*) with it. This example configures *eth1* to use *br-ex*:

Example

```
parameter_defaults:
  NeutronTypeDrivers: 'vxlan,flat,vlan'
  NeutronNetworkVLANRanges: 'physnet1:171,172'
  NeutronBridgeMappings: 'datacentre:br-ex,tenant:br-int'
```

5. Run the **openstack overcloud deploy** command and include the core templates and the environment files, including this new environment file.



IMPORTANT

The order of the environment files is important because the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-neutron-
environment.yaml
```

Verification steps

1. Create the external networks as type **vlan**, and associate them with the configured **physical_network**.

When you create the external networks, use the **--shared** option so that users in other project can share the external networks and can connect VM instances directly.

Run the following example command to create two networks: one for VLAN 171, and another for VLAN 172:

Example

```
$ openstack network create \
  --provider-network-type vlan \
  --external \
```

```

--provider-physical-network physnet1 \
--provider-segment 171 \
--share \
provider-vlan171

$ openstack network create \
--provider-network-type vlan \
--external \
--provider-physical-network physnet1 \
--provider-segment 172 \
--share \
provider-vlan172

```

2. Create a number of subnets and configure them to use the external network.

You can use either **openstack subnet create** or the dashboard to create these subnets. Ensure that the external subnet details you have received from your network administrator are correctly associated with each VLAN.

In this example, VLAN 171 uses subnet **10.65.217.0/24** and VLAN 172 uses **10.65.218.0/24**:

Example

```

$ openstack subnet create \
--network provider-171 \
--subnet-range 10.65.217.0/24 \
--dhcp \
--gateway 10.65.217.254 \
subnet-provider-171

$ openstack subnet create \
--network provider-172 \
--subnet-range 10.65.218.0/24 \
--dhcp \
--gateway 10.65.218.254 \
subnet-provider-172

```

Additional resources

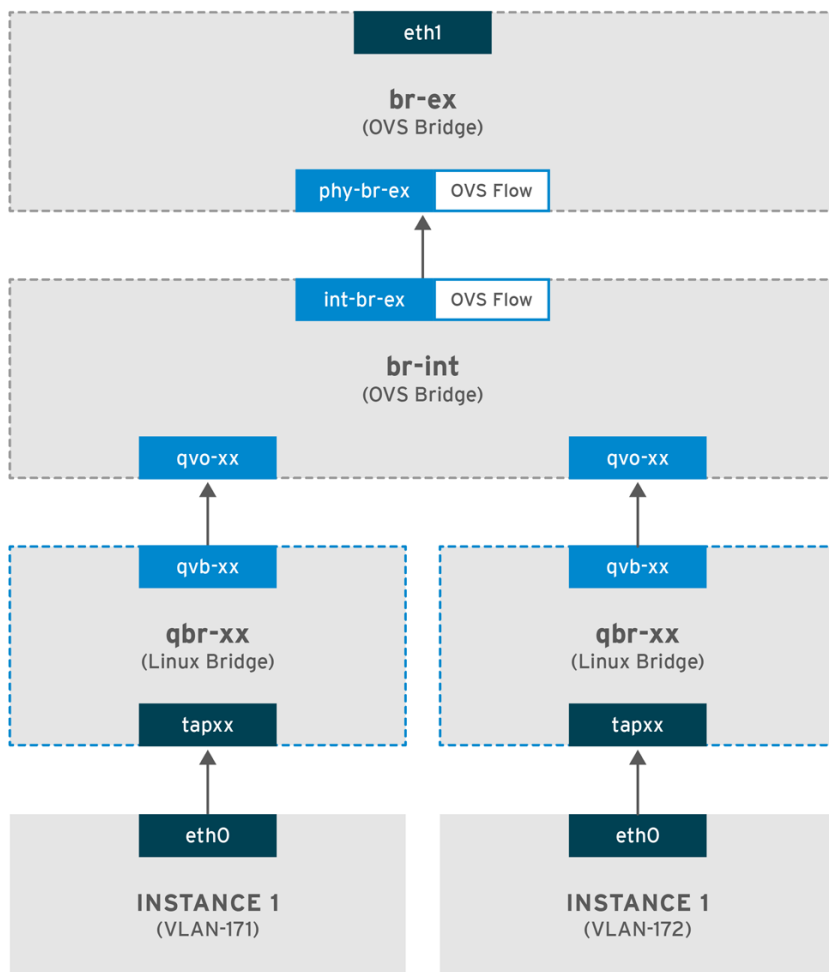
- [Custom network interface templates](#) in the *Advanced Overcloud Customization* guide
- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide
- [network create](#) in the *Command Line Interface Reference*
- [subnet create](#) in the *Command Line Interface Reference*

7.7. HOW DOES THE VLAN PROVIDER NETWORK PACKET FLOW WORK?

This section describes in detail how traffic flows to and from an instance with VLAN provider network configuration.

The flow of outgoing traffic in a VLAN provider network

The following diagram describes the packet flow for traffic leaving an instance and arriving directly to a VLAN provider external network. This example uses two instances attached to the two VLAN networks (171 and 172). After you configure *br-ex*, add a physical interface to it, and spawn an instance to a Compute node, the resulting configuration of interfaces and bridges resembles the configuration in the following diagram:



OPENSTACK_450456_0617

1. Packets leaving the *eth0* interface of the instance arrive at the linux bridge *qbr-xx* connected to the instance.
2. *qbr-xx* is connected to *br-int* using veth pair *qvbxx* \leftrightarrow *qvoxxx*.
3. *qvbxx* is connected to the linux bridge *qbr-xx* and *qvoxx* is connected to the Open vSwitch bridge *br-int*.

Example configuration of *qbr-xx* on the Linux bridge.

This example features two instances and two corresponding linux bridges:

```
# brctl show
bridge name bridge id STP enabled interfaces
qbr84878b78-63 8000.e6b3df9451e0 no qvb84878b78-63
tap84878b78-63

qbr86257b61-5d 8000.3a3c888eeae6 no qvb86257b61-5d
tap86257b61-5d
```

The configuration of `qvoxx` on `br-int`:

```

        options: {peer=phy-br-ex}
    Port "qvo86257b61-5d"
        tag: 3

    Interface "qvo86257b61-5d"
    Port "qvo84878b78-63"
        tag: 2
    Interface "qvo84878b78-63"

```

- **qvoxx** is tagged with the internal VLAN tag associated with the VLAN provider network. In this example, the internal VLAN tag 2 is associated with the VLAN provider network **provider-171** and VLAN tag 3 is associated with VLAN provider network **provider-172**. When the packet reaches `qvoxx`, the this VLAN tag is added to the packet header.
- The packet is then moved to the `br-ex` OVS bridge using patch-peer **int-br-ex** ↔ **phy-br-ex**. Example patch-peer on `br-int`:

```

Bridge br-int
    fail_mode: secure
    Port int-br-ex
        Interface int-br-ex
            type: patch
            options: {peer=phy-br-ex}

```

Example configuration of the patch peer on `br-ex`:

```

Bridge br-ex
    Port phy-br-ex
        Interface phy-br-ex
            type: patch
            options: {peer=int-br-ex}
    Port br-ex
        Interface br-ex
            type: internal

```

- When this packet reaches `phy-br-ex` on `br-ex`, an OVS flow inside `br-ex` replaces the internal VLAN tag with the actual VLAN tag associated with the VLAN provider network.

The output of the following command shows that the port number of `phy-br-ex` is **4**:

```

# ovs-ofctl show br-ex
4(phy-br-ex): addr:32:e7:a1:6b:90:3e
  config: 0
  state: 0
  speed: 0 Mbps now, 0 Mbps max

```

The following command shows any packet that arrives on `phy-br-ex` (**in_port=4**) which has VLAN tag 2 (**dl_vlan=2**). Open vSwitch replaces the VLAN tag with 171 (**actions=mod_vlan_vid:171,NORMAL**) and forwards the packet to the physical interface. The command also shows any packet that arrives on `phy-br-ex` (**in_port=4**) which has VLAN tag 3 (**dl_vlan=3**). Open vSwitch replaces the VLAN tag with 172 (**actions=mod_vlan_vid:172,NORMAL**) and forwards the packet to the physical interface. The `neutron-openvswitch-agent` adds these rules.


```
# ovs-ofctl dump-flows br-ex
NXST_FLOW reply (xid=0x4):
NXST_FLOW reply (xid=0x4):
  cookie=0x0, duration=6527.527s, table=0, n_packets=29211, n_bytes=2725576, idle_age=0,
  priority=1 actions=NORMAL
  cookie=0x0, duration=2939.172s, table=0, n_packets=117, n_bytes=8296, idle_age=58,
  priority=4,in_port=4,dl_vlan=3 actions=mod_vlan_vid:172,NORMAL
  cookie=0x0, duration=6111.389s, table=0, n_packets=145, n_bytes=9368, idle_age=98,
  priority=4,in_port=4,dl_vlan=2 actions=mod_vlan_vid:171,NORMAL
  cookie=0x0, duration=6526.675s, table=0, n_packets=82, n_bytes=6700, idle_age=2462,
  priority=2,in_port=4 actions=drop
```

- This packet is then forwarded to physical interface *eth1*.

The flow of incoming traffic in a VLAN provider network

The following example flow was tested on a Compute node using VLAN tag 2 for provider network provider-171 and VLAN tag 3 for provider network provider-172. The flow uses port 18 on the integration bridge br-int.

Your VLAN provider network may require a different configuration. Also, the configuration requirement for a network may differ between two different Compute nodes.

The output of the following command shows *int-br-ex* with port number 18:

```
# ovs-ofctl show br-int
18(int-br-ex): addr:fe:b7:cb:03:c5:c1
  config: 0
  state: 0
  speed: 0 Mbps now, 0 Mbps max
```

The output of the following command shows the flow rules on br-int.

```
# ovs-ofctl dump-flows br-int
NXST_FLOW reply (xid=0x4):
  cookie=0x0, duration=6770.572s, table=0, n_packets=1239, n_bytes=127795, idle_age=106,
  priority=1 actions=NORMAL

  cookie=0x0, duration=3181.679s, table=0, n_packets=2605, n_bytes=246456, idle_age=0,
  priority=3,in_port=18,dl_vlan=172 actions=mod_vlan_vid:3,NORMAL

  cookie=0x0, duration=6353.898s, table=0, n_packets=5077, n_bytes=482582, idle_age=0,
  priority=3,in_port=18,dl_vlan=171 actions=mod_vlan_vid:2,NORMAL

  cookie=0x0, duration=6769.391s, table=0, n_packets=22301, n_bytes=2013101, idle_age=0,
  priority=2,in_port=18 actions=drop

  cookie=0x0, duration=6770.463s, table=23, n_packets=0, n_bytes=0, idle_age=6770, priority=0
  actions=drop
```

Incoming flow example

This example demonstrates the the following br-int OVS flow:

```
cookie=0x0, duration=3181.679s, table=0, n_packets=2605, n_bytes=246456, idle_age=0,
priority=3,in_port=18,dl_vlan=172 actions=mod_vlan_vid:3,NORMAL
```

- A packet with VLAN tag 172 from the external network reaches the *br-ex* bridge via *eth1* on the physical node.
- The packet moves to *br-int* via the patch-peer **phy-br-ex <-> int-br-ex**.
- The packet matches the flow's criteria (**in_port=18,dl_vlan=172**).
- The flow actions (**actions=mod_vlan_vid:3,NORMAL**) replace the VLAN tag 172 with internal VLAN tag 3 and forwards the packet to the instance with normal Layer 2 processing.

Additional resources

- [Troubleshooting instance-physical network connections on VLAN provider networks](#)

7.8. TROUBLESHOOTING INSTANCE-PHYSICAL NETWORK CONNECTIONS ON VLAN PROVIDER NETWORKS

Refer to the packet flow described in "How does the VLAN provider network packet flow work?" when troubleshooting connectivity in a VLAN provider network. In addition, review the following configuration options:

Procedure

1. Verify that physical network name is used consistently. In this example, **physnet1** is used consistently while creating the network, and within the **bridge_mapping** configuration:

```
# grep bridge_mapping /etc/neutron/plugins/ml2/openvswitch_agent.ini
bridge_mappings = physnet1:br-ex

# openstack network show provider-vlan171
...
| provider:physical_network | physnet1
...
```

2. Confirm that the network was created as **external**, is type **vlan**, and uses the correct **segmentation_id** value:

```
# openstack network show provider-vlan171
...
| provider:network_type   | vlan           |
| provider:physical_network | physnet1       |
| provider:segmentation_id | 171            |
...
```

3. Run **ovs-vsctl show** and verify that *br-int* and *br-ex* are connected using the patch-peer **int-br-ex <-> phy-br-ex**.

This connection is created while restarting *neutron-openvswitch-agent*, provided that the **bridge_mapping** is correctly configured in **/etc/neutron/plugins/ml2/openvswitch_agent.ini**. Recheck the *bridge_mapping* setting if this is not created even after restarting the service.

4. To review the flow of outgoing packets, run **ovs-ofctl dump-flows br-ex** and **ovs-ofctl dump-flows**

br-int, and verify that the flows map the internal VLAN IDs to the external VLAN ID (*segmentation_id*). For incoming packets, map the external VLAN ID to the internal VLAN ID.

This flow is added by the neutron OVS agent when you spawn an instance to this network for the first time. If this flow is not created after spawning the instance, ensure that the network is created as **vlan**, is **external**, and that the **physical_network** name is correct. In addition, re-check the **bridge_mapping** settings.

5. Finally, re-check the *ifcfg-br-ex* and *ifcfg-ethx* configuration. Ensure that *br-ex* includes port *ethX*, and that both *ifcfg-br-ex* and *ifcfg-ethx* have an **UP** flag in the output of the **ip a** command.

For example, the following output shows that *eth1* is a port in *br-ex*:

```
Bridge br-ex
  Port phy-br-ex
    Interface phy-br-ex
      type: patch
      options: {peer=int-br-ex}
  Port "eth1"
    Interface "eth1"
```

The following command shows that *eth1* has been added as a port, and that the kernel is configured to move all packets from the interface to the OVS bridge *br-ex*. This is demonstrated by the entry: **master ovs-system**.

```
# ip a
5: eth1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq master ovs-system state
UP qlen 1000
```

Additional resources

- [How does the VLAN provider network packet flow work?](#)

7.9. ENABLING MULTICAST SNOOPING FOR PROVIDER NETWORKS IN AN ML2/OVS DEPLOYMENT

To prevent flooding multicast packets to every port in a Red Hat OpenStack Platform (RHOSP) provider network, you must enable multicast snooping. In RHOSP deployments that use the Modular Layer 2 plug-in with the Open vSwitch mechanism driver (ML2/OVS), you do this by adding the appropriate Puppet variable to a custom environment file and running the **openstack overcloud deploy** command.



IMPORTANT

You should thoroughly test and understand any multicast snooping configuration before applying it to a production environment. Misconfiguration can break multicasting or cause erratic network behavior.

Prerequisites

- Your configuration must only use ML2/OVS provider networks.
- Your physical routers must also have IGMP snooping enabled.

That is, the physical router must send IGMP query packets on the provider network to solicit regular IGMP reports from multicast group members to maintain the snooping cache in OVS (and for physical networking).

- An RHOSP Networking service security group rule must be in place to allow inbound IGMP to the VM instances (or port security disabled).
In this example, a rule is created for the **ping_ssh** security group:

Example

```
$ openstack security group rule create --protocol igmp --ingress ping_ssh
```

Procedure

1. On the undercloud host, logged in as the stack user, create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-ovs-environment.yaml
```

TIP

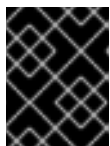
The Orchestration service (heat) uses a set of plans called templates to install and configure your environment. You can customize aspects of the overcloud with a custom environment file, which is a special type of template that provides customization for your heat templates.

2. In the YAML environment file under the **ExtraConfig** section for the appropriate role, set the Puppet variable, **igmp_snooping_enable**, to **true**.

Example

If the role used is **ComputeOvsDpdk**, then the lines that you add to the custom environment file would be:

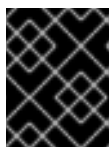
```
parameter_defaults:
  ComputeOvsDpdkExtraConfig:
    neutron::agents::ml2::ovs::igmp_snooping_enable: true
```



IMPORTANT

Ensure that you add a whitespace character between the single colon (:) and the value.

3. Run the **openstack overcloud deploy** command and include the core heat templates, environment files, and this new custom environment file.



IMPORTANT

The order of the environment files is important as the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-ovs-
environment.yaml
```

Verification steps

- Verify that the multicast snooping is enabled.

Example

```
# sudo ovs-vsctl list bridge br-int
```

Sample output

```
...
mcast_snooping_enable: true
...
other_config: {mac-table-size="50000", mcast-snooping-disable-flood-unregistered=True}
...
```

Additional resources

- [Neutron](#) in *Component, Plug-In, and Driver Support in Red Hat OpenStack Platform*
- [Environment Files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

7.10. ENABLING MULTICAST IN AN ML2/OVN DEPLOYMENT

To support multicast traffic, modify the deployment's security configuration to allow multicast traffic to reach the virtual machine (VM) instances in the multicast group. To prevent multicast traffic flooding, enable IGMP snooping.



IMPORTANT

Test and understand any multicast snooping configuration before applying it to a production environment. Misconfiguration can break multicasting or cause erratic network behavior.

Prerequisites

- An OpenStack deployment with the ML2/OVN mechanism driver.

Procedure

1. Configure security to allow multicast traffic to the appropriate VM instances. For instance, create a pair of security group rules to allow IGMP traffic from the IGMP querier to enter and exit the VM instances, and a third rule to allow multicast traffic.

Example

A security group *mySG* allows IGMP traffic to enter and exit the VM instances.

```
openstack security group rule create --protocol igmp --ingress mySG
openstack security group rule create --protocol igmp --egress mySG
```

Another rule allows multicast traffic to reach VM instances.

