VSPERF Test Specification

Release 1.0.0

OPNFV

November 26, 2015

CONTENTS

CHAPTER

ONE

CHARACTERIZE VSWITCH PERFORMANCE FOR TELCO NFV USE CASES LEVEL TEST DESIGN

Table of Contents

- CHARACTERIZE VSWITCH PERFORMANCE FOR TELCO NFV USE CASES LEVEL TEST DESIGN
- 1. Introduction
  - 1.1. Document identifier
  - 1.2. Scope
  - 1.3. References
- 2. Details of the Level Test Design
  - 2.1. Features to be tested
- 2.2. Approach
  - 2.2.1 Test Categories
  - 2.2.2 Deployment Scenarios
    - Physical port → vSwitch → physical port
    - Physical port → vSwitch → VNF → vSwitch → physical port
    - Physical port → vSwitch → VNF → vSwitch → VNF → vSwitch → physical port
    - Physical port → vSwitch → VNF
    - VNF → vSwitch → physical port
    - VNF → vSwitch → VNF → vSwitch
    - HOST 1(Physical port → virtual switch → VNF → virtual switch →
    - Physical port) → HOST 2(Physical port → virtual switch → VNF →
    - virtual switch → Physical port)
  - 2.2.3 General Methodology:
    - 2.2.3.1 Default Test Parameters
    - 2.2.3.2 Flow Classification
    - 2.2.3.3 Test Priority
    - 2.2.3.4 SUT Setup
      - 2.2.3.4.1 Port Configuration
      - 2.2.3.4.2 Frame Formats
      - Frame formats Layer 2 (data link layer) protocols
      - Layer 3 (network layer) protocols
      - Layer 4 (transport layer) protocols
      - Layer 5 (application layer) protocols
      - 2.2.3.4.3 Packet Throughput
      - 2.2.3.4.4 System isolation and validation
      - System Isolation
      - System Validation
      - Methodology to benchmark Platform/VNF forwarding capability
    - 2.2.4 RFCs for testing virtual switch performance
      - RFC 1242 Benchmarking Terminology for Network Interconnection
      - RFC 2544 Benchmarking Methodology for Network Interconnect Devices
      - RFC 2285 Benchmarking Terminology for LAN Switching Devices
      - RFC 2889 Benchmarking Methodology for LAN Switching
      - RFC 3918 Methodology for IP Multicast Benchmarking
      - RFC 4737 Packet Reordering Metrics
      - RFC 5481 Packet Delay Variation Applicability Statement
      - RFC 6201 Device Reset Characterization
    - 2.2.5 Details of the Test Report
  - 2.3. Test identification
    - 2.3.1 Throughput tests
      - Test ID: LTD.Throughput.RFC2544.PacketLossRatio
      - Test ID: LTD.Throughput.RFC2544.PacketLossRatioFrameModification
      - Test ID: LTD.Throughput.RFC2544.Profile
      - Test ID: LTD.Throughput.RFC2544.SystemRecoveryTime
      - Test ID: LTD.Throughput.RFC2544.BackToBackFrames
      - Test ID: LTD.Throughput.RFC2889.MaxForwardingRateSoak
      - Test ID: LTD.Throughput.RFC2889.MaxForwardingRateSoakFrameModification
      - Test ID: LTD.Throughput.RFC2889.PacketDelayVariation
      - Test ID: LTD.Throughput.RFC2889.MaxForwardingRate
      - Test ID: LTD.Throughput.RFC2889.ForwardPressure
      - Test ID: LTD.Throughput.RFC2889.ErrorFramesFiltering

CHAPTER

TWO

1. INTRODUCTION

The objective of the OPNFV project titled **"Characterize vSwitch Performance for Telco NFV Use Cases"**, is to evaluate a virtual switch to identify its suitability for a Telco Network Function Virtualization (NFV) environment. The intention of this Level Test Design (LTD) document is to specify the set of tests to carry out in order to objectively measure the current characteristics of a virtual switch in the Network Function Virtualization Infrastructure (NFVI) as well as the test pass criteria. The detailed test cases will be defined in *Section 2*, preceded by the *Document identifier* and the *Scope*.

This document is currently in draft form.

2.1 1.1. Document identifier

The document id will be used to uniquely identify versions of the LTD. The format for the document id will be: OPNFV_vswitchperf_LTD_ver_NUM_MONTH_YEAR_STATUS, where by the status is one of: draft, reviewed, corrected or final. The document id for this version of the LTD is: OPNFV_vswitchperf_LTD_ver_1.6_Jan_15_DRAFT.

2.2 1.2. Scope

The main purpose of this project is to specify a suite of performance tests in order to objectively measure the current packet transfer characteristics of a virtual switch in the NFVI. The intent of the project is to facilitate testing of any virtual switch. Thus, a generic suite of tests shall be developed, with no hard dependencies to a single implementation. In addition, the test case suite shall be architecture independent.

The test cases developed in this project shall not form part of a separate test framework, all of these tests may be inserted into the Continuous Integration Test Framework and/or the Platform Functionality Test Framework - if a vSwitch becomes a standard component of an OPNFV release.

2.3 1.3. References

- RFC 1242 Benchmarking Terminology for Network Interconnection Devices
- RFC 2544 Benchmarking Methodology for Network Interconnect Devices
- RFC 2285 Benchmarking Terminology for LAN Switching Devices
- RFC 2889 Benchmarking Methodology for LAN Switching Devices
- RFC 3918 Methodology for IP Multicast Benchmarking
- RFC 4737 Packet Reordering Metrics

- RFC 5481 Packet Delay Variation Applicability Statement
- RFC 6201 Device Reset Characterization

CHAPTER

THREE

2. DETAILS OF THE LEVEL TEST DESIGN

This section describes the features to be tested (*cf. 2.1*), the test approach (*cf. 2.2*); it also identifies the sets of test cases or scenarios (*cf. 2.3*) along with the pass/fail criteria (*cf. 2.4*) and the test deliverables (*cf. 2.5*).

