Considerations for Benchmarking Network Performance in Containerized Infrastructure

draft-dcn-bmwg-containerized-infra-13

Minh-Ngoc Tran (Soongsil University), Sridhar Rao (The Linux Foundation),

Jangwon Lee, Younghan Kim (Soongsil University)

Updates from v11 to v13 (current)

Updates are based on IETF 117 meeting comments (requesting clearer draft scope)

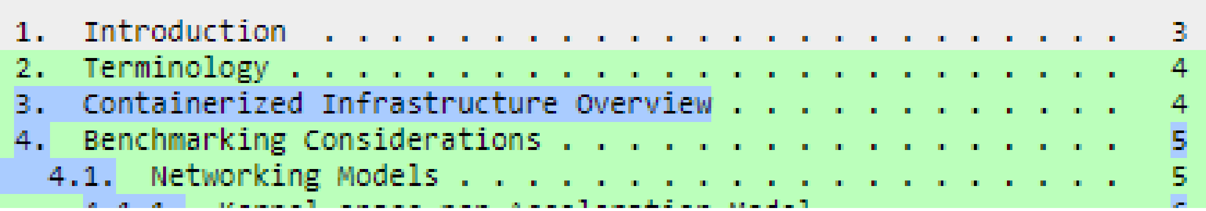
and previous unaddressed reviews of version 10 from Vratko and Gabor (most comments has been addressed in v11 presented at IETF 117)

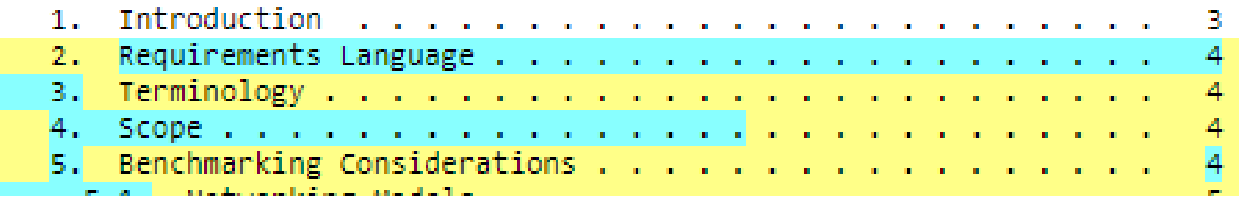
NEW

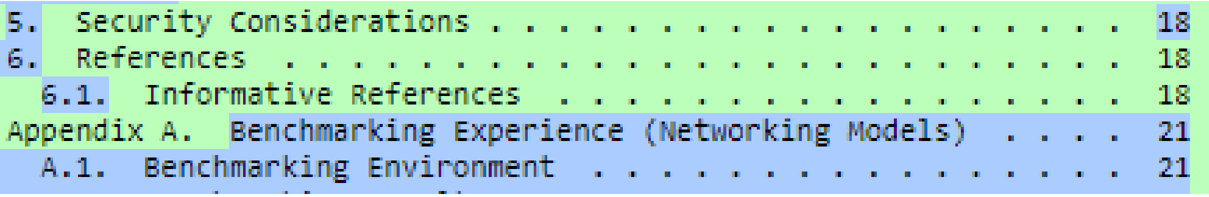
- Add new Scope section
- Specify new Benchmarking parameters according to each container Resources Configuration consideration

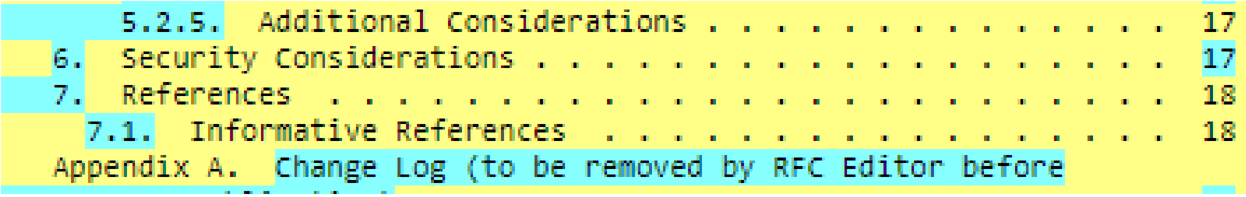
MERGE/REMOVE

- Merge duplicate information in Containerized Infrastructure Overview section to Introduction Section
- Remove Benchmarking Appendixes







