Support of OVS-DPDK in Airship

**Project Technical Lead:** Georg Kunz. Elected 1/17/19.

**Project Committer Details:**

Initial Committers for a project will be specified at project creation. Committers have the right to commit code to the source code management system for that project.

A Contributor may be promoted to a Committer by the project’s Committers after demonstrating a history of contributions to that project.

Candidates for the project’s Project Technical Leader will be derived from the Committers of the Project. Candidates must self nominate by marking “Y” in the Self Nominate column below by Jan. 16th. Voting will take place January 17th.

Only Committers for a project are eligible to vote for a project’s Project Technical Lead.

Please see Akraino Technical Community Document section 3.1.3 for more detailed information.

<table>
<thead>
<tr>
<th>Committer</th>
<th>Company</th>
<th>Contact Info</th>
<th>Committer Bio</th>
<th>Committer Picture</th>
<th>Self Nominate for PTL (Y/N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew Wilkinson</td>
<td>Ericsson</td>
<td><a href="mailto:andrew.wilkinson@ericsson.com">andrew.wilkinson@ericsson.com</a></td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Georg Kunz</td>
<td>Ericsson</td>
<td><a href="mailto:georg.kunz@ericsson.com">georg.kunz@ericsson.com</a></td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>Cheng Li</td>
<td>Intel</td>
<td><a href="mailto:cheng1.li@intel.com">cheng1.li@intel.com</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rihab Banday</td>
<td>Ericsson</td>
<td><a href="mailto:rhab.banday@ericsson.com">rhab.banday@ericsson.com</a></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Use Case Details:**

Feature project proposers: Ericsson, Nokia and Radysis (confirm) and (others - confirm).

The following BP proposals require support of OVS-DPDK in Airship:

- Network Cloud : OVS-DPDK Unicycle Dell Blueprint Proposal
- Radio Edge Cloud
- Edge Video Processing
- (any other dependent BPs?)

**Initial list of high-level working items**

The table below provides a list of high-level work items required to enable support for OVS-DPDK in Airship. This list is not considered to be complete but a starting point for design discussions in the feature project. Please feel free to add / modify / extend work items.

<table>
<thead>
<tr>
<th>Task name</th>
<th>Description</th>
<th>Airship component</th>
<th>Implementation</th>
<th>Upstream reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>create openvswitch agent chart</td>
<td>create helm chart for openvswitch agent</td>
<td>openstack-helm</td>
<td>in place and being deployed when OVS is enabled</td>
<td></td>
</tr>
<tr>
<td>create ovs-dpdk chart</td>
<td>create helm chart for openvswitch dpdk container</td>
<td>openstack-helm</td>
<td>Extend existing openvswitch chart with config parameters for DPDK.</td>
<td><a href="https://review.openstack.org/#/c/826894/">https://review.openstack.org/#/c/826894/</a></td>
</tr>
<tr>
<td>deploy neutron openvswitch agent</td>
<td>ensure chart of openvswitch agent is deployed</td>
<td>treasuremap</td>
<td>In place and being deployed when OVS is enabled</td>
<td></td>
</tr>
</tbody>
</table>
### DPDK host config: enable 1G hugepages
- Modify kernel cmdline to enable 1G hugepages (hugepages=1G hugepages=8 iommu=pt intel_iommu=on)
- **Drydock**
  - Already available:
    - Define the number of available hugepages in the node's HardwareProfile:
      https://github.com/openstack/airship-drydock/blob/master/python/tests/yaml_samples/deckhand_fullsite.yaml#L427
    - Define the kernel paramters for enabling hugepages in the node's BaremetalNode configuration (kernel_param section):
      https://github.com/openstack/airship-drydock/blob/master/python/tests/yaml_samples/deckhand_fullsite.yaml#L366
- **Storyboard**: https://storyboard.openstack.org/#!/story/2004790

### DPDK host config: mount hugepages
- Mount hugepages into local file system on system boot (hardcoded mount point)
- **Divingbell**
  - Two alternatives available:
    - Utilize auto-mount capabilities of Ubuntu
      https://www.apt-browse.org/browse/ubuntu/xenial/main/i386/systemd/229-4ubuntu4/file
      /lib/systemd/system/dev-hugepages.mount
    - Issue: does not allow fine-grained control of mount-options (e.g. specify the size of the hugepages if there are multiple available)
    - Deploy Divingbell daemons on compute nodes with a given chart configuration (values.yaml).
- **Storyboard**: https://storyboard.openstack.org/#!/story/2004790

