WHITE PAPER

Communications Service Providers
Characterization and Benchmarking

Dell EMC* and Intel Demonstrate NSB for Next-Gen Data Centers

Network Services Benchmarking (NSB), part of the OPNFV Yardstick* project, allows evaluation of network functions virtualization infrastructure and virtual network functions to ensure conformance to performance requirements. Dell EMC and Intel devised a test using PROX* to show NSB capabilities using PowerEdge* Servers and Dell Networking Switches.

Introduction

Communications service providers (CommSPs) are expanding their use of network functions virtualization (NFV) because the decoupling of software from specialized hardware means lowered costs as fixed-function appliances are replaced by software running on Intel® architecture CPU-based servers such as the Dell EMC* PowerEdge* product line.

Some of the challenges of this new model have included a lack of service assurance and performance management capabilities and the lack of common standards and industry-accepted benchmarks for conformance to carrier-grade performance requirements. This has made it difficult to evaluate the NFV infrastructure (NFVI) or the virtual network functions (VNFs) for consistent performance.

In response to these challenges, the Linux Foundation* has started the Network Services Benchmarking (NSB) project under its Open Platform for NFV (OPNFV)* Yardstick* program to help CommSPs better dimension their virtual network workloads and determine the stress vectors. The use of NSB allows CommSPs to develop a common reference set of performance benchmarks for each class of VNFs.

Intel and its Intel® Network Builders ecosystem partner Dell EMC support the development of NSB and have conducted an in-depth test to detail how the testing tools work in a real deployment scenario on PowerEdge.

Introduction to Yardstick and NSB

NSB is a part of the OPNFV Yardstick project that offers a framework for characterization and benchmarking of VNFs, NFVI, and network services.

The goal of NSB is to extend the OPNFV Yardstick framework to perform real-world VNF and NFVI characterization and benchmarking with repeatable and deterministic methods. It allows users (service developers, VNF developers, NFVI developers, and others) to identify performance bottlenecks that may be encountered in VNF development or deployment to achieve overall system performance. By using NSB software and an external traffic generator, a service developer can benchmark and test many different parameters in an automated fashion, while keeping a record of all the results.

NSB can be used to characterize both VNF network performance metrics and NFVI resource utilization statistics under various configurations and user workloads. The NSB code base allows operators to gather key performance indicator (KPI) data such as:

Table of Contents

- NFVI KPIs: CPU statistics, memory bandwidth, Open vSwitch (OVS)*, and Data Plane Development Kit (DPDK) stats
- Network KPIs: Throughput, latency, jitter, session scale, and others
- VNF KPIs: Packets in, packets dropped, packets forwarded, and others

For both VNF and NFVI performance benchmarking, NSB can be applied in native Linux bare metal environments, standalone virtual environments (PCI pass-through, SR-IOV, OVS-DPDK), or in OpenStack* or other managed virtualized environments. NSB offers the capability to interact with both hardware and software-based external traffic generators for triggering and validating traffic patterns according to user-defined traffic profiles.

NSB Test: Hardware and Software

Of all of the operating environments, NSB is especially important to the performance benchmarking of NFV environments. Using virtualization adds performance overhead to systems, which makes it important for the CommSP to understand the impact of that overhead in terms of performance variability. For the tests described in this document, a special test VNF called Packet pROcessing eXecution engine (PROX) was used. PROX is part of the OPNVF's samplevnf project and utilizes DPDK. PROX implements a suite of test cases and displays the key metric measurements of the test suite on a Grafana* GUI dashboard.

Dell EMC PowerEdge* R740XD servers were used for this NSB test. These servers utilize the latest Intel® Xeon® Scalable processors. The 2U, 2-socket PowerEdge R740XD is based on a new architecture that features more storage with scalable storage performance and data set processing.