```
openstack security group rule create --protocol udp mySG
```

As an alternative to setting security group rules, some operators choose to selectively disable port security on the network. If you choose to disable port security, consider and plan for any related security risks.

- Set the heat parameter **NeutronEnableIcmpSnooping: True** in an environment file on the undercloud node. For instance, add the following lines to *ovn-extras.yaml*.

Example

```
parameter_defaults:
  NeutronEnableIcmpSnooping: True
```

- Include the environment file in the **openstack overcloud deploy** command with any other environment files that are relevant to your environment and deploy the overcloud.

```
$ openstack overcloud deploy \
--templates \
...
-e <other_overcloud_environment_files> \

-e ovn-extras.yaml \
...
```

Replace **<other_overcloud_environment_files>** with the list of environment files that are part of your existing deployment.

Verification steps

- Verify that the multicast snooping is enabled. List the northbound database *Logical_Switch* table.

```
$ ovn-nbctl list Logical_Switch
```

Sample output

```
_uuid      : d6a2fbcd-aaa4-4b9e-8274-184238d66a15
other_config : {mcast_flood_unregistered="false", mcast_snoop="true"}
...
```

The Networking Service (neutron) *igmp_snooping_enable* configuration is translated into the *mcast_snoop* option set in the *other_config* column of the *Logical_Switch* table in the OVN Northbound Database. Note that *mcast_flood_unregistered* is always "false".

2. Show the IGMP groups.

```
$ ovn-sbctl list IGMP_group
```

Sample output

```
_uuid : 2d6cae4c-bd82-4b31-9c63-2d17cbeadc4e
address : "225.0.0.120"
chassis : 34e25681-f73f-43ac-a3a4-7da2a710ecd3
datapath : eaf0f5cc-a2c8-4c30-8def-2bc1ec9dcabc
ports : [5eaf9dd5-ee5-4749-ac60-4c1451901c56, 8a69efc5-38c5-48fb-bbab-30f2bf9b8d45]
...
```

Additional resources

- [Neutron](#) in *Component, Plug-In, and Driver Support in Red Hat OpenStack Platform*
- [Environment Files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

7.11. ENABLING COMPUTE METADATA ACCESS

Instances connected as described in this chapter are directly attached to the provider external networks, and have external routers configured as their default gateway. No OpenStack Networking (neutron) routers are used. This means that *neutron* routers cannot be used to proxy metadata requests from instances to the *nova-metadata* server, which may result in failures while running *cloud-init*. However, this issue can be resolved by configuring the dhcp agent to proxy metadata requests. You can enable this functionality in `/etc/neutron/dhcp_agent.ini`. For example:

```
enable_isolated_metadata = True
```

7.12. FLOATING IP ADDRESSES

You can use the same network to allocate floating IP addresses to instances, even if the floating IPs are already associated with private networks. The addresses that you allocate as floating IPs from this network are bound to the *qrouter-xxx* namespace on the Network node, and perform *DNAT-SNAT* to the associated private IP address. In contrast, the IP addresses that you allocate for direct external network access are bound directly inside the instance, and allow the instance to communicate directly with external network.

CHAPTER 8. CONFIGURING PHYSICAL SWITCHES FOR OPENSTACK NETWORKING

This chapter documents the common physical switch configuration steps required for OpenStack Networking. Vendor-specific configuration is included for certain switches.

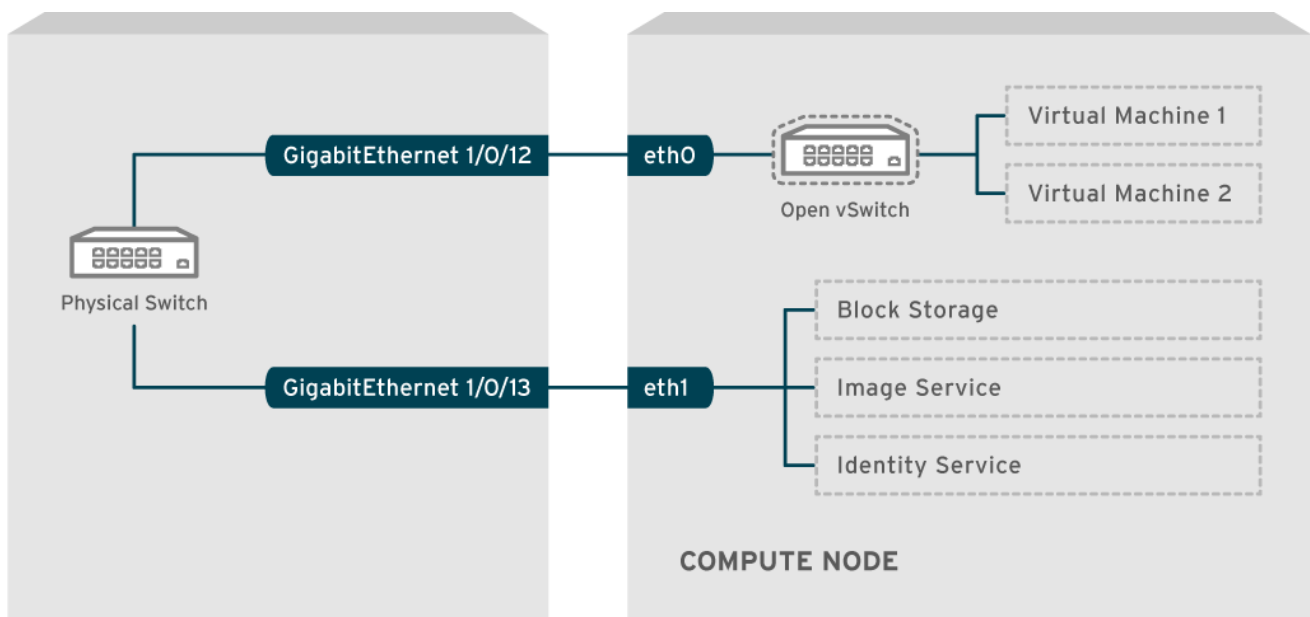
8.1. PLANNING YOUR PHYSICAL NETWORK ENVIRONMENT

The physical network adapters in your OpenStack nodes carry different types of network traffic, such as instance traffic, storage data, or authentication requests. The type of traffic these NICs carry affects how you must configure the ports on the physical switch.

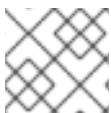
First, you must decide which physical NICs on your Compute node you want to carry which types of traffic. Then, when the NIC is cabled to a physical switch port, you must configure the switch port to allow trunked or general traffic.

For example, the following diagram depicts a Compute node with two NICs, eth0 and eth1. Each NIC is cabled to a Gigabit Ethernet port on a physical switch, with eth0 carrying instance traffic, and eth1 providing connectivity for OpenStack services:

Figure 8.1. Sample network layout



OPENSTACK_377160_1115



NOTE

This diagram does not include any additional redundant NICs required for fault tolerance.

For information on network interface bonds, see the [Network Interface Bonding](#) chapter of the [Advanced Overcloud Customization](#) guide.

8.2. CONFIGURING A CISCO CATALYST SWITCH

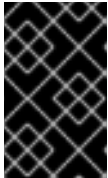
8.2.1. About trunk ports

With OpenStack Networking you can connect instances to the VLANs that already exist on your physical

network. The term *trunk* is used to describe a port that allows multiple VLANs to traverse through the same port. Using these ports, VLANs can span across multiple switches, including virtual switches. For example, traffic tagged as VLAN110 in the physical network reaches the Compute node, where the 8021q module directs the tagged traffic to the appropriate VLAN on the vSwitch.

8.2.2. Configuring trunk ports for a Cisco Catalyst switch

- If using a Cisco Catalyst switch running Cisco IOS, you might use the following configuration syntax to allow traffic for VLANs 110 and 111 to pass through to your instances. This configuration assumes that your physical node has an ethernet cable connected to interface GigabitEthernet1/0/12 on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

```
interface GigabitEthernet1/0/12
description Trunk to Compute Node
spanning-tree portfast trunk
switchport trunk encapsulation dot1q
switchport mode trunk
switchport trunk native vlan 2
switchport trunk allowed vlan 2,110,111
```

Use the following list to understand these parameters:

Field	Description
interface GigabitEthernet1/0/12	The switch port that the NIC of the X node connects to. Ensure that you replace the GigabitEthernet1/0/12 value with the correct port value for your environment. Use the show interface command to view a list of ports.
description Trunk to Compute Node	A unique and descriptive value that you can use to identify this interface.
spanning-tree portfast trunk	If your environment uses STP, set this value to instruct Port Fast that this port is used to trunk traffic.
switchport trunk encapsulation dot1q	Enables the 802.1q trunking standard (rather than ISL). This value varies depending on the configuration that your switch supports.
switchport mode trunk	Configures this port as a trunk port, rather than an access port, meaning that it allows VLAN traffic to pass through to the virtual switches.

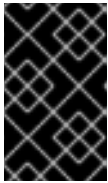
Field	Description
switchport trunk native vlan 2	Set a native VLAN to instruct the switch where to send untagged (non-VLAN) traffic.
switchport trunk allowed vlan 2,110,111	Defines which VLANs are allowed through the trunk.

8.2.3. About access ports

Not all NICs on your Compute node carry instance traffic, and so you do not need to configure all NICs to allow multiple VLANs to pass through. Access ports require only one VLAN, and might fulfill other operational requirements, such as transporting management traffic or Block Storage data. These ports are commonly known as access ports and usually require a simpler configuration than trunk ports.

8.2.4. Configuring access ports for a Cisco Catalyst switch

- Using the example from the [Figure 8.1, "Sample network layout"](#) diagram, GigabitEthernet1/0/13 (on a Cisco Catalyst switch) is configured as an access port for **eth1**. In this configuration, your physical node has an ethernet cable connected to interface GigabitEthernet1/0/12 on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

```
interface GigabitEthernet1/0/13
description Access port for Compute Node
switchport mode access
switchport access vlan 200
spanning-tree portfast
```

These settings are described below:

Field	Description
interface GigabitEthernet1/0/13	The switch port that the NIC of the X node connects to. Ensure that you replace the GigabitEthernet1/0/12 value with the correct port value for your environment. Use the show interface command to view a list of ports.
description Access port for Compute Node	A unique and descriptive value that you can use to identify this interface.

Field	Description
switchport mode access	Configures this port as an access port, rather than a trunk port.
switchport access vlan 200	Configures the port to allow traffic on VLAN 200. You must configure your Compute node with an IP address from this VLAN.
spanning-tree portfast	If using STP, set this value to instruct STP not to attempt to initialize this as a trunk, allowing for quicker port handshakes during initial connections (such as server reboot).

8.2.5. About LACP port aggregation

You can use LACP to bundle multiple physical NICs together to form a single logical channel. Also known as 802.3ad (or bonding mode 4 in Linux), LACP creates a dynamic bond for load-balancing and fault tolerance. You must configure LACP at both physical ends: on the physical NICs, and on the physical switch ports.

8.2.6. Configuring LACP on the physical NIC

1. Edit the `/home/stack/network-environment.yaml` file:

```
- type: linux_bond
  name: bond1
  mtu: 9000
  bonding_options:{get_param: BondInterfaceOvsOptions};
  members:
    - type: interface
      name: nic3
      mtu: 9000
      primary: true
    - type: interface
      name: nic4
      mtu: 9000
```

2. Configure the Open vSwitch bridge to use LACP:

```
BondInterfaceOvsOptions:
  "mode=802.3ad"
```

For information on configuring network bonds, see the [Network Interface Bonding chapter of the Advanced Overcloud Customization guide](#).

8.2.7. Configuring LACP for a Cisco Catalyst switch

In this example, the Compute node has two NICs using VLAN 100:

Procedure

1. Physically connect both NICs on the Compute node to the switch (for example, ports 12 and 13).
2. Create the LACP port channel:

```
interface port-channel1
  switchport access vlan 100
  switchport mode access
  spanning-tree guard root
```

3. Configure switch ports 12 (Gi1/0/12) and 13 (Gi1/0/13):

```
sw01# config t
Enter configuration commands, one per line. End with CNTL/Z.

sw01(config) interface GigabitEthernet1/0/12
  switchport access vlan 100
  switchport mode access
  speed 1000
  duplex full
  channel-group 10 mode active
  channel-protocol lacp

interface GigabitEthernet1/0/13
  switchport access vlan 100
  switchport mode access
  speed 1000
  duplex full
  channel-group 10 mode active
  channel-protocol lacp
```

4. Review your new port channel. The resulting output lists the new port-channel **Po1**, with member ports **Gi1/0/12** and **Gi1/0/13**:

```
sw01# show etherchannel summary
<snip>

Number of channel-groups in use: 1
Number of aggregators:          1

Group Port-channel Protocol Ports
-----+-----+-----+-----
1   Po1(SD)      LACP   Gi1/0/12(D) Gi1/0/13(D)
```



NOTE

Remember to apply your changes by copying the running-config to the startup-config: **copy running-config startup-config**.

8.2.8. About MTU settings

You must adjust your MTU size for certain types of network traffic. For example, jumbo frames (9000 bytes) are required for certain NFS or iSCSI traffic.

**NOTE**

You must change MTU settings from end-to-end on all hops that the traffic is expected to pass through, including any virtual switches.

Additional resources

- [Configuring maximum transmission unit \(MTU\) settings](#)

8.2.9. Configuring MTU settings for a Cisco Catalyst switch

Complete the steps in this example procedure to enable jumbo frames on your Cisco Catalyst 3750 switch.

1. Review the current MTU settings:

```
sw01# show system mtu

System MTU size is 1600 bytes
System Jumbo MTU size is 1600 bytes
System Alternate MTU size is 1600 bytes
Routing MTU size is 1600 bytes
```

2. MTU settings are changed switch-wide on 3750 switches, and not for individual interfaces. Run the following commands to configure the switch to use jumbo frames of 9000 bytes. You might prefer to configure the MTU settings for individual interfaces, if your switch supports this feature.

```
sw01# config t
Enter configuration commands, one per line. End with CNTL/Z.

sw01(config)# system mtu jumbo 9000
Changes to the system jumbo MTU will not take effect until the next reload is done
```

**NOTE**

Remember to save your changes by copying the running-config to the startup-config: **copy running-config startup-config**.

3. Reload the switch to apply the change.

**IMPORTANT**

Reloading the switch causes a network outage for any devices that are dependent on the switch. Therefore, reload the switch only during a scheduled maintenance period.

```
sw01# reload
Proceed with reload? [confirm]
```

4. After the switch reloads, confirm the new jumbo MTU size. The exact output may differ depending on your switch model. For example, **System MTU** might apply to non-Gigabit interfaces, and **Jumbo MTU** might describe all Gigabit interfaces.

```
sw01# show system mtu

System MTU size is 1600 bytes
System Jumbo MTU size is 9000 bytes
System Alternate MTU size is 1600 bytes
Routing MTU size is 1600 bytes
```

8.2.10. About LLDP discovery

The **ironic-python-agent** service listens for LLDP packets from connected switches. The collected information can include the switch name, port details, and available VLANs. Similar to Cisco Discovery Protocol (CDP), LLDP assists with the discovery of physical hardware during the director introspection process.

8.2.11. Configuring LLDP for a Cisco Catalyst switch

Procedure

1. Run the **lldp run** command to enable LLDP globally on your Cisco Catalyst switch:

```
sw01# config t
Enter configuration commands, one per line. End with CNTL/Z.

sw01(config)# lldp run
```

2. View any neighboring LLDP-compatible devices:

```
sw01# show lldp neighbor
Capability codes:
  (R) Router, (B) Bridge, (T) Telephone, (C) DOCSIS Cable Device
  (W) WLAN Access Point, (P) Repeater, (S) Station, (O) Other

Device ID      Local Intf  Hold-time  Capability  Port ID
DEP42037061562G3  Gi1/0/11   180        B,T         422037061562G3:P1

Total entries displayed: 1
```



NOTE

Remember to save your changes by copying the running-config to the startup-config: **copy running-config startup-config**.

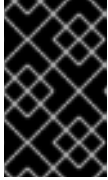
8.3. CONFIGURING A CISCO NEXUS SWITCH

8.3.1. About trunk ports

With OpenStack Networking you can connect instances to the VLANs that already exist on your physical network. The term *trunk* is used to describe a port that allows multiple VLANs to traverse through the same port. Using these ports, VLANs can span across multiple switches, including virtual switches. For example, traffic tagged as VLAN110 in the physical network reaches the Compute node, where the 8021q module directs the tagged traffic to the appropriate VLAN on the vSwitch.

8.3.2. Configuring trunk ports for a Cisco Nexus switch

- If using a Cisco Nexus you might use the following configuration syntax to allow traffic for VLANs 110 and 111 to pass through to your instances. This configuration assumes that your physical node has an ethernet cable connected to interface **Ethernet1/12** on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

```
interface Ethernet1/12
description Trunk to Compute Node
switchport mode trunk
switchport trunk allowed vlan 2,110,111
switchport trunk native vlan 2
end
```

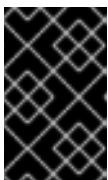
8.3.3. About access ports

Not all NICs on your Compute node carry instance traffic, and so you do not need to configure all NICs to allow multiple VLANs to pass through. Access ports require only one VLAN, and might fulfill other operational requirements, such as transporting management traffic or Block Storage data. These ports are commonly known as access ports and usually require a simpler configuration than trunk ports.

8.3.4. Configuring access ports for a Cisco Nexus switch

Procedure

- Using the example from the [Figure 8.1, "Sample network layout"](#) diagram, Ethernet1/13 (on a Cisco Nexus switch) is configured as an access port for **eth1**. This configuration assumes that your physical node has an ethernet cable connected to interface **Ethernet1/13** on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

```
interface Ethernet1/13
description Access port for Compute Node
switchport mode access
switchport access vlan 200
```

8.3.5. About LACP port aggregation

You can use LACP to bundle multiple physical NICs together to form a single logical channel. Also known as 802.3ad (or bonding mode 4 in Linux), LACP creates a dynamic bond for load-balancing and fault tolerance. You must configure LACP at both physical ends: on the physical NICs, and on the

physical switch ports.

8.3.6. Configuring LACP on the physical NIC

1. Edit the `/home/stack/network-environment.yaml` file:

```
- type: linux_bond
  name: bond1
  mtu: 9000
  bonding_options:{get_param: BondInterfaceOvsOptions};
  members:
    - type: interface
      name: nic3
      mtu: 9000
      primary: true
    - type: interface
      name: nic4
      mtu: 9000
```

2. Configure the Open vSwitch bridge to use LACP:

```
BondInterfaceOvsOptions:
  "mode=802.3ad"
```

For information on configuring network bonds, see the [Network Interface Bonding chapter of the Advanced Overcloud Customization](#) guide.

8.3.7. Configuring LACP for a Cisco Nexus switch

In this example, the Compute node has two NICs using VLAN 100:

Procedure

1. Physically connect the Compute node NICs to the switch (for example, ports 12 and 13).
2. Confirm that LACP is enabled:

```
(config)# show feature | include lacp
lacp          1      enabled
```

3. Configure ports 1/12 and 1/13 as access ports, and as members of a channel group. Depending on your deployment, you can deploy trunk interfaces rather than access interfaces.

For example, for Cisco UCI the NICs are virtual interfaces, so you might prefer to configure access ports exclusively. Often these interfaces contain VLAN tagging configurations.

