CNTT and VSPERF

Testing and Testcases

The CNTT Reference Model specification contains some information and requirements where VSPERF (and other projects) need to respond and enhance their current capabilities to address the implied testing more fully.

For example, the current (9/11) reference model master .MD has:

4.1.3.1 Exposed performance metrics

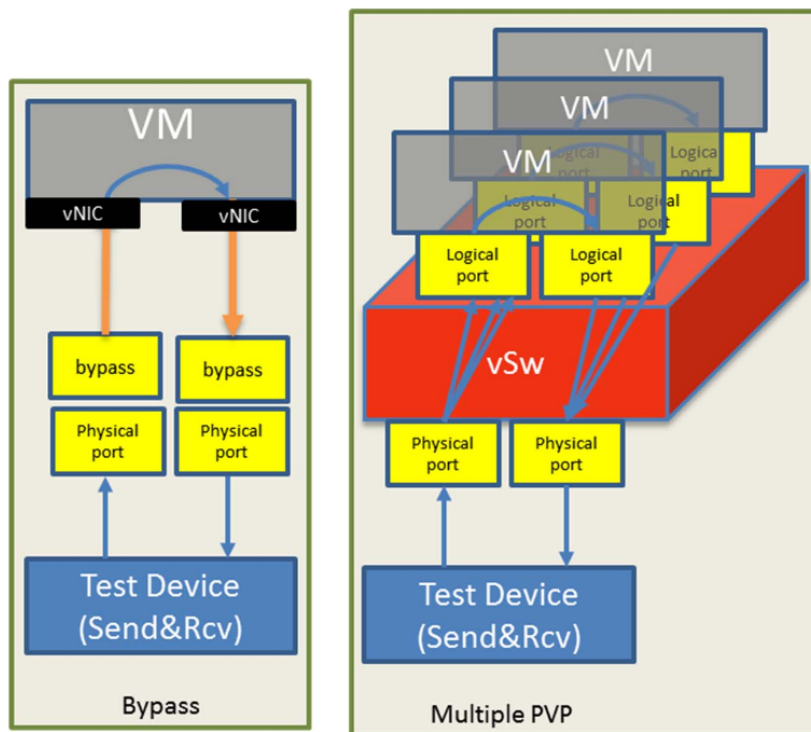The following shows performance metrics per VNF-C, vNIC or vCPU.

Ref	NFVI metric	Unit	Definition/Notes
e.nfvi.per.met.001	Network throughput	bits/s or packets/s	Max throughput per vNIC (as aligned with ETSI GS NFV-TST 009 [2])
e.nfvi.per.met.002	Network latency	second	Max round trip time to vNIC (as aligned with ETSI GS NFV-TST 009 [2])
e.nfvi.per.met.003	Network Delay Variation	second	Max packet delay variation (a.k.a., jitter) of round trip time to vNIC (as aligned with ETSI GS NFV-TST 009 [2])
e.nfvi.per.met.004	Simultaneous active flows	number	Max simultaneous active L4 flows per vNIC before a new flow is dropped
e.nfvi.per.met.005	New flows rate	flows/s	Max new L4 flow rate per vNIC
e.nfvi.per.met.006	Storage throughput	bytes/s or IO/s	Max throughput per virtual block storage unit assigned to VNF-C
e.nfvi.per.met.007	Processing capacity	test-specific	Processing capacity test-specific score per vCPU

Table 4-4: Exposed performance metrics of NFVI.

UPDATE: Oct 2, 2019   Note that Table 4-4 will likely move to Chapter 8 of the Reference Model, because it includes Benchmarks that are externally measured. This is certainly within the spirit of "exposed" metrics, but it seems that the contributors really want measurements to **monitor** the performance exposed to VNFs. However, each of the above metrics will need to be replaced by their **passively observable** counterparts, and many of these metrics will come from ETSI NFV TST008.

Per VNF-C, per vNIC, and per vCPU have implications on VSPERF testing features, and may require new methods in TST 009

For example, the Test Setups in TST 009 begin to address the needed Per-V* testing, in Section 6.2,

Figure 6.2-2: Additional Data plane Test Setups in a single host: Bypass and Multiple PVP

where the Multiple PVP configuration can be used to assess performance of 1, 2, 4, 8, 12, 16, parallel VMs (or Pods), and a characteristic curve can be provided using the same resources and TST001 workload profile as the VNF of interest (if this specialization is needed). This setup assumes that the Benchmarking Goals and Use cases (Section 6.1 of TST009) require pairs of ports and bidirectional testing. Single port unidirectional testing is also possible with Multiple PVP.