3.1 2.1. Features to be tested

Characterizing virtual switches (i.e. Device Under Test (DUT) in this document) includes measuring the following performance metrics:

- **Throughput** as defined by RFC1242: The maximum rate at which **none** of the offered frames are dropped by the DUT. The maximum frame rate and bit rate that can be transmitted by the DUT without any error should be recorded. Note there is an equivalent bit rate and a specific layer at which the payloads contribute to the bits. Errors and improperly formed frames or packets are dropped.
- **Packet delay** introduced by the DUT and its cumulative effect on E2E networks. Frame delay can be measured equivalently.
- **Packet delay variation**: measured from the perspective of the VNF/application. Packet delay variation is sometimes called "jitter". However, we will avoid the term "jitter" as the term holds different meaning to different groups of people. In this document we will simply use the term packet delay variation. The preferred form for this metric is the PDV form of delay variation defined in RFC5481. The most relevant measurement of PDV considers the delay variation of a single user flow, as this will be relevant to the size of end-system buffers to compensate for delay variation. The measurement system's ability to store the delays of individual packets in the flow of interest is a key factor that determines the specific measurement method. At the outset, it is ideal to view the complete PDV distribution. Systems that can capture and store packets and their delays have the freedom to calculate the reference minimum delay and to determine various quantiles of the PDV distribution accurately (in post-measurement processing routines). Systems without storage must apply algorithms to calculate delay and statistical measurements on the fly. For example, a system may store temporary estimates of the mimimum delay and the set of (100) packets with the longest delays during measurement (to calculate a high quantile, and update these sets with new values periodically. In some cases, a limited number of delay histogram bins will be available, and the bin limits will need to be set using results from repeated experiments. See section 8 of RFC5481.
- **Packet loss** (within a configured waiting time at the receiver): All packets sent to the DUT should be accounted for.
- **Burst behaviour**: measures the ability of the DUT to buffer packets.
- **Packet re-ordering**: measures the ability of the device under test to maintain sending order throughout transfer to the destination.
- **Packet correctness**: packets or Frames must be well-formed, in that they include all required fields, conform to length requirements, pass integrity checks, etc.

- **Availability and capacity** of the DUT i.e. when the DUT is fully "up" and connected:
  - Includes power consumption of the CPU (in various power states) and system.
  - Includes CPU utilization.
  - Includes the number of NIC interfaces supported.
  - Includes headroom of VM workload processing cores (i.e. available for applications).

CHAPTER

FOUR

2.2. APPROACH

In order to determine the packet transfer characteristics of a virtual switch, the tests will be broken down into the following categories:

4.1 2.2.1 Test Categories

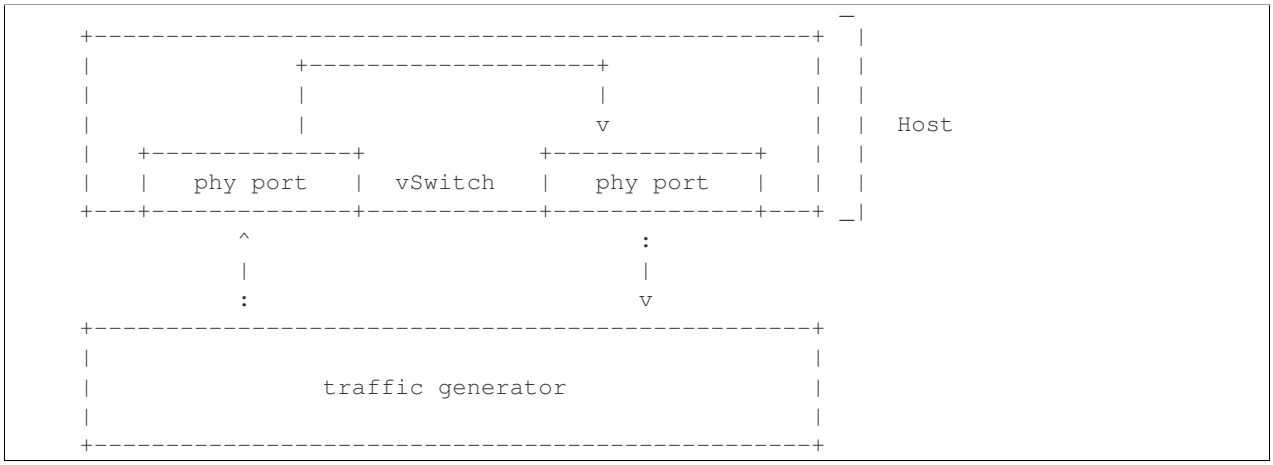
- **Throughput Tests** to measure the maximum forwarding rate (in frames per second or fps) and bit rate (in Mbps) for a constant load (as defined by RFC1242) without traffic loss.
- **Packet and Frame Delay Tests** to measure average, min and max packet and frame delay for constant loads.
- **Stream Performance Tests** (TCP, UDP) to measure bulk data transfer performance, i.e. how fast systems can send and receive data through the virtual switch.
- **Request/Response Performance** Tests (TCP, UDP) the measure the transaction rate through the virtual switch.
- **Packet Delay Tests** to understand latency distribution for different packet sizes and over an extended test run to uncover outliers.
- **Scalability Tests** to understand how the virtual switch performs as the number of flows, active ports, complexity of the forwarding logic's configuration... it has to deal with increases.
- **Control Path and Datapath Coupling** Tests, to understand how closely coupled the datapath and the control path are as well as the effect of this coupling on the performance of the DUT.
- **CPU and Memory Consumption Tests** to understand the virtual switch's footprint on the system, this includes:
  - CPU utilization
  - Cache utilization
  - Memory footprint
  - Time To Establish Flows Tests.
- **Noisy Neighbour Tests**, to understand the effects of resource sharing on the performance of a virtual switch.

Note: some of the tests above can be conducted simultaneously where the combined results would be insightful, for example Packet/Frame Delay and Scalability.

