



NFV FEATURES IN KUBERNETES

KURALAMUDHAN RAMAKRISHNAN
SENIOR SOFTWARE ENGINEER
SOFTWARE DEFINED DATA CENTER SOLUTION GROUP(SDSC)
INTEL

INPUTS FROM:

INTEL:

DANA NEHAMA

IVAN COUGHLAN

KATE MULHALL

RED HAT:

FENG PAN

DOUGHLAS SMITH

BILLY MC FALL

ABOUT THIS SESSION

•After this session you will :

- Understand how to spin-up a Kubernetes cluster, concepts on Kubernetes Networking and compute
- Learn how DPDK, OVS, VPP, SRIOV are implemented in Kubernetes context
- Get hands-on deployment of simulated Cloud Native NFV infrastructure

Target Audience – Technical at all levels

*Addressing barriers on the path to Cloud Native
Step 0 – Containers and looking at the Industry gaps*

CONTAINERS NETWORKING - TECHNOLOGY OVERVIEW

CONTAINERS NETWORKING DEPLOYMENTS CONSIDERATIONS

MULTIPLE DEPLOYMENT MODELS

CLOUD NATIVE
COMPUTING FOUNDATION

VNFS

vCMTS

vIMS

vEPC

vCPE

vSBC

NFV ORCHESTRATION



OVN
(Open Virtual Network)

ONAP



kubernetes

OPEN
DAYLIGHT

Open Source
MANO



openstack



KURYR
an OpenStack Community Project

NFVI-NETWORK



Open vSwitch



OpenFlow



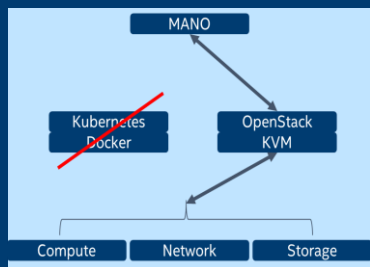
DPDK
DATA PLANE DEVELOPMENT KIT

SR-IOV



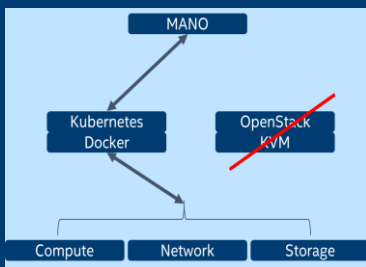
CNI

VM



Containers

BARE METAL

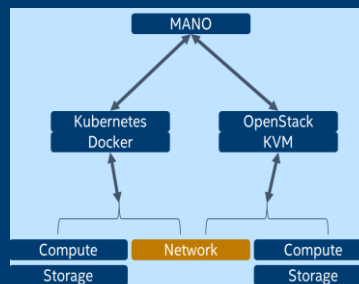


VM

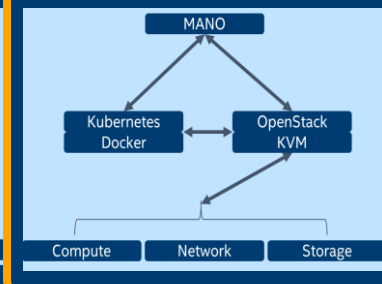
&

Containers

HYBRID



UNIFIED



ADDRESS KEY CHALLENGES IN CONTAINERS BARE METAL

CHALLENGES BEING ADDRESSED

SOLUTION

Multitenant	Kubernetes Networking	🔒 ➡
High performance	Data Plane (E-W)	🔒 ➡
High performance	Data Plane Acceleration	🔒 ➡
High availability	Fail-over and high availability of networking	🔒 ➡
Ability to request/allocate platform capabilities		🔒 ➡
CPU Pinning and Scheduling	Resource Management	🔒 ➡
Enhance Platform Awareness	(EPA)	🔒 ➡
Discovery, Advertise, schedule and manage devices with K8s		🔒 ➡
Guarantee NUMA node resource alignment		🔒 ➡
Platform telemetry	Telemetry	🔒 ➡