The Test Device/Traffic Generator needs to generate flows designed to pass through the vSwitch and reach the VM, where the processing needed to return the flows takes place.

vNICs themselves can be monitored for octets transferred, and ultimately the limiting bottleneck of each multiple parallel VM setup should be determined.

Below are some considerations and questions, where "the present document" refers to TST009:
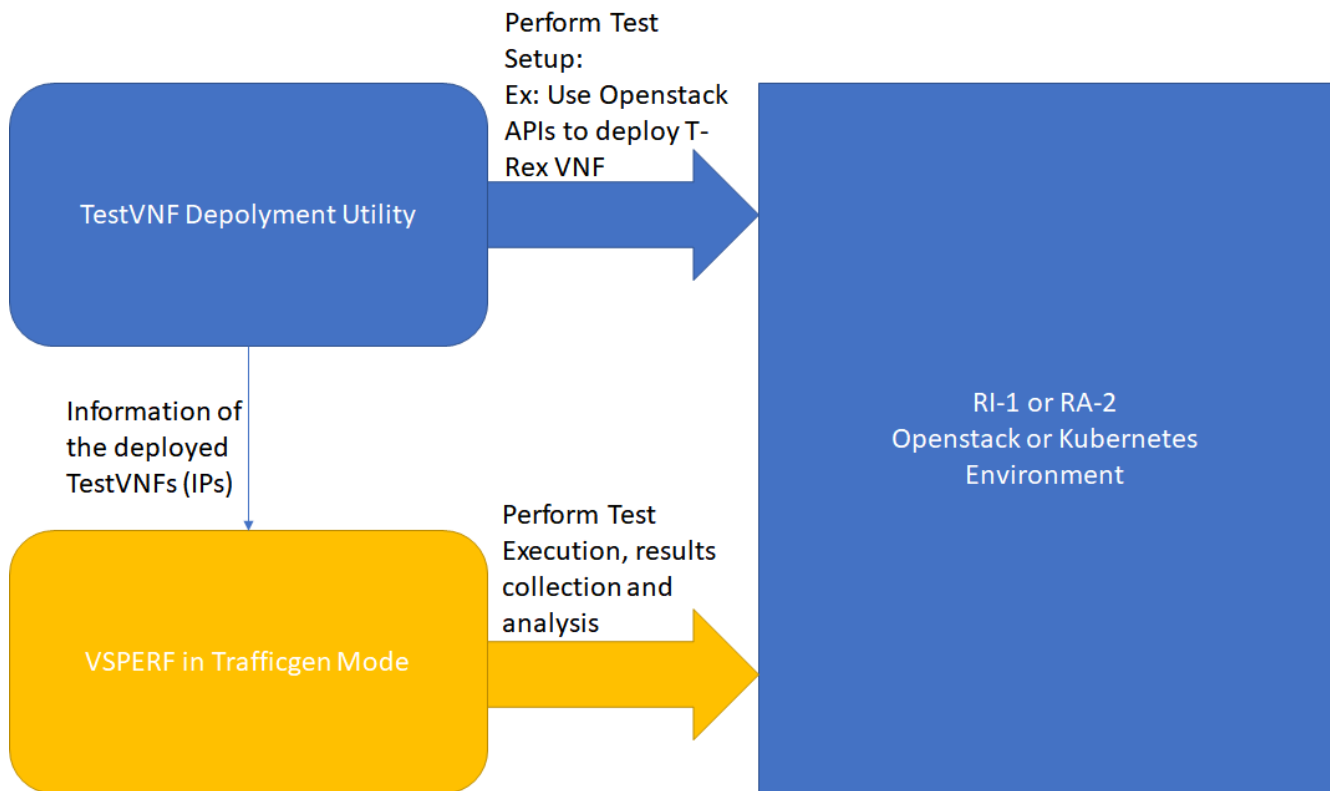
1. The number of VMs in the setup is a parameter for Throughput, Latency and Delay Variation.
2. The number of VMs in series on each parallel path through the DUT (PVVP means 2, see Figure 6.2-1 of the present document).
3. The number of simultaneous L2/L3/L4 flows is a parameter of Throughput, Latency, and Delay Variation benchmark testing. This implies a fourth dimension. Good to clarify this in the present doc.
4. All of the Test setups in clause 6.2 employ 2 physical NICs. Are there any single NIC or single vNIC VNFs envisioned (for "telco" deployment) that are critical/need to be tested immediately?
5. For the VMs themselves, should the testing employ reference VNFs/VNF-Cs with specific workload profiles, as described in ETSI NFV TST001, Clause 6 **Table 6.2 has VNF to Workload mapping**, See table 6.3 below for the **Workload to Metric** Categories (what to measure with reference VNF that match the Workload combinations, because "we" never get to test with the commercial VNFs so we need surrogates) :

