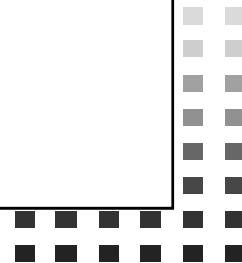




BMWG – Containerized Infrastructure Benchmarking

**IETF 113 Hackathon
March 19-20, 2022**



Hackathon Plan

- Our draft main goal is to figure out container networking performance impacts by various resource options.
 - Draft:
Considerations for Benchmarking Network Performance in Containerized Infrastructures
<https://tools.ietf.org/html/draft-dcn-bmwg-containerized-infra>
 - Two main features
 - Verify container network performance with various network models
 - Verifying performance impacts depending on different configuration settings

Hackathon Plan

What we have done so far in Hackathon

- ✓ OvS DPDK, VPP, SR-IOV single node benchmarking (4.1,2,4)
- ✓ CPU Isolation Impact VPP (5.1)
- ✓ NUMA Alignment Impact (5.1)
 - ✓ Single pods
 - ✓ Multi pods (5.3)
- ✓ VPP multi pod benchmarking (5.3)

In this hackathon

- Combined SRIOV VPP multi pods (service chain) (5.3)
- Number of VNFs in service chain impact (5.3)

<u>4.</u>	<u>Networking Models in Containerized Infrastructure</u>	<u>8</u>
<u>4.1.</u>	<u>Kernel-space vSwitch Model</u>	<u>9</u>
<u>4.2.</u>	<u>User-space vSwitch Model</u>	<u>10</u>
<u>4.3.</u>	<u>eBPF Acceleration Model</u>	<u>10</u>
<u>4.4.</u>	<u>Smart-NIC Acceleration Model</u>	<u>12</u>
<u>4.5.</u>	<u>Model Combination</u>	<u>13</u>
<u>5.</u>	<u>Performance Impacts</u>	<u>14</u>
<u>5.1.</u>	<u>CPU Isolation / NUMA Affinity</u>	<u>14</u>
<u>5.2.</u>	<u>Hugepages</u>	<u>15</u>
<u>5.3.</u>	<u>Service Function Chaining</u>	<u>15</u>
<u>5.4.</u>	<u>Additional Considerations</u>	<u>16</u>

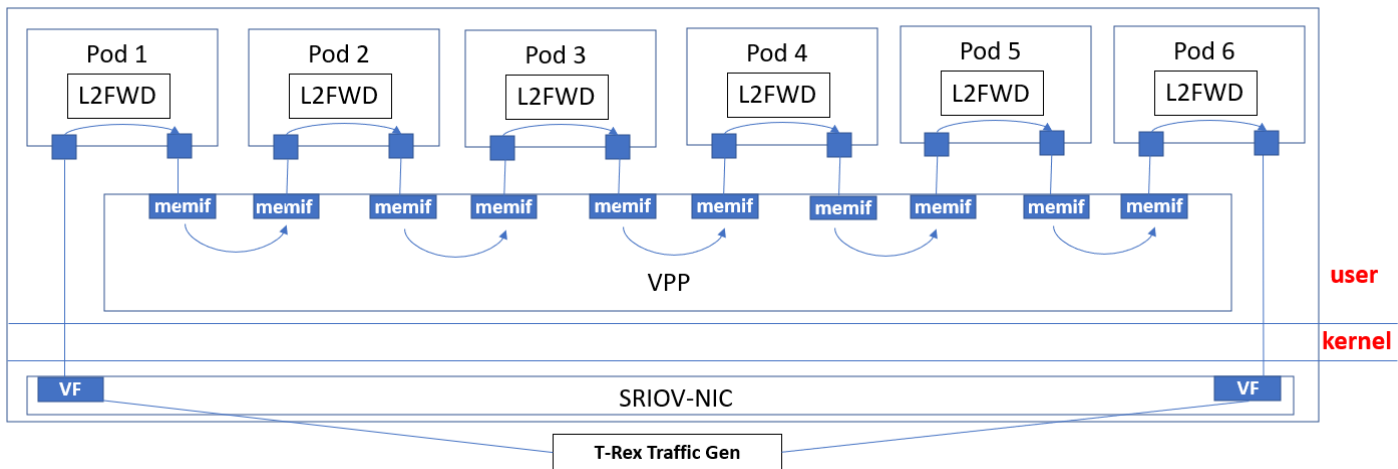
BMWG – Containerized Infrastructure Benchmarking