The servers feature embedded intelligence and automation from Dell EMC's integrated Dell EMC Remote Access Controller 9 (iDRAC9)* and other technologies within Dell EMC's OpenManage Server Administrator (OMSA)* portfolio, which monitors the health status of servers from Dell EMC. Combined, these management innovations simplify the server's lifecycle from deployment to retirement.

BIOS settings are important to the performance of the solution and in this test were configured for maximum performance. P-state and C-state have been disabled for this test. Intel® Hyper-Threading Technology (Intel® HT Technology) was enabled and power management was set to performance.

Software

The servers were deployed with Ubuntu* Server (v16.04) as the host operating system and OpenStack Pike* as the virtualization environment with libvirt* for virtual machine (VM) management and QEMU* for VM monitoring. The container environment was based on Kolla-Ansible* (5.0.3), with the NSB (commit id: 70fa1fcf) code running in its own container. Open vSwitch (OVS 2.9.3) facilitated the virtual network and delivered connectivity between the VNFs and the physical 25 G Ethernet ports. The DPDK 17.11.4 was utilized for enhanced packet throughput.

To get optimal performance, poll mode driver (PMD) cores for OVS were isolated; in addition, the interrupt request (IRQ) of the management interface was pinned to cores 0-3 (i.e., the non-isolated cores), and NUMA balancing was disabled.

OVS was configured to use four cores per each NUMA node (from the range isolated before), 1024 MB memory for DPDK from each NUMA node, and one receiving and one transmitting queue for each OVS port associated with physical interface. OVS rules were configured by OpenStack, resulting in many rules being configured (e.g., some Internet Control Message Protocol-related rules, rules for adding/stripping VLANs, etc.).

All VMs used the same OpenStack version, defined as one NUMA node, one CPU socket, 10 CPU cores, large memory page size, dedicated CPU policy, and isolated thread policy. Each VM was configured to use 10 virtual CPUs, pinned to 10 physical cores from the same NUMA node (i.e., VM1 only uses cores from NUMA 0 and VM2 from NUMA 1). Each VM was configured with 10 GB of memory. Within the VM, cores were isolated in a similar way as the host hypervisor, and eight times 1 GB huge pages were allocated.