```
interface Ethernet1/13
  description Access port for Compute Node
  switchport mode access
  switchport access vlan 200
  channel-group 10 mode active

interface Ethernet1/13
  description Access port for Compute Node
```



```
switchport mode access
switchport access vlan 200
channel-group 10 mode active
```

8.3.8. About MTU settings

You must adjust your MTU size for certain types of network traffic. For example, jumbo frames (9000 bytes) are required for certain NFS or iSCSI traffic.



NOTE

You must change MTU settings from end-to-end on all hops that the traffic is expected to pass through, including any virtual switches.

Additional resources

- [Configuring maximum transmission unit \(MTU\) settings](#)

8.3.9. Configuring MTU settings for a Cisco Nexus 7000 switch

Apply MTU settings to a single interface on 7000-series switches.

Procedure

- Run the following commands to configure interface 1/12 to use jumbo frames of 9000 bytes:

```
interface ethernet 1/12
mtu 9216
exit
```

8.3.10. About LLDP discovery

The **ironic-python-agent** service listens for LLDP packets from connected switches. The collected information can include the switch name, port details, and available VLANs. Similar to Cisco Discovery Protocol (CDP), LLDP assists with the discovery of physical hardware during the director introspection process.

8.3.11. Configuring LLDP for a Cisco Nexus 7000 switch

Procedure

- You can enable LLDP for individual interfaces on Cisco Nexus 7000-series switches:

```
interface ethernet 1/12
lldp transmit
lldp receive
no lACP suspend-individual
no lACP graceful-convergence

interface ethernet 1/13
lldp transmit
```

```
lldp receive
no lACP suspend-individual
no lACP graceful-convergence
```



NOTE

Remember to save your changes by copying the running-config to the startup-config:
copy running-config startup-config.

8.4. CONFIGURING A CUMULUS LINUX SWITCH

8.4.1. About trunk ports

With OpenStack Networking you can connect instances to the VLANs that already exist on your physical network. The term *trunk* is used to describe a port that allows multiple VLANs to traverse through the same port. Using these ports, VLANs can span across multiple switches, including virtual switches. For example, traffic tagged as VLAN110 in the physical network reaches the Compute node, where the 8021q module directs the tagged traffic to the appropriate VLAN on the vSwitch.

8.4.2. Configuring trunk ports for a Cumulus Linux switch

This configuration assumes that your physical node has transceivers connected to switch ports swp1 and swp2 on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

Procedure

- Use the following configuration syntax to allow traffic for VLANs 100 and 200 to pass through to your instances.

```
auto bridge
iface bridge
bridge-vlan-aware yes
bridge-ports glob swp1-2
bridge-vids 100 200
```

8.4.3. About access ports

Not all NICs on your Compute node carry instance traffic, and so you do not need to configure all NICs to allow multiple VLANs to pass through. Access ports require only one VLAN, and might fulfill other operational requirements, such as transporting management traffic or Block Storage data. These ports are commonly known as access ports and usually require a simpler configuration than trunk ports.

8.4.4. Configuring access ports for a Cumulus Linux switch

This configuration assumes that your physical node has an ethernet cable connected to the interface on the physical switch. Cumulus Linux switches use **eth** for management interfaces and **swp** for access/trunk ports.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

Procedure

- Using the example from the [Figure 8.1, “Sample network layout”](#) diagram, **swp1** (on a Cumulus Linux switch) is configured as an access port.

```

auto bridge
iface bridge
  bridge-vlan-aware yes
  bridge-ports glob swp1-2
  bridge-vids 100 200

auto swp1
iface swp1
  bridge-access 100

auto swp2
iface swp2
  bridge-access 200

```

8.4.5. About LACP port aggregation

You can use LACP to bundle multiple physical NICs together to form a single logical channel. Also known as 802.3ad (or bonding mode 4 in Linux), LACP creates a dynamic bond for load-balancing and fault tolerance. You must configure LACP at both physical ends: on the physical NICs, and on the physical switch ports.

8.4.6. About MTU settings

You must adjust your MTU size for certain types of network traffic. For example, jumbo frames (9000 bytes) are required for certain NFS or iSCSI traffic.



NOTE

You must change MTU settings from end-to-end on all hops that the traffic is expected to pass through, including any virtual switches.

Additional resources

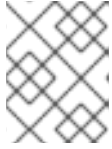
- [Configuring maximum transmission unit \(MTU\) settings](#)

8.4.7. Configuring MTU settings for a Cumulus Linux switch

Procedure

- This example enables jumbo frames on your Cumulus Linux switch.

```
auto swp1
iface swp1
mtu 9000
```



NOTE

Remember to apply your changes by reloading the updated configuration: **sudo ifreload -a**

8.4.8. About LLDP discovery

The **ironic-python-agent** service listens for LLDP packets from connected switches. The collected information can include the switch name, port details, and available VLANs. Similar to Cisco Discovery Protocol (CDP), LLDP assists with the discovery of physical hardware during the director introspection process.

8.4.9. Configuring LLDP for a Cumulus Linux switch

By default, the LLDP service `lldpd` runs as a daemon and starts when the switch boots.

Procedure

- To view all LLDP neighbors on all ports/interfaces, run the following command:

```
cumulus@switch$ netshow lldp
Local Port Speed Mode Remote Port Remote Host Summary
-----
eth0 10G Mgmt ===== swp6 mgmt-sw IP: 10.0.1.11/24
swp51 10G Interface/L3 ===== swp1 spine01 IP: 10.0.0.11/32
swp52 10G Interface/L ===== swp1 spine02 IP: 10.0.0.11/32
```

8.5. CONFIGURING A EXTREME EXOS SWITCH

8.5.1. About trunk ports

With OpenStack Networking you can connect instances to the VLANs that already exist on your physical network. The term *trunk* is used to describe a port that allows multiple VLANs to traverse through the same port. Using these ports, VLANs can span across multiple switches, including virtual switches. For example, traffic tagged as VLAN110 in the physical network reaches the Compute node, where the 8021q module directs the tagged traffic to the appropriate VLAN on the vSwitch.

8.5.2. Configuring trunk ports on an Extreme Networks EXOS switch

If using an X-670 series switch, refer to the following example to allow traffic for VLANs 110 and 111 to pass through to your instances.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

Procedure

- This configuration assumes that your physical node has an ethernet cable connected to interface 24 on the physical switch. In this example, DATA and MNGT are the VLAN names.

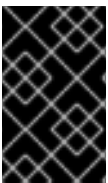
```
#create vlan DATA tag 110
#create vlan MNGT tag 111
#configure vlan DATA add ports 24 tagged
#configure vlan MNGT add ports 24 tagged
```

8.5.3. About access ports

Not all NICs on your Compute node carry instance traffic, and so you do not need to configure all NICs to allow multiple VLANs to pass through. Access ports require only one VLAN, and might fulfill other operational requirements, such as transporting management traffic or Block Storage data. These ports are commonly known as access ports and usually require a simpler configuration than trunk ports.

8.5.4. Configuring access ports for an Extreme Networks EXOS switch

This configuration assumes that your physical node has an ethernet cable connected to interface **10** on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

Procedure

- In this configuration example, on a Extreme Networks X-670 series switch, **10** is used as an access port for **eth1**.

```
create vlan VLANNAME tag NUMBER
configure vlan Default delete ports PORTSTRING
configure vlan VLANNAME add ports PORTSTRING untagged
```

For example:

```
#create vlan DATA tag 110
#configure vlan Default delete ports 10
#configure vlan DATA add ports 10 untagged
```

8.5.5. About LACP port aggregation

You can use LACP to bundle multiple physical NICs together to form a single logical channel. Also known as 802.3ad (or bonding mode 4 in Linux), LACP creates a dynamic bond for load-balancing and

fault tolerance. You must configure LACP at both physical ends: on the physical NICs, and on the physical switch ports.

8.5.6. Configuring LACP on the physical NIC

1. Edit the `/home/stack/network-environment.yaml` file:

```
- type: linux_bond
  name: bond1
  mtu: 9000
  bonding_options:{get_param: BondInterfaceOvsOptions};
  members:
    - type: interface
      name: nic3
      mtu: 9000
      primary: true
    - type: interface
      name: nic4
      mtu: 9000
```

2. Configure the Open vSwitch bridge to use LACP:

```
BondInterfaceOvsOptions:
  "mode=802.3ad"
```

For information on configuring network bonds, see the [Network Interface Bonding](#) chapter of the [Advanced Opencloud Customization](#) guide.

8.5.7. Configuring LACP on an Extreme Networks EXOS switch

Procedure

- In this example, the Compute node has two NICs using VLAN 100:

```
enable sharing MASTERPORT grouping ALL_LAG_PORTS lacp
configure vlan VLANNNAME add ports PORTSTRING tagged
```

For example:

```
#enable sharing 11 grouping 11,12 lacp
#configure vlan DATA add port 11 untagged
```



NOTE

You might need to adjust the timeout period in the LACP negotiation script. For more information, see

https://gtacknowledge.extremenetworks.com/articles/How_To/LACP-configured-ports-interfere-with-PXE-DHCP-on-servers

8.5.8. About MTU settings

You must adjust your MTU size for certain types of network traffic. For example, jumbo frames (9000 bytes) are required for certain NFS or iSCSI traffic.



NOTE

You must change MTU settings from end-to-end on all hops that the traffic is expected to pass through, including any virtual switches.

Additional resources

- [Configuring maximum transmission unit \(MTU\) settings](#)

8.5.9. Configuring MTU settings on an Extreme Networks EXOS switch

Procedure

- Run the commands in this example to enable jumbo frames on an Extreme Networks EXOS switch and configure support for forwarding IP packets with 9000 bytes:

```
enable jumbo-frame ports PORTSTRING
configure ip-mtu 9000 vlan VLANNNAME
```

Example

```
# enable jumbo-frame ports 11
# configure ip-mtu 9000 vlan DATA
```

8.5.10. About LLDP discovery

The **ironic-python-agent** service listens for LLDP packets from connected switches. The collected information can include the switch name, port details, and available VLANs. Similar to Cisco Discovery Protocol (CDP), LLDP assists with the discovery of physical hardware during the director introspection process.

8.5.11. Configuring LLDP settings on an Extreme Networks EXOS switch

Procedure

- In this example, LLDP is enabled on an Extreme Networks EXOS switch. **11** represents the port string:

```
enable lldp ports 11
```

8.6. CONFIGURING A JUNIPER EX SERIES SWITCH

8.6.1. About trunk ports

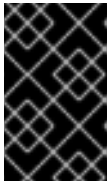
With OpenStack Networking you can connect instances to the VLANs that already exist on your physical network. The term *trunk* is used to describe a port that allows multiple VLANs to traverse through the same port. Using these ports, VLANs can span across multiple switches, including virtual switches. For

example, traffic tagged as VLAN110 in the physical network reaches the Compute node, where the 8021q module directs the tagged traffic to the appropriate VLAN on the vSwitch.

8.6.2. Configuring trunk ports for a Juniper EX Series switch

Procedure

- If using a Juniper EX series switch running Juniper JunOS, use the following configuration syntax to allow traffic for VLANs 110 and 111 to pass through to your instances. This configuration assumes that your physical node has an ethernet cable connected to interface ge-1/0/12 on the physical switch.



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

```
ge-1/0/12 {
  description Trunk to Compute Node;
  unit 0 {
    family ethernet-switching {
      port-mode trunk;
      vlan {
        members [110 111];
      }
      native-vlan-id 2;
    }
  }
}
```

8.6.3. About access ports

Not all NICs on your Compute node carry instance traffic, and so you do not need to configure all NICs to allow multiple VLANs to pass through. Access ports require only one VLAN, and might fulfill other operational requirements, such as transporting management traffic or Block Storage data. These ports are commonly known as access ports and usually require a simpler configuration than trunk ports.

8.6.4. Configuring access ports for a Juniper EX Series switch

This example on a Juniper EX series switch, shows **ge-1/0/13** as an access port for **eth1**.

+



IMPORTANT

These values are examples. You must change the values in this example to match those in your environment. Copying and pasting these values into your switch configuration without adjustment can result in an unexpected outage.

Procedure

This configuration assumes that your physical node has an ethernet cable connected to interface **ge-1/0/13** on the physical switch.

```
+
ge-1/0/13 {
  description Access port for Compute Node
  unit 0 {
    family ethernet-switching {
      port-mode access;
      vlan {
        members 200;
      }
      native-vlan-id 2;
    }
  }
}
```

8.6.5. About LACP port aggregation

You can use LACP to bundle multiple physical NICs together to form a single logical channel. Also known as 802.3ad (or bonding mode 4 in Linux), LACP creates a dynamic bond for load-balancing and fault tolerance. You must configure LACP at both physical ends: on the physical NICs, and on the physical switch ports.

8.6.6. Configuring LACP on the physical NIC

1. Edit the `/home/stack/network-environment.yaml` file:

```
- type: linux_bond
  name: bond1
  mtu: 9000
  bonding_options:{get_param: BondInterfaceOvsOptions};
  members:
    - type: interface
      name: nic3
      mtu: 9000
      primary: true
    - type: interface
      name: nic4
      mtu: 9000
```

2. Configure the Open vSwitch bridge to use LACP:

```
BondInterfaceOvsOptions:
  "mode=802.3ad"
```

For information on configuring network bonds, see the [Network Interface Bonding](#) chapter of the [Advanced Openstack Customization](#) guide.

8.6.7. Configuring LACP for a Juniper EX Series switch

In this example, the Compute node has two NICs using VLAN 100.

Procedure

1. Physically connect the Compute node's two NICs to the switch (for example, ports 12 and 13).
2. Create the port aggregate:

```
chassis {
  aggregated-devices {
    ethernet {
      device-count 1;
    }
  }
}
```

3. Configure switch ports 12 (ge-1/0/12) and 13 (ge-1/0/13) to join the port aggregate **ae1**:

```
interfaces {
  ge-1/0/12 {
    gigheter-options {
      802.3ad ae1;
    }
  }
  ge-1/0/13 {
    gigheter-options {
      802.3ad ae1;
    }
  }
}
```



NOTE

For Red Hat OpenStack Platform director deployments, in order to PXE boot from the bond, you must configure one of the bond members as lACP force-up to ensure that only one bond member comes up during introspection and first boot. The bond member that you configure with lACP force-up must be the same bond member that has the MAC address in *instackenv.json* (the MAC address known to ironic must be the same MAC address configured with force-up).

4. Enable LACP on port aggregate **ae1**:

```
interfaces {
  ae1 {
    aggregated-ether-options {
      lACP {
        active;
      }
    }
  }
}
```

5. Add aggregate **ae1** to VLAN 100:

```
interfaces {
  ae1 {
```

```

vlan-tagging;
native-vlan-id 2;
unit 100 {
    vlan-id 100;
}
}
}

```

- Review your new port channel. The resulting output lists the new port aggregate **ae1** with member ports **ge-1/0/12** and **ge-1/0/13**:

```

> show lacp statistics interfaces ae1

Aggregated interface: ae1
LACP Statistics: LACP Rx LACP Tx Unknown Rx Illegal Rx
ge-1/0/12 0 0 0 0
ge-1/0/13 0 0 0 0

```



NOTE

Remember to apply your changes by running the **commit** command.

8.6.8. About MTU settings

You must adjust your MTU size for certain types of network traffic. For example, jumbo frames (9000 bytes) are required for certain NFS or iSCSI traffic.



NOTE

You must change MTU settings from end-to-end on all hops that the traffic is expected to pass through, including any virtual switches.

Additional resources

- [Configuring maximum transmission unit \(MTU\) settings](#)

8.6.9. Configuring MTU settings for a Juniper EX Series switch

This example enables jumbo frames on your Juniper EX4200 switch.



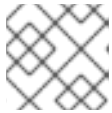
NOTE

The MTU value is calculated differently depending on whether you are using Juniper or Cisco devices. For example, **9216** on Juniper would equal to **9202** for Cisco. The extra bytes are used for L2 headers, where Cisco adds this automatically to the MTU value specified, but the usable MTU will be 14 bytes smaller than specified when using Juniper. So in order to support an MTU of **9000** on the VLANs, the MTU of **9014** would have to be configured on Juniper.

Procedure

- For Juniper EX series switches, MTU settings are set for individual interfaces. These commands configure jumbo frames on the **ge-1/0/14** and **ge-1/0/15** ports:

```
set interfaces ge-1/0/14 mtu 9216
set interfaces ge-1/0/15 mtu 9216
```

**NOTE**

Remember to save your changes by running the **commit** command.

2. If using a LACP aggregate, you will need to set the MTU size there, and not on the member NICs. For example, this setting configures the MTU size for the ae1 aggregate:

```
set interfaces ae1 mtu 9216
```

8.6.10. About LLDP discovery

The **ironic-python-agent** service listens for LLDP packets from connected switches. The collected information can include the switch name, port details, and available VLANs. Similar to Cisco Discovery Protocol (CDP), LLDP assists with the discovery of physical hardware during the director introspection process.

8.6.11. Configuring LLDP for a Juniper EX Series switch

You can enable LLDP globally for all interfaces, or just for individual ones.

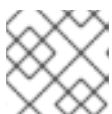
Procedure

- Use the following to enable LLDP globally on your Juniper EX 4200 switch:

```
lldp {
  interface all {
    enable;
  }
}
```

- Use the following to enable LLDP for the single interface **ge-1/0/14**:

```
lldp {
  interface ge-1/0/14 {
    enable;
  }
}
```

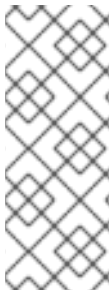
**NOTE**

Remember to apply your changes by running the **commit** command.

CHAPTER 9. CONFIGURING MAXIMUM TRANSMISSION UNIT (MTU) SETTINGS

9.1. MTU OVERVIEW

OpenStack Networking can calculate the largest possible maximum transmission unit (MTU) size that you can apply safely to instances. The MTU value specifies the maximum amount of data that a single network packet can transfer; this number is variable depending on the most appropriate size for the application. For example, NFS shares might require a different MTU size to that of a VoIP application.

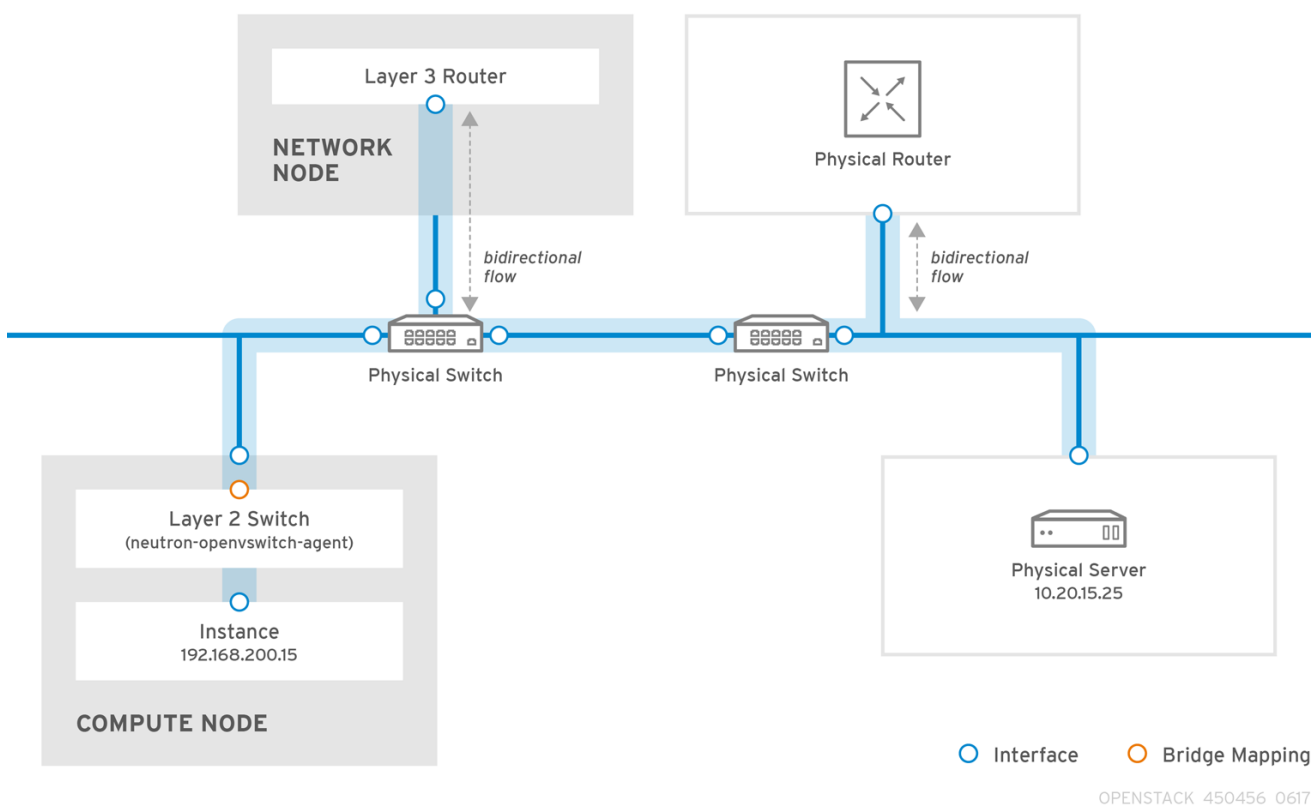


NOTE

You can use the **openstack network show <network_name>** command to view the largest possible MTU values that OpenStack Networking calculates. **net-mtu** is a neutron API extension that is not present in some implementations. The MTU value that you require can be advertised to DHCPv4 clients for automatic configuration, if supported by the instance, as well as to IPv6 clients through Router Advertisement (RA) packets. To send Router Advertisements, the network must be attached to a router.

You must configure MTU settings consistently from end-to-end. This means that the MTU setting must be the same at every point the packet passes through, including the VM, the virtual network infrastructure, the physical network, and the destination server.

For example, the circles in the following diagram indicate the various points where an MTU value must be adjusted for traffic between an instance and a physical server. You must change the MTU value for every interface that handles network traffic to accommodate packets of a particular MTU size. This is necessary if traffic travels from the instance *192.168.200.15* through to the physical server *10.20.15.25*:



Inconsistent MTU values can result in several network issues, the most common being random packet loss that results in connection drops and slow network performance. Such issues are problematic to

troubleshoot because you must identify and examine every possible network point to ensure it has the correct MTU value.

9.2. CONFIGURING MTU SETTINGS IN DIRECTOR

This example demonstrates how to set the MTU using the NIC config templates. You must set the MTU on the bridge, bond (if applicable), interface(s), and VLAN(s):

```
-
  type: ovs_bridge
  name: br-isolated
  use_dhcp: false
  mtu: 9000 # <--- Set MTU
  members:
  -
    type: ovs_bond
    name: bond1
    mtu: 9000 # <--- Set MTU
    ovs_options: {get_param: BondInterfaceOvsOptions}
    members:
    -
      type: interface
      name: ens15f0
      mtu: 9000 # <--- Set MTU
      primary: true
    -
      type: interface
      name: enp131s0f0
      mtu: 9000 # <--- Set MTU
  -
    type: vlan
    device: bond1
    vlan_id: {get_param: InternalApiNetworkVlanID}
    mtu: 9000 # <--- Set MTU
    addresses:
    -
      ip_netmask: {get_param: InternalApiIpSubnet}
  -
    type: vlan
    device: bond1
    mtu: 9000 # <--- Set MTU
    vlan_id: {get_param: TenantNetworkVlanID}
    addresses:
    -
      ip_netmask: {get_param: TenantIpSubnet}
```

9.3. REVIEWING THE RESULTING MTU CALCULATION

You can view the calculated MTU value, which is the largest possible MTU value that instances can use. Use this calculated MTU value to configure all interfaces involved in the path of network traffic.

```
# openstack network show <network>
```

CHAPTER 10. CONFIGURING QUALITY OF SERVICE (QOS) POLICIES

With Red Hat OpenStack Platform network quality-of-service (QoS) policies, an OpenStack administrator can offer varying service levels by applying rate limits to egress traffic for instances. As a result of implementing a QoS policy, any traffic that exceeds the specified rate is consequently dropped.

10.1. QOS POLICY SCOPE

You can apply QoS policies to individual ports, or to a particular tenant network, where ports with no specific policy attached inherit the policy.

10.2. CREATING AND APPLYING A QOS POLICY AND RULE

To create a quality of service (QoS) policy and rule and apply the policy to a port, complete the following steps:

1. If the **qos** extension is not already enabled for OpenStack Networking in `/etc/neutron/plugins/ml2/<agent_name>_agent.ini`, then follow these steps:
 - a. Create a custom Heat environment file, and add the following lines:

```
parameter_defaults:
  NeutronSriovAgentExtensions: 'qos'
```



IMPORTANT

YAML files are extremely sensitive about where in the file a parameter is placed. Make sure that **parameter_defaults:** starts in the first column (no leading whitespace characters), and your parameter value pair starts in column three (the parameter has two whitespace characters in front of it).

- b. Run the Red Hat OpenStack Platform director command, **openstack overcloud deploy** and include all your current environment files and this custom new one. For more information, see "[Modifying the overcloud environment](#)" in the *Director Installation and Usage* guide.
2. Review the list of tenants and determine the id of where you need to create QoS policy:

```
# openstack project list
+-----+-----+
| ID                | Name  |
+-----+-----+
| 4b0b98f8c6c040f38ba4f7146e8680f5 | auditors |
| 519e6344f82e4c079c8e2eabb690023b | services |
| 80bf5732752a41128e612fe615c886c6 | demo   |
| 98a2f53c20ce4d50a40dac4a38016c69 | admin  |
+-----+-----+
```

3. Create a QoS policy named **bw-limiter** in the **admin** tenant:

```
# openstack network qos policy create --project 98a2f53c20ce4d50a40dac4a38016c69 bw-
limiter
```

4. Configure the policing rules for the **bw-limiter** policy:

```
# openstack network qos rule create --type bandwidth-limit --max-kbps 3000 --max-burst-
kbits 30000 bw-limiter
```

5. Configure a neutron port to apply the **bw-limiter** policy:

```
# openstack port set --qos-policy bw-limiter port_name_or_id
```

6. Review the QoS rule. For example:

```
# openstack network qos policy show 9be535c3-daa2-4d7b-88ea-e8de16

+-----+-----+
| Field      | Value                               |
+-----+-----+
| id         | 9be535c3-daa2-4d7b-88ea-e8de16 |
| rule_type  | bandwidth_limit                   |
| description |                                     |
| max_kbps   | 3000                               |
| max_burst_kbps | 300                               |
+-----+-----+
```

These values allow you to configure the policing algorithm accordingly:

- **max_kbps** - the maximum rate (in Kbps) that the instance is allowed to send.
- **max_burst_kbps** - the maximum amount of data (in kbits) that the port could send in a instant if the token buffer was full. The token buffer is replenished at a "max_kbps" rate.

10.3. DSCP MARKING FOR EGRESS TRAFFIC

You can use differentiated services code point (DSCP) to implement quality-of-service (QoS) policies on your network by embedding relevant values in the IP headers. The Networking service (neutron) QoS policies can use DSCP marking to manage egress traffic on neutron ports and networks.

Procedure

1. If you are:
 - using ML2/OVN, skip to step 2.
 - using ML2/OVS without a tunneling protocol, skip to step 2.
 - using ML2/OVS with a tunneling protocol (VXLAN and GRE), then perform the following steps:
 - a. Log in to the undercloud host as the **stack** user.

```
$ source ~/stackrc
```


- b. Create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-neutron-environment.yaml
```

- c. In the YAML environment file under **parameter_defaults**, add the following lines:

```
parameter_defaults:
  ControllerExtraConfig:
    neutron::config::server_config:
      agent/dscp_inherit:
        value: true
```

When **dscp_inherit** is **true**, the Networking service copies the DSCP value of the inner header to the outer header.

- d. Run the deployment command and include the core heat templates, environment files, and this new custom environment file.



IMPORTANT

The order of the environment files is important as the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/ \
  neutron-ovs.yaml \
-e /home/stack/templates/my-neutron-environment.yaml
```

2. Source your credentials file.

Example

```
$ source ~/overcloudrc
```

3. Create a new QoS policy:

Example

```
openstack network qos policy create --project 98a2f53c20ce4d50a40dac4a38016c69 qos-
web-servers
```

4. Create a DSCP rule and apply it to a policy.

Example

In this example, a DSCP rule is created using DSCP mark **18** and is applied to the **qos-web-servers** policy:

```
openstack network qos rule create --type dscp-marking --dscp-mark 18 qos-web-servers
```

Sample output

```
Created a new dscp_marking_rule:
+-----+-----+
| Field | Value |
+-----+-----+
| dscp_mark | 18 |
| id | d7f976ec-7fab-4e60-af70-f59bf88198e6 |
+-----+-----+
```

- You can change the DSCP value assigned to a rule.

Example

```
openstack network qos rule set --dscp-mark 22 qos-web-servers d7f976ec-7fab-4e60-af70-f59bf88198e6
```

- You can delete a DSCP rule.

Example

```
openstack network qos rule delete qos-web-servers d7f976ec-7fab-4e60-af70-f59bf88198e6
```

Verification

- Confirm that the DSCP rule (d7f976ec-7fab-4e60-af70-f59bf88198e6) is applied to the QoS policy (qos-web-servers).

Example

```
openstack network qos rule list qos-web-servers
```

Sample output

```
+-----+-----+
| dscp_mark | id |
+-----+-----+
| 18 | d7f976ec-7fab-4e60-af70-f59bf88198e6 |
+-----+-----+
```

Additional resources

- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide
- [network qos rule create](#) in the *Command Line Interface Reference*
- [network qos rule set](#) in the *Command Line Interface Reference*

- [network qos rule delete](#) in the *Command Line Interface Reference*
- [network qos rule list](#) in the *Command Line Interface Reference*

10.4. RBAC FOR QOS POLICIES

You can add a role-based access control (RBAC) for quality-of-service (QoS) policies. As a result, you can now make QoS policies available to certain projects.

For example, you can now create a QoS policy that allows for lower-priority network traffic, and have it only apply to certain projects. Run the following command to assign the **bw-limiter** policy to the project, **demo**:

```
# openstack network rbac create --type qos_policy --target-project  
80bf5732752a41128e612fe615c886c6 --action access_as_shared rbac_name
```

CHAPTER 11. CONFIGURING BRIDGE MAPPINGS

This chapter contains information about configuring bridge mappings in Red Hat OpenStack Platform.

11.1. OVERVIEW OF BRIDGE MAPPINGS

A bridge mapping associates a physical network name (an interface label) to a bridge created with OVS or OVN. In this example, the physical name (datacentre) is mapped to the external bridge (br-ex):

```
bridge_mappings = datacentre:br-ex
```

Bridge mappings allow provider network traffic to reach the physical network. Traffic leaves the provider network from the **qg-xxx** interface of the router and arrives at **br-int**. For OVS, a patch port between **br-int** and **br-ex** then allows the traffic to pass through the bridge of the provider network and out to the physical network. OVN creates a patch port on a hypervisor only when there is a VM bound to the hypervisor that requires the port.

You configure bridge mappings on the network node on which the router is scheduled. Router traffic can egress using the correct physical network, as represented by the provider network.

11.2. TRAFFIC FLOW

Each external network is represented by an internal VLAN ID, which is tagged to the router **qg-xxx** port. When a packet reaches **phy-br-ex**, the **br-ex** port strips the VLAN tag and moves the packet to the physical interface and then to the external network.

The return packet from the external network arrives on **br-ex** and moves to **br-int** using **phy-br-ex <-> int-br-ex**. When the packet is going through **br-ex** to **br-int**, the packet's external vlan ID is replaced by an internal vlan tag in **br-int**, and this allows **qg-xxx** to accept the packet.

In the case of egress packets, the packet's internal vlan tag is replaced with an external vlan tag in br-ex (or in the external bridge that is defined in the **network_vlan_ranges** parameter).

11.3. CONFIGURING BRIDGE MAPPINGS

Red Hat OpenStack Platform (RHOSP) director uses predefined NIC templates to install and configure your initial networking configuration.

You can customize aspects of your initial networking configuration, such as bridge mappings, by using the **NeutronBridgeMappings** parameter in a customized environment file. You call the environment file in the **openstack overcloud deploy** command.

Prerequisites

- You must configure bridge mappings on the network node on which the router is scheduled.
- For both ML2/OVS and ML2/OVN DVR configurations, you must configure bridge mappings for the compute nodes, too.

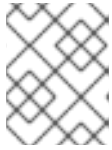
Procedure

1. Create a custom environment file and add the **NeutronBridgeMappings** heat parameter with values that are appropriate for your site.

■

```
parameter_defaults:
  NeutronBridgeMappings: "datacentre:br-ex,tenant:br-tenant"
```

The **NeutronBridgeMappings** heat parameter associates a physical name (**datacentre**) to a bridge (**br-ex**).



NOTE

When the **NeutronBridgeMappings** parameter is not used, the default maps the external bridge on hosts (**br-ex**) to a physical name (**datacentre**).

2. To apply this configuration, deploy the overcloud, adding your custom environment file to the stack along with your other environment files.

```
(undercloud) $ openstack overcloud deploy --templates \
-e [your environment files]
-e /home/stack/templates/<custom-environment-file>.yaml
```

3. You are ready for the next steps, which are the following:
 - a. Using the network VLAN ranges, create the provider networks that represent the corresponding external networks. (You use the physical name when creating neutron provider networks or floating IP networks.)
 - b. Connect the external networks to your project networks with router interfaces.

Additional resources

- [Network environment parameters](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

11.4. MAINTAINING BRIDGE MAPPINGS FOR OVS

After removing any OVS bridge mappings, you must perform a subsequent cleanup to ensure that the bridge configuration is cleared of any associated patch port entries. You can perform this operation in the following ways:

- Manual port cleanup - requires careful removal of the superfluous patch ports. No outages of network connectivity are required.
- Automated port cleanup - performs an automated cleanup, but requires an outage, and requires that the necessary bridge mappings be re-added. Choose this option during scheduled maintenance windows when network connectivity outages can be tolerated.



NOTE

When OVN bridge mappings are removed, the OVN controller automatically cleans up any associated patch ports.

11.4.1. Cleaning up OVS patch ports manually

After removing any OVS bridge mappings, you must also remove the associated patch ports.

Prerequisites

- The patch ports that you are cleaning up must be Open vSwitch (OVS) ports.
- A system outage is **not** required to perform a manual patch port cleanup.
- You can identify the patch ports to cleanup by their naming convention:
 - In **br-\$external_bridge** patch ports are named **phy-<external bridge name>** (for example, **phy-br-ex2**).
 - In **br-int** patch ports are named **int-<external bridge name>** (for example, **int-br-ex2**).

Procedure

1. Use **ovs-vsctl** to remove the OVS patch ports associated with the removed bridge mapping entry:

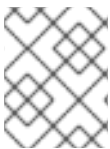
```
# ovs-vsctl del-port br-ex2 datacentre
# ovs-vsctl del-port br-tenant tenant
```

2. Restart **neutron-openvswitch-agent**:

```
# service neutron-openvswitch-agent restart
```

11.4.2. Cleaning up OVS patch ports automatically

After removing any OVS bridge mappings, you must also remove the associated patch ports.



NOTE

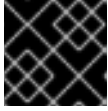
When OVN bridge mappings are removed, the OVN controller automatically cleans up any associated patch ports.

Prerequisites

- The patch ports that you are cleaning up must be Open vSwitch (OVS) ports.
- Cleaning up patch ports automatically with the **neutron-ovs-cleanup** command causes a network connectivity outage, and should be performed only during a scheduled maintenance window.
- Use the flag **--ovs_all_ports** to remove all patch ports from **br-int**, cleaning up tunnel ends from **br-tun**, and patch ports from bridge to bridge.
- The **neutron-ovs-cleanup** command unplugs all patch ports (instances, qdhcp/qrouter, among others) from all OVS bridges.

Procedure

1. Run the **neutron-ovs-cleanup** command with the **--ovs_all_ports** flag.



IMPORTANT

Performing this step will result in a total networking outage.