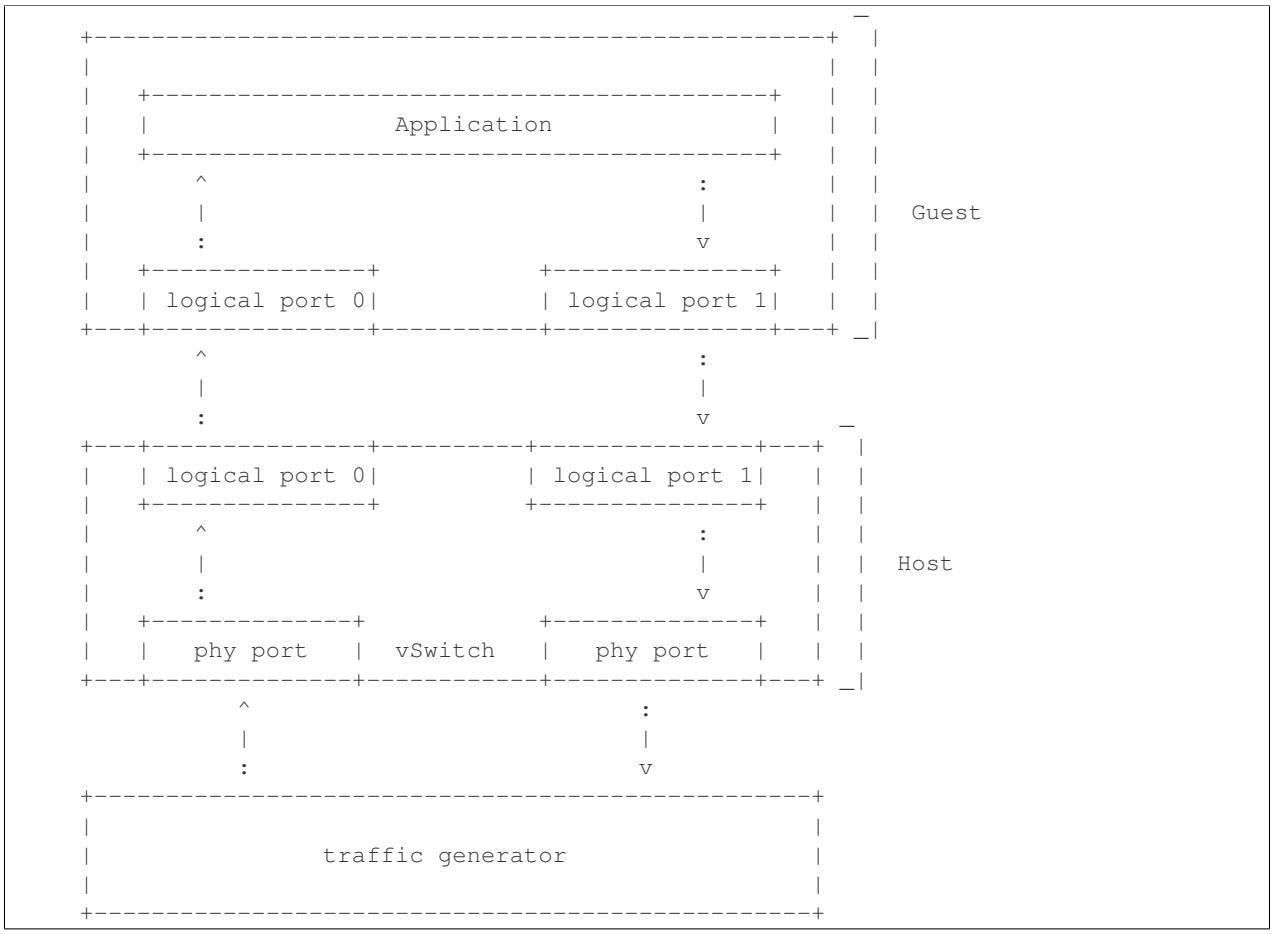
4.2 2.2.2 Deployment Scenarios

The following represents possible deployments which can help to determine the performance of both the virtual switch and the datapath into the VNF:

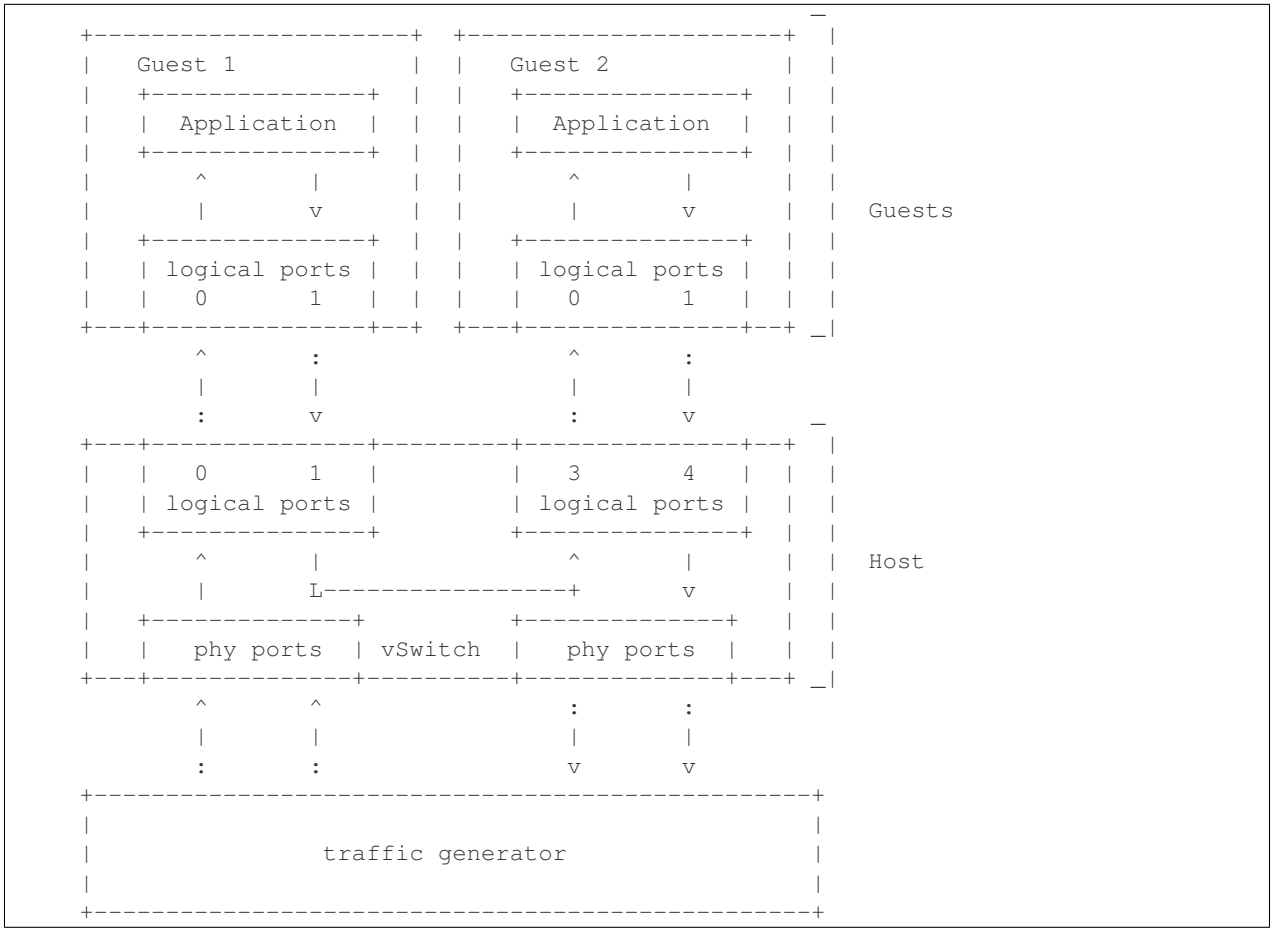
4.2.1 Physical port → vSwitch → physical port