- Champion(s)
 - Younghun Kim <younghak at ssu.ac.kr>
 - Minh-Ngoc Tran <mipearluka1307 at dcn.ssu.ac.kr>
 - Hokeun Lim <limhk at dcn.ssu.ac.kr>
- Project(s)
 - Benchmarking performance of SR-IOV and VPP combination
 - SR-IOV for North-South Traffic
 - VPP for East-West Traffic
 - Benchmarking number of C-VNFs impact in service function chain scenario
- Specification(s)
 - <https://datatracker.ietf.org/doc/html/draft-dcn-bmwg-containerized-infra>

What got done

- Benchmarking Scenario – **SRIOV and VPP** Combination

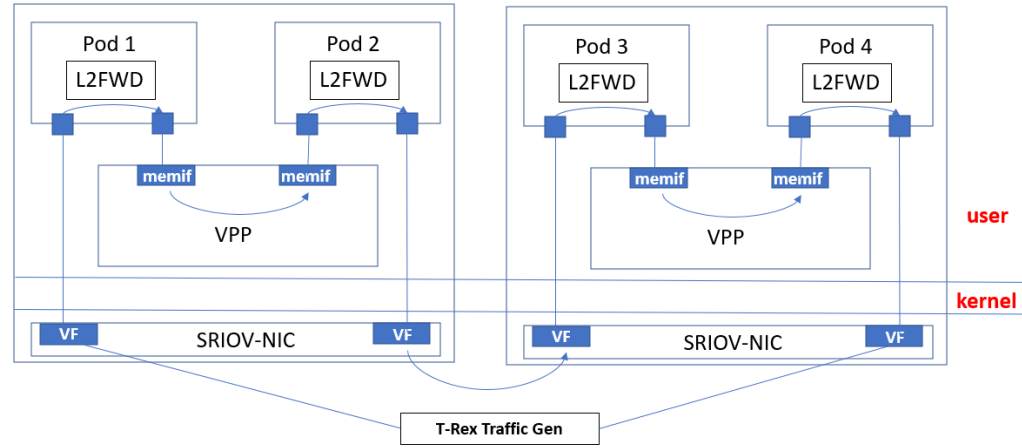
- **Test impact of number of VNFs:** In each scenario, the number of C-VNFs is increasing for each test
- Single node: 1node-2pods/4pods/6pods (1n2p, 1n4p, 1n6p)



Single node scenario

What got done

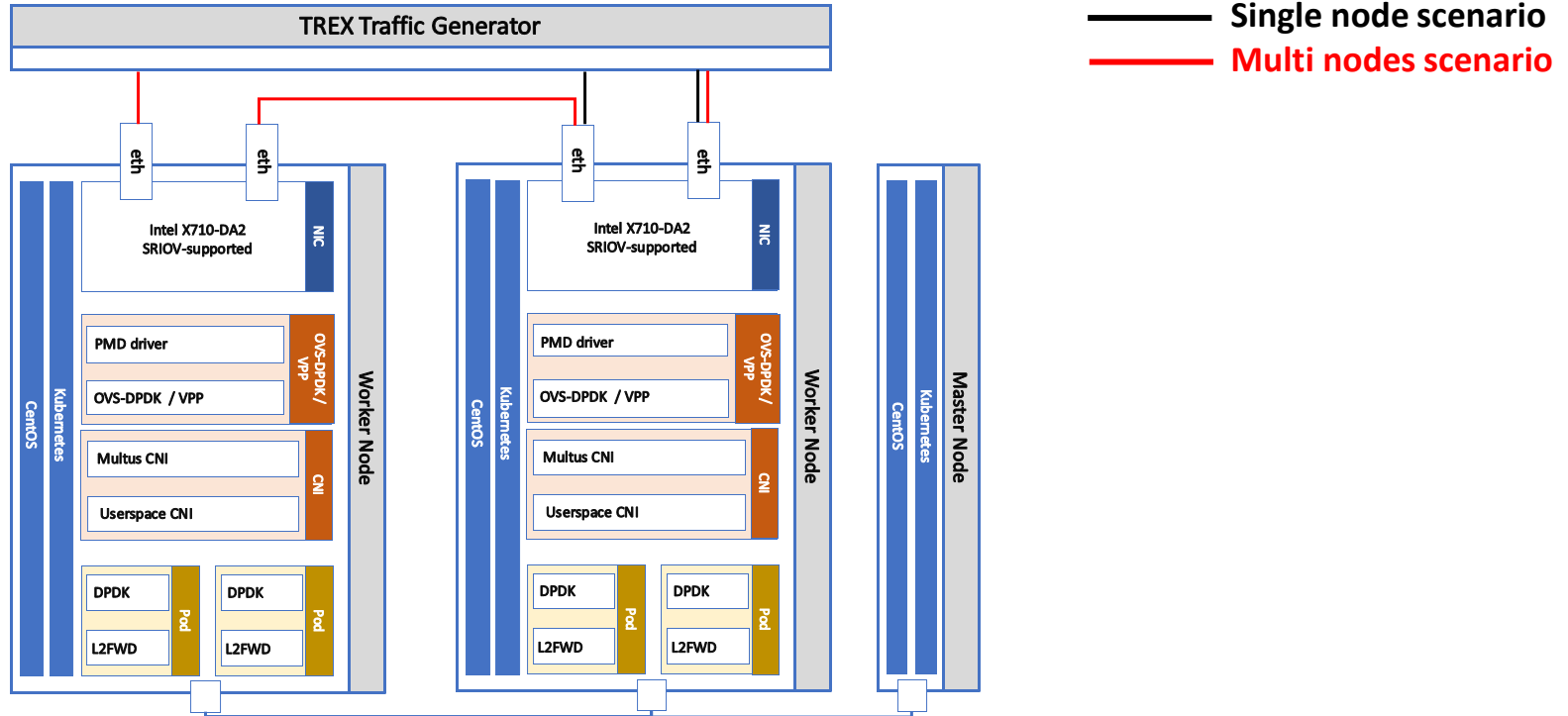
- Benchmarking Scenario – **SRIOV and VPP Combination**
 - **Test impact of number of VNFs:** In each scenario, the number of C-VNFs is increasing for each test
 - Multi node: 2node-4pods(2n4p) - 2 pods each node