kubernetes CNI		SOFTWARE AVAILABILITY*
🔑	MULTUS	Open Source: CNI plug-in – V3.1 Aug '18 Upstream K8s: TBD
🔑	USERSPACE CNI	Open Source: CNI plug-in – V1.1 July '18
🔑	SR-IOV DPDK	Open Source: CNI plug-in – V0.3 May 2018
🔑	BOND-CNI	Open Source: CNI plug-in – V1.0 Dec '17
🔑	Node Feature Discovery	Open Source: v0.2.0 May 2018 Upstream K8: Incubation Graduation TBD
🔑	CPU Manager for Kubernetes	Open Source: V1.2 .2 Feb'18 Upstream K8: Phase 1 - V1.8 Sept '17
🔑	Native Huge page support for Kubernetes	Upstream K8: V1.8 Sept '17
🔑	Device Plugin SRIOV & QAT	Open Source: SR-IOV Jul '18, QAT Aug'18
🔑	NUMA Manager	Upstream: Working PoC with proposal Upstream K8: TBD
🔑	collectd	Upstream collectd: V5.7.2 June '17 ;5.8.0 ((Q4 2017 date TBD)

Open Source: Available on github <https://github.com/intel> & <https://github.com/redhat-nfvpe>
NFD at <https://github.com/kubernetes-incubator/node-feature-discovery>

MULTIPLE NETWORK INTERFACES FOR VNFs

PROBLEM

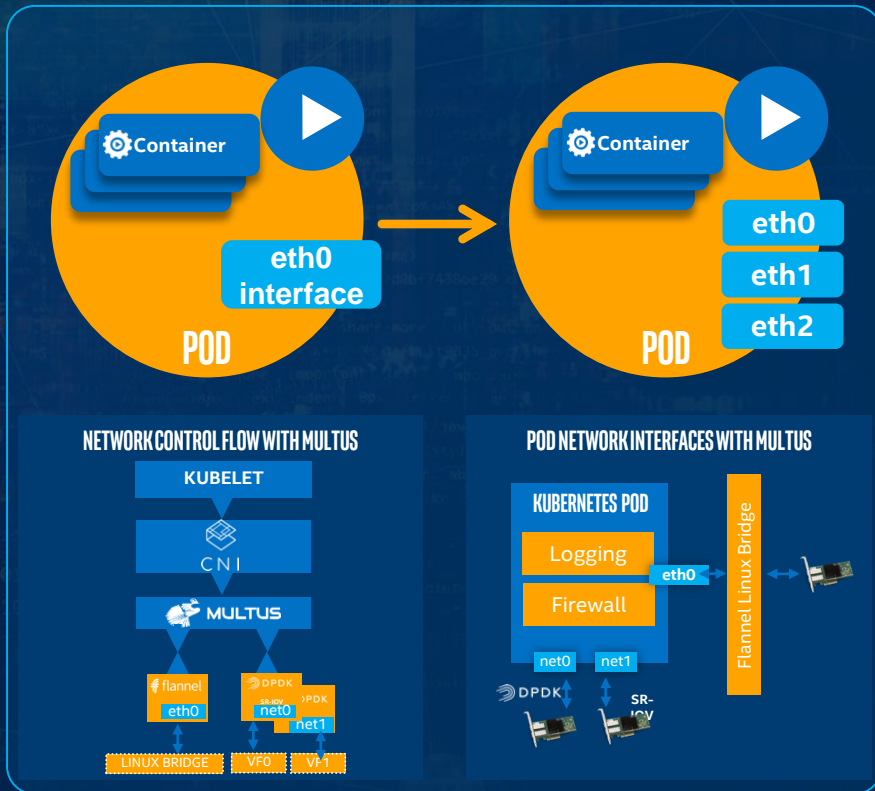
- Kubernetes support only one Network interface – “eth0”
- In NFV use cases, it is required to provide multiple network interfaces to the virtualized operating environment of the VNF

USE CASES

- Functional separation of control and data network planes link aggregation/bonding for redundancy of the network
- Support for implementation of different network SLAs
- Network segregation and Security

REFERENCE

- Multus CNI – Founded by Intel & a collaboration projects between Intel & Red Hat <https://github.com/intel/multus-cni> (v3.1 Aug 18)
- Native Kubernetes - Mailing list with details on discussions: <https://groups.google.com/forum/#!forum/kubernetes-sig-network>



