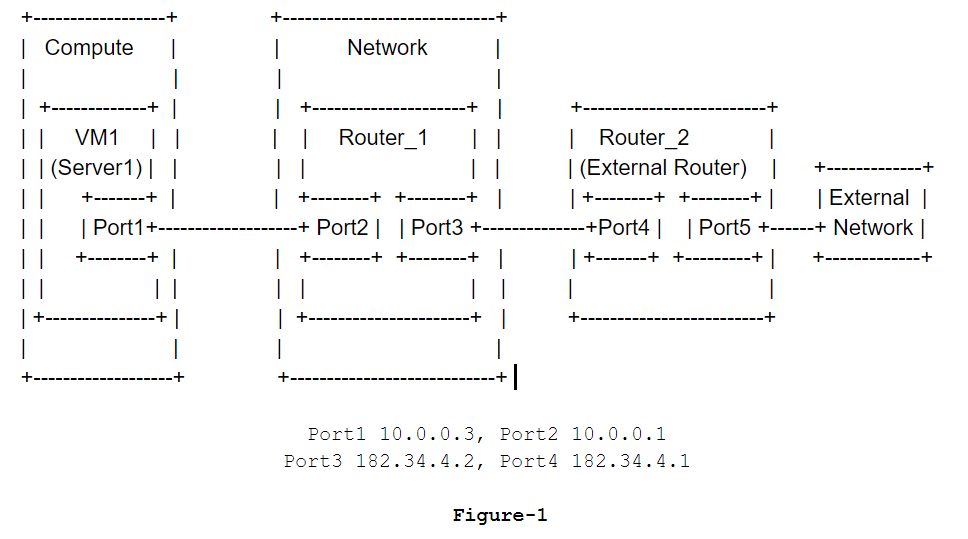
ADD L2POP SUPPORT FOR

FLOATINGIP RESOURCES

Layer 2 population (l2pop) mechanism driver implements the ML2 driver to improve open source plugins overlay implementations (VXLAN with Linux bridge and GRE/VXLAN with OVS)**[1]**. L2pop avoid the broadcast in MAC learning and ARP resolution by prepopulate the bridge forwarding table**[2]**. However l2pop doesn’t have support for floating IP resources. If the floating IP resources can be prepopulated, the broadcast for scenarios when two VM instances, residing on different networks, communicate via their respective FIP addresses can be avoided.

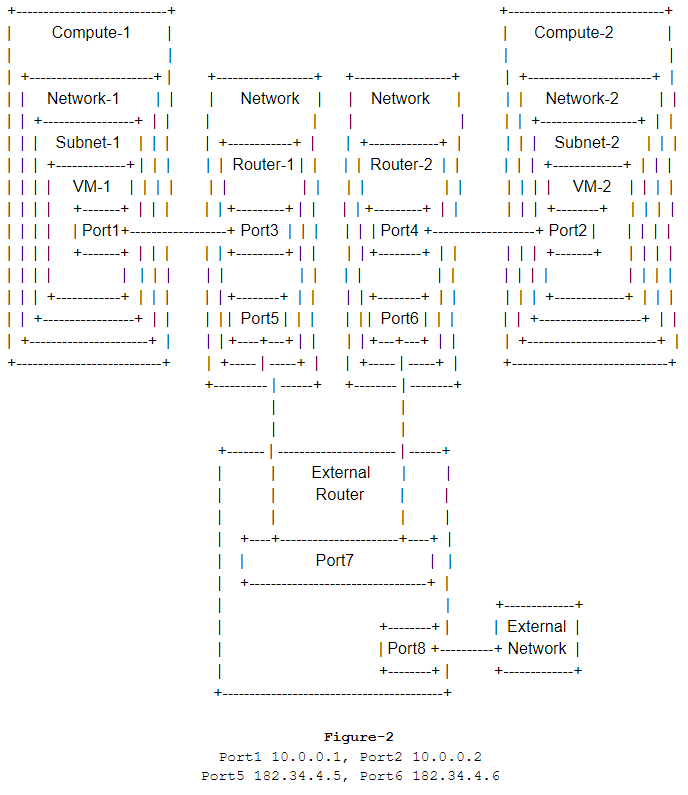
Problem Description

Figure-1 illustrates the scenarios for floating IP. The IP address “182.34.4.2” in Port3 is the IP address in floating IP resources. When the client in the external network tries to access VM1, the destination IP “182.34.4.2” is replaced by the IP address “10.0.0.3” of the Port1. Notice that the IP address of Port1 is the IP address of VM1. When VM1 send packets back to client, the source IP “10.0.0.3” will be replaced by “182.34.4.2”.



When users want to use floating IP, they should first require allocating a floating IP from floating IP pool or choose a pre allocated floating IP. Allocating floating IP will create a new port and also updates the new port. Creating port will never trigger l2pop but updating port will trigger l2pop on the condition that the port’s new status is ACTIVE or DOWN. However the status of the port for floating IP will always be N/A. Thus allocating floating IP won’t trigger l2pop.