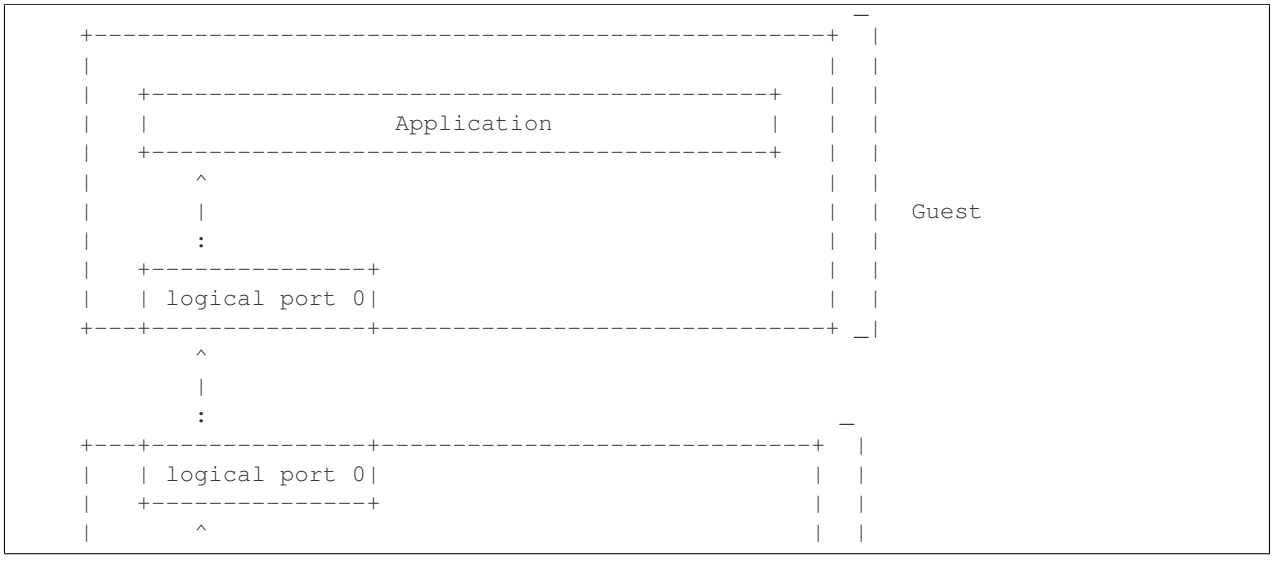
4.2.2 Physical port → vSwitch → VNF → vSwitch → physical port