USERSPACE CNI PLUGIN

PROBLEM

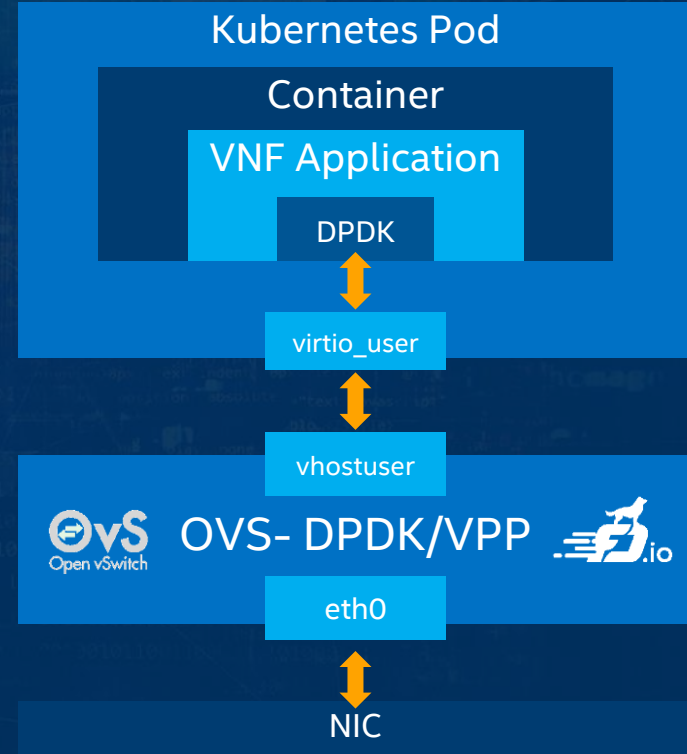
- No Container Networking with software acceleration for NFV, particularly for East – West Traffic

SOLUTION

- Virtio_user/vhost_user performance better than VETH pairs
- Supports VPP as well as DPDK OVS
- Vhost_user CNI plugin enables K8s to leverage data plane acceleration

REFERENCE

- Userspace CNI - Founded by Intel, and a Collaboration projects between Intel & Red Hat <https://github.com/intel/vhost-user-net-plugin> (V1.1 July '18)



DPDK - SR-IOV CNI PLUGIN

PROBLEM

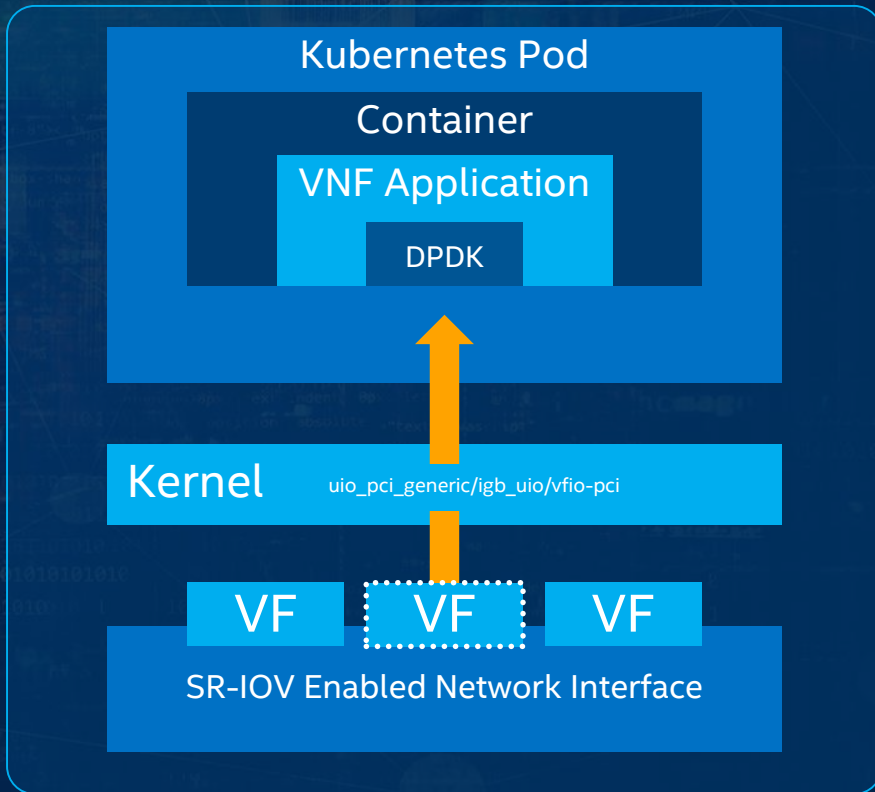
- Lack of support for physical platform resource isolation
- No guaranteed network IO performance
- No support for Data Plane Networking