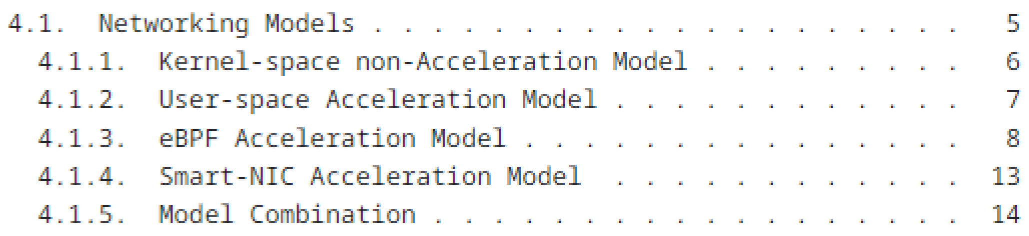


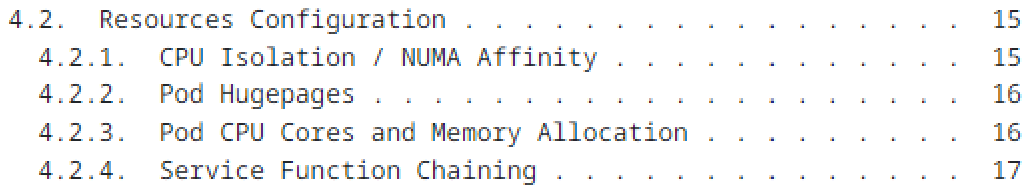
Scope

Based on IETF 117 meeting comments: want to see a clearer scope of the draft, clear differences against previous works

- Add a separate Scope section for the draft

- The primary scope of this document is to fill in the gaps of previous BMWG's NFV benchmarking consideration works ([RFC8172] and [RFC8204]) when applying to containerized NFV infrastructure.
- The consideration gaps are:
  - Different network models/topologies configured by container network interfaces (especially the extended Berkeley Packet Filter model which was not mentioned in previous documents)
  - Resources configuration for containers.





Specify new Benchmarking Parameters

To show clear additional benchmarking parameters against related RFCs

- **CPU Isolation / NUMA Affinity**
    - Selected CPU Isolation level
    - NUMA cores allocation to pod
- **Pod Hugepages**
    - Pod's hugepage size
- **Pod CPU cores and Memory allocation**
    - Pod's CPU core allocation
    - Pod's RAM allocation
- **Service Function Chaining**
    - Number of CNFs/pod
    - Selected CNI plugin

Merge/Remove

To address Vratko’s comment about grouping Introduction and Containerized Infrastructure Overview section

- Merge duplicate information in Containerized Infrastructure Overview section to Introduction Section

About comments regarding benchmarking results:

Benchmarking results are just proof of concepts for the proposed considerations while developing the draft. At this stable stage, these information can be removed

- Remove Benchmarking Appendixes

Next Steps

- As we have addressed the comments after several reviews, we would like to call for adoption of the draft again as a working group draft

Backup Slides

Previous addressed review comments at IETF 117

Detailed Updates (1)

Introduction and Overview inconsistency with remain draft contents

Vratko's review: veth is not a general concept

- Use "container network plugin" as the general container networking mechanism
  - In terms of networking, to route traffic between containers which are isolated in different network namespaces, virtual ethernet (vETH) interface pairs are used to create …
  - ➢ In terms of networking, to route traffic between containers which are isolated in different network namespaces, a container network plugin is required. This network plugin creates …

Vratko's review: CNI is specific to Kubernetes, there are other container orchestration services.

- Kubernetes is the main and most popular orchestration platform nowadays, so we can use Kubernetes' Container Network Interface (CNI) for the draft. All of networking models in the draft require Kubernetes CNIs:
  - Kernel-Space non-Acceleration: normal Kubernetes CNI (i.e. flannel)
  - User-Space Acceleration: Userspace K8s CNI
  - eBPF Acceleration: Cilium/AFXDP/Userspace K8s CNI
  - Smart-NIC Acceleration: SR-IOV K8s CNI

Vratko's review: List of networking model might inevitably be incomplete

- The Networking Model consideration list in the draft does not list out container networking techniques. It is a list of all possible categories
- Any additional technique can fall into one of the considered categories in the draft

Detailed Updates (2)

Resources Configuration Additions

Vratko’s review: Some resources can also be applied to VM-VNF

- Update in Draft: Mentions about NUMA and CPU Isolation can also be applied for VM, others are specific to pod

Vratko’s review: Most consideration also applies for other SUT components, not only NFV DUTs

- Added in Draft

Vratko’s review: Nosiy neighbor is not the only use-case of Resource Isolation practices,

- Added in Draft

Vratko’s review: Recommend varying non-DUT resources in Resource Isolation benchmarking

- Under test and will be consider to add

Gábor’s review: Hugepage size value of 2MB and 1GB

- Updated to current standard.