```
# /usr/bin/neutron-ovs-cleanup
--config-file /etc/neutron/plugins/ml2/openvswitch_agent.ini
--log-file /var/log/neutron/ovs-cleanup.log --ovs_all_ports
```

2. Restore connectivity by redeploying the overcloud.

When you rerun the **openstack overcloud deploy** command, your bridge mapping values are reapplied.



NOTE

After a restart, the OVS agent does not interfere with any connections that are not present in `bridge_mappings`. So, if you have **br-int** connected to **br-ex2**, and **br-ex2** has some flows on it, removing **br-int** from the `bridge_mappings` configuration does not disconnect the two bridges when you restart the OVS agent or the node.

Additional resources

- [Network environment parameters](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

CHAPTER 12. VLAN-AWARE INSTANCES

12.1. VLAN TRUNKS AND VLAN TRANSPARENT NETWORKS

Instances can send and receive VLAN-tagged traffic over a single virtual NIC. This is particularly useful for NFV applications (VNFs) that expect VLAN-tagged traffic, allowing a single virtual NIC to serve multiple customers or services.

Instances can send and receive VLAN-tagged traffic over a single vNIC. This is particularly useful for NFV applications (VNFs) that expect VLAN-tagged traffic, allowing a single vNIC to serve multiple customers or services.

For example, the project data network can use VLANs, or tunneling (VXLAN/GRE) segmentation, while the instances see the traffic tagged with VLAN IDs. As a result, network packets are tagged just before they are injected to the instance and do not need to be tagged throughout the entire network.

To implement VLAN-tagged traffic, create a parent port and attach the new port to an existing neutron network. When you attach the new port, OpenStack Networking adds a trunk connection to the parent port you created. Next, create subports. These subports connect VLANs to instances, which allow connectivity to the trunk. Within the instance operating system, you must also create a sub-interface that tags traffic for the VLAN associated with the subport.

12.2. REVIEWING THE TRUNK PLUG-IN

During a Red Hat openStack deployment, the trunk plug-in is enabled by default. You can review the configuration on the controller nodes:

- On the controller node, confirm that the trunk plug-in is enabled in the `/var/lib/config-data/neutron/etc/neutron/neutron.conf` file:

```
service_plugins=router,qos,trunk
```

12.3. CREATING A TRUNK CONNECTION

To implement trunks for VLAN-tagged traffic, create a parent port and attach the new port to an existing neutron network. When you attach the new port, OpenStack Networking adds a trunk connection to the parent port you created. Next, create subports. These subports connect VLANs to instances, which allow connectivity to the trunk. Within the instance operating system, you must also create a sub-interface that tags traffic for the VLAN associated with the subport.

1. Identify the network that requires the trunk port connection. This would be the network that will contain the instance that requires access to the trunked VLANs. In this example, this is the *public* network:

```
openstack network list
+-----+-----+-----+
| ID                | Name  | Subnets                |
+-----+-----+-----+
| 82845092-4701-4004-add7-838837837621 | private | 434c7982-cd96-4c41-a8c9-
b93adbdcdb197 |
| 8d8bc6d6-5b28-4e00-b99e-157516ff0050 | public  | 3fd811b4-c104-44b5-8ff8-
7a86af5e332c |
+-----+-----+-----+
```


2. Create the parent trunk port, and attach it to the network that the instance connects to. In this example, create a neutron port named `parent-trunk-port` on the `public` network. This trunk is the *parent* port, as you can use it to create *subports*.

```

openstack port create --network public parent-trunk-port
+-----+-----+
| Field      | Value                                     |
+-----+-----+
| admin_state_up | UP                                       |
| allowed_address_pairs |                                         |
| binding_host_id |                                         |
| binding_profile |                                         |
| binding_vif_details |                                         |
| binding_vif_type | unbound                                 |
| binding_vnic_type | normal                                 |
| created_at    | 2016-10-20T02:02:33Z                   |
| description   |                                         |
| device_id    |                                         |
| device_owner  |                                         |
| extra_dhcp_opts |                                         |
| fixed_ips     | ip_address='172.24.4.230', subnet_id='dc608964-9af3-4fed-9f06-6d3844fb9b9b' |
| headers      |                                         |
| id           | 20b6fdf8-0d43-475a-a0f1-ec8f757a4a39   |
| mac_address  | fa:16:3e:33:c4:75                     |
| name         | parent-trunk-port                       |
| network_id   | 871a6bd8-4193-45d7-a300-dcb2420e7cc3   |
| project_id   | 745d33000ac74d30a77539f8920555e7     |
| project_id   | 745d33000ac74d30a77539f8920555e7     |
| revision_number | 4                                       |
| security_groups | 59e2af18-93c6-4201-861b-19a8a8b79b23  |
| status       | DOWN                                    |
| updated_at   | 2016-10-20T02:02:33Z                   |
+-----+-----+

```

3. Create a trunk using the port that you created in step 2. In this example the trunk is named **parent-trunk**.

```

openstack network trunk create --parent-port parent-trunk-port parent-trunk
+-----+-----+
| Field      | Value                                     |
+-----+-----+
| admin_state_up | UP                                       |
| created_at    | 2016-10-20T02:05:17Z                   |
| description   |                                         |
| id           | 0e4263e2-5761-4cf6-ab6d-b22884a0fa88 |
| name         | parent-trunk                           |
| port_id      | 20b6fdf8-0d43-475a-a0f1-ec8f757a4a39 |
| revision_number | 1                                       |
| status       | DOWN                                    |
| sub_ports    |                                         |
| tenant_id    | 745d33000ac74d30a77539f8920555e7     |
| updated_at   | 2016-10-20T02:05:17Z                   |
+-----+-----+

```

- View the trunk connection:

```

openstack network trunk list
+-----+-----+-----+-----+
| ID                | Name      | Parent Port          | Description |
+-----+-----+-----+-----+
| 0e4263e2-5761-4cf6-ab6d-b22884a0fa88 | parent-trunk | 20b6fdf8-0d43-475a-a0f1-ec8f757a4a39 |
+-----+-----+-----+-----+

```

- View the details of the trunk connection:

```

openstack network trunk show parent-trunk
+-----+-----+
| Field      | Value                |
+-----+-----+
| admin_state_up | UP                  |
| created_at   | 2016-10-20T02:05:17Z |
| description   |                      |
| id           | 0e4263e2-5761-4cf6-ab6d-b22884a0fa88 |
| name         | parent-trunk        |
| port_id      | 20b6fdf8-0d43-475a-a0f1-ec8f757a4a39 |
| revision_number | 1                   |
| status       | DOWN                |
| sub_ports    |                      |
| tenant_id    | 745d33000ac74d30a77539f8920555e7 |
| updated_at   | 2016-10-20T02:05:17Z |
+-----+-----+

```

12.4. ADDING SUBPORTS TO THE TRUNK

- Create a neutron port.

This port is a subport connection to the trunk. You must also specify the MAC address that you assigned to the parent port:

```

openstack port create --network private --mac-address fa:16:3e:33:c4:75 subport-trunk-port
+-----+-----+-----+-----+
| Field      | Value                |
+-----+-----+-----+-----+
| admin_state_up | UP                  |
| allowed_address_pairs |                    |
| binding_host_id |                    |
| binding_profile |                    |
| binding_vif_details |                    |
| binding_vif_type | unbound             |
| binding_vnic_type | normal              |
| created_at   | 2016-10-20T02:08:14Z |
| description   |                    |
| device_id    |                    |
| device_owner  |                    |
| extra_dhcp_opts |                    |
| fixed_ips    | ip_address='10.0.0.11', subnet_id='1a299780-56df-4c0b-a4c0-c5a612cef2e8' |
| headers      |                    |
+-----+-----+-----+-----+

```

```

| id          | 479d742e-dd00-4c24-8dd6-b7297fab3ee9 |
| mac_address | fa:16:3e:33:c4:75 |
| name        | subport-trunk-port |
| network_id  | 3fe6b758-8613-4b17-901e-9ba30a7c4b51 |
| project_id  | 745d33000ac74d30a77539f8920555e7 |
| project_id  | 745d33000ac74d30a77539f8920555e7 |
| revision_number | 4 |
| security_groups | 59e2af18-93c6-4201-861b-19a8a8b79b23 |
| status      | DOWN |
| updated_at  | 2016-10-20T02:08:15Z |
+-----+-----+-----+-----+

```



NOTE

If you receive the error **HttpException: Conflict**, confirm that you are creating the subport on a different network to the one that has the parent trunk port. This example uses the public network for the parent trunk port, and private for the subport.

- Associate the port with the trunk (**parent-trunk**), and specify the VLAN ID (**55**):

```

openstack network trunk set --subport port=subport-trunk-port,segmentation-
type=vlan,segmentation-id=55 parent-trunk

```

12.5. CONFIGURING AN INSTANCE TO USE A TRUNK

You must configure the VM instance operating system to use the MAC address that the Red Hat OpenStack Platform (RHOSP) Networking service (neutron) assigned to the subport. You can also configure the subport to use a specific MAC address during the subport creation step.

Prerequisites

- If you are performing live migrations of your Compute nodes, ensure that the RHOSP Networking service RPC response timeout is appropriately set for your RHOSP deployment. The RPC response timeout value can vary between sites and is dependent on the system speed. The general recommendation is to set the value to at least 120 seconds per/100 trunk ports. The best practice is to measure the trunk port bind process time for your RHOSP deployment, and then set the RHOSP Networking service RPC response timeout appropriately. Try to keep the RPC response timeout value low, but also provide enough time for the RHOSP Networking service to receive an RPC response. For more information, see [Section 12.6, "Configuring Networking service RPC timeout"](#).

Procedure

- Review the configuration of your network trunk, using the **network trunk** command.

Example

```

$ openstack network trunk list

```

Sample output

```

+-----+-----+-----+-----+

```

ID	Name	Parent Port	Description
0e4263e2-5761-4cf6-ab6d-b22884a0fa88	parent-trunk	20b6fdf8-0d43-475a-a0f1-ec8f757a4a39	

Example

```
$ openstack network trunk show parent-trunk
```

Sample output

Field	Value
admin_state_up	UP
created_at	2021-10-20T02:05:17Z
description	
id	0e4263e2-5761-4cf6-ab6d-b22884a0fa88
name	parent-trunk
port_id	20b6fdf8-0d43-475a-a0f1-ec8f757a4a39
revision_number	2
status	DOWN
sub_ports	port_id='479d742e-dd00-4c24-8dd6-b7297fab3ee9', segmentation_id='55', segmentation_type='vlan'
tenant_id	745d33000ac74d30a77539f8920555e7
updated_at	2021-08-20T02:10:06Z

2. Create an instance that uses the parent **port-id** as its vNIC.

Example

```
openstack server create --image cirros --flavor m1.tiny --security-group default --key-name sshaccess --nic port-id=20b6fdf8-0d43-475a-a0f1-ec8f757a4a39 testInstance
```

Sample output

Property	Value
OS-DCF:diskConfig	MANUAL
OS-EXT-AZ:availability_zone	
OS-EXT-SRV-ATTR:host	-
OS-EXT-SRV-ATTR:hostname	testinstance
OS-EXT-SRV-ATTR:hypervisor_hostname	-
OS-EXT-SRV-ATTR:instance_name	
OS-EXT-SRV-ATTR:kernel_id	
OS-EXT-SRV-ATTR:launch_index	0
OS-EXT-SRV-ATTR:ramdisk_id	
OS-EXT-SRV-ATTR:reservation_id	r-juqco0e1
OS-EXT-SRV-ATTR:root_device_name	-
OS-EXT-SRV-ATTR:user_data	-
OS-EXT-STS:power_state	0

```

| OS-EXT-STS:task_state      | scheduling      |
| OS-EXT-STS:vm_state       | building       |
| OS-SRV-USG:launched_at    | -              |
| OS-SRV-USG:terminated_at  | -              |
| accessIPv4                |                |
| accessIPv6                |                |
| adminPass                  | uMyL8PnZRBwQ  |
| config_drive               |                |
| created                    | 2021-08-20T03:02:51Z |
| description                 | -              |
| flavor                     | m1.tiny (1)    |
| hostId                     |                |
| host_status                |                |
| id                          | 88b7aede-1305-4d91-a180-67e7eac |
|                             | 8b70d          |
| image                       | cirros (568372f7-15df-4e61-a05f |
|                             | -10954f79a3c4) |
| key_name                    | sshaccess      |
| locked                      | False          |
| metadata                   | {}             |
| name                        | testInstance   |
| os-extended-volumes:volumes_attached | []            |
| progress                    | 0              |
| security_groups             | default        |
| status                       | BUILD          |
| tags                         | []             |
| tenant_id                   | 745d33000ac74d30a77539f8920555e |
|                             | 7              |
| updated                     | 2021-08-20T03:02:51Z |
| user_id                     | 8c4aea738d774967b4ef388eb41fef5 |
|                             | e              |
+-----+-----+

```

Additional resources

- [Configuring Networking service RPC timeout](#)

12.6. CONFIGURING NETWORKING SERVICE RPC TIMEOUT

There can be situations when you must modify the Red Hat OpenStack Platform (RHOSP) Networking service (neutron) RPC response timeout. For example, live migrations for Compute nodes that use trunk ports can fail if the timeout value is too low.

The RPC response timeout value can vary between sites and is dependent on the system speed. The general recommendation is to set the value to at least 120 seconds per/100 trunk ports.

If your site uses trunk ports, the best practice is to measure the trunk port bind process time for your RHOSP deployment, and then set the RHOSP Networking service RPC response timeout appropriately. Try to keep the RPC response timeout value low, but also provide enough time for the RHOSP Networking service to receive an RPC response.

By using a manual hieradata override, **rpc_response_timeout**, you can set the RPC response timeout value for the RHOSP Networking service.

Procedure

1. On the undercloud host, logged in as the stack user, create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-modules-environment.yaml
```

TIP

The RHOSP Orchestration service (heat) uses a set of plans called *templates* to install and configure your environment. You can customize aspects of the overcloud with a *custom environment file*, which is a special type of template that provides customization for your heat templates.

2. In the YAML environment file under **ExtraConfig**, set the appropriate value (in seconds) for **rpc_response_timeout**. (The default value is 60 seconds.)

Example

```
parameter_defaults:
  ExtraConfig:
    neutron::rpc_response_timeout: 120
```



NOTE

The RHOSP Orchestration service (heat) updates all RHOSP nodes with the value you set in the custom environment file, however this value only impacts the RHOSP Networking components.

3. Run the **openstack overcloud deploy** command and include the core heat templates, environment files, and this new custom environment file.