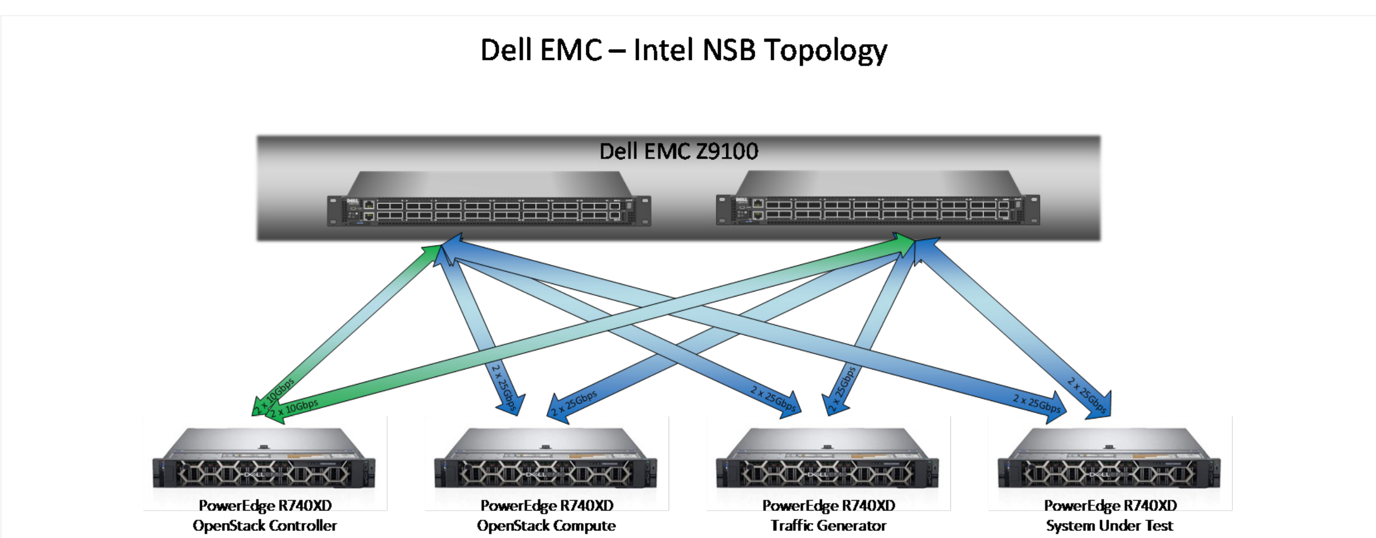

Figure 1. Dell EMC and Intel NSB Topology[1]

NSB Test: Set Up

The test required separate servers configured as a compute node, a control node, and a jump node. To simplify setup, the OpenStack infrastructure was deployed using Kolla-Ansible on the compute node, which was used for the performance testing with all PROX VNFs spawned on it. To avoid any influence on performance, all non-compute related tasks were handled by a separate controller node, which provided control plane services during the tests.

The compute and controller nodes and the NSB jump node were connected to the same Ethernet management switch. Additionally, the compute node was connected to a traffic generator using four 25 Gbps network interfaces that made up the data plane. The 4 x 25 GbE connection between the traffic generator and the compute node was realized by a set of Dell EMC Networking switches working in a spine-and-leaf architecture (Figure 1). The switches were divided into four VLANs, while an additional layer of switching was provided by OVS running on the compute node.

Optimization/Tuning

Optimizing and tuning NFV deployments is a common way to improve performance and was done for these NSB tests. Tuning can be done on various levels from hardware up to test case file configuration. Optimization parameters depend on many factors, and they can be part of NFVI or specific to VNFs being tested. Also, different traffic profiles require different optimization parameters in one or more of these areas.

Data Flows

The logical connection diagram and test traffic flow is depicted in Figure 2. The test generator created Ethernet/IP/UDP frames with one flow (one IP source, one IP destination, one UDP source and one UDP destination) per 25 GbE port. When the system under test (SUT) is configured with one receive queue per port, one flow per port usually gives better per-core performance; hence the one flow test is usually very useful to show the upper bound of the performance. Ideally, the test should be repeated with more flows (e.g., up to 100,000 flows) to reflect more realistic scenarios and measure how the performance is scaling with the number of cores.

The throughput reported was the maximum throughput with a tolerated packet loss of 0.001%. It was achieved by binary search; the generator maintained the load for 30 seconds; then it decreased the load when the amount of loss packets was higher than the tolerated loss, and increased the load when the amount of loss was lower than the tolerated loss.