Detailed Updates (3)

Benchmarking Appendixes

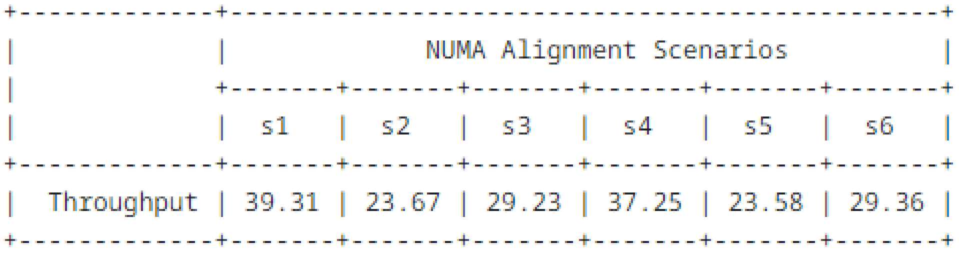
Vratko suggestion about input/output/result reporting, and putting Benchmarking result as a main draft section

- We provided Benchmarking results in our draft as a kind of “proof-of-concept” for verifying the benchmarking considerations proposed in our draft.
- Containerized Networking Benchmarking Reporting standard is out of scope of our draft

Gábor’s review: Packet Frame size values are not provided

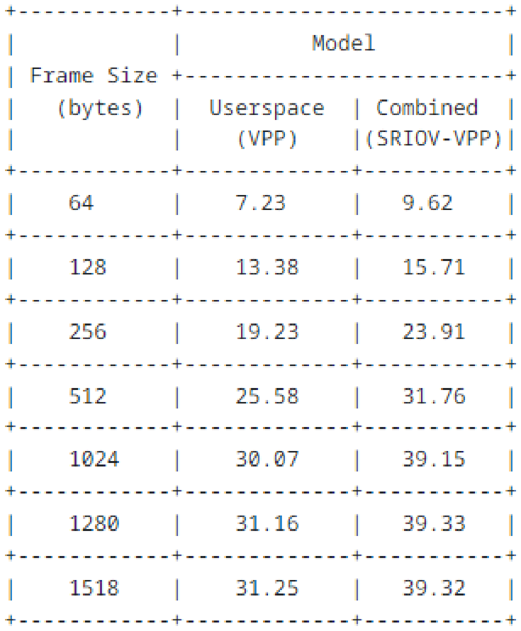
- Updated the benchmarking results with frame size values

(as in example figures here)



	NUMA Alignment Scenarios					
	s1	s2	s3	s4	s5	s6
Throughput	39.31	23.67	29.23	37.25	23.58	29.36

Figure 17: Different resource configurations 1518-byte packet size's zero packet loss throughput test result in single pod scenario (Gbps)



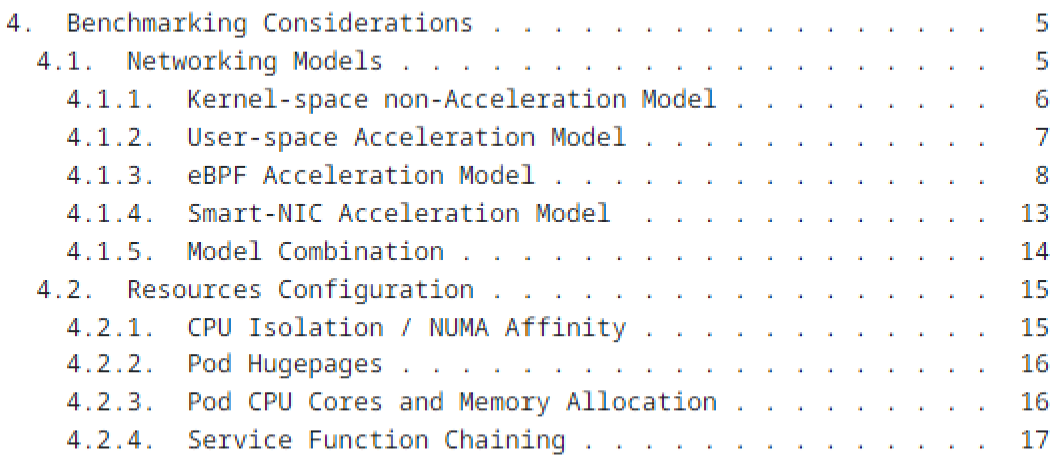
Frame Size	Model	
(bytes)	Userspace (VPP)	Combined (SRIOV-VPP)
64	7.23	9.62
128	13.38	15.71
256	19.23	23.91
512	25.58	31.76
1024	30.07	39.15
1280	31.16	39.33
1518	31.25	39.32

Figure 16: Networking Model Combination Zero Packet Loss Throughput Test Results (Gbps)

Detailed Updates (3)

Benchmarking Appendixes

Update our latest Benchmarking Results and re-organize to align with proposed consideration in draft



➡