

Known Behavior

- « Testing and D... (tungsten-fabric-operator/testing-and-debugging-tf-operator.html)

Known Behavior

This section lists known limitations with this release. Known Behavior in Tungsten Fabric Release 2011

- In rare cases, sFlow node provisioning fails while initializing kafka container. If this scenario happens during provisioning, redeploying will bring up the sFlow nodes.
- When creating any new user-defined namespace on Openshift-4.x/Contrail, by default SNAT is enabled and so all the pods part of this namespace by default can reach internet servers. As a workaround, explicitly configure the Contrail annotations on the namespace as "opencontrail.org/ip_fabric_snat": "false".
- Telemetry KPI display for Junos EVO devices are not supported.
- The BGP routes widget under **Fabrics > Ports > ``Leaf network device``** account for routes from inet.0 table only.
- Gatewayless forwarding feature is not supported on Netronome.
- While upgrading Tungsten Fabric Release 19xx with RHOSP13 to Tungsten Fabric Release 2011 with RHOSP16.1, FFU upgrade of compute node fails with the error— 'RPC failed at server. Insufficient access: Insufficient 'add' privilege to add the entry 'krbprincipalname=qemu/compute-0-ffu.internalapi.nodel8.local@NODEL8.LOCAL,cn=services,cn=accounts,dc=nodel8,dc=local' .

Perform the following steps to apply the patch:

- On undercloud:

1. Apply the patch as given at <https://review.opendev.org/c/openstack/tripleo-heat-templates/+764064/3/deployment/nova/novajoin-container-puppet.yaml> (<https://review.opendev.org/c/openstack/tripleo-heat-templates/+764064/3/deployment/nova/novajoin-container-puppet.yaml>).

2. Deploy undercloud by running `openstack undercloud install` command.

3. Ensure the novajoin_notifier container is up and not constantly restarting.

```
podman ps |grep novajoin
```

Validate the status by checking `/var/log/containers/novajoin/*` logs.

TLS-E can not work properly without the novajoin containers running.

4. Ensure that the transport_url for novajoin does not contain guest user.

```
grep ^transport_url /var/lib/config-data/puppet-generated/novajoin/etc/novajoin/join.conf
```

- On compute (kernel/dpdk) node:

1. Log in to knit admin.

```
heat-admin@overcloud-contraildpdk-0-ffu ~]$ kinit admin
```

Use FreeIPA password for admin@NODEL8.LOCAL (mailto:admin%40NODEL8.LOCAL).

2. Add DNS record of compute node.

```
[heat-admin@overcloud-contraildpdk-0-ffu ~]$ ipa dnsrecord-add
```

```
Record name: overcloud-contraildpdk-0-ffu
Zone name: internalapi.nodel8.local
Please choose a type of DNS resource record to be added
The most common types for this type of zone are: A, AAAA
DNS resource record type: a
A IP Address: 192.168.xx.18
Record name: overcloud-contraildpdk-0-ffu
A record: 192.168.xx.18 #( IP address of compute node from open-
stack server list)
```

3. Add IPA service.

```
[heat-admin@overcloud-contraildpdk-0-ffu ~]$ ipa service-add qemu/overcloud-
contraildpdk-0-ffu.internalapi.nodel8.local@NODEL8.LOCAL
```

```
Principal name: qemu/overcloud-contraildpdk-0-ffu.internalapi.nodel8.local@NODEL
8.LOCAL
Principal alias: qemu/overcloud-contraildpdk-0-ffu.internalapi.nodel8.local@NODE
L8.LOCAL
Managed by: overcloud-contraildpdk-0-ffu.internalapi.nodel8.local
```

4. Add host to the IPA service.

```
[heat-admin@overcloud-contraildpdk-0-ffu ~]$ ipa service-add-host --hosts
overcloud-contraildpdk-0-ffu.nodel8.local qemu/overcloud-contraildpdk-0-
ffu.internalapi.nodel8.local@NODEL8.LOCAL
```

```
Principal name: qemu/overcloud-contraildpdk-0-ffu.internalapi.node18.local@NODEL8.LOCAL
Principal alias: qemu/overcloud-contraildpdk-0-ffu.internalapi.node18.local@NODEL8.LOCAL
Managed by: overcloud-contraildpdk-0-ffu.internalapi.node18.local, overcloud-contraildpdk-0-ffu.node18.local
Number of members added 1
```

- In Tungsten Fabric, the logical router (LR) does not support dynamic next-hop port-mirroring, when Juniper headers are enabled. The Juniper header is not supported in port-mirroring as VXLAN is not the tunnel type used for the dynamic next hop in this case.
- BFD session takes longer time to come up after the agent is restarted when MAC/IP is enabled in the VN and associated with BFD health check with target-ip set to "all".
- After one of HA Master nodes failover (or) vrouter restart, further user PODs creation might fail without getting IP-address.