IMPORTANT

The order of the environment files is important as the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-modules-
environment.yaml
```

Additional resources

- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

12.7. UNDERSTANDING TRUNK STATES

- **ACTIVE:** The trunk is working as expected and there are no current requests.
- **DOWN:** The virtual and physical resources for the trunk are not in sync. This can be a temporary state during negotiation.
- **BUILD:** There has been a request and the resources are being provisioned. After successful completion the trunk returns to **ACTIVE**.
- **DEGRADED:** The provisioning request did not complete, so the trunk has only been partially provisioned. It is recommended to remove the subports and try again.
- **ERROR:** The provisioning request was unsuccessful. Remove the resource that caused the error to return the trunk to a healthier state. Do not add more subports while in the **ERROR** state, as this can cause more issues.

CHAPTER 13. CONFIGURING RBAC POLICIES

13.1. OVERVIEW OF RBAC POLICIES

Role-based access control (RBAC) policies in OpenStack Networking allow granular control over shared *neutron* networks. OpenStack Networking uses a RBAC table to control sharing of *neutron* networks among projects, allowing an administrator to control which projects are granted permission to attach instances to a network.

As a result, cloud administrators can remove the ability for some projects to create networks and can instead allow them to attach to pre-existing networks that correspond to their project.

13.2. CREATING RBAC POLICIES

This example procedure demonstrates how to use a role-based access control (RBAC) policy to grant a project access to a shared network.

1. View the list of available networks:

```
# openstack network list
+-----+-----+-----+
| id                | name      | subnets                |
+-----+-----+-----+
| fa9bb72f-b81a-4572-9c7f-7237e5fcabd3 | web-servers | 20512ffe-ad56-4bb4-b064-2cb18fecc923 192.168.200.0/24 |
| bcc16b34-e33e-445b-9fde-dd491817a48a | private    | 7fe4a05a-4b81-4a59-8c47-82c965b0e050 10.0.0.0/24 |
| 9b2f4feb-fee8-43da-bb99-032e4aaf3f85 | public     | 2318dc3b-cff0-43fc-9489-7d4cf48aaab9 172.24.4.224/28 |
+-----+-----+-----+
```

2. View the list of projects:

```
# openstack project list
+-----+-----+
| ID                | Name      |
+-----+-----+
| 4b0b98f8c6c040f38ba4f7146e8680f5 | auditors  |
| 519e6344f82e4c079c8e2eabb690023b | services  |
| 80bf5732752a41128e612fe615c886c6 | demo      |
| 98a2f53c20ce4d50a40dac4a38016c69 | admin     |
+-----+-----+
```

3. Create a RBAC entry for the **web-servers** network that grants access to the *auditors* project (**4b0b98f8c6c040f38ba4f7146e8680f5**):

```
# openstack network rbac create --type network --target-project
4b0b98f8c6c040f38ba4f7146e8680f5 --action access_as_shared web-servers
Created a new rbac_policy:
+-----+-----+-----+
| Field      | Value                |
+-----+-----+-----+
| action     | access_as_shared     |
| id         | 314004d0-2261-4d5e-bda7-0181fcf40709 |
+-----+-----+-----+
```



```

| object_id   | fa9bb72f-b81a-4572-9c7f-7237e5fcabd3 |
| object_type | network                               |
| target_project | 4b0b98f8c6c040f38ba4f7146e8680f5 |
| project_id   | 98a2f53c20ce4d50a40dac4a38016c69 |
+-----+-----+

```

As a result, users in the *auditors* project can connect instances to the **web-servers** network.

13.3. REVIEWING RBAC POLICIES

1. Run the **openstack network rbac list** command to retrieve the ID of your existing role-based access control (RBAC) policies:

```

# openstack network rbac list
+-----+-----+-----+-----+
| id                | object_type | object_id                |
+-----+-----+-----+-----+
| 314004d0-2261-4d5e-bda7-0181fcf40709 | network    | fa9bb72f-b81a-4572-9c7f-7237e5fcabd3 |
| bbab1cf9-edc5-47f9-ae3-a413bd582c0a | network    | 9b2f4feb-fee8-43da-bb99-032e4aaf3f85 |
+-----+-----+-----+-----+

```

2. Run the **openstack network rbac-show** command to view the details of a specific RBAC entry:

```

# openstack network rbac show 314004d0-2261-4d5e-bda7-0181fcf40709
+-----+-----+
| Field      | Value                |
+-----+-----+
| action     | access_as_shared    |
| id         | 314004d0-2261-4d5e-bda7-0181fcf40709 |
| object_id  | fa9bb72f-b81a-4572-9c7f-7237e5fcabd3 |
| object_type | network              |
| target_project | 4b0b98f8c6c040f38ba4f7146e8680f5 |
| project_id  | 98a2f53c20ce4d50a40dac4a38016c69 |
+-----+-----+

```

13.4. DELETING RBAC POLICIES

1. Run the **openstack network rbac list** command to retrieve the ID of your existing role-based access control (RBAC) policies:

```

# openstack network rbac list
+-----+-----+-----+-----+
| id                | object_type | object_id                |
+-----+-----+-----+-----+
| 314004d0-2261-4d5e-bda7-0181fcf40709 | network    | fa9bb72f-b81a-4572-9c7f-7237e5fcabd3 |
| bbab1cf9-edc5-47f9-ae3-a413bd582c0a | network    | 9b2f4feb-fee8-43da-bb99-032e4aaf3f85 |
+-----+-----+-----+-----+

```

- Run the **openstack network rbac delete** command to delete the RBAC, using the ID of the RBAC that you want to delete:

```
# openstack network rbac delete 314004d0-2261-4d5e-bda7-0181fcf40709
Deleted rbac_policy: 314004d0-2261-4d5e-bda7-0181fcf40709
```

13.5. GRANTING RBAC POLICY ACCESS FOR EXTERNAL NETWORKS

You can grant role-based access control (RBAC) policy access to external networks (networks with gateway interfaces attached) using the **--action access_as_external** parameter.

Complete the steps in the following example procedure to create a RBAC for the web-servers network and grant access to the engineering project (c717f263785d4679b16a122516247deb):

- Create a new RBAC policy using the **--action access_as_external** option:

```
# openstack network rbac create --type network --target-project
c717f263785d4679b16a122516247deb --action access_as_external web-servers
Created a new rbac_policy:
+-----+-----+
| Field      | Value                               |
+-----+-----+
| action     | access_as_external                 |
| id         | ddef112a-c092-4ac1-8914-c714a3d3ba08 |
| object_id  | 6e437ff0-d20f-4483-b627-c3749399bdca |
| object_type | network                             |
| target_project | c717f263785d4679b16a122516247deb |
| project_id  | c717f263785d4679b16a122516247deb |
+-----+-----+
```

As a result, users in the engineering project are able to view the network or connect instances to it:

```
$ openstack network list
+-----+-----+-----+
| id                | name      | subnets                               |
+-----+-----+-----+
| 6e437ff0-d20f-4483-b627-c3749399bdca | web-servers | fa273245-1eff-4830-b40c-57eaeac9b904 192.168.10.0/24 |
+-----+-----+-----+
```

CHAPTER 14. CONFIGURING DISTRIBUTED VIRTUAL ROUTING (DVR)

14.1. UNDERSTANDING DISTRIBUTED VIRTUAL ROUTING (DVR)

When you deploy Red Hat OpenStack Platform you can choose between a centralized routing model or DVR.

Each model has advantages and disadvantages. Use this document to carefully plan whether centralized routing or DVR better suits your needs.

DVR is enabled by default in new ML2/OVN deployments and disabled by default in new ML2/OVS deployments. The Heat template for the OpenStack Networking (neutron) API (**deployment/neutron/neutron-api-container-puppet.yaml**) contains a parameter to enable and disable Distributed Virtual Routing (DVR). To disable DVR, use the following in an environment file:

```
parameter_defaults:
  NeutronEnableDVR: false
```

14.1.1. Overview of Layer 3 routing

The Red Hat OpenStack Platform Networking service (neutron) provides routing services for project networks. Without a router, VM instances in a project network can communicate with other instances over a shared L2 broadcast domain. Creating a router and assigning it to a project network allows the instances in that network to communicate with other project networks or upstream (if an external gateway is defined for the router).

14.1.2. Routing flows

Routing services in OpenStack can be categorized into three main flows:

- **East-West routing** - routing of traffic between different networks in the same project. This traffic does not leave the OpenStack deployment. This definition applies to both IPv4 and IPv6 subnets.
- **North-South routing with floating IPs** - Floating IP addressing is a one-to-one network address translation (NAT) that can be modified and that floats between VM instances. While floating IPs are modeled as a one-to-one association between the floating IP and a Networking service (neutron) port, they are implemented by association with a Networking service router that performs the NAT translation. The floating IPs themselves are taken from the uplink network that provides the router with external connectivity. As a result, instances can communicate with external resources (such as endpoints on the internet) or the other way around. Floating IPs are an IPv4 concept and do not apply to IPv6. It is assumed that the IPv6 addressing used by projects uses Global Unicast Addresses (GUAs) with no overlap across the projects, and therefore can be routed without NAT.
- **North-South routing without floating IPs** (also known as *SNAT*) - The Networking service offers a default port address translation (PAT) service for instances that do not have allocated floating IPs. With this service, instances can communicate with external endpoints through the router, but not the other way around. For example, an instance can browse a website on the internet, but a web browser outside cannot browse a website hosted within the instance. SNAT

is applied for IPv4 traffic only. In addition, Networking service networks that are assigned GUAs prefixes do not require NAT on the Networking service router external gateway port to access the outside world.

14.1.3. Centralized routing

Originally, the Networking service (neutron) was designed with a centralized routing model where a project's virtual routers, managed by the neutron L3 agent, are all deployed in a dedicated node or cluster of nodes (referred to as the Network node, or Controller node). This means that each time a routing function is required (east/west, floating IPs or SNAT), traffic would traverse through a dedicated node in the topology. This introduced multiple challenges and resulted in sub-optimal traffic flows. For example:

- Traffic between instances flows through a Controller node - when two instances need to communicate with each other using L3, traffic has to hit the Controller node. Even if the instances are scheduled on the same Compute node, traffic still has to leave the Compute node, flow through the Controller, and route back to the Compute node. This negatively impacts performance.
- Instances with floating IPs receive and send packets through the Controller node - the external network gateway interface is available only at the Controller node, so whether the traffic is originating from an instance, or destined to an instance from the external network, it has to flow through the Controller node. Consequently, in large environments the Controller node is subject to heavy traffic load. This would affect performance and scalability, and also requires careful planning to accommodate enough bandwidth in the external network gateway interface. The same requirement applies for SNAT traffic.

To better scale the L3 agent, the Networking service can use the L3 HA feature, which distributes the virtual routers across multiple nodes. In the event that a Controller node is lost, the HA router will failover to a standby on another node and there will be packet loss until the HA router failover completes.

14.2. DVR OVERVIEW

Distributed Virtual Routing (DVR) offers an alternative routing design. DVR isolates the failure domain of the Controller node and optimizes network traffic by deploying the L3 agent and schedule routers on every Compute node. DVR has these characteristics:

- East-West traffic is routed directly on the Compute nodes in a distributed fashion.
- North-South traffic with floating IP is distributed and routed on the Compute nodes. This requires the external network to be connected to every Compute node.
- North-South traffic without floating IP is not distributed and still requires a dedicated Controller node.
- The L3 agent on the Controller node uses the **dvr_snat** mode so that the node serves only SNAT traffic.
- The neutron metadata agent is distributed and deployed on all Compute nodes. The metadata proxy service is hosted on all the distributed routers.

14.3. DVR KNOWN ISSUES AND CAVEATS

**NOTE**

For Red Hat OpenStack Platform 13, do not use DVR unless your kernel version is at least **kernel-3.10.0-514.1.1.el7**

- Support for DVR is limited to the ML2 core plug-in and the Open vSwitch (OVS) mechanism driver or ML2/OVN mechanism driver. Other back ends are not supported.
- On both OVS and OVN DVR deployments, network traffic for the Red Hat OpenStack Platform Load-balancing service (octavia) goes through the Controller and network nodes, instead of the compute nodes.
- With an ML2/OVS mechanism driver network back end and DVR, it is possible to create VIPs. However, the IP address assigned to a bound port using **allowed_address_pairs**, should match the virtual port IP address (/32).
If you use a CIDR format IP address for the bound port **allowed_address_pairs** instead, port forwarding is not configured in the back end, and traffic fails for any IP in the CIDR expecting to reach the bound IP port.
- SNAT (source network address translation) traffic is not distributed, even when DVR is enabled. SNAT does work, but all ingress/egress traffic must traverse through the centralized Controller node.
- IPv6 traffic is not distributed, even when DVR is enabled. IPv6 routing does work, but all ingress/egress traffic must traverse through the centralized Controller node. If you use IPv6 routing extensively, do not use DVR.
- DVR is not supported in conjunction with L3 HA. If you use DVR with Red Hat OpenStack Platform 13 director, L3 HA is disabled. This means that routers are still scheduled on the Network nodes (and load-shared between the L3 agents), but if one agent fails, all routers hosted by this agent fail as well. This affects only SNAT traffic. The **allow_automatic_l3agent_failover** feature is recommended in such cases, so that if one network node fails, the routers are rescheduled to a different node.
- DHCP servers, which are managed by the neutron DHCP agent, are not distributed and are still deployed on the Controller node. The DHCP agent is deployed in a highly available configuration on the Controller nodes, regardless of the routing design (centralized or DVR).
- To work with floating IPs, each Compute node requires an interface on the *External* network. In addition, each Compute node requires one additional IP address. This is due to the implementation of the external gateway port and the floating IP network namespace.
- VLAN, GRE, and VXLAN are all supported for project data separation. When you use GRE or VXLAN, you must enable the L2 Population feature. The Red Hat OpenStack Platform director enforces L2 Population during installation.

14.4. SUPPORTED ROUTING ARCHITECTURES

Red Hat OpenStack Platform (RHOSP) supports both centralized, high-availability (HA) routing and distributed virtual routing (DVR) in the RHOSP versions listed:

- RHOSP centralized HA routing support began in RHOSP 8.
- RHOSP distributed routing support began in RHOSP 12.

14.5. DEPLOYING DVR WITH ML2 OVS

To deploy and manage distributed virtual routing (DVR) in an ML2/OVS deployment, you configure settings in heat templates and environment files.

You use heat template settings to provision host networking:

- Configure the interface connected to the physical network for external network traffic on both the Compute and Controller nodes.
- Create a bridge on Compute and Controller nodes, with an interface for external network traffic.

You also configure the Networking service (neutron) to match the provisioned networking environment and allow traffic to use the bridge.

The default settings are provided as guidelines only. They are not expected to work in production or test environments which may require customization for network isolation, dedicated NICs, or any number of other variable factors. In setting up an environment, you need to correctly configure the bridge mapping type parameters used by the L2 agents and the external facing bridges for other agents, such as the L3 agent.

The following example procedure shows how to configure a proof-of-concept environment using the typical defaults.

Procedure

1. Verify that the value for **OS::TripleO::Compute::Net::SoftwareConfig** matches the value of **OS::TripleO::Controller::Net::SoftwareConfig** in the file *overcloud-resource-registry.yaml* or in an environment file included in the deployment command.

This value names a file, such as *net_config_bridge.yaml*. The named file configures Neutron bridge mappings for external networks Compute node L2 agents. The bridge routes traffic for the floating IP addresses hosted on Compute nodes in a DVR deployment. Normally, you can find this filename value in the network environment file that you use when deploying the overcloud, such as *environments/net-multiple-nics.yaml*.



NOTE

If you customize the network configuration of the Compute node, you may need to add the appropriate configuration to your custom files instead.

2. Verify that the Compute node has an external bridge.
 - a. Make a local copy of the **openstack-tripleo-heat-templates** directory.
 - b. **\$ cd <local_copy_of_templates_directory>**
 - c. Run the **process-templates** script to render the templates to a temporary output directory:

```
$ ./tools/process-templates.py -r <roles_data.yaml> \
-n <network_data.yaml> -o <temporary_output_directory>
```

- d. Check the role files in **<temporary_output_directory>/network/config**.

3. If needed, customize the Compute template to include an external bridge that matches the Controller nodes, and name the custom file path in **OS::TripleO::Compute::Net::SoftwareConfig** in an environment file.
4. Include the `environments/services/neutron-ovs-dvr.yaml` file in the deployment command when deploying the overcloud:

```
$ openstack overcloud deploy --templates -e /usr/share/openstack-tripleo-heat-templates/environments/services/neutron-ovs-dvr.yaml
```

5. Verify that L3 HA is disabled.



NOTE

The external bridge configuration for the L3 agent was deprecated in Red Hat OpenStack Platform 13 and removed in Red Hat OpenStack Platform 15.

14.6. MIGRATING CENTRALIZED ROUTERS TO DISTRIBUTED ROUTING

This section contains information about upgrading to distributed routing for Red Hat OpenStack Platform deployments that use L3 HA centralized routing.

Procedure

1. Upgrade your deployment and validate that it is working correctly.
2. Run the director stack update to configure DVR.
3. Confirm that routing functions correctly through the existing routers.
4. You cannot transition an L3 HA router to *distributed* directly. Instead, for each router, disable the L3 HA option, and then enable the distributed option:

- a. Disable the router:

Example

```
$ openstack router set --disable router1
```

- b. Clear high availability:

Example

```
$ openstack router set --no-ha router1
```

- c. Configure the router to use DVR:

Example

```
$ openstack router set --distributed router1
```

- d. Enable the router:

Example

