BMWG – Containerized Infrastructure Benchmarking

IETF 116 Hackathon

March 25-26, 2023

Yokohama, Japan

Hackathon Plan

- Our hackathon activities provide benchmarking references to support our draft:
  - Draft:
    Considerations for Benchmarking Network Performance in Containerized Infrastructures
    https://tools.ietf.org/html/draft-dcn-bmwg-containerized-infra
  - This draft discusses all possible factors that can affect container network performance:
    - Different networking model based on packet acceleration techniques
    - Different resources configuration settings

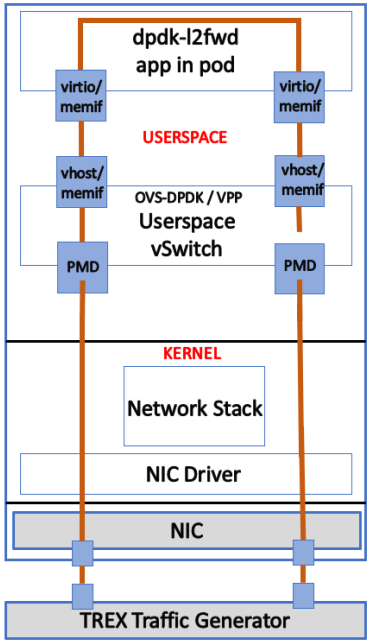
Hackathon Plan

Verify and Summarize all the container network performance impact results caused by the mentioned factors in all of our previous hackathon (111-116)

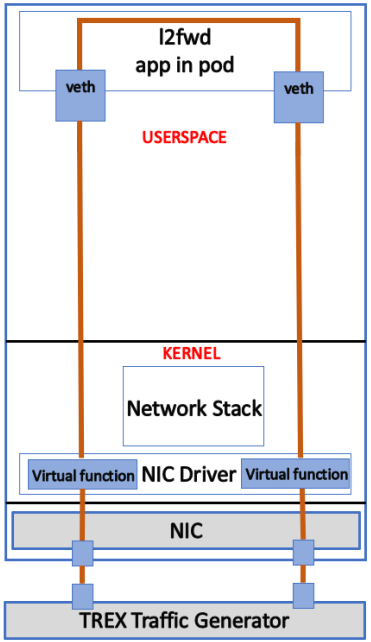
- **Networking models**
  - ✔ Userspace acceleration (OVS-DPDK, VPP)
  - ✔ SRIOV
  - ✔ eBPF
  - ✔ Combined models
- **Resources configuration settings**
  - ✔ NUMA allocation
  - ✔ CPU, Memory allocation, hugepages
  - ✔ Services Function chaining

What got done

- Different networking models based on packet acceleration techniques



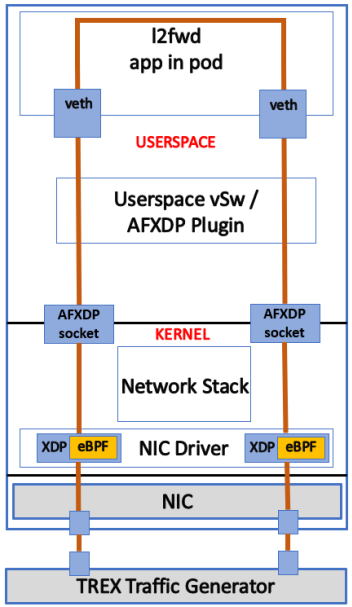
Userspace Acceleration
(OVS-DPDK, VPP vSwitch)