Multi nodes scenario

What got done

- Benchmarking Testbed Topology



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)**

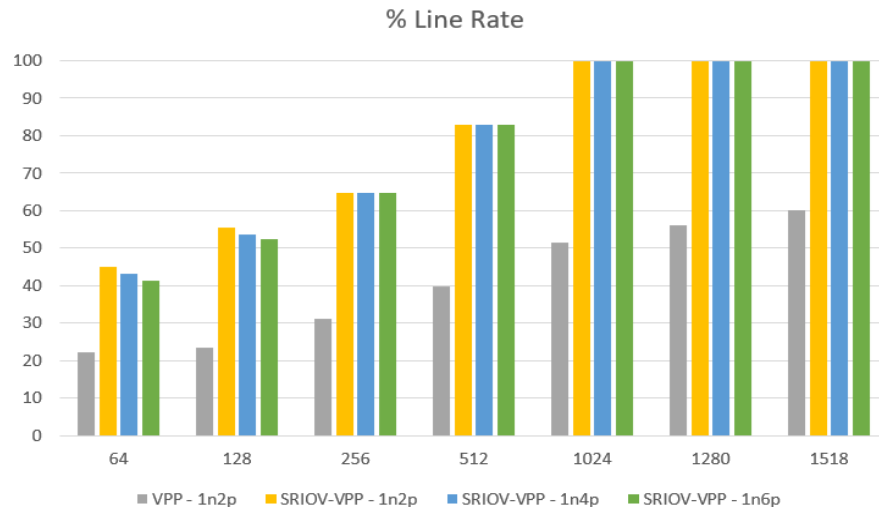
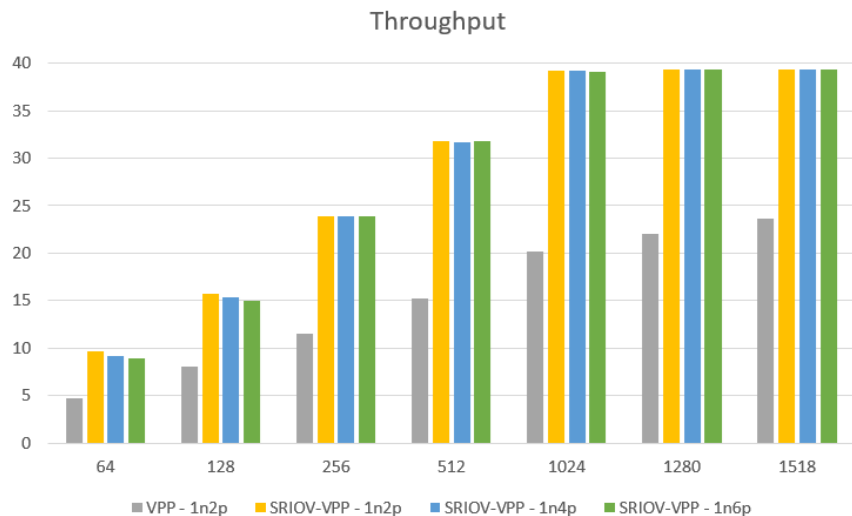
Name	T-Rex
Version	2.92
Benchmark method	T-Rex Non Drop Rate application (accepted percentage of drop rate is less than 0.1%)

