SR-IOV with VLAN and Avi Vantage (OpenStack No-Access) Integration in DPDK

Avi Technical Reference (v20.1)

SR-IOV with VLAN and Avi Vantage (OpenStack No-Access) Integration in DPDK

view online

Overview

This guide explains the configuration aspects of Avi Vantage in OpenStack no-Access mode (DPDK for the SE) and using SR-IOV and VLAN from OpenStack for better performance.

Note: This feature is supported starting Avi Vantage version 18.2.5.

SR-IOV specification defines a standardized mechanism to virtualize PCIe devices. This mechanism can virtualize a single PCIe Ethernet Controller to appear as multiple PCIe devices. Each device can be directly assigned to an instance, bypassing the hypervisor and virtual switch layer. As a result, you can achieve low latency and near-line wire speed.

PF-Physical Function: The physical Ethernet Controller that supports SR-IOV.

VF-Virtual Function: The virtual PCIe device created from a physical Ethernet Controller.

This guide outlines steps for `ixgbe-vf` driver that supports the following NICs: * 82599 * X520 * X540 * X550 * X552

The following are the limitations in OpenStack:

- For creating SR-IOV ports, Horizon should not be used.
- SR-IOV is not integrated into the OpenStack Dashboard (Horizon). You should use the CLI or API to configure SR-IOV interfaces.
- Attaching SR-IOV ports to existing servers is not currently supported. In this case, the Avi Vantage cloud type is OpenStack No-Access.

Enabling SR-IOV

The specific steps for equivalent OpenStack nodes are mentioned below:

1. Create Virtual Functions (Compute)
2. Allowlist PCI devices in nova-compute (Compute)
3. Configure neutron-server (Controller)
4. Configure nova-scheduler (Controller)
5. Enable neutron sriov-agent (Compute)

Creating Virtual Functions (Compute)

The following are the steps to create virtual functions:

1. Ensure that SR-IOV and VT-d are enabled in BIOS.
2. Enable IOMMU in Linux by adding `intel_iommu=on` to the kernel parameters, for instance, using GRUB.
3. On each compute node, create the VFs via the PCI SYS interface:

```
echo '8' > /sys/class/net/eth3/device/sriov_numvfs
```

4. A network interface can be used both for PCI passthrough, using the PF, and SR-IOV, using the VFs. If the PF is used, the VF number stored in the `sriov_numvfs` file is lost. If the PF is attached again to the operating system, the number of VFs assigned to this interface will be zero. To keep the number of VFs always assigned to this interface, modify the interfaces configuration file by adding an `ifup` script command.

5. The maximum number of VFs a PF can support:

```
cat /sys/class/net/eth3/device/sriov_totalvfs63
```

6. In Ubuntu, modify `/etc/network/interfaces` file as follows:

```
auto eth3
iface eth3 inet dhcp
pre-up echo '4' > /sys/class/net/eth3/device/sriov_numvfs
```

7. In Red Hat, modify `/sbin/ifup-local` file as follows:

```
#!/bin/sh
if [[ "$1" == "eth3" ]]
then
echo '4' > /sys/class/net/eth3/device/sriov_numvfs
fi
```

8. Verify if the VFs are created:

```
root@sriov:~# lspci | grep "Ethernet"
03:10.1 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:10.3 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:10.5 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:10.7 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:11.1 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:11.3 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:11.5 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
03:11.7 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
```

9. Persist created VFs on reboot:

```
echo "echo '7' > /sys/class/net/eth3/device/sriov_numvfs" >> /etc/rc.local
```

Allowlisting PCI Devices nova-compute (Compute)

Make the change in `/etc/nova/nova.conf` file as follows:

```
pci_passthrough_whitelist = { "devname": "ens1f1", "physical_network": "provider"}
```

This informs the compute service that all VFs belonging to `ens1f1` are allowed to be passed through to instances and belong to the network provider.

Configuring neutron-server (Controller)

Add `sriovnicswitch` as mechanism driver.

Make changes in `/etc/neutron/plugins/ml2/ml2_conf.ini` files. * mechanism_drivers = openvswitch, sriovnicswitch * restart neutron-server service

Configuring nova-scheduler (Controller)

Make changes under default section in `/etc/nova/nova.conf` file.

```
scheduler_default_filters = RetryFilter, AvailabilityZoneFilter, RamFilter, ComputeFilter, ComputeCapabilitiesFilter, I
scheduler_available_filters = nova.scheduler.filters.all_filters
```

Enabling Neutron sriov-agent (Compute)

Make changes in `/etc/neutron/plugins/ml2/sriov_agent.ini` file.

```
[sriov_nic]
physical_device_mappings = provider:ens1f1
exclude_devices = [securitygroup]
firewall_driver = neutron.agent.firewall.NoopFirewallDriver
```

Launching Instances with SR-IOV Ports

1. Get network ID to create SR-IOV (in this case there is a network provider1 which is already created and is used for the SE's management as well as data).