As a workaround, find HA master nodes control-pods which are not in sync with respect to "new user-pod" and to restart them. Perform the following steps:

1. Log in to 3 HA masters and find the crictl pod with name "control" and log in to it verify the command output of "curl -cert /etc/certificates/server-key-localhost -insecure https://localhost:8083/Snh_IFMapTableShowReq?table_name=virtual-machine (https://localhost:8083/Snh_IFMapTableShowReq?table_name=virtual-machine)" showing the name of the latest user pod which failed.
 2. Restart those control PODs which are not in sync.
- In a Kubernetes and OpenStack joint setup, vrouter-agent restart sometimes leads to an unauthorized operation error. To resolve the issue, restart the vrouter-agent again.
 - In L2 DCI mode, if selected fabrics have the same overlay ASN numbers, overlay iBGP is used between fabric devices for L2 DCI mode. In this case, the border device (physical router) marked with DCI-gateway RB role (routing and bridging) must also have RR (route reflector) RB role assigned. Without RR RB role, Overlay iBGP session won't stretch Layer 2 tenant virtual network across the fabric's leaf devices. So, we recommend that for L2 DCI Mode, ensure the physical router device is marked as DCI Gateway RB role along with RR role.
 - While upgrading Tungsten Fabric Release 19xx with RHOSP13 to Tungsten Fabric Release 2011 with RHOSP16.1, overcloud nodes transition to ERROR state after upgrading the undercloud. As a workaround, apply a patch as mentioned at https://bugzilla.redhat.com/show_bug.cgi?id=1850929 (https://bugzilla.redhat.com/show_bug.cgi?id=1850929).
 - Security Groups cannot be used on QFX10K interfaces.
 - During deployment we see race condition, due to which ipa-client installation on compute nodes fails. This is an issue with Red Hat. As a workaround, before deployment starts, modify the following file to add sleep of 400 seconds on undercloud.

```
sudo vi /usr/share/ansible/roles/tripleo-kernel/tasks/kernelargs.yml - name: DBG
debug:
  msg: "sleep 400 sec if reboot_required == {{ reboot_required }}"
- name: DBG sleep
  shell: sleep 400
  when:
    - reboot_required is defined and reboot_required# then
sudo find / -name kernelargs.yml
# to find all such files on undercloud and in containers because I am not sure which exactly is used (from host or from container)
# and overwriting such files in containers like
sudo cp /usr/share/ansible/roles/tripleo-kernel/tasks/kernelargs.yml /var/lib/containers/storage/overlay/6dc6b96b1392e5302b63156fa093525e17131bef1203cad005a911ad09241f5a/diff/usr/share/ansible/roles/tripleo-kernel/tasks/kernelargs.yml
```

- The vRouter to vRouter encryption feature is beta quality and should be used for future product capability demonstrations only.
- On DPDK compute, memory of the VMs are mapped to only one numa. VM creation fails after the hugepages in that numa are exhausted if it is launched with hw:mem_page_size='any' flavor. As a workaround, use the hw:mem_page_size='large' flavor instead to avoid the issue.
- In case of RHOSP16 deployment with TLS, XMPP connection down is seen post deployment completion. While this is a cosmetic issue and does not impact functionality, as a workaround, restart the vRouter agent container on all compute nodes to update status.
- In DPDK1911 with X710 NIC performance degrades due to mbuf leak if tx and rx are configured. Intel recommends configuring atleast 1K tx and rx descriptors on Fortville NICs for better and consistent performance, but they seem to have a degrading effect on X710 NIC.
- Tungsten Fabric WebUI doesn't work for System/Node status monitoring. As a workaround, check using CLI on the relevant nodes. This will not impact functionality.
- QFX5120 cannot be used as border leaf role in SP style for CRB role.
- On a DPDK compute, if contrail-vrouter-agent crashes or if contrail-vrouter-agent is restarted in a scaled setup with many sub-interfaces, all the sub-interfaces and their parent interface may become inactive. As a workaround, stop / start the instances whose interfaces are down.
- In an OpenStack HA setup provisioned using Kolla and OpenStack Rocky, if you shut down all the servers at the same time and bring them up later, the Galera cluster fails. To recover the Galera cluster, follow these steps:

1. Edit the `/etc/kolla/mariadb/galera.cnf` file to remove the `wsrep` address on one of the controllers as shown here.

```
wsrep_cluster_address = gcomm://  
#wsrep_cluster_address = gcomm://10.x.x.8:4567,10.x.x.10:4567,10.x.x.11:4567
```

Note

If all the controllers are shut down in the managed scenario at the same time, you must select the controller that was shut down last.