After getting the available floating IP, users can require to associate the VM and the floating IP. Associating the IP address of the VM with the floating IP only updates the floating IP object and doesn’t update the port for floating IP. Thus associating the VM and the floating IP won’t trigger l2pop.



Based on the above analysis, the FDB of floating IP is not pre populated. In the following use case, ARP request will be sent out to query the MAC address related to the specific floating IP. Figure-2 illustrates an environment (DVR disabled) that there are two network nodes and two compute nodes.

Use Case 1

1. Tenant-1 creates Network-1 and Tenant-2 creates Network-2.
2. Tenant-1 creates Subnet-1 belonging to Network-1 and Tenant-2 creates Subnet-2 belonging to Network-2.
3. Tenant-1 creates Router-1 and links the Router-1 to the external provider network. Tenant-2 creates Router-2 and links the Router-2 to the external provider network.
4. Tenant-1 links the Router-1 to the Subnet-1 and Tenant-2 links Router-2 to the Subnet-2.
5. Tenant-1 creates VM-1 in Subnet-1 and Tenant-2 creates VM-2 in Subnet-2.
6. Tenant-1 creates FloatingIP-1 and Tenant-2 creates FloatingIP-2.
7. Tenant-1 associates VM-1 with FloatingIP-1 and Tenant2 associates VM-2 with FloatingIP-2.
8. VM-1 communicates VM-2 by floating IP. For example: VM-1 pings FloatingIP-2 for the first time.

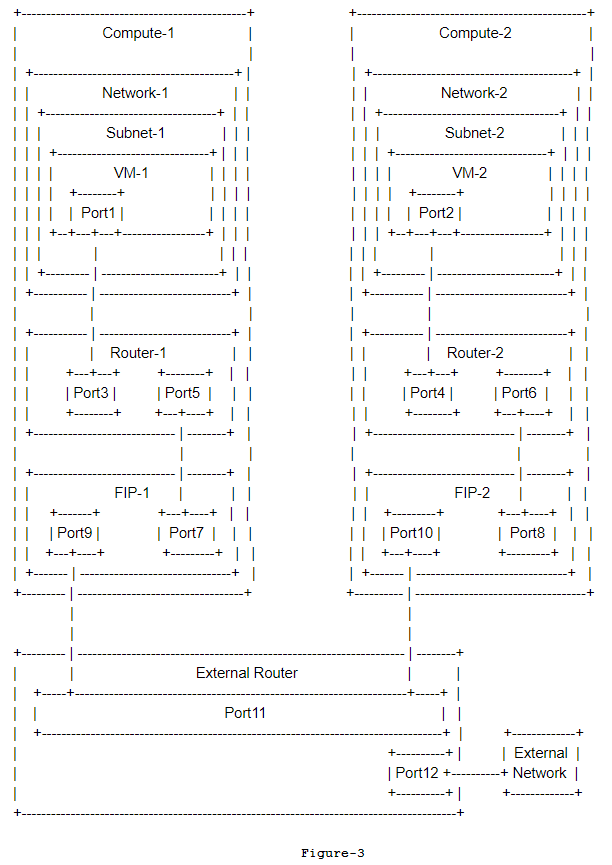
VM-1 has the IP address “10.0.0.1” which is the IP address of Port1 and VM-2 has the IP address “10.0.0.2” which is the IP address of Port2. FloatingIP-1 has the IP address “182.34.4.5” which is the IP address of Port5 and FloatingIP-2 has the IP address “182.34.4.6” which is the IP address of Port6. The MAC address of Router-1 gateway is “fa:16:3e:1b:ee:2b” and the MAC address of Router-2 gateway is “fa:16:3e:ab:cf:34”.

VM-1 and VM-2 reside on different networks and communicate via their respective floating IP addresses. When VM-1 pings FloatingIP-2 for the first time, it needs to know the MAC address for FloatingIP-2. Thus ARP request is sent out. And the MAC address “fa:16:3e:ab:cf:34” of Router-2 gateway will be answered. If the FDB is pre-populated, the ARP request can be avoided. For non-DVR use case, (host\_ip, router\_gateway\_mac\_address, floating\_ip\_address) should be pre-populated.