   ```
   net_id=neutron net-show provider1 | grep "\ id\ " | awk '{ print $4 }'
   ```

2. Create the SR-IOV port. `vnic_type=direct` is used here.

   ```
   port_id=neutron port-create $net_id --name sriov_port_1 --binding:vnic_type direct |
   ```

```
port_id=neutron port-create $net_id --name sriov_port_2 --binding:vnic_type direct | grep
" id " | awk '{ print $4 }'
```

3. Port-Show of the ports indicates that the respective ports have been created as SR-IOV ports:

   ```
   root@sriov:~# neutron port-show 3d6ef025-77ae-4136-a8b6-c48955fe5d2f
   ```

   ```
   +-----------------------+--------------------------------------------------------------------------------+
   | Field                 | Value                                                                          |
   ```

```
+-------------------------+--------------------------------------------------------------------------------+
| admin_state_up          | True                                                                           |
| allowed_address_pairs   |                                                                                |
| binding:host_id         | sriov                                                                          |
| binding:profile         | {"pci_slot": "0000:03:10.7", "physical_network": "provider",
| binding:vif_details     | {"port_filter": false, "vlan": "0"}                                            |
| binding:vif_type        | hw_veb                                                                         |
| binding:vnic_type       | direct                                                                         |
| created_at              | 2019-04-05T04:32:15Z                                                           |
| description             |                                                                                |
| device_id               | 6cf94dd4-c6f3-4d49-84bd-639f40ed1b5e                                           |
| device_owner            | compute:nova                                                                   |
| extra_dhcp_opts         |                                                                                |
| fixed_ips               | {"subnet_id": "03e68028-ecec-4fb4-9c03-546ed14bf3c4", "ip_address":
| id                      | 3d6ef025-77ae-4136-a8b6-c48955fe5d2f                                           |
| mac_address             | fa:16:3e:af:d1:f6                                                              |
| name                    | sriov_port_2                                                                   |
| network_id              | c4260d3e-f275-4097-96be-03751495f291                                           |
| port_security_enabled   | True                                                                           |
| project_id              | dbe81cf9baa8492288456cbb295a529e                                               |
| revision_number         | 18                                                                             |
| security_groups         | 643c4bb8-7236-47ec-a91d-9038be4774cb                                           |
| status                  | ACTIVE                                                                         |
| tags                    |                                                                                |
| tenant_id               | dbe81cf9baa8492288456cbb295a529e                                               |
| updated_at              | 2019-04-05T04:39:02Z                                                           |
+-------------------------+--------------------------------------------------------------------------------+
```

4. `root@sriov:~# neutron port-show a9fe3f1a-2e1c-4c3f-9fa0-1c03ab29d2c0`

```
+-----------------------------+--------------------------------------------------------------------------------+
| Field                       | Value                                                                          |
+-----------------------------+--------------------------------------------------------------------------------+
| admin_state_up              | True                                                                           |
| allowed_address_pairs       |                                                                                |
| binding:host_id             | sriov                                                                          |
| binding:profile             | {"pci_slot": "0000:03:11.1", "physical_network": "provider",
| binding:vif_details         | {"port_filter": false, "vlan": "0"}                                            |
| binding:vif_type            | hw_veb                                                                         |
| binding:vnic_type           | direct                                                                         |
| created_at                  | 2019-04-05T04:32:06Z                                                           |
| description                 |                                                                                |
| device_id                   | 6cf94dd4-c6f3-4d49-84bd-639f40ed1b5e                                           |
| device_owner                | compute:nova                                                                   |
| extra_dhcp_opts             |                                                                                |
| fixed_ips                   | {"subnet_id": "03e68028-ecec-4fb4-9c03-546ed14bf3c4", "ip_address":
| id                          | a9fe3f1a-2e1c-4c3f-9fa0-1c03ab29d2c0                                           |
| mac_address                 | fa:16:3e:db:61:0a                                                              |
| name                        | sriov_port_1                                                                   |
| network_id                  | c4260d3e-f275-4097-96be-03751495f291                                           |
```

```
| port_security_enabled         | True                                                                          |
| project_id                    | dbe81cf9baa8492288456cbb295a529e                                              |
| revision_number               | 19                                                                            |
| security_groups               | 643c4bb8-7236-47ec-a91d-9038be4774cb                                          |
| status                        | ACTIVE                                                                        |
| tags                          |                                                                               |
| tenant_id                     | dbe81cf9baa8492288456cbb295a529e                                              |
| updated_at                    | 2019-04-05T04:39:02Z                                                          |
+-------------------------------+-------------------------------------------------------------------------------+
```

5. An OpenStack Controller is brought up in No-Access mode and se.qcow2 image is pushed to glance (As illustrated in Step 3 of Installing Avi Vantage into No Access OpenStack Cloud)

6. Create an instance with two NIC's (Data and Management) using se.qcow2 image (As illustrated in Step 4 of Installing Avi Vantage into No Access OpenStack Cloud) as follows:

```
openstack server create --flavor m1.se --image AVi-se-18.2.2-9224 --port a9fe3f1a-2e1c-4c3f-9fa0-1c03ab29d2c0 -
```

7. Run `/opt/avi/init_system.py` script and make sure that the SE can connect to the Controller. (As illustrated in Step 15 of Installing Avi Vantage into No Access OpenStack Cloud)
8. Ensure that the SE data vNIC?s (login to the SE) have come up as SR-IOV VF?s and not as VIRTIO interfaces.

```
root@10-140-81-213:~# lspci | grep "Ethernet"
00:04.0 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
00:05.0 Ethernet controller: Intel Corporation X540 Ethernet Controller Virtual Function (rev 01)
```

9. For the SE data vNIC ports to run DPDK on top of the SR-IOV ports, login to the Controller shell (CLI Access) and under SE-group properties (this will affect all SE's under that SE group); make the change for the NIC's to be up in DPDK mode by making the following changes:

```
[admin:avi-ctrlr]: serviceenginegroup> se_use_dpdk
[admin:avi-ctrlr]: serviceenginegroup> se_dpdk_pmd
```

Note: The SE's need to be rebooted for the change to take effect.