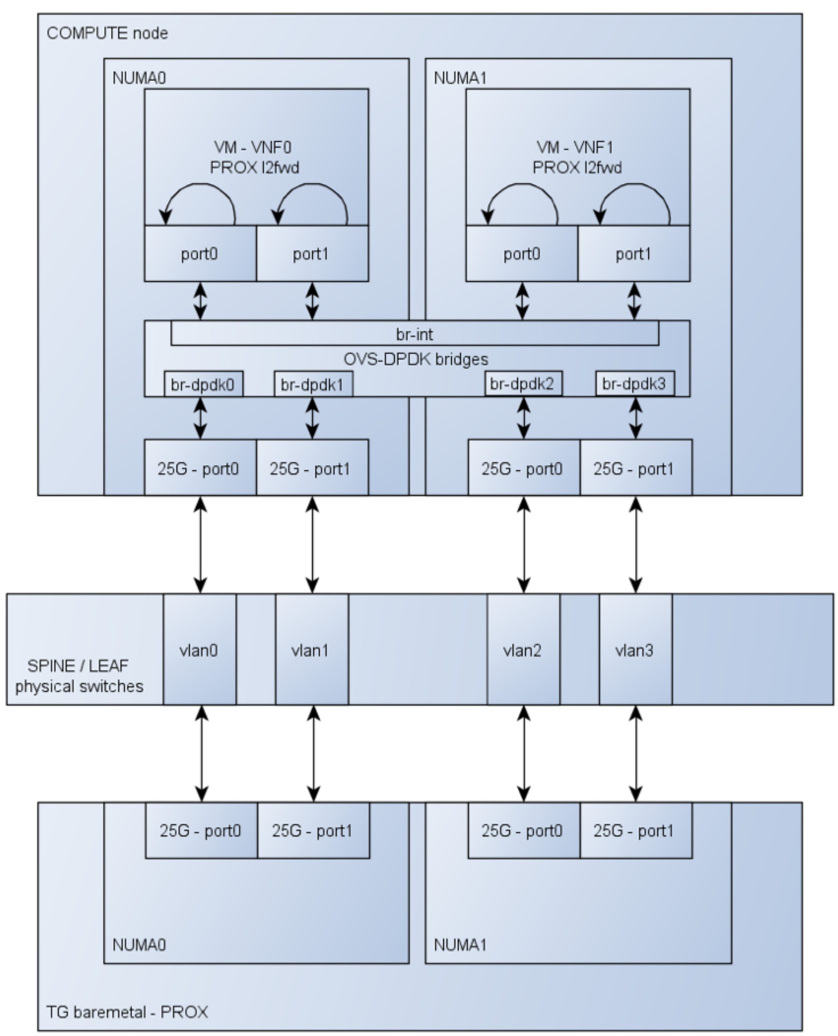

Figure 2. Logical connections between traffic generator and the OpenStack compute node.[1]

PACKET SIZE [BYTES]	LINE RATE [%]	DROP [%]	4-PORT AGGREGATE THROUGHPUT [MPPS]	THEORETICAL 4-PORT AGGREGATE MAX. THROUGHPUT [MPPS]
64	20.02	0.00046	29.79	148.81
128	31.64	0.00081	26.72	84.46
256	52.05	0.00068	23.57	45.29
512	80.96	0.00079	19.02	23.50
1024	89.55	0.00067	10.72	11.97
1280	97.17	0.00058	9.34	9.62
1518	96.68	0.00038	7.85	8.13

Table 1. NSB performance results for PROX l2fwd VNF in OpenStack environment[2]

Results of NSB Tests

Table 1 shows NSB performance results obtained for PROX l2fwd VNFs deployed in OpenStack environment. Traffic was generated on a bare metal host using four 25 Gbps ports. The throughput shown in the last column of Table 1 is a total value for all four ports. OVS is handling traffic in both directions on each of four ports; however, the numbers shown above represent only receiving port.

End results depend on many factors involving hardware, system architecture, network topology, software components, and traffic profile. Each factor or optimization area should be optimized in a way that considers the dependencies between them. This means that the optimization of one area is done with consideration of the performance of adjacent areas and how they influence each other.

The results presented in Table 1 show that the servers attained OVS throughput limits of 7.5 Mpps per port (aggregate throughput of 29.79 Mpps for four ports) for 64 byte frame size in each direction, or 20% of line rate.[2] At 1518 bytes, this aggregate data flow reached 7.85 Mpps, which is more than 96% of line rate.[2] Real network traffic is a mix of packet sizes that is often skewed to the larger packet sizes.

The performance was obtained utilizing a configuration without any manual OVS rules. Proper NUMA placement and vCPU pinning was key to achieving this throughput. Configuration and optimizations, such as IRQ pinning and disabling NUMA balancing, allow the data plane cores to be only minimally interrupted by these processes. Setting the number of descriptors in OVS, as well as the queue size in QEMU, helped in making sure that the packets received during those small interrupts could be properly buffered (instead of being lost). Hence the VNF could achieve the high throughput with a low number of dropped packets.