SOLUTION

- Allows SR-IOV support in Kubernetes via a CNI plugin
- Supports two modes of operation:
 1. SR-IOV: SR-IOV VFs are allocated to pod network namespace
 2. DPDK: SR-IOV VFs are bounded to DPDK drivers in user space

REFERENCE

- SRIOV CNI - A Collaboration projects between Intel & Red Hat
<https://github.com/intel/sriov-cni>(V0.3 May '18)



DEVICE PLUGINS OVERVIEW

WHY?

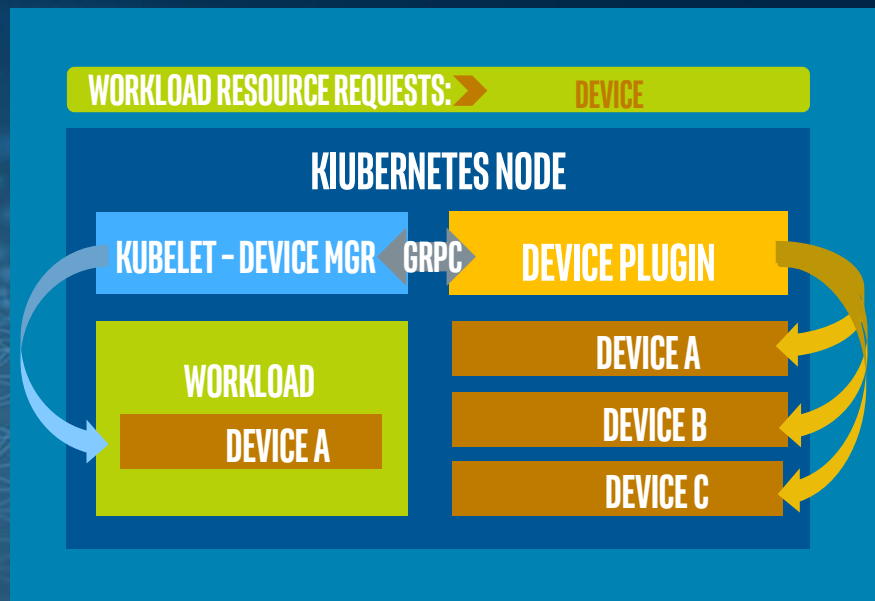
- Device vendors have to write custom Kubernetes code in order to integrate their device with the ecosystem
- Results in multiple vendors maintaining custom code making it difficult for a customer to consume

HOW?

- Provide a device plugin framework which enables vendors to advertise, schedule and setup devices with native Kubernetes integration
- Device Plugins are easily deployed and workload device requests are made via extended resource requests in the Pod Specification

BENEFITS

Enables effective resource utilization



REFERENCE

<https://kubernetes.io/docs/concepts/cluster-administration/device-plugins/>

KUBERNETES RESOURCE MANAGEMENT

PROBLEM STATEMENT

Kubernetes clusters are deployed on a wide array of heterogeneous environments with different hardware resources

Today, CPU and Memory are the core resources orchestrated by Kubernetes. Workloads have a wide variety of hardware resource requirements as well as CPU and Memory but Kubernetes is agnostic to these.

GOAL

Introduce a broader array of resources representing cluster abilities to cater for the wide range of workloads being deployed using Kubernetes

SRIOV NETWORK DEVICE PLUGIN

PROBLEM

- SRIOV CNI has no mechanism to expose number of VFs available in a host.
- Allocated network resources are not guaranteed from same NUMA node.

