Dataplane Networking journey in Containers

Gary Loughnane – gary.loughnane@intel.com
Kuralamudhan Ramakrishnan – kuralamudhan.ramakrishnan@intel.com
DPDK Summit Userspace - Dublin- 2017
Discussion topics

- Container Deployment Models
- Container Bare Metal Reference Architecture
- Container Unified Infrastructure Reference Architecture
Network Cloudification
Multiple Deployment Models

NFVi- Network

VM
Containers
Bare Metal

Hybrid

Unified

NFV Orchestration

CLOUD NATIVE COMPUTING FOUNDATION
Network Cloudification
Multiple Deployment Models – Today Discussion Focus

NFVi - Network

NFV Orchestration

Bare metal

Unified

CLOUD NATIVE COMPUTING FOUNDATION
Address Container Networking Industry Gaps
Intel Containers NFV Reference Architectures
(Experience Kits)

**Identified Gaps**

- Containers
- Compute & Networking

**Resolve, Integrate**

- Existing Solution
- SW Contributions
- Demo Integration

**Communicate**

- Container BM Ref. Architecture
  - Rel. 1 Oct. 2017
- Container UI Ref. Architecture
  - Rel. 1 Feb 2017

**Open Source Projects**

- **VNF**
  - Kubernetes v1.5.0
  - Docker v1.12
- **MULTUS**
  - Multilink Support
- **SR-IOV CNI PLUGIN**
  - Resource Isolation
- **DPDK CNI PLUGIN**
  - Data Plane Support
- **NODE FEATURE DISCOVERY**
  - Data Center Heterogeneity
- **CPU MANAGER for K8s**
  - CPU Core Pinning
- **RESOURCE MANAGEMENT**
  - NUMA Awareness
  - Huge pages
  - QAT

**Best Practice Guidelines**

- Experience kits (Examples)

**NEW SOLUTIONS: WIP**

- NUMA Awareness
- Huge pages
- QAT

**Container BM**

- Data Plane Scale

**Container UI**

- EPA - CMK
- Kuryr
Network Cloudification
Multiple Deployment Models – Today Discussion Focus

**VNFS**
- vEPC
- vCPE
- vNAT
- vRNC
- vIMS
- vHLR
- vNode
- vSGSN
- vMME
- vIDS
- vFirewall

**NFVi - Network**

**NFV Orchestration**
- Onap
- Kubernetes
- OpenStack
- OpenDayLight
- MANO

**Bare metal**
- VM
- Containers
- Compute
- Network
- Storage

**Unified**
- VM
- Containers
- Compute
- Network
- Storage
Industry challenges in containers
Bare Metal

Multiple network interfaces for VNFs
Support for high performance Data Plane (N-S)
Support for high performance Data Plane (E-W)
Ability to request/allocate platform capabilities
Support for CPU Core-Pinning for K8s pods
Dynamic Huge Page allocation

Node Feature Discovery
CPU Manager for Kubernetes
Native Huge page support for Kubernetes
Kubernetes networks via Container Network Interface (CNI)

- Container Orchestration Engines
- Container Runtime
- Container Network Interface
- Container Network Interface Plugins
Containers can "mix and match" with CNI as the API.
DPDK - SRIOV CNI Plugin

PROBLEM
Lack of support for physical platform resource isolation
No guaranteed network IO performance
No support for Data Plane Networking

SOLUTION
Allows SRIOV support in Kubernetes via a CNI plugin
Intel contributor and maintainer of SR-IOV CNI plugin
Supports two modes of operation:
SR-IOV: SR-IOV VFs are allocated to pod network namespace
DPDK: SR-IOV VFs are bounded to DPDK drivers in the userspace

REFERENCE
https://github.com/Intel-Corp/sriov-cni
Vhost user CNI Plugin

PROBLEM
No Container Networking with software acceleration for NFV particularly for East–West Traffic

SOLUTION
Virtio_user/ vhost_user gives boosted performance than VETH pairs
Support VPP as well as DPDK OVS
Vhost_user CNI plugin enables K8s to leverage data plane acceleration

REFERENCE
https://github.com/intel/vhost-user-net-plugin
Network Cloudification
Multiple Deployment Models – Today Discussion Focus

NFVi – Network

VNFs:
- vEPC
- vCPE
- vVLAN
- vIMS
- vRNC
- vHLR
- vSGSN
- vMME
- vIDS
- vIDS

NFV Orchestration

CLOUD NATIVE COMPUTING FOUNDATION

Unified

VM
Containers

Bare metal

MANO

Compute  Network  Storage

Kubernetes Docker
OpenStack KVM

Compute  Network  Storage

Kubernetes Docker
OpenStack KVM

Compute  Network  Storage

Kubernetes Docker
OpenStack KVM
Industry challenges in Nested Containers

- Network performance penalties for Container in VM
- Scaling Data plane networking
- Multiple network interfaces for VNFs
- Support for CPU Core-Pinning for K8s pods

Kubernetes

Kuryr-Kubernetes

MULTUS

CPU Manager for Kubernetes
OBJECTIVES
One Virtual Machine to many Containers
Target: 1k Containers per VM
Container Data Plane performance

USE CASE
Elasticity and scalability of containerized VNF application in VM

BENEFITS
- VT-x ring de-privileging to move the VM and Container into userspace, making it accessible to the userspace vSwitch with just a single copy.
- Standard Virtio interface that supports both interrupt and poll modes, VNF and Cloud based applications.
- Standard Vhost shared memory interface between DPDK vSwitch and VNF.

SOLUTIONS
- Enabling DPDK in containers using VIRTIO
- Using Kuryr–Kubernetes, orchestrate the Dataplane networking – DPDK workload using the vSwitch
Kuryr – Kubernetes with Dataplane Networking with DPDK based vSwitch

PROBLEM (Current Status)
No support for Data Plane Networking for nested containers in Unified Infrastructure RA

SOLUTION (WIP)
Working on the PoC for the development of DPDK net plugin support in kuryr CNI
DPDK: Virtio are bounded to DPDK drivers in the userspace

REFERENCE
https://blueprints.launchpad.net/kuryr-kubernetes/+spec/nested-dpdk-support
Questions?

Gary Loughnane
gary.loughnane@intel.com
Kuralamudhan Ramakrishnan
kuralamudhan.ramakrishnan@intel.com