- **Software**

Operating System	CentOS Linux 7
Linux Kernel Version	3.10.0-1160.31.1.el7.x86_64
GCC version	gcc version 4.8.5 20150623 (Red Hat 4.8.5-44)
DPDK version	20.11
Hugepages	1Gi

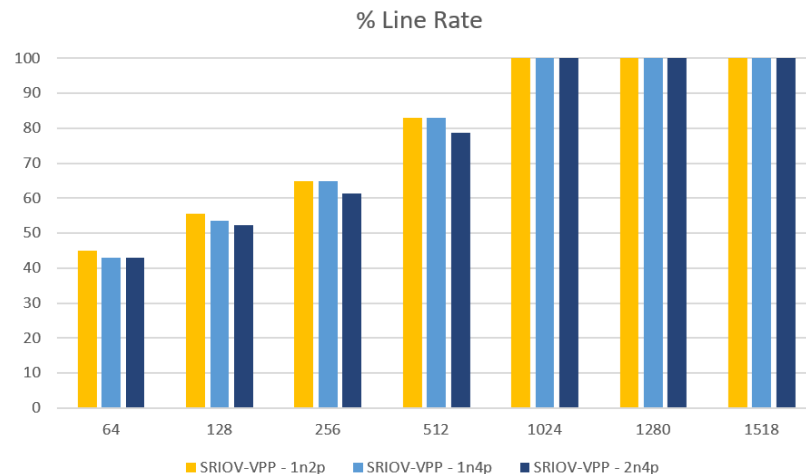
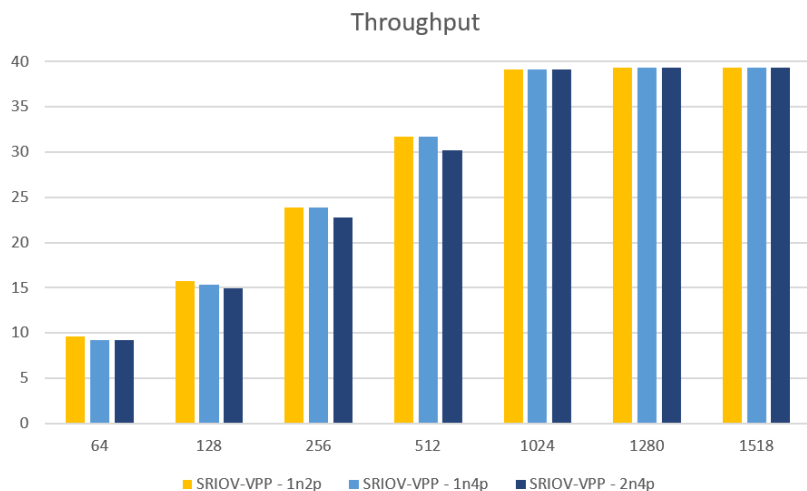
What we learned

- Benchmarking Performance Results – Single node SRIOV-VPP service chain
 1. SRIOV-VPP performed significantly better than VPP only (packets through VPP need go through vSwitch, no need with SRIOV)
 2. Increase number of pod slightly reduce throughput 2% at small packet size (64,128)



What we learned

- Benchmarking Performance Results – Multi-nodes SRIOV-VPP service chain
 1. Throughput in multi-nodes scenario is slightly smaller than single-node with smaller packet size (<512) due to increasing in number of pod (4 pods total in multi-nodes > 2 pods in single node)



What we learned

- Future Works
 - Inter-node network technique impact testing (L2 underlay, tunnel end point VXLAN, BGP)
 - eBPF/XDP with/without offload benchmarking
 - Update our draft

Wrap Up

Team members:

Younghan Kim (SSU)

Minh Ngoc Tran(SSU)

Hokeun Lim (SSU)

Phuong Bac Ta (SSU)

Jangwon Lee (SSU)

Git repo:

<https://github.com/huyng14/bmwg-container-networking>

IETF Korea Forum:

Remote Hackathon Event at Busan

with I2NSF, IPWAVE (SKKU team)