### DPDK host config: make hugepage mount point a config option
- Make mount point config option for e.g. use by helm charts
- **Openstack-helm**
  - Config option of helm chart (see patchset). Need to figure out how to pass overrides to chart.
    - https://review.openstack.org/#/c/626894/

### Specify PCI IDs of NICs for use by DPDK
- Specify in site config which PCI IDs (NICs) should be decided to DPDK
- **Openstack-helm**
  - Add parameters to OVS helm chart
    - https://review.openstack.org/#/c/626894/

### DPDK host config: install DPDK kernel modules and tools on host OS
- Either install host OS dpdk package or build from source
- **Drydock or divingbell**
  - Depends on DPDK driver we want to use: igb_uio, uio_generic_pci or vfio-pci. Only for igb_uio a custom build kernel module is needed.
    - https://review.openstack.org/#/c/626894/

### DPDK host config: load DPDK kernel module in host OS
- Load dpdk kernel modules uio and igb_uio during host boot-up
- **Drydock or divingbell**
  - Extend existing openvswitch-vswhitchd init container
    - https://review.openstack.org/#/c/626894/

### DPDK host config: bind NICs to DPDK
- Use dpdk-devbind to bind specified NICs to DPDK
- **Drydock or divingbell**
  - Look into re-using and/or adapting a tool used by kolla-ansible
    - https://review.openstack.org/#/c/626894/

### DPDK host config: enable hugepage support for Kubernetes
- Enable hugepage support for k8s kubelet via feature-gate option
- **Promenade**
  - Hugepages is a beta feature since K8s 1.10 and enabled by default.
  - https://review.openstack.org/#/c/626894/

### Ensure communication between OVS agent and OVS
- Ensure common socket configuration
- **Openstack-helm**
  - Already in place
  - https://review.openstack.org/#/c/626894/

### Adapt OVS bridge configuration for OVS setup
- Work items:
  - Create bridges with datapath type netdev
  - Add physical interface to physical bridge (br-phy)
- **Openstack-helm**
  - Look into re-using and/or adapting a tool used by kolla-ansible
    - https://review.openstack.org/#/c/626894/

### Adapt Neutron (ovs-agent) configuration
- Adapt neutron.conf and ml2 plugin config
- **Openstack-helm**
  - Extend ml2 plugin configuration of neutron in openstack-helm
    - https://review.openstack.org/#/c/643284/

### Adapt IP address assignment to OVS DPDK bridge
- Make sure that correct IPs get assigned to the OVS bridges running DPDK.
  - Every dpdk bridge needs a separate IP address for the tunnel endpoints.
  - **Openstack-helm**

---

**Links**
- https://github.com/openstack/airship-drydock/blob/master/python/tests/yaml_samples/deckhand_fullsite.yaml#L427
- https://github.com/openstack/airship-drydock/blob/master/python/tests/yaml_samples/deckhand_fullsite.yaml#L366
- https://storyboard.openstack.org/#!/story/2004790
- https://www.apt-browse.org/browse/ubuntu/xenial/main/i386/systemd/229-4ubuntu4/file
  /lib/systemd/system/dev-hugepages.mount
<table>
<thead>
<tr>
<th>Task Description</th>
<th>Description</th>
<th>Repo</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>create docker image with dpdk-enabled OVS</td>
<td>update the image build scripts of openstack-helm to include DPDK in the OVS image</td>
<td>openstack-helm-images</td>
<td><a href="https://review.openstack.org/#/c/650152/">https://review.openstack.org/#/c/650152/</a></td>
</tr>
<tr>
<td>update nova and neutron images with newer version of OVS</td>
<td>OVS in current images is outdated</td>
<td>openstack-helm-images</td>
<td><a href="https://review.opendev.org/#/c/651060/">https://review.opendev.org/#/c/651060/</a></td>
</tr>
<tr>
<td>update site configuration to deploy ovs-dpdk</td>
<td>create a site configuration which actually deploys ovs-dpdk as data plane</td>
<td>treasuremap</td>
<td>enable openvswitch chart group</td>
</tr>
<tr>
<td>add LAG support to DPDK configuration</td>
<td>configure LAG support on DPDK NICs in OVS</td>
<td>openstack-helm</td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- done / available
- ongoing
- blocking