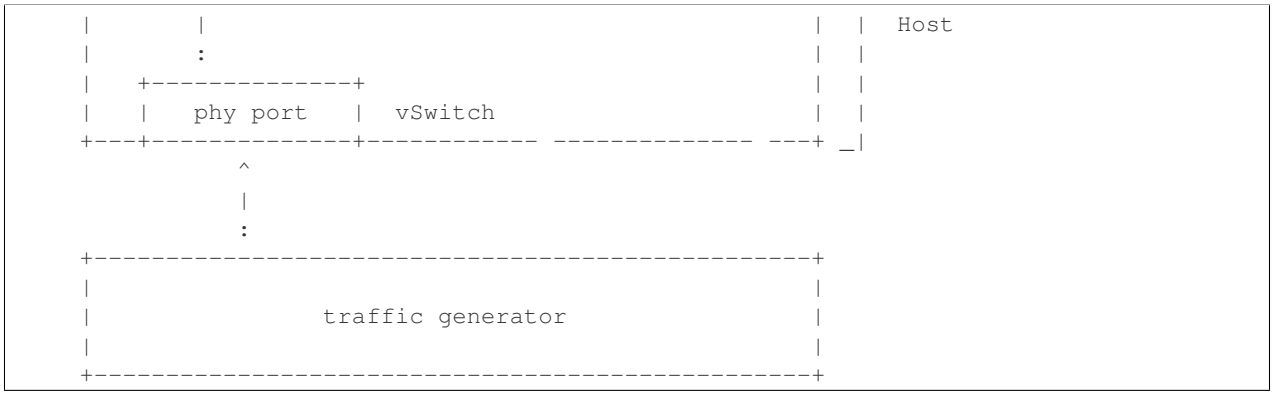


4.2.3 Physical port → vSwitch → VNF → vSwitch → VNF → vSwitch → physical port

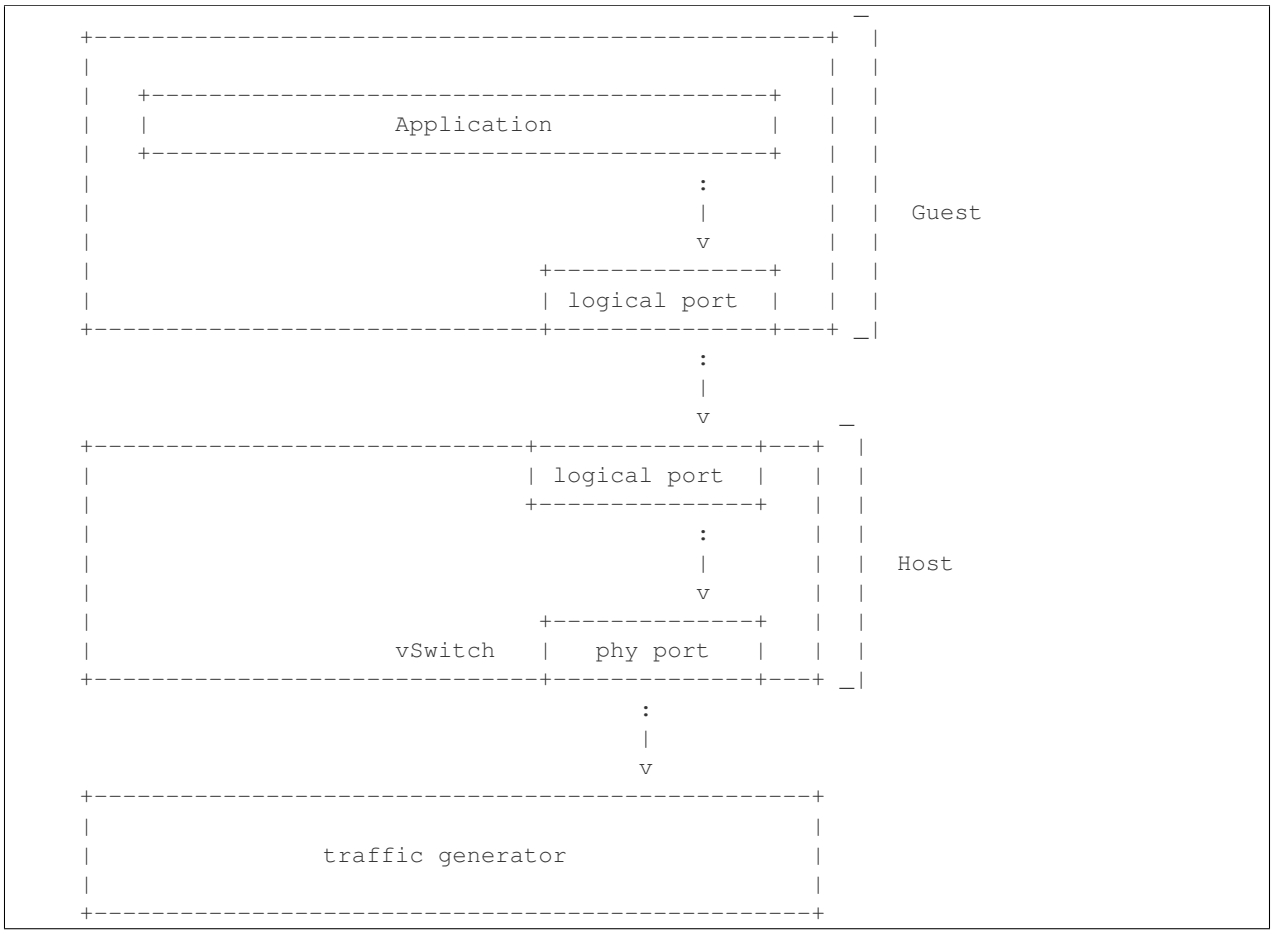


4.2.4 Physical port → vSwitch → VNF

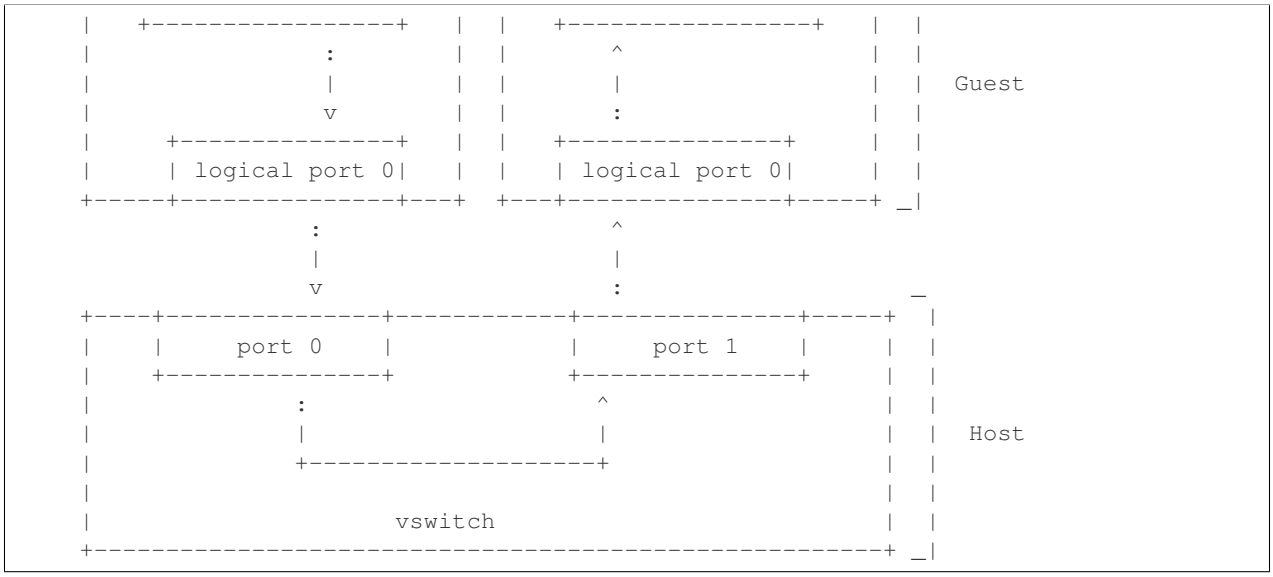




4.2.5 VNF → vSwitch → physical port



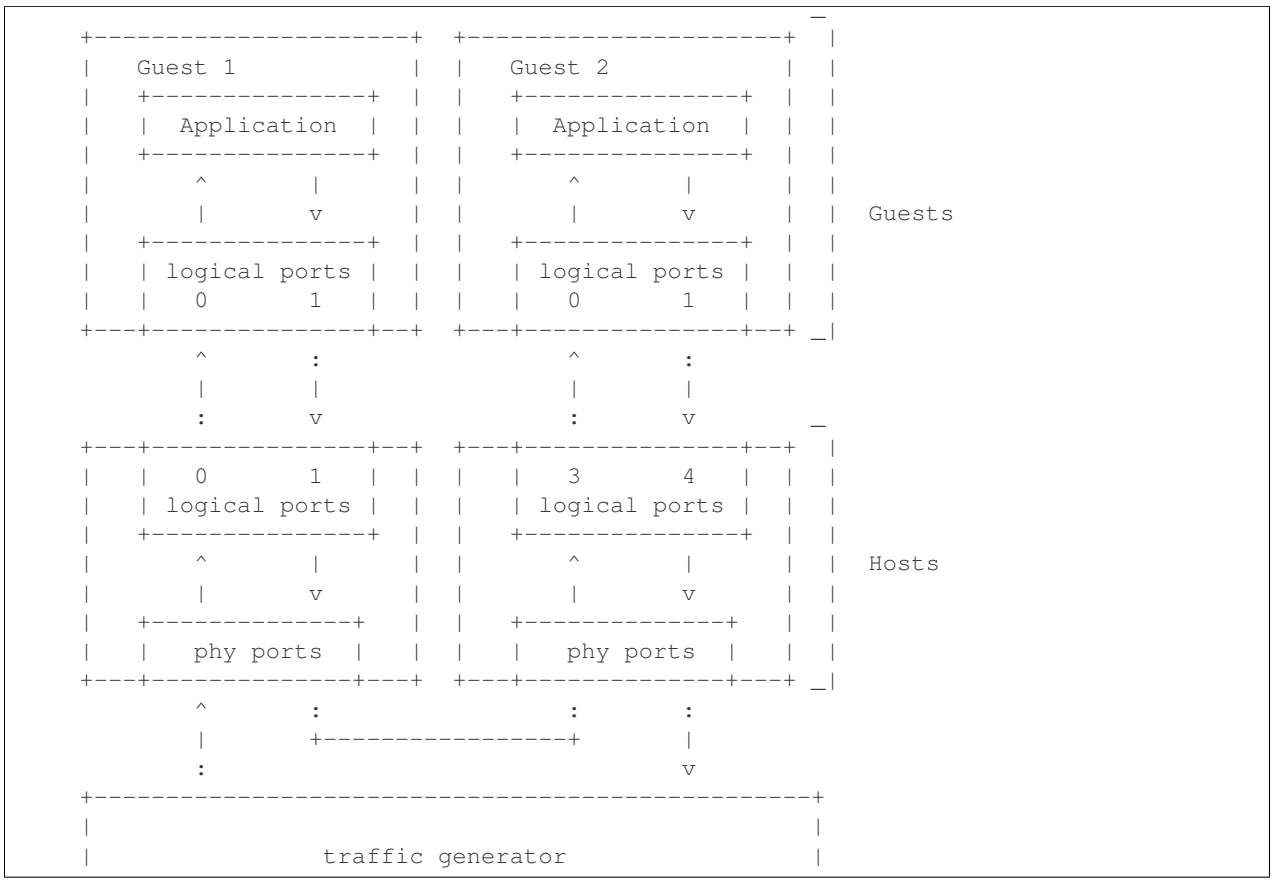
4.2.6 VNF → vSwitch → VNF → vSwitch



4.2.7 HOST 1(Physical port → virtual switch → VNF → virtual switch →

4.2.8 Physical port) → HOST 2(Physical port → virtual switch → VNF →

4.2.9 virtual switch → Physical port)



```
    |                                                   |
    +---------------------------------------------------+
```

Note: For tests where the traffic generator and/or measurement receiver are implemented on VM and connected to the virtual switch through vNIC, the issues of shared resources and interactions between the measurement devices and the device under test must be considered.

Note: Some RFC 2889 tests require a full-mesh sending and receiving pattern involving more than two ports. This possibility is illustrated in the Physical port → vSwitch → VNF → vSwitch → VNF → vSwitch → physical port diagram above (with 2 sending and 2 receiving ports, though all ports could be used bi-directionally).

Note: When Deployment Scenarios are used in RFC 2889 address learning or cache capacity testing, an additional port from the vSwitch must be connected to the test device. This port is used to listen for flooded frames.

4.3 2.2.3 General Methodology:

To establish the baseline performance of the virtual switch, tests would initially be run with a simple workload in the VNF (the recommended simple workload VNF would be DPDK's testpmd application forwarding packets in a VM or vloop_vnf a simple kernel module that forwards traffic between two network interfaces inside the virtualized environment while bypassing the networking stack). Subsequently, the tests would also be executed with a real Telco workload running in the VNF, which would exercise the virtual switch in the context of higher level Telco NFV use cases, and prove that its underlying characteristics and behaviour can be measured and validated. Suitable real Telco workload VNFs are yet to be identified.