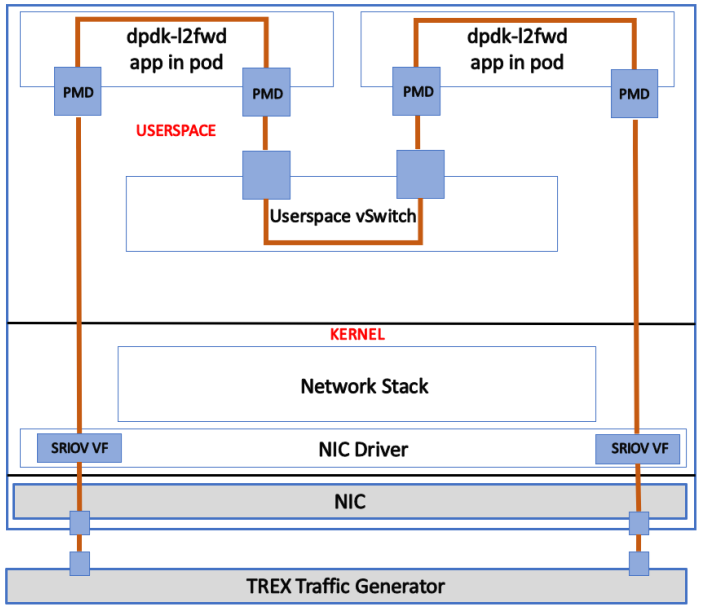
SRIOV Acceleration

What got done

- Different networking models based on packet acceleration techniques



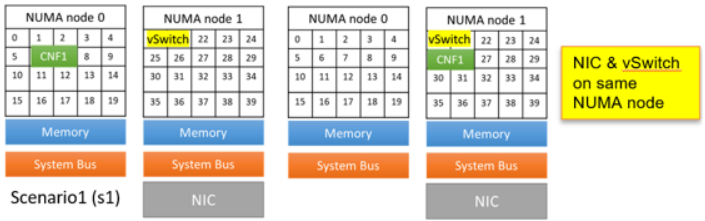
eBPF Acceleration
(using AFXDP)

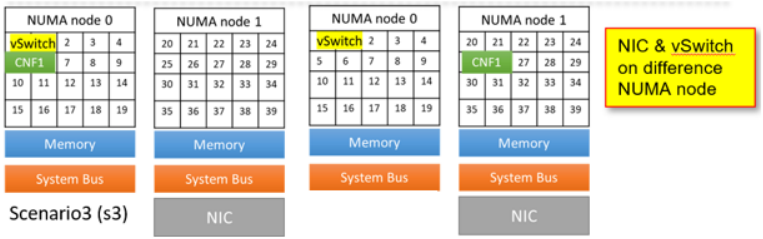


Combined Model Acceleration
(SR-IOV + OVS-DPDK/VPP)

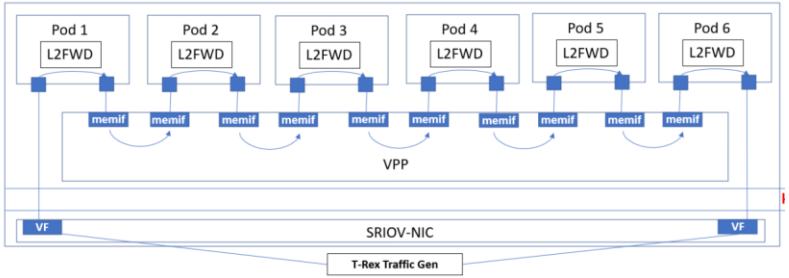
What got done

- Different Resources Configuration Settings
  - **NUMA allocation (C-VNF, vSwitch, NIC same/different nodes NUMA)**





  - **Number of CPU cores, memory allocation**
    - CPU, memory allocation to C-VNF/pod via –request attribute in yaml file
    - CPU cores allocation to vSwitches in their startup configuration
  - **Service function chaining**
    - Impact of number of C-VNFs in a service chain



What got done

- Benchmarking Configuration
- **Hardware – Worker Node**

CPU	Intel(R) Xeon(R) Gold 5220R CPU @ 2.20GHz 48 CPU cores * 2 NUMA nodes
Memory	256GB: 32GB x 4DIMMs x 2 NUMA nodes @ 2400MHz
NIC	**Intel Corporation Ethernet Network Adapter X71-40Gbps**
Microcode	0x5003102
Intel NIC Device ID	0x1572
Intel NIC Firmware version	6.01 0x800035cf 1.1747.0
BIOS setting	CPU Power and Performance Policy <Performance> CPU C-state Disabled CPU P-state Disabled Intel(R) Hyper-Threading Tech Enabled Turbo Boost Disabled