```
█ $ openstack router set --enable router1
```

- e. Confirm that distributed routing functions correctly.

Additional resources

- [Deploying DVR with ML2 OVS](#)

CHAPTER 15. CONFIGURE LOAD BALANCING-AS-A-SERVICE WITH THE NETWORKING LBAASV2 API

15.1. OVERVIEW OF LBaaS

Load Balancing-as-a-Service (LBaaS) enables OpenStack Networking to distribute incoming requests evenly between designated instances. Complete the steps in this section to configure OpenStack Networking to use LBaaS with the Open vSwitch (OVS) plug-in.

Load Balancing-as-a-Service (LBaaS) enables OpenStack Networking to distribute incoming requests evenly between designated instances. This ensures the workload is shared predictably among instances, and allows more effective use of system resources. Incoming requests are distributed using one of the following load balancing methods:

- **Round robin** - Rotates requests evenly between multiple instances.
- **Source IP** - Requests from a unique source IP address are consistently directed to the same instance.
- **Least connections** - Allocates requests to the instance with the least number of active connections.

Table 15.1. LBaaS features

Feature	Description
Monitors	LBaaS provides availability monitoring with the PING, TCP, HTTP and HTTPS GET methods. Monitors determine whether pool members are available to handle requests.
Management	LBaaS is managed using a variety of tool sets. The REST API is available for programmatic administration and scripting. Users perform administrative management of load balancers through either the CLI (neutron) or the OpenStack dashboard.
Connection limits	Ingress traffic can be shaped with connection limits. This feature allows workload control and can also assist with mitigating DoS (Denial of Service) attacks.
Session persistence	LBaaS supports session persistence by ensuring incoming requests are routed to the same instance within a pool of multiple instances. LBaaS supports routing decisions based on cookies and source IP address.



NOTE

LBaaS is currently supported only with IPv4 addressing.

15.2. OPENSTACK NETWORKING AND LBAAS TOPOLOGY

OpenStack Networking (neutron) services can be broadly classified into two categories.

Neutron API server - This service runs the OpenStack Networking API server, which has the main responsibility of providing an API for end users and services to interact with OpenStack Networking. This server also has the responsibility of interacting with the underlying database to store and retrieve tenant network, router, and loadbalancer details, among others.

Neutron Agents - These are the services that deliver various network functionality for OpenStack Networking.

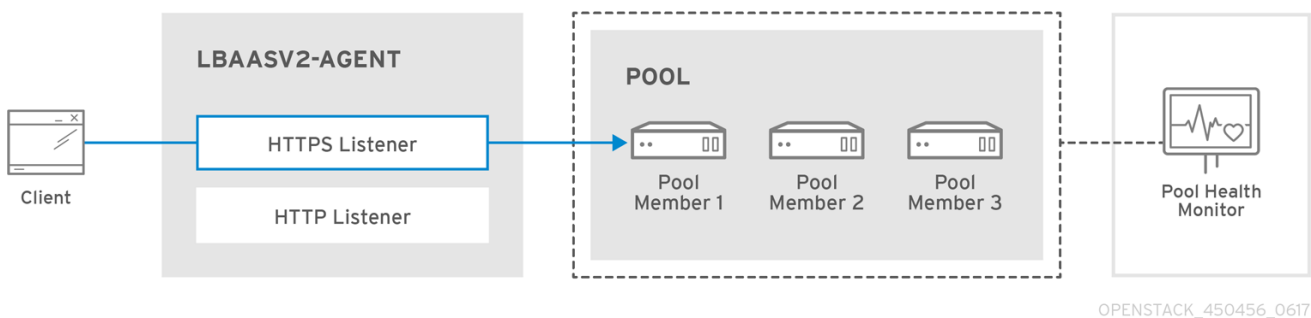
- **neutron-dhcp-agent** - manages DHCP IP addressing for tenant private networks.
- **neutron-l3-agent** - facilitates layer 3 routing between tenant private networks, the external network, and other networks.



NOTE

The **neutron-lbaasv2-agent** (with HAProxy) is deprecated. For the preferred load-balancing reference implementation with octavia, see the [Using Octavia for Load Balancing-as-a-Service](#) guide.

The following diagram shows the flow of HTTPS traffic through to a pool member:



15.2.1. Support Status of LBaaS

- LBaaS v1 API was removed in version 10.
- LBaaS v2 API is deprecated and removed after Red Hat OpenStack Platform 13. (Octavia is the replacement.)
- LBaaS deployment is not currently supported in Red Hat OpenStack Platform director.

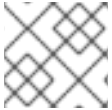


NOTE

The **neutron-lbaasv2-agent** (with HAProxy) is deprecated. For the preferred load-balancing reference implementation with Octavia, see the [Using Octavia for Load Balancing-as-a-Service](#) guide. The neutron-lbaas RPMs are still available to support the API for third-party plugin support.

15.3. CONFIGURING LBAAS

This procedure configures OpenStack Networking (neutron) to use LBaaS with the Open vSwitch (OVS) plugin.



NOTE

Perform these steps on nodes running the **neutron-server** service:

Procedure

On the Controller node (API Server):

1. Enable LBaaS:

```
# yum install openstack-neutron-lbaas -y
```

2. Add the LBaaS tables to the neutron database:

```
$ neutron-db-manage --subproject neutron-lbaas --config-file /var/lib/config-  
data/neutron/etc/neutron/neutron.conf --config-file /var/lib/config-  
data/neutron/etc/neutron/plugins/ml2/ml2_conf.ini upgrade head
```

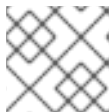
3. Change the service provider in **/var/lib/config-data/neutron/etc/neutron/neutron_lbaas.conf**. In the **[service providers]** section, comment out (**#**) all entries except for this entry:

```
service_provider=LOADBALANCERV2:Haproxy:neutron_lbaas.drivers.haproxy.plugin_driver.H  
aproxyOnHostPluginDriver:default
```

4. In **/var/lib/config-data/neutron/etc/neutron/neutron.conf**, confirm that you have the LBaaS v2 plugin configured in **service_plugins**:

```
service_plugins=neutron_lbaas.services.loadbalancer.plugin.LoadBalancerPluginv2
```

You can also expect to see any other plugins you have previously added.



NOTE

If you have **lbaasv1** configured, replace it with the above setting for **lbaasv2**.

5. In **/var/lib/config-data/neutron/etc/neutron/lbaas_agent.ini**, add the following to the **[DEFAULT]** section:

```
ovs_use_veth = False  
interface_driver =neutron.agent.linux.interface.OVSInterfaceDriver
```

6. In **/var/lib/config-data/neutron/etc/neutron/services_lbaas.conf**, add the following to the **[haproxy]** section:

```
user_group = haproxy
```

- a. Comment out any other **device driver** entries.

**NOTE**

If the **l3-agent** is in a failed mode, see the **l3_agent** log files. You may need to edit `/var/lib/config-data/neutron/etc/neutron/neutron.conf` and comment out certain values in **[DEFAULT]**, and uncomment the corresponding values in **oslo_messaging_rabbit**, as described in the log file.

7. Configure the LbaaS services, and review their status:

a. Stop the **lbaasv1** services and start **lbaasv2**:

```
# systemctl disable neutron-lbaas-agent.service
# systemctl stop neutron-lbaas-agent.service
# systemctl mask neutron-lbaas-agent.service
# systemctl enable neutron-lbaasv2-agent.service
# systemctl start neutron-lbaasv2-agent.service
```

b. Review the status of **lbaasv2**:

```
# systemctl status neutron-lbaasv2-agent.service
```

c. Restart **neutron-server** and check the status:

```
# systemctl restart neutron-server.service
# systemctl status neutron-server.service
```

d. Check the **Loadbalancerv2** agent:

```
$ openstack network agent list
```

CHAPTER 16. PROJECT NETWORKING WITH IPV6

16.1. IPV6 SUBNET OPTIONS

When you create IPv6 subnets in a Red Hat OpenStack Platform (RHOSP) project network you can specify address mode and Router Advertisement mode to obtain a particular result as described in the following table.



NOTE

RHOSP does not support IPv6 prefix delegation in ML2/OVN deployments. You must set the Global Unicast Address prefix manually.

RA Mode	Address Mode	Result
ipv6_ra_mode=not set	ipv6-address-mode=slaac	<p>The instance receives an IPv6 address from the external router (not managed by OpenStack Networking) using Stateless Address Autoconfiguration (SLAAC).</p> <div style="display: flex; align-items: center;"> <div> <p>NOTE</p> <p>OpenStack Networking supports only EUI-64 IPv6 address assignment for SLAAC. This allows for simplified IPv6 networking, as hosts self-assign addresses based on the base 64-bits plus the MAC address. You cannot create subnets with a different netmask and <i>address_assign_type</i> of SLAAC.</p> </div> </div>
ipv6_ra_mode=not set	ipv6-address-mode=dhcpv6-stateful	<p>The instance receives an IPv6 address and optional information from OpenStack Networking (dnsmasq) using DHCPv6 stateful.</p>

RA Mode	Address Mode	Result
ipv6_ra_mode=not set	ipv6-address-mode=dhcpv6-stateless	The instance receives an IPv6 address from the external router using SLAAC, and optional information from OpenStack Networking (dnsmasq) using DHCPv6 stateless .
ipv6_ra_mode=slaac	ipv6-address-mode=not-set	The instance uses SLAAC to receive an IPv6 address from OpenStack Networking (radvd).
ipv6_ra_mode=dhcpv6-stateful	ipv6-address-mode=not-set	The instance receives an IPv6 address and optional information from an external DHCPv6 server using DHCPv6 stateful .
ipv6_ra_mode=dhcpv6-stateless	ipv6-address-mode=not-set	The instance receives an IPv6 address from OpenStack Networking (radvd) using SLAAC, and optional information from an external DHCPv6 server using DHCPv6 stateless .
ipv6_ra_mode=slaac	ipv6-address-mode=slaac	The instance receives an IPv6 address from OpenStack Networking (radvd) using SLAAC .
ipv6_ra_mode=dhcpv6-stateful	ipv6-address-mode=dhcpv6-stateful	The instance receives an IPv6 address from OpenStack Networking (dnsmasq) using DHCPv6 stateful , and optional information from OpenStack Networking (dnsmasq) using DHCPv6 stateful .
ipv6_ra_mode=dhcpv6-stateless	ipv6-address-mode=dhcpv6-stateless	The instance receives an IPv6 address from OpenStack Networking (radvd) using SLAAC , and optional information from OpenStack Networking (dnsmasq) using DHCPv6 stateless .

16.2. CREATE AN IPV6 SUBNET USING STATEFUL DHCPV6

You can create an IPv6 subnet in a Red Hat OpenStack (RHOSP) project network.

For example, you can create an IPv6 subnet using Stateful DHCPv6 in network named database-servers in a project named QA.

Procedure

1. Retrieve the project ID of the Project where you want to create the IPv6 subnet. These values are unique between OpenStack deployments, so your values differ from the values in this example.

```
# openstack project list
+-----+
| ID              | Name  |
+-----+
| 25837c567ed5458fbb441d39862e1399 | QA   |
| f59f631a77264a8eb0defc898cb836af | admin |
| 4e2e1951e70643b5af7ed52f3ff36539 | demo  |
| 8561dff8310e4cd8be4b6fd03dc8acf5 | services |
+-----+
```

2. Retrieve a list of all networks present in OpenStack Networking (neutron), and note the name of the network where you want to host the IPv6 subnet:

```
# openstack network list
+-----+
| id              | name          | subnets          |
+-----+
| 8357062a-0dc2-4146-8a7f-d2575165e363 | private      | c17f74c4-db41-4538-af40-48670069af70 10.0.0.0/24 |
| 31d61f7d-287e-4ada-ac29-ed7017a54542 | public       | 303ced03-6019-4e79-a21c-1942a460b920 172.24.4.224/28 |
| 6aff6826-4278-4a35-b74d-b0ca0cbba340 | database-servers |
+-----+
```

3. Include the project ID, network name, and ipv6 address mode in the **openstack subnet create** command:

```
# openstack subnet create --ip-version 6 --ipv6-address-mode dhcpv6-stateful --project 25837c567ed5458fbb441d39862e1399 --network database-servers --subnet-range fdf8:f53b:82e4::53/125 subnet_name
```

Created a new subnet:

```
+-----+
| Field          | Value                                               |
+-----+
| allocation_pools | {"start": "fdf8:f53b:82e4::52", "end": "fdf8:f53b:82e4::56"} |
| cidr            | fdf8:f53b:82e4::53/125                             |
| dns_nameservers |                                                       |
| enable_dhcp     | True                                                |
| gateway_ip      | fdf8:f53b:82e4::51                                 |
| host_routes     |                                                       |
| id              | cdfc3398-997b-46eb-9db1-ebbd88f7de05             |
| ip_version      | 6                                                    |
| ipv6_address_mode | dhcpv6-stateful                                     |
| ipv6_ra_mode    |                                                       |
```

```

| name          |
| network_id   | 6aff6826-4278-4a35-b74d-b0ca0cbba340 |
| tenant_id    | 25837c567ed5458fbb441d39862e1399   |
+-----+

```

Validation steps

1. Validate this configuration by reviewing the network list. Note that the entry for *database-servers* now reflects the newly created IPv6 subnet:

```

# openstack network list
+-----+-----+-----+-----+
-----+
| id          | name          | subnets          |
+-----+-----+-----+-----+
-----+
| 6aff6826-4278-4a35-b74d-b0ca0cbba340 | database-servers | cdfc3398-997b-46eb-9db1-ebbd88f7de05 | fdf8:f53b:82e4::50/125 |
| 8357062a-0dc2-4146-8a7f-d2575165e363 | private          | c17f74c4-db41-4538-af40-48670069af70 | 10.0.0.0/24 |
| 31d61f7d-287e-4ada-ac29-ed7017a54542 | public           | 303ced03-6019-4e79-a21c-1942a460b920 | 172.24.4.224/28 |
+-----+-----+-----+-----+
-----+

```

Result

As a result of this configuration, instances that the QA project creates can receive a DHCP IPv6 address when added to the *database-servers* subnet:

```

# openstack server list
+-----+-----+-----+-----+-----+-----+
-----+
| ID          | Name          | Status | Task State | Power State | Networks          |
+-----+-----+-----+-----+-----+-----+
-----+
| fad04b7a-75b5-4f96-aed9-b40654b56e03 | corp-vm-01 | ACTIVE | -          | Running    | database-servers=fdf8:f53b:82e4::52 |
+-----+-----+-----+-----+-----+-----+
-----+

```

Additional resources

To find the Router Advertisement mode and address mode combinations to achieve a particular result in an IPv6 subnet, see [IPv6 subnet options](#) in the [Networking Guide](#).

CHAPTER 17. MANAGING PROJECT QUOTAS

17.1. CONFIGURING PROJECT QUOTAS

OpenStack Networking (neutron) supports the use of quotas to constrain the number of resources created by tenants/projects.

Procedure

- You can set project quotas for various network components in the `/var/lib/config-data/neutron/etc/neutron/neutron.conf` file.
For example, to limit the number of routers that a project can create, change the **quota_router** value:

```
quota_router = 10
```

In this example, each project is limited to a maximum of 10 routers.

For a listing of the quota settings, see sections that immediately follow.

17.2. L3 QUOTA OPTIONS

Here are quota options available for layer 3 (L3) networking:

- **quota_floatingip** - The number of floating IPs available to a project.
- **quota_network** - The number of networks available to a project.
- **quota_port** - The number of ports available to a project.
- **quota_router** - The number of routers available to a project.
- **quota_subnet** - The number of subnets available to a project.
- **quota_vip** - The number of virtual IP addresses available to a project.

17.3. FIREWALL QUOTA OPTIONS

Here are quota options available for managing firewalls for projects:

- **quota_firewall** - The number of firewalls available to a project.
- **quota_firewall_policy** - The number of firewall policies available to a project.
- **quota_firewall_rule** - The number of firewall rules available to a project.

17.4. SECURITY GROUP QUOTA OPTIONS

The Networking service quota engine manages security groups and security group rules, and it is not possible to set all quotas to zero before creating the default security group (and the two default security group rules that accepts all egress traffic for IPv4 and IPv6). When you create a new project, the Networking service does not create the default security group until a network or a port is created, or until you list the security group or the security group rules.

Here are quota options available for managing the number of security groups that projects can create:

- **quota_security_group** - The number of security groups available to a project.
- **quota_security_group_rule** - The number of security group rules available to a project.

17.5. MANAGEMENT QUOTA OPTIONS

Here are additional options available to administrators for managing quotas for projects:

- **default_quota*** - The default number of resources available to a project.
- **quota_health_monitor*** - The number of health monitors available to a project.
Health monitors do not consume resources, however the quota option is available because OpenStack Networking considers health monitors as resource consumers.
- **quota_member** - The number of pool members available to a project.
Pool members do not consume resources, however the quota option is available because OpenStack Networking considers pool members as resource consumers.
- **quota_pool** - The number of pools available to a project.

CHAPTER 18. CONFIGURING FIREWALL-AS-A-SERVICE (FWAAS)

18.1. OVERVIEW OF FIREWALL-AS-A-SERVICE (FWAAS)

The Firewall-as-a-Service (FWaaS) plug-in adds perimeter firewall management to OpenStack Networking (neutron). FWaaS uses iptables to apply firewall policy to all virtual routers within a project, and supports one firewall policy and logical firewall instance for each project.

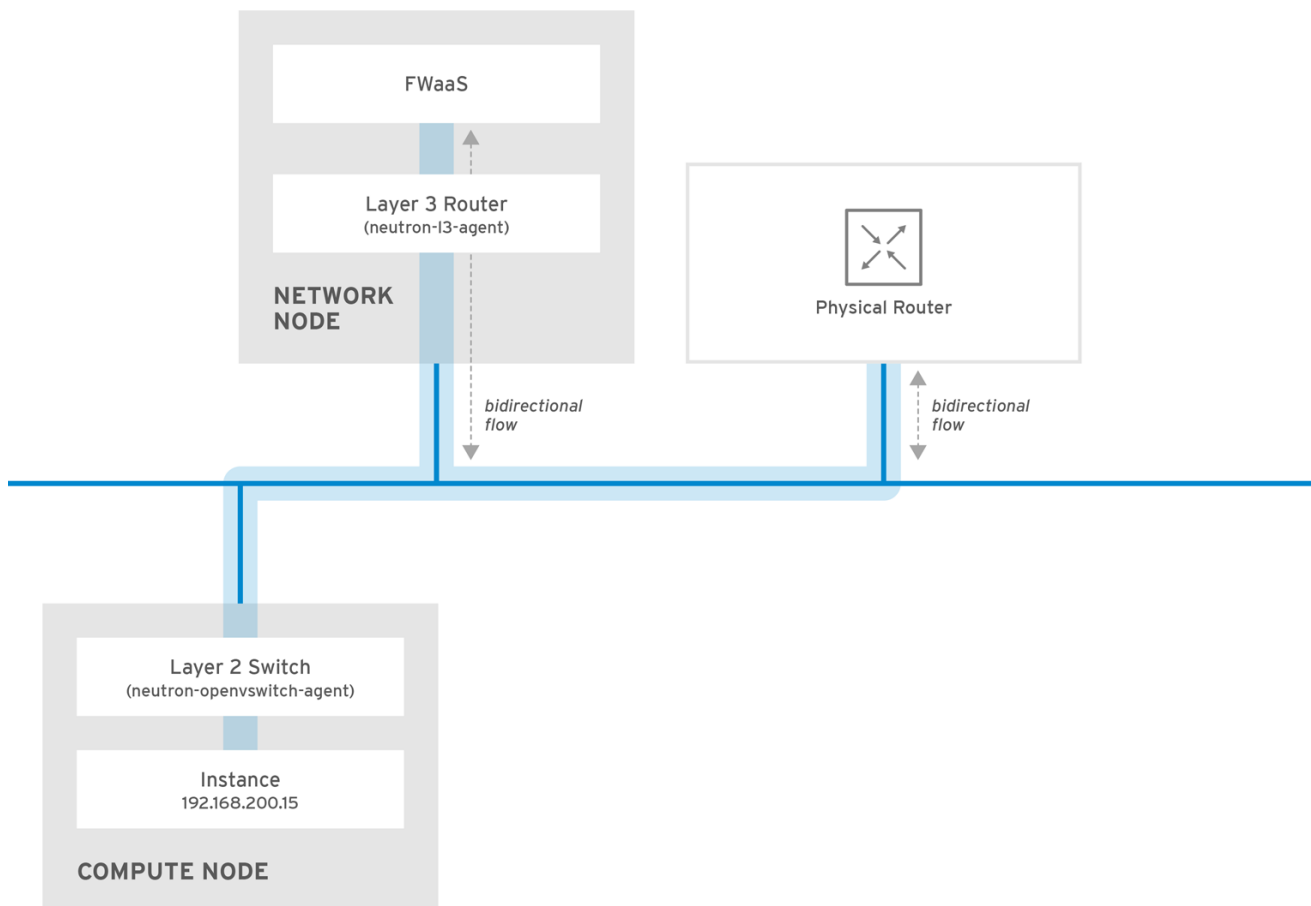
FWaaS operates at the perimeter by filtering traffic at the OpenStack Networking (neutron) router. This distinguishes FWaaS from security groups, which operate at the instance level.



NOTE

FWaaS is currently in Technology Preview; untested operation is not recommended. FWaaS is not available in future releases.

The following example diagram illustrates the flow of ingress and egress traffic for the VM2 instance:



OPENSTACK_450456_0617

Figure 1. FWaaS architecture

18.2. ENABLING FIREWALL-AS-A-SERVICE (FWAAS)

1. Install the FWaaS packages:

```
# dnf install openstack-neutron-fwaas python-neutron-fwaas
```

2. Enable the FWaaS plugin in the `/var/lib/config-data/neutron/etc/neutron/neutron.conf` file:

```
service_plugins = neutron.services.firewall.fwaas_plugin.FirewallPlugin
```

3. Configure FWaaS in the `fwaas_driver.ini` file:

```
[fwaas]
driver = neutron.services.firewall.drivers.linux.iptables_fwaas.IptablesFwaasDriver
enabled = True

[service_providers]
service_provider =
LOADBALANCER:Haproxy:neutron_lbaas.services.loadbalancer.drivers.haproxy.plugin_driver.
HaproxyOnHostPluginDriver:default
```

4. Enable the FWaaS dashboard management option in the `local_settings.py` file, usually located on the Controller node:

```
/usr/share/openstack-dashboard/openstack_dashboard/local/local_settings.py
'enable_firewall' = True
```

5. Restart **neutron-server** to apply the changes.