4.3.1 2.2.3.1 Default Test Parameters

The following list identifies the default parameters for suite of tests:

- Reference application: Simple forwarding or Open Source VNF.
- Frame size (bytes): 64, 128, 256, 512, 1024, 1280, 1518, 2K, 4k OR Packet size based on use-case (e.g. RTP 64B, 256B) OR Mix of packet sizes as maintained by the Functest project <https://wiki.opnfv.org/traffic_profile_management>.
- Reordering check: Tests should confirm that packets within a flow are not reordered.
- Duplex: Unidirectional / Bidirectional. Default: Full duplex with traffic transmitting in both directions, as network traffic generally does not flow in a single direction. By default the data rate of transmitted traffic should be the same in both directions, please note that asymmetric traffic (e.g. downlink-heavy) tests will be mentioned explicitly for the relevant test cases.
- Number of Flows: Default for non scalability tests is a single flow. For scalability tests the goal is to test with maximum supported flows but where possible will test up to 10 Million flows. Start with a single flow and scale up. By default flows should be added sequentially, tests that add flows simultaneously will explicitly call out their flow addition behaviour. Packets are generated across the flows uniformly with no burstiness.
- Traffic Types: UDP, SCTP, RTP, GTP and UDP traffic.
- Deployment scenarios are:
- Physical → virtual switch → physical.
- Physical → virtual switch → VNF → virtual switch → physical.
- Physical → virtual switch → VNF → virtual switch → VNF → virtual switch → physical.
- Physical → virtual switch → VNF.

- VNF → virtual switch → Physical.
- VNF → virtual switch → VNF.

Tests MUST have these parameters unless otherwise stated. **Test cases with non default parameters will be stated explicitly**.

Note: For throughput tests unless stated otherwise, test configurations should ensure that traffic traverses the installed flows through the virtual switch, i.e. flows are installed and have an appropriate time out that doesn't expire before packet transmission starts.

4.3.2 2.2.3.2 Flow Classification

Virtual switches classify packets into flows by processing and matching particular header fields in the packet/frame and/or the input port where the packets/frames arrived. The vSwitch then carries out an action on the group of packets that match the classification parameters. Thus a flow is considered to be a sequence of packets that have a shared set of header field values or have arrived on the same port and have the same action applied to them. Performance results can vary based on the parameters the vSwitch uses to match for a flow. The recommended flow classification parameters for L3 vSwitch performance tests are: the input port, the source IP address, the destination IP address and the Ethernet protocol type field. It is essential to increase the flow time-out time on a vSwitch before conducting any performance tests that do not measure the flow set-up time. Normally the first packet of a particular flow will install the flow in the vSwitch which adds an additional latency, subsequent packets of the same flow are not subject to this latency if the flow is already installed on the vSwitch.

4.3.3 2.2.3.3 Test Priority

Tests will be assigned a priority in order to determine which tests should be implemented immediately and which tests implementations can be deferred.

Priority can be of following types: - Urgent: Must be implemented immediately. - High: Must be implemented in the next release. - Medium: May be implemented after the release. - Low: May or may not be implemented at all.

4.3.4 2.2.3.4 SUT Setup

The SUT should be configured to its "default" state. The SUT's configuration or set-up must not change between tests in any way other than what is required to do the test. All supported protocols must be configured and enabled for each test set up.

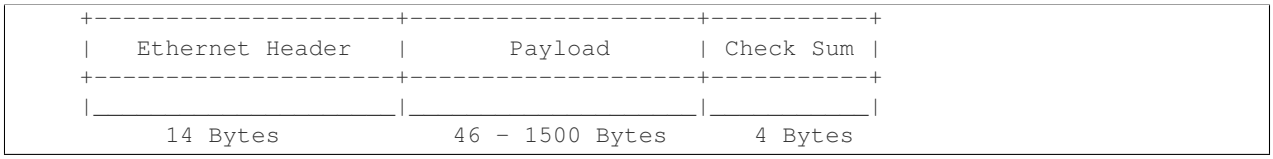
2.2.3.4.1 Port Configuration

The DUT should be configured with n ports where n is a multiple of 2. Half of the ports on the DUT should be used as ingress ports and the other half of the ports on the DUT should be used as egress ports. Where a DUT has more than 2 ports, the ingress data streams should be set-up so that they transmit packets to the egress ports in sequence so that there is an even distribution of traffic across ports. For example, if a DUT has 4 ports 0(ingress), 1(ingress), 2(egress) and 3(egress), the traffic stream directed at port 0 should output a packet to port 2 followed by a packet to port 3. The traffic stream directed at port 1 should also output a packet to port 2 followed by a packet to port 3.

2.2.3.4.2 Frame Formats

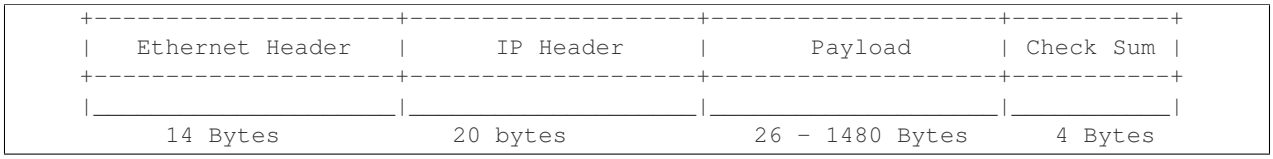
Frame formats Layer 2 (data link layer) protocols