- **Traffic Generator : T-Rex (v2.92)**

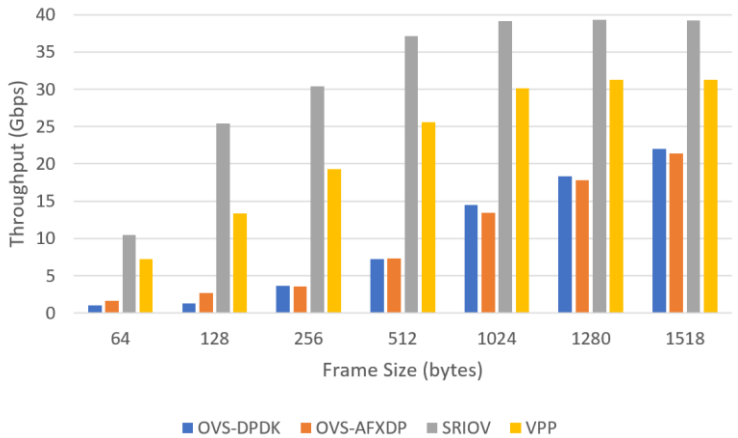
Name	T-Rex
Version	2.92
Benchmark method	**T-Rex Non Drop Rate application** (accepted percentage of drop rate is less than 0.1%)

- **Software**

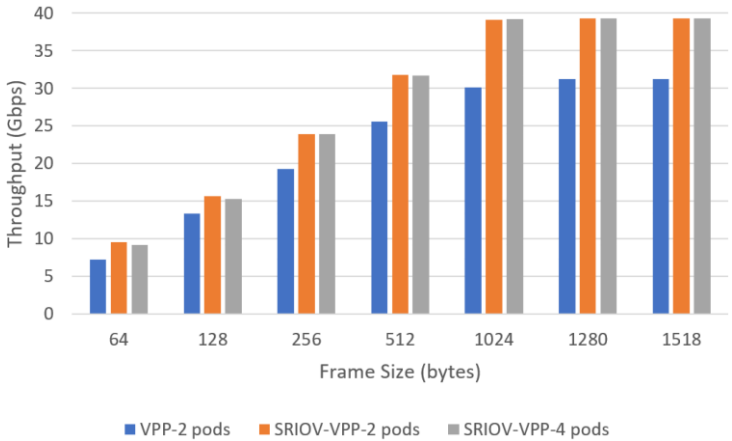
Operating System	**Ubuntu 22.04**
Linux Kernel Version	5.15
GCC version	gcc version 4.8.5 20150623 (Red Hat 4.8.5-44)
DPDK version	**22.11.1**
Hugepages	1Gi

What we learned

- Different performance caused by different networking models
  - (OVS-DPDK < VPP < SRIOV)
  - VPP < Combined VPP-SRIOV (multi-pod)



