목차

PCI Passthrough / SR-IOV 3
PCI Passthrough 3
Openstack SR-IOV 3
알려진 제한사항 3
Network Interface별 사전 준비 3
Mellanox CONNECTX-4/CONNECTX-5/CONNECTX-6 SR-IOV 4
Intel X550 SR-IOV 4
SR-IOV 구성 4
kolla-ansible 설정 4
globals.yml 설정 4
neutron 설정 4
nova 설정 5

PCI Passthrough / SR-IOV

PCI Passthrough

nova.conf내에 아래 내용을 추가한다. 여러개의 장치를 Passthrough할 경우에는 []로 묶어서 목록 형태로 나열한다.

```
[filter_scheduler]
enabled_filters = AvailabilityZoneFilter, ComputeFilter,
ComputeCapabilitiesFilter, ImagePropertiesFilter,
ServerGroupAntiAffinityFilter, ServerGroupAffinityFilter,
PciPassthroughFilter
available_filters = nova.scheduler.filters.all_filters

[pci]
passthrough_whitelist = [{ "devname": "ens6", "physical_network": "physnet2"
}, { "vendor_id": "10de", "product_id": "13c2" }]
alias = { "vendor_id":"10de", "product_id":"13c2", "device_type":"type-PF",
"name":"GTX970" }
```

Openstack SR-IOV

- https://docs.openstack.org/neutron/latest/admin/config-sriov.html

알려진 제한사항

QoS(Quality of Service) 사용 시 max_burst_kbps(버스트 오버 max_kbps)는 지원되지 않습니다. 또한 max_kbps는 Mbps로 반올림됩니다.

SR-IOV를 사용하는 경우 **보안 그룹**이 지원되지 않으므로 **방화벽 드라이버를 비활성화**해야 합니다. 이 작업은 neutron.conf파일에서 수행할 수 있습니다 .

```
[securitygroup]
firewall_driver = noop
```

SR-IOV는 OpenStack Dashboard(horizon)에 통합되지 않습니다. 사용자는 CLI 또는 API를 사용하여 SR-IOV 인터페이스를 구성해야 합니다.

Network Interface별 사전 준비

Mellanox CONNECTX-4/CONNECTX-5/CONNECTX-6 SR-IOV

- https://enterprise-support.nvidia.com/s/article/HowTo-Configure-SR-IOV-for-ConnectX-4-ConnectX-5-ConnectX-6-with-KVM-Ethernet

Intel X550 SR-IOV

- https://forum.proxmox.com/threads/enabling-sr-iov-for-intel-nic-x550-t2-on-proxmox-6.56677/

SR-IOV 구성

1. 커널 파라메터 설정 `IOMMU`

```
grubby --update-kernel=ALL --args="intel_iommu=on iommu=pt"
grubby --update-kernel=ALL --args="amd_iommu=on iommu=pt"

dracut -f -v
```

2. SR-IOV device 설정

```
echo "echo '8' > /sys/class/net/ens6/device/sriov_numvfs" >> /etc/rc.local
```

kolla-ansible 설정

globals.yml 설정

`/etc/kolla/globals.yml`

```
...
enable_neutron_sriov: "yes"
...
```

neutron 설정

`/etc/kolla/config/neutron/tp-compute3/sriov_agent.ini` `tp-compute3` 노드의 장치를 `SR-IOV` 할당하려고 할때

```
[sriov_nic]
physical_device_mappings = sriovnet1:ens6
exclude_devices =
```

```
[securitygroup]
firewall_driver = neutron.agent.firewall.NoopFirewallDriver
```

/etc/kolla/config/neutron/ml2_conf.ini

```
[ml2]
type_drivers = flat,vlan,vxlan,geneve
tenant_network_types = vxlan
mechanism_drivers = ovn,openvswitch,l2population,sriovnicswitch
extension_drivers = port_security

[ml2_type_vlan]
network_vlan_ranges = physnet1,sriovnet1

[ml2_type_flat]
flat_networks = physnet1,sriovnet1

[ml2_type_vxlan]
vni_ranges = 1:1000
```

nova 설정

/etc/kolla/config/nova.conf에 아래 설정 추가

- GPU Passthrough

```
[pci]
passthrough_whitelist = { "vendor_id": "10de", "product_id":"20f1" }
```

- SR-IOV

```
[pci]
passthrough_whitelist = { "devname": "ens6", "physical_network": "physnet2"
}
```

이렇게 하면 Compute 서비스에 ens6에 속한 모든 VF가 인스턴스로 전달될 수 있으며 공급자 네트워크 physnet2에 속한다고 알려줍니다.

다른방법으로 아래처럼 id를 직접 기입해도 됨, 반드시 PF장치의 값을 넣어야 한다. VF장치값을 넣으면 안됨.

```
passthrough_whitelist = { "vendor_id": "8086", "product_id":"1563",
"physical_network": "physnet2" }
```

여러 장치를 Passthrough 할경우엔 아래와같이 적용

```
[pci]
passthrough_whitelist = [{ "devname": "ens6", "physical_network": "physnet2"
}, { "vendor_id": "10de", "product_id": "13c2" }]
alias = { "vendor_id":"10de", "product_id":"13c2", "device_type":"type-PF",
"name":"GTX970" }
```

3. Launching instances with SR-IOV ports

```
openstack network create --provider-physical-network sriovnet1 \
    --provider-network-type vlan --provider-segment 319 \
    sriov-net

openstack subnet pool create --share --pool-prefix 10.33.19.0/24 \
--default-prefix-length 24 shared-default-subnetpool-v4

openstack subnet create --network sriov-net \
    --subnet-pool shared-default-subnetpool-v4 \
    sriov-subnet

net_id=$(openstack network show sriov-net -c id -f value)

openstack port create --network $net_id --vnic-type direct sriov-port

port_id=$(openstack port show sriov-port -c id -f value)

openstack server create --availability-zone sriov --flavor c1_r2_d30 --image
rhel8 \
    --nic port-id=$port_id test-sriov
```

From:
https://atl.kr/dokuwiki/ - **AllThatLinux!**

Permanent link:
https://atl.kr/dokuwiki/doku.php/pci_passthrough_sr-iov

Last update: **2023/12/04 07:13**