SOLUTION

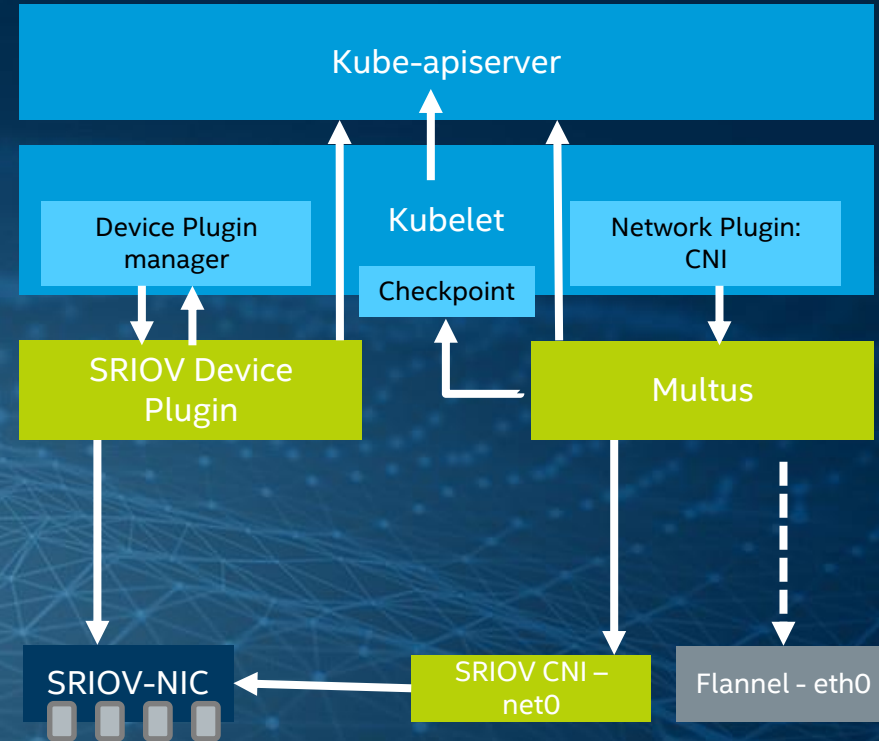
- SRIOV Network device plugin provides a mechanism to advertise SRIOV network capabilities of a host.
- VNF application that demands high performance NICs as resource requests then can be scheduled to this host by K8s scheduler.
- SRIOV device plugin allocates required VFs meeting NUMA alignment.

REFERENCE

SRIOV Network device plugin

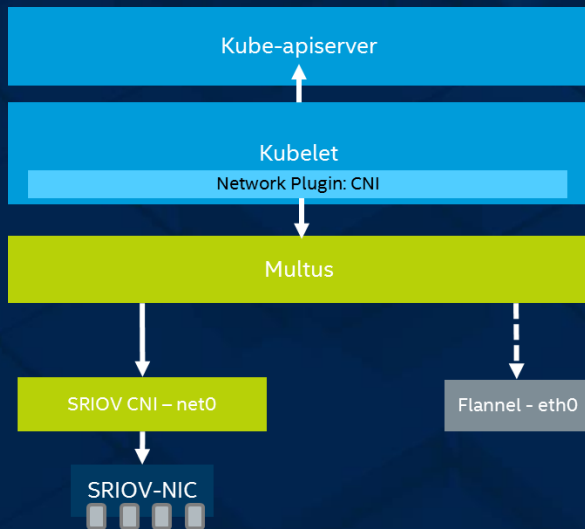
A Collaboration project between Intel & Red Hat