```
# systemctl restart neutron-server
```

18.3. CONFIGURING FIREWALL-AS-A-SERVICE (FWAAS)

First, create the firewall rules and create a policy to contain them, then create a firewall and apply the policy:

1. Create a firewall rule:

```
$ neutron firewall-rule-create --protocol <tcp|udp|icmp|any> --destination-port <port-range> --
action <allow|deny>
```

The CLI requires a protocol value. If the rule is protocol agnostic, you can use the value **any**.

2. Create a firewall policy:

```
$ neutron firewall-policy-create --firewall-rules "<firewall-rule IDs or names separated by
space>" myfirewallpolicy
```

The order of the rules that you specify in the **neutron firewall-policy-create** command is important. You can create an empty firewall policy and add rules later, either with the update operation (when adding multiple rules) or with the insert-rule operations (when adding a single rule).

**NOTE**

FWaaS always adds a default deny all rule at the lowest precedence of each policy. Consequently, a firewall policy with no rules blocks all traffic by default.

18.4. CREATING FIREWALLS

- Use the **openstack security group create** command, to create a firewall:

```
$ openstack security group create <firewall-policy-uuid>
```

The firewall remains in **PENDING_CREATE** state until you create an OpenStack Networking router and attach an interface.

CHAPTER 19. CONFIGURING ALLOWED ADDRESS PAIRS

19.1. OVERVIEW OF ALLOWED ADDRESS PAIRS

An *allowed address pair* is when you identify a specific MAC address, IP address, or both to allow network traffic to pass through a port regardless of the subnet. When you define allowed address pairs, you are able to use protocols like VRRP (Virtual Router Redundancy Protocol) that float an IP address between two VM instances to enable fast data plane failover.



NOTE

The allowed-address pairs extension is currently supported only by the ML2 and Open vSwitch plug-ins.

You define allowed address pairs using the Red Hat OpenStack Platform command-line client **openstack port** command.



IMPORTANT

Be aware that you should not use the default security group with a wider IP address range in an allowed address pair. Doing so can allow a single port to bypass security groups for all other ports within the same network.

For example, this command impacts all ports in the network and bypasses all security groups:

```
# openstack port set --allowed-address
mac_address=3e:37:09:4b,ip_address=0.0.0.0/0
9e67d44eab334f07bf82fa1b17d824b6
```



NOTE

With an ML2/OVN mechanism driver network back end, it is possible to create VIPs. However, the IP address assigned to a bound port using **allowed_address_pairs**, should match the virtual port IP address (/32).

If you use a CIDR format IP address for the bound port **allowed_address_pairs** instead, port forwarding is not configured in the back end, and traffic fails for any IP in the CIDR expecting to reach the bound IP port.

Additional resources

- [port command](#) in the *Command Line Interface Reference*
- [Creating a port and allowing one address pair](#)
- [Adding allowed address pairs](#)

19.2. CREATING A PORT AND ALLOWING ONE ADDRESS PAIR

Creating a port with an allowed address pair enables network traffic to flow through the port regardless of the subnet.

Prerequisites

- You are using an ML2/OVS plug-in.



IMPORTANT

Do not use the default security group with a wider IP address range in an allowed address pair. Doing so can allow a single port to bypass security groups for all other ports within the same network.

Procedure

- Use the following command to create a port and allow one address pair:

```
# openstack port create <port-name> --network <network> --allowed-address mac_address=
<mac-address>,ip_address=<ip-cidr>
```

Additional resources

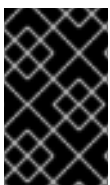
- [port command](#) in the *Command Line Interface Reference*
- [Overview of allowed address pairs](#)
- [Adding allowed address pairs](#)

19.3. ADDING ALLOWED ADDRESS PAIRS

You can add an allowed address pair to a port to enable network traffic to flow through the port regardless of the subnet.

Prerequisites

- You are using an ML2/OVS plug-in.



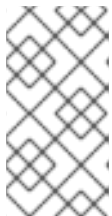
IMPORTANT

Do not use the default security group with a wider IP address range in an allowed address pair. Doing so can allow a single port to bypass security groups for all other ports within the same network.

Procedure

- Use the following command to add allowed address pairs:

```
# openstack port set <port-uuid> --allowed-address mac_address=
<mac_address>,ip_address=<ip_cidr>
```



NOTE

You cannot set an allowed-address pair that matches the **mac_address** and **ip_address** of a port. This is because such a setting has no effect since traffic matching the **mac_address** and **ip_address** is already allowed to pass through the port.

Additional resources

- [port command](#) in the *Command Line Interface Reference*
- [Overview of allowed address pairs](#)
- [Creating a port and allowing one address pair](#)

CHAPTER 20. CONFIGURING LAYER 3 HIGH AVAILABILITY (HA)

20.1. RHOSP NETWORKING SERVICE WITHOUT HIGH AVAILABILITY (HA)

Red Hat OpenStack Platform (RHOSP) Networking service deployments without any high availability (HA) features are vulnerable to physical node failures.

In a typical deployment, projects create virtual routers, which are scheduled to run on physical Networking service Layer 3 (L3) agent nodes. This becomes an issue when you lose an L3 agent node and the dependent virtual machines subsequently lose connectivity to external networks. Any floating IP addresses are also unavailable. In addition, connectivity is lost between any networks that the router hosts.

20.2. OVERVIEW OF LAYER 3 HIGH AVAILABILITY (HA)

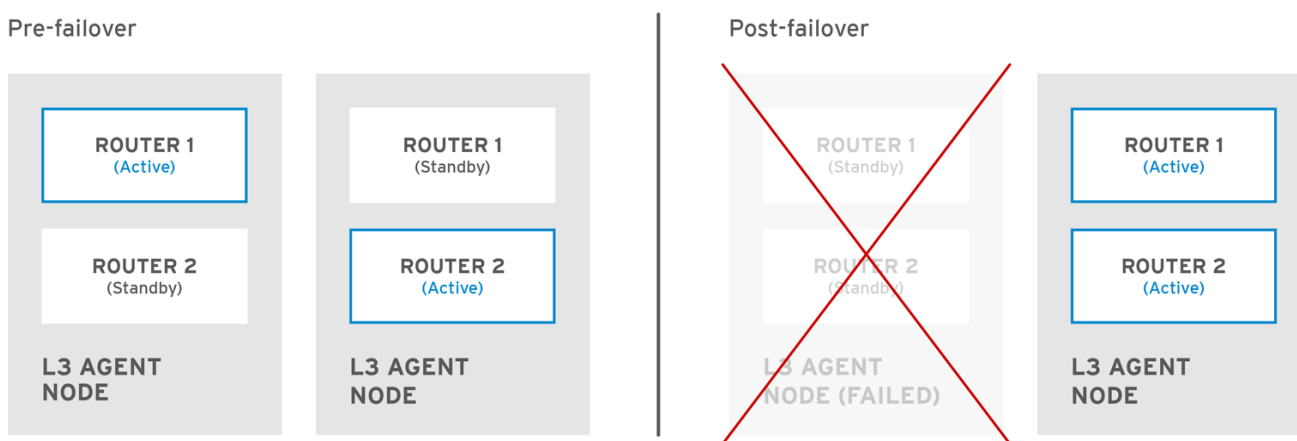
This active/passive high availability (HA) configuration uses the industry standard VRRP (as defined in RFC 3768) to protect project routers and floating IP addresses. A virtual router is randomly scheduled across multiple Red Hat OpenStack Platform (RHOSP) Networking service nodes, with one designated as the *active* router, and the remainder serving in a *standby* role.



NOTE

To deploy Layer 3 (L3) HA, you must maintain similar configuration on the redundant Networking service nodes, including floating IP ranges and access to external networks.

In the following diagram, the active **Router1** and **Router2** routers are running on separate physical L3 Networking service agent nodes. L3 HA has scheduled backup virtual routers on the corresponding nodes, ready to resume service in the case of a physical node failure. When the L3 agent node fails, L3 HA reschedules the affected virtual router and floating IP addresses to a working node:



OPENSTACK_450456_0617

During a failover event, instance TCP sessions through floating IPs remain unaffected, and migrate to the new L3 node without disruption. Only SNAT traffic is affected by failover events.

The L3 agent is further protected when in an active/active HA mode.

Additional resources

- [Virtual Router Redundancy Protocol \(VRRP\)](#)

20.3. LAYER 3 HIGH AVAILABILITY (HA) FAILOVER CONDITIONS

Layer 3 (L3) high availability (HA) for the Red Hat OpenStack Platform (RHOSP) Networking service automatically reschedules protected resources in the following events:

- The Networking service L3 agent node shuts down or otherwise loses power because of a hardware failure.
- The L3 agent node becomes isolated from the physical network and loses connectivity.



NOTE

Manually stopping the L3 agent service does not induce a failover event.

20.4. PROJECT CONSIDERATIONS FOR LAYER 3 HIGH AVAILABILITY (HA)

Red Hat OpenStack Platform (RHOSP) Networking service Layer 3 (L3) high availability (HA) configuration occurs in the back end and is invisible to the project. Projects can continue to create and manage their virtual routers as usual, however there are some limitations to be aware of when designing your L3 HA implementation:

- L3 HA supports up to 255 virtual routers per project.
- Internal VRRP messages are transported within a separate internal network, created automatically for each project. This process occurs transparently to the user.
- When implementing high availability (HA) routers on ML2/OVS, each L3 agent spawns **haproxy** and **neutron-keepalived-state-change-monitor** processes for each router. Each process consumes approximately 20MB of memory. By default, each HA router resides on three L3 agents and consumes resources on each of the nodes. Therefore, when sizing your RHOSP networks, ensure that you have allocated enough memory to support the number of HA routers that you plan to implement.

20.5. HIGH AVAILABILITY (HA) CHANGES TO THE RHOSP NETWORKING SERVICE

The Red Hat OpenStack Platform (RHOSP) Networking service (neutron) API has been updated to allow administrators to set the **--ha=True/False** flag when creating a router, which overrides the default configuration of **I3_ha** in **/var/lib/config-data/neutron/etc/neutron/neutron.conf**.

- **High availability (HA) changes to neutron-server:**
 - Layer 3 (L3) HA assigns the active role randomly, regardless of the scheduler used by the Networking service (whether random or leastrouter).
 - The database schema has been modified to handle allocation of virtual IP addresses (VIPs) to virtual routers.
 - A transport network is created to direct L3 HA traffic.
- **HA changes to the Networking service L3 agent:**

- A new keepalived manager has been added, providing load-balancing and HA capabilities.
- IP addresses are converted to VIPs.

20.6. ENABLING LAYER 3 HIGH AVAILABILITY (HA) ON RHOSP NETWORKING SERVICE NODES

During installation, Red Hat OpenStack Platform (RHOSP) director enables high availability (HA) for virtual routers by default when you have at least two RHOSP Controllers and are not using distributed virtual routing (DVR). Using an RHOSP Orchestration service (heat) parameter, **max_l3_agents_per_router**, you can set the maximum number of RHOSP Networking service Layer 3 (L3) agents on which an HA router is scheduled.

Prerequisites

- Your RHOSP deployment does not use DVR.
- You have at least two RHOSP Controllers deployed.

Procedure

1. Log in to the undercloud as the stack user, and source the **stackrc** file to enable the director command line tools.

Example

```
$ source ~/stackrc
```

2. Create a custom YAML environment file.

Example

```
$ vi /home/stack/templates/my-neutron-environment.yaml
```

TIP

The Orchestration service (heat) uses a set of plans called *templates* to install and configure your environment. You can customize aspects of the overcloud with a *custom environment file*, which is a special type of template that provides customization for your heat templates.

3. Set the **NeutronL3HA** parameter to **true** in the YAML environment file. This ensures HA is enabled even if director did not set it by default.

```
parameter_defaults:
  NeutronL3HA: 'true'
```

4. Set the maximum number of L3 agents on which an HA router is scheduled. Set the **max_l3_agents_per_router** parameter to a value between the minimum and total number of network nodes in your deployment. (A zero value indicates that the router is scheduled on every agent.)

Example

-

```
parameter_defaults:
  NeutronL3HA: 'true'
  ControllerExtraConfig:
    neutron::server::max_l3_agents_per_router: 2
```

In this example, if you deploy four Networking service nodes, only two L3 agents protect each HA virtual router: one active, and one standby.

If you set the value of **max_l3_agents_per_router** to be greater than the number of available network nodes, you can scale out the number of standby routers by adding new L3 agents. For every new L3 agent node that you deploy, the Networking service schedules additional standby versions of the virtual routers until the **max_l3_agents_per_router** limit is reached.

5. Run the **openstack overcloud deploy** command and include the core heat templates, environment files, and this new custom environment file.



IMPORTANT

The order of the environment files is important because the parameters and resources defined in subsequent environment files take precedence.

Example

```
$ openstack overcloud deploy --templates \
-e [your-environment-files] \
-e /usr/share/openstack-tripleo-heat-templates/environments/services/my-neutron-
environment.yaml
```



NOTE

When **NeutronL3HA** is set to **true**, all virtual routers that are created default to HA routers. When you create a router, you can override the HA option by including the **--no-ha** option in the **openstack router create** command:

```
# openstack router create --no-ha
```

Additional resources

- [Environment files](#) in the *Advanced Overcloud Customization* guide
- [Including Environment Files in Overcloud Creation](#) in the *Advanced Overcloud Customization* guide

20.7. REVIEWING HIGH AVAILABILITY (HA) RHOSP NETWORKING SERVICE NODE CONFIGURATIONS

Procedure

- Run the **ip address** command within the virtual router namespace to return a high availability (HA) device in the result, prefixed with *ha-*.

```
# ip netns exec qrouter-b30064f9-414e-4c98-ab42-646197c74020 ip address
```

```
<snip>  
2794: ha-45249562-ec: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc  
noqueue state DOWN group default  
link/ether 12:34:56:78:2b:5d brd ff:ff:ff:ff:ff:ff  
inet 169.254.0.2/24 brd 169.254.0.255 scope global ha-54b92d86-4f
```

With Layer 3 HA enabled, virtual routers and floating IP addresses are protected against individual node failure.

CHAPTER 21. IDENTIFYING VIRTUAL DEVICES WITH TAGS

21.1. OVERVIEW OF VIRTUAL DEVICE TAGGING

If you launch an instance with multiple network interfaces or block devices, you can use device tagging to communicate the intended role of each device to the instance operating system. Tags are assigned to devices at instance boot time, and are available to the instance operating system through the metadata API and the configuration drive (if enabled).

The tags are set using the following parameters:

- **--block-device tag=device metadata**
- **--nic tag=device metadata**

21.2. TAGGING VIRTUAL DEVICES

Procedure

- To tag virtual devices, use the tag parameters, **--block-device** and **--nic**, when creating instances.

Here is an example:

```
$ nova boot test-vm --flavor m1.tiny --image cirros \
--nic net-id=55411ca3-83dd-4036-9158-bf4a6b8fb5ce,tag=nfv1 \
--block-device id=b8c9bef7-aa1d-4bf4-a14d-17674b370e13,bus=virtio,tag=database-server
NFVappServer
```

The resulting tags are added to the existing instance metadata and are available through both the metadata API, and on the configuration drive.

In this example, the following devices section populates the metadata:

Sample contents of the *meta_data.json* file:

```
{
  "devices": [
    {
      "type": "nic",
      "bus": "pci",
      "address": "0030:00:02.0",
      "mac": "aa:00:00:00:01",
      "tags": ["nfv1"]
    },
    {
      "type": "disk",
      "bus": "pci",
      "address": "0030:00:07.0",
      "serial": "disk-vol-227",
      "tags": ["database-server"]
    }
  ]
}
```

The device tag metadata is available using **GET /openstack/latest/meta_data.json** from the metadata API.

If the configuration drive is enabled, and mounted under */configdrive* in the instance operating system, the metadata is also present in */configdrive/openstack/latest/meta_data.json*.