2. Docker start mariadb on the controller on which you edited the file.
 3. Wait for a couple of minutes, ensure that the mariadb container is not restarting, and then Docker start mariadb on the remaining controllers.
 4. Restore the `/etc/kolla/mariadb/galera.cnf` file changes and restart the mariadb container on the previously selected controller.
- VLAN tag does not work with Mellanox CX5 cards with DPDK 19.11.
 - Monitoring/Operations page crashes with “Cannot read property ‘className’ of undefined”. As a workaround, refresh the page to display the content properly.
 - In case of BMS to BMS EVPN “Transparent” service chaining, Tunneled packet sent out of Transparent service instance to QFX have vlan-id and hence Traffic from left-bms to right-bms gets dropped since the inner header of the tunneled packet has vlan-id info which is internal to vRouter and QFX is not aware of the vlan-id so the packet gets dropped by the switch.
 - Under Security Groups, the entry appearing with `__no_rule__` can be ignored.
 - IPv6 ipam subnet option “enable_dhcp” is always ignored.
 - Updating VLAN-ID on a VPG in an enterprise style fabric is not supported. As a workaround, delete and recreate the fabric.
 - In Octavia Load Balancer, traffic destined to the Floating IP of the load balancer VM does not get directed to the backend VMs. Traffic destined to the actual VM IP of the Load Balancer VM will work fine.
 - vRouter offload with Mellanox NIC cards does not work. However the DPDK on Mellanox NICs without offload is supported.
 - In release 2003, the Virtual Port Group create workflow will not pre-populate the VLAN-ID with the existing value that was defined with the first VPG for a given virtual network. The field is editable unlike in previous releases. This issue occurs in a fabric that was provisioned with the **Fabric-wide VLAN-ID significance** checkbox enabled.
 - DPDK vRouter with MLNX CX5 takes about 10 minutes and also lcore crash is seen. This happens once during initial installation.
 - AppFormix Flows does not show up for multi homed devices on the fabric
 - In Fortville X710 NIC: With TX and RX buffers performance degrade is observed as mbufs gets exhausted.
 - During upgrade of DPDK computes deployed with OOO Heat Templates in RHOSP environment, vRouter core-dumps are observed. This is due to the sequence in which the services are started during upgrade and does not have impact on cluster operation.
 - Onboarding of multiple BMS in parallel on SP-style fabric does not work. While bringing up a BMS using the Life Cycle Management workflow, sometimes on faster servers the re-image does not go through and instance not moved from ironic vn to tenant vn. This is because if the PXE boot request from the BMS is sent before the routes are converged between the BMS port and the TFTP service running in Contrail nodes. As a workaround, the servers can be rebooted or the BIOS in the servers can be configured to have a delayed boot.
 - BMS LCM with fabric set with `enterprise_style=True` is not supported. By default, `enterprise_style` is set to `False`. Avoid using `enterprise_style=True` if the fabric object onboards the BMS LCM instance.
 - For deleting compute nodes, the UI workflow will not work. Instead, update the `instances.yaml` with “ENABLE_DESTROY: True” and “roles:” (leave it empty) and run the following playbooks.

```
ansible-playbook -i inventory/ -e orchestrator=openstack --tags nova playbooks/install_openstack.yml  
ansible-playbook -i inventory/ -e orchestrator=openstack playbooks/install_contrail.yml
```

For example:

```
global_configuration:  
  ENABLE_DESTROY: True  
...  
...  
instances:  
...  
...  
  srvr5:  
    provider: bms  
    ip: 19x.xxx.x.55  
    roles:  
...  
...
```

- VNI update on a LR doesn't update the RouteTable. As a workaround, delete the LogicalRouter and create a new LogicalRouter with the new VNI.
- Additional links cannot be appended to service templates used to create PNF service chaining. If there is a need to add additional links, the service template needs to be deleted and re-added again.
- BMS movement across TORs is not supported. To move BMS across TORs the whole VPG needs to be moved. That means if there are more than one BMS associated to one VPG, and one of the BMS need to be moved, the whole VPG need to be deleted and re-configured as per the new association.
- Multicast traffic originated from type-6 incapable QFX devices are duplicated by vRouters.
- JCB-187287 High Availability provisioning of Kubernetes master is not supported.
- JCB-184776 When the vRouter receives the head fragment of an ICMPv6 packet, the head fragment is immediately enqueued to the assembler. The flow is created as hold flow and then trapped to the agent. If fragments corresponding to this head fragment are already in the assembler or if new fragments arrive immediately after the head fragment, the assembler releases them to flow module. Fragments get enqueued in the hold queue if agent does not write flow action by the time the assembler releases fragments to the flow module. A maximum of three fragments are enqueued in the hold queue at a time. The remaining fragments are dropped from the assembler to the flow module.

As a workaround, the head fragment is enqueued to assembler only after flow action is written by agent. If the flow is already present in non-hold state, it is immediately enqueued to assembler.

- JCB-177787 In DPDK vRouter use cases such as SNAT and LBaaS that require netns, jumbo MTU cannot be set. Maximum MTU allowed: <=1500.