Comparison between
Userspace Model vs eBPF-AFXDP vs SRIOV

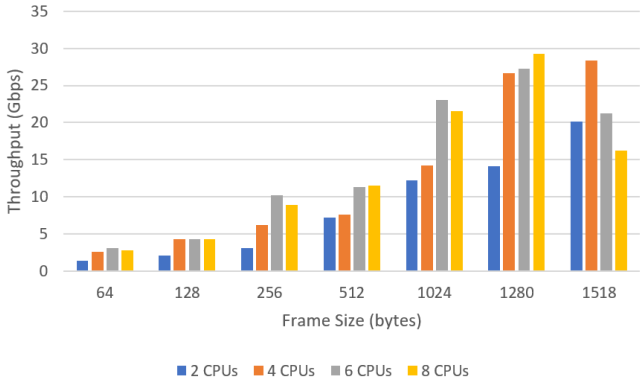


Combined model performance with impact of
different number of C-VNF

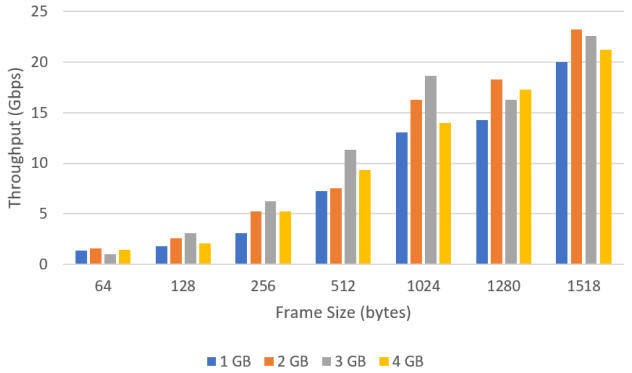
What we learned

- Different performance caused by different CPU Memory allocation to pod
  - Might only affect when using DPDK version 20 or below
  - With higher DPDK version, we witness no impact to network performance when changing resources allocation

***Results tested on OVS-DPDK*



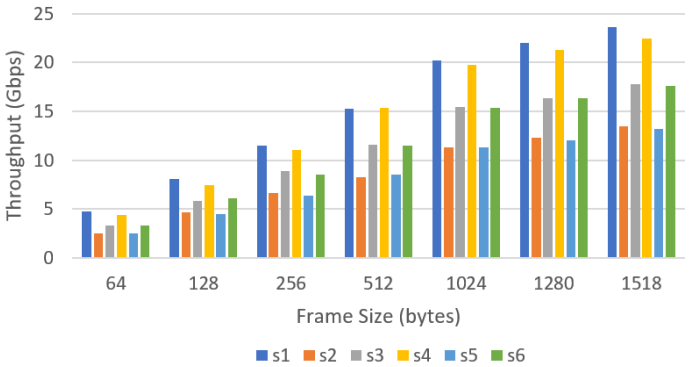
Different Pod's CPU cores allocation performances



Different Pod's memory allocation performances

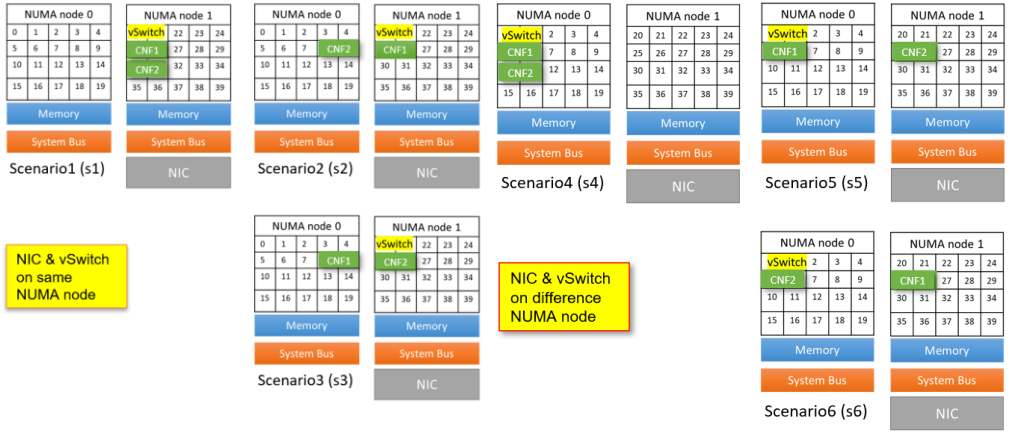
What we learned

- Different performance caused by different NUMA alignment
  - Between Pods – vSwitch - NIC



Different NUMA alignment performances

**Results tested on OVS-DPDK



NUMA alignment scenarios

Future Works

- After this benchmarking verification, we finalize our draft and submit for WG draft adoption
- We would like to welcome any questions, comments to the draft

Wrap Up

Team members:

On-site

Younghan Kim (SSU)

Minh Ngoc Tran(SSU)

Jangwon Lee (SSU)

Remote

Van Binh Duong (SSU)

Git repo:

https://github.com/mipearlska/bmwg-container-networking

DCN LAB Distributed Cloud and Network Research Laboratory

Soongsil University, South Korea