The figure-3 illustrates an environment (DVR enabled). The use case is listed below:

Use Case 2

1. Tenant-1 creates Network-1 and Tenant-2 creates Network-2.
2. Tenant-1 creates Subnet-1 belonging to Network-1 and Tenant-2 creates Subnet-2 belonging to Network-2.
3. Tenant-1 creates a distributed virtual router DVR-1 and Router-1 is the distributed router on compute-1. Tenant-2 creates a distributed virtual router DVR-2 and Router-2 is the distributed router on compute-2.
4. Tenant-1 links the DVR-1 to the external provider network. Tenant-2 links the DVR-2 to the external provider network.
5. Tenant-1 links the DVR-1 to the Subnet-1 and Tenant-2 links DVR-2 to the Subnet-2.
6. Tenant-1 creates VM-1 in Subnet-1 and Tenant-2 creates VM-2 in Subnet-2.
7. Tenant-1 creates FloatingIP-1 and Tenant-2 creates FloatingIP-2.
8. Tenant-1 associates VM-1 with FloatingIP-1 and Tenant2 associates VM-2 with FloatingIP-2.
9. VM-1 communicates VM-2 by floating IP. For example: VM-1 pings FloatingIP-2 for the first time.



VM-1 has the IP address “10.0.0.1” which is the IP address of Port1 and VM-2 has the IP address “10.0.0.2” which is the IP address of Port2. FloatingIP-1 has the IP address “182.34.4.5” which is the IP address of Port5 and FloatingIP-2 has the IP address “182.34.4.6” which is the IP address of Port6. When a floating IP is attached to a VM, the L3 agent will create a FIP namespace (If one does not already exist) for the external network that the FIP belongs to. After the step 8, FIP-1 and FIP-2 will be created. Port9 is FIP-1 floating IP agent gateway and Port10 is FIP-2 floating IP agent gateway. The MAC address of Port9 is “fa:16:3e:e9:87:24” and the MAC address of Port10 is “fa:16:3e:e7:86:db”.

VM-1 and VM-2 reside on different networks and communicate via their respective floating IP addresses. When VM-1 pings FloatingIP-2 for the first time, it needs to know the MAC address for FloatingIP-2. Thus ARP request is sent out. And the MAC address “fa:16:3e:e7:86:db” of FIP-2 floating IP agent gateway will be answered. If the FDB is prepopulated, the ARP request can be avoided. For DVR use case, (host\_ip, floatingip\_agent\_gateway\_mac\_address, floating\_ip\_address) should be prepopulated.

Proposed Change

The idea is that advertising the FDBs for floating IP when the FIP status changes to "ACTIVE" and withdraw the FDBs for floating IP whenever the status is set to "DOWN" or the resource is deleted or disassociated.

Function \_notify\_fip\_status will be added to send an event after updating floating IPs.  Function \_delete\_floatingip should be modified to use \_notify\_fip\_status to notify the floating IP status. Function \_update\_fip\_assoc is used by function \_create\_floatingip and \_update\_floatingip. Both functions will send events after updating floating IP with the result from \_update\_fip\_assoc. Thus we can just insert the status of floating IP into the result of \_update\_fip\_assoc. Function disassociate\_floatingips already sends event out after updating floating IP. However after disassociating floating IPs, the router\_id and fixed\_port\_id both are None. Thus need to insert floating IP status, last\_known\_router\_id and last\_fixed\_port\_id into the event.

A class L3RouterL2PopMixin should be added to process the event sent after updating floating IP. This class should subscribe the event and register the callback to l2pop mechanism driver to extend the l2pop FDBs (Depends on RFE: <https://bugs.launchpad.net/neutron/+bug/1793653>). Thus Callback l3\_fdb\_extend\_func should be added. Function handle\_fip\_status\_update should be added to process the event and send the l2pop FDBs out. For non-DVR, (host\_ip, router\_gateway\_mac\_address, floating\_ip\_address) should be prepopulated. For DVR, (host\_ip, floatingip\_agent\_gateway\_mac\_address, floating\_ip\_address) should be prepopulated. What’s more, some other functions used by the functions described above should be added.

All changes can be viewed through the link below:

<https://review.openstack.org/#/c/611261/>

https://review.openstack.org/#/c/611284/

Data Model Impact

None

REST API Impact

None

Command Line Client Impact

None

Other Impact

None

Other Deployer Impact

None

Performance Impact

Performance testing should be conducted to see test the overhead of adding more information to FDB.

Implementation

Assignee(s)

Work Items

* Add function \_notify\_fip\_status to neutron/db/l3\_db.py and modify some existing functions in l3\_db.py.
* Add class L3RouterL2PopMixin to neutron/services/l3\_router/service\_providers/l2pop.py
* Add related tests.

Dependencies

Need the RFE “Enable other subprojects to extend l2pop FDB information” merged in OpenStack. This RFE enables other subprojects to extend l2pop FDB information. We need to add floating IP related information to l2pop FDB. Thus we need to use the ability provided by the RFE. The link for RFE is below:

https://bugs.launchpad.net/neutron/+bug/1793653

Testing

Unit tests are necessary.

Documentation Impact

None.

References

**[1]** <https://github.com/openstack/neutron/tree/master/neutron/plugins/ml2/drivers/l2pop>

**[2]** <https://wiki.openstack.org/wiki/L2population_blueprint>