Table 6.3: VNF workload operations and metric categories

	Compute Performance/ Speed	Compute Capacity/Scale	Compute Reliability/ Availability	Network Performance/ Speed	Network Capacity/Scale	Network Reliability/ Availability	Storage Performance/ Speed	Storage Capacity/ Scale	Storage Reliability/ Availability
Data TX/RX	✓			✓	✓	✓			
Data session initialization/ termination				✓	✓	✓	✓		
Signal processing	✓		✓						
Switching/ forwarding	✓		✓	✓	✓	✓			
Routing	✓	✓		✓	✓	✓			
Data session/flow accounting	✓	✓							
Pattern matching	✓	✓							
Encrypt / decrypt	✓	✓							
H-QoS	✓		✓	✓					
Encapsulate/ de-capsulate	✓	✓							
Compression	✓	✓							
Control session initialization/ termination	✓	✓			✓	✓			
Control session management	✓	✓	✓						
Access Control/ Authentication	✓	✓	✓						
Disk Read/Write		✓		✓			✓	✓	✓

- ETSI NFV INF010 recommended several metrics for the Infrastructure "service quality" provided to the VNFs instantiated there. This was an early ETSI NFV Group Specification (December 2014), and it summarizes the service quality metrics envisioned at that time in Table 1. One of the key metrics is VM Stalls, with event Frequency and Duration, which would be characterized using the methods described in clause 12.4 (Long Duration Testing) in the present document, but there is no(?) current CNTT RM requirement that matches this metric. Both ETSI NFV GS INF010 and GS TST001 envision metrics to capture this area of infrastructure behaviour.
- Power Efficiency appears in CNTT RM Table 4-10. There was IETF Benchmarking Methodology Working Group work on this topic, but it did not advance beyond the proposal stage: https://tools.ietf.org/html/draft-manral-bmwg-power-usage-04
- what else?

Test Setup

Containers - CNF

Networking Performance Benchmarking

Intern Project: Container Networking Testing and Benchmarking

VSPERF-Container Networking Benchmarking and Testing

Workload Characterization

Understanding the performance of containerized applications under various scenarios is critical for cloud native migration. One option is a black-box approach for profiling various (compute/network intensive) containerized applications under different CPU and Memory constraints. Profiling, refers to study how sensitive the container is to the shared resource availability. The approach is black-box because it does not rely on any specific adaptions or any agents within containers. The main components of this approach are Resource Limiter, Metrics collection, Data Analytics, and Traffic/Load-Gens. For every combinations of CPU and Memory, which are varied in pre-defined steps (10%, 25%, 50%, 75%, 90%), the sensitivity of the application container to this each combination is calculated. A model will be built to categorize the sensitivity of the application.

Prototype Test and Evaluation

Following the September 2019 CNTT workshop, it seemed agreeable to the OPNFV test project participants to conduct some tests on a surrogate for "reference implementation" using the tools we have, and what we can easily enhance with the CNTT Reference Model evaluations in-mind. The main goal is to look for GAPS between the Reference Model Requirements and the Test Tool Capabilities, decide the disposition/resolution of the Gaps (work to add test capability or de-scope some requirements if enough info is available). There is no schedule yet to perform this testing, but there do seem to resources coming available, and it seemed agreeable that the next meet-up would be a Hackfest for the OPNFV testing community.