The results were obtained using eight OVS cores (four handling the four physical NIC ports, and four handling the four virtual ports). The test generator produced only one flow per physical NIC port. If multiple flows were received from the test generator, then the number of receive queues and the number of OVS cores could be increased to improve the performance.

NSB Optimization Parameters for Virtualization Contexts

The following list gives an overview of virtualization parameters that can be tuned for better performance:

- Open vSwitch (OVS)*
  - Poll mode driver cores
  - Packet queues
  - Descriptors
  - Rebalance
- QEMU/libvirt*
  - vCPU pinning
  - Queue size
  - Number of queues
  - NUMA placement
- Nova compute service
  - vCPU pinning
- Aggregation zones
  - VM disk image
  - Flavor

The use of OpenStack was also a factor as OpenStack configures OVS, resulting in more OVS rules than would be configured in a manual setup (manual setup usually contains eight rules for this kind of setup, connecting each physical port to a virtual port and back). These additional rules have a performance impact.

Hyperthreading and sibling CPU cores can give different results depending how they are used. There is always a dilemma as to how many CPU cores should be given to OVS. Better network performance comes from dedicating many cores to OVS, but this means fewer cores will be available for other virtual machines. For this test, four physical cores out of 16 were used per NUMA node in order to achieve a balance of good performance and enough available cores for other VNFs. In the case where more cores were assigned to the VMs, it would be possible to configure OVS with only four physical CPU cores (two per NUMA node) handling eight OVS ports. In this application, enabling hyperthreading can provide additional performance gains.

Conclusion

NSB can be used to measure NFVI performance using simple network-centric workloads, and to compare the performance impacts from different optimization options. Utilizing NSB and proper configuration, these tests show that high network throughput was achieved on Dell EMC PowerEdge R740XD using Intel Xeon Scalable processors.

Learn More

Read more about Network Services Benchmarking in the white paper "Network Services Benchmarking: Accelerating the Virtualization of the Network": https://builders.intel.com/docs/networkbuilders/network-services-benchmarking-accelerating-the-virtualization-of-the-network.pdf

About Dell EMC

Dell EMC, a part of Dell Technologies, enables organizations to modernize, automate, and transform their data center using industry-leading converged infrastructure, servers, storage and data protection technologies. This provides a trusted foundation for businesses to transform IT, through the creation of a hybrid cloud, and transform their business through the creation of cloud-native applications and big data solutions. Dell EMC services customers across 180 countries—including 99 percent of the Fortune 500—with the industry's most comprehensive and innovative portfolio from edge to core to cloud.

About® Intel Network Builders

Intel Network Builders is an ecosystem of infrastructure, software, and technology vendors coming together with communications service providers and end users to accelerate the adoption of solutions based on network functions virtualization (NFV) and software defined networking (SDN) in telecommunications and data center networks. The program offers technical support, matchmaking, and co-marketing opportunities to help facilitate joint collaboration through to the trial and deployment of NFV and SDN solutions. Learn more at http://networkbuilders.intel.com.

[1] Figure provided courtesy of Dell EMC.

[2] Testing conducted by Dell EMC as of Oct. 2018. Configurations: The bare metal server was a Dell EMC PowerEdge R740XD with 192GB of RAM and two 16-cores Intel Xeon Gold 6130 CPU (32 logical cores per server) running at 2.10 GHz. Servers were equipped with 2x Intel® Ethernet Network Adapters XXV710 25 GbE. The jump node server was a VMware ESXi* virtual machine using four cores on a server with an Intel Xeon Gold 6130 processor running at 2.10 GHz with 4 GB of memory. The software environment included Ubuntu Server (v16.04), OpenStack Pike, Kolla-Ansible (5.0.3), NSB (commit id: 70fa1fcf), DPDK (17.11.4), and Open vSwitch (2.9.3). CPU microcode: 0xb00002e.