<https://github.com/intel/sriov-network-device-plugin>

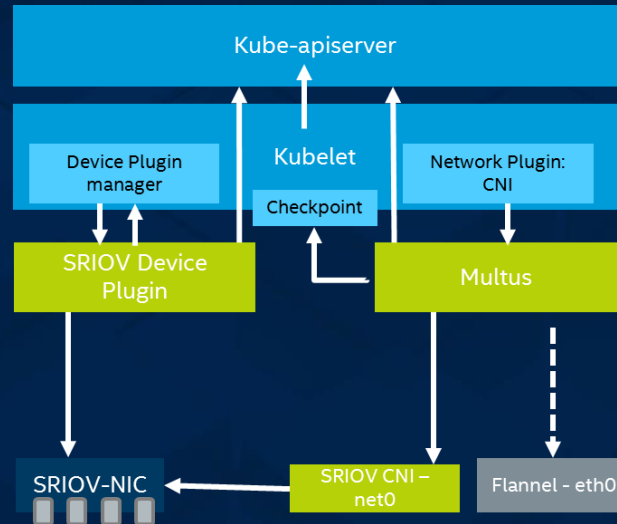


SRIOV NETWORK DEVICE PLUGIN

CNI Model



CNI & Device plugin approach pluggable models

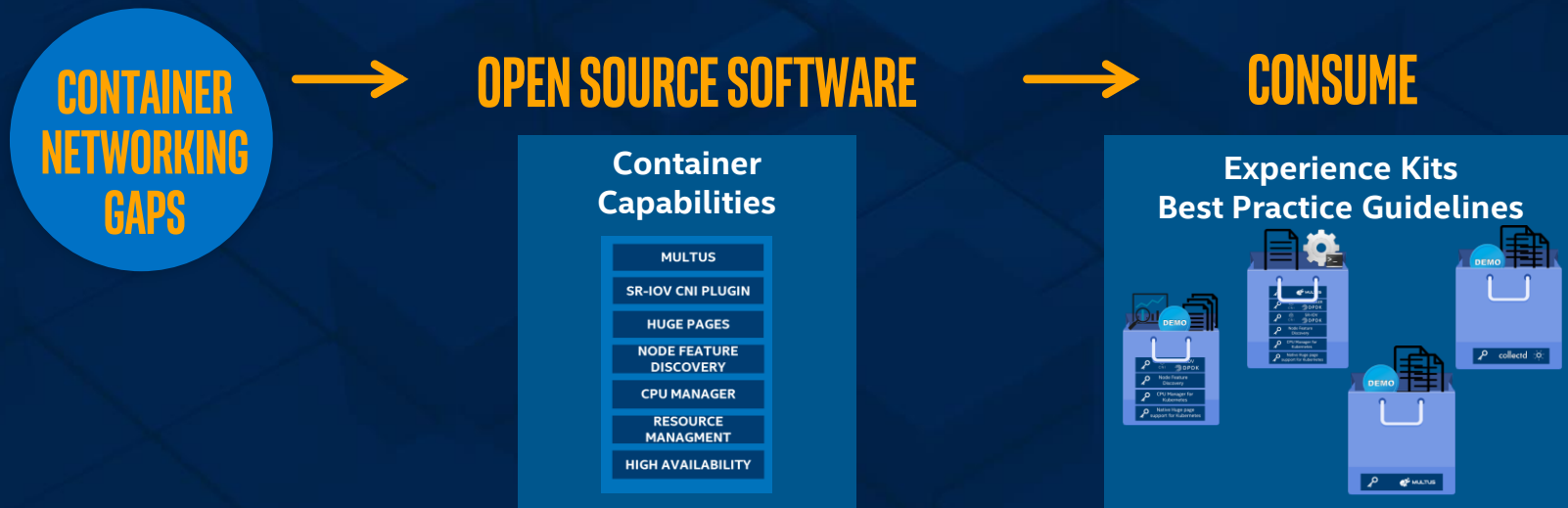


Challenge resolved with new approach:

The CNI interface has no mechanism to enable k8s to:

- Schedule consumable network resources e.g. SR-IOV VFs
- Guarantee NUMA node resource alignment
- Restrict access to allocated devices only

INTEL & RED HAT ENABLING CONTAINER NETWORKING COMMERCIAL DEPLOYMENT



Red Hat & Intel is addressing key challenges to using containers for NFV use cases

Many of these have been open sourced already

Explore more information available on Intel's Network Builders site

<https://networkbuilders.intel.com/network-technologies/container-experience-kits>

<https://www.okd.io/>



Q&A THANK YOU