- Ethernet II

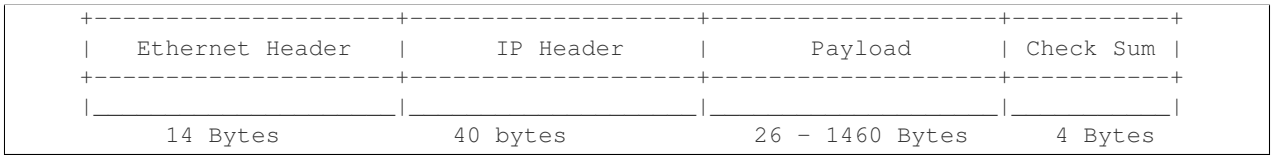


Layer 3 (network layer) protocols

- IPv4

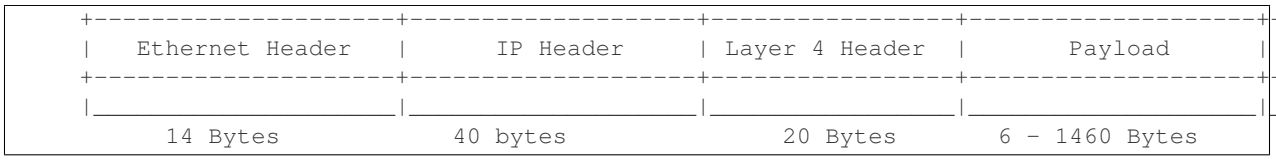


- IPv6



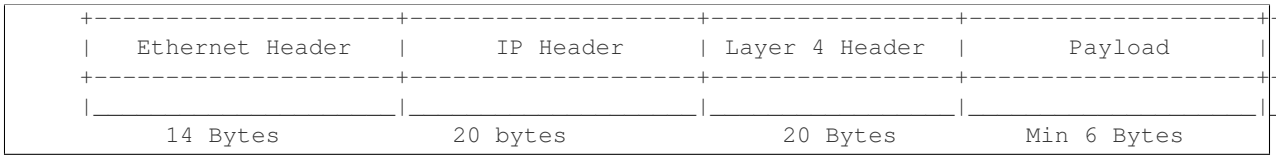
Layer 4 (transport layer) protocols

- TCP
- UDP
- SCTP



Layer 5 (application layer) protocols

- RTP
- GTP



2.2.3.4.3 Packet Throughput

There is a difference between an Ethernet frame, an IP packet, and a UDP datagram. In the seven-layer OSI model of computer networking, packet refers to a data unit at layer 3 (network layer). The correct term for a data unit at layer 2 (data link layer) is a frame, and at layer 4 (transport layer) is a segment or datagram.

Important concepts related to 10GbE performance are frame rate and throughput. The MAC bit rate of 10GbE, defined in the IEEE standard 802 .3ae, is 10 billion bits per second. Frame rate is based on the bit rate and frame format definitions. Throughput, defined in IETF RFC 1242, is the highest rate at which the system under test can forward the offered load, without loss.

The frame rate for 10GbE is determined by a formula that divides the 10 billion bits per second by the preamble + frame length + inter-frame gap.

The maximum frame rate is calculated using the minimum values of the following parameters, as described in the IEEE 802 .3ae standard:

- Preamble: 8 bytes * 8 = 64 bits
- Frame Length: 64 bytes (minimum) * 8 = 512 bits
- Inter-frame Gap: 12 bytes (minimum) * 8 = 96 bits

Therefore, Maximum Frame Rate (64B Frames) = MAC Transmit Bit Rate / (Preamble + Frame Length + Inter-frame Gap) = 10,000,000,000 / (64 + 512 + 96) = 10,000,000,000 / 672 = 14,880,952.38 frame per second (fps)

2.2.3.4.4 System isolation and validation

A key consideration when conducting any sort of benchmark is trying to ensure the consistency and repeatability of test results between runs. When benchmarking the performance of a virtual switch there are many factors that can affect the consistency of results. This section describes these factors and the measures that can be taken to limit their effects. In addition, this section will outline some system tests to validate the platform and the VNF before conducting any vSwitch benchmarking tests.

System Isolation

When conducting a benchmarking test on any SUT, it is essential to limit (and if reasonable, eliminate) any noise that may interfere with the accuracy of the metrics collected by the test. This noise may be introduced by other hardware or software (OS, other applications), and can result in significantly varying performance metrics being collected between consecutive runs of the same test. In the case of characterizing the performance of a virtual switch, there are a number of configuration parameters that can help increase the repeatability and stability of test results, including:

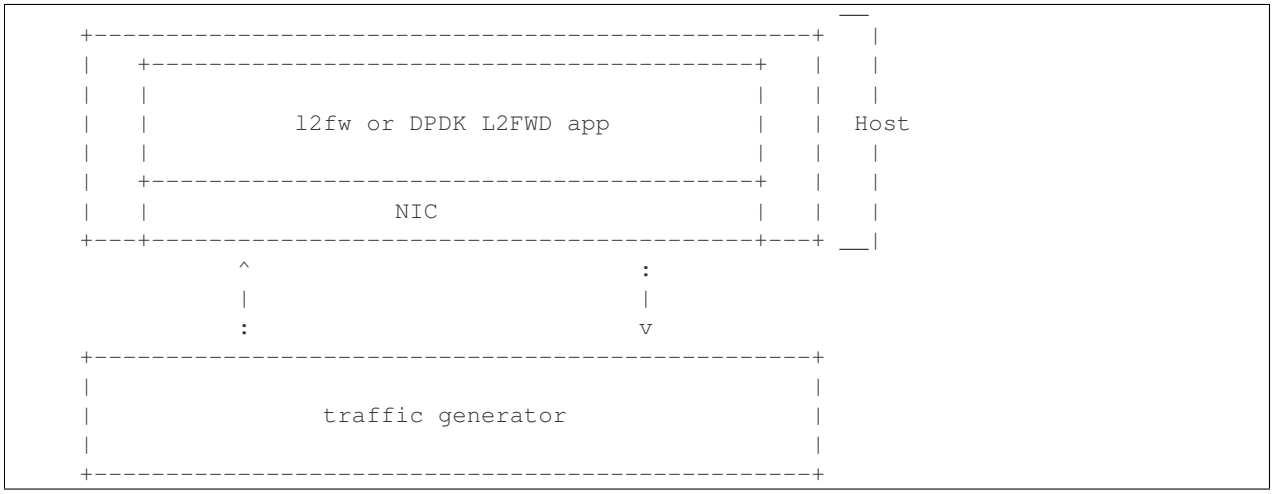
- OS/GRUB configuration:
  - maxcpus = n where n >= 0; limits the kernel to using 'n' processors. Only use exactly what you need.
  - isolcpus: Isolate CPUs from the general scheduler. Isolate all CPUs bar one which will be used by the OS.
  - use taskset to affinitize the forwarding application and the VNFs onto isolated cores. VNFs and the vSwitch should be allocated their own cores, i.e. must not share the same cores. vCPUs for the VNF should be affinitized to individual cores also.
  - Limit the amount of background applications that are running and set OS to boot to runlevel 3. Make sure to kill any unnecessary system processes/daemons.
  - Only enable hardware that you need to use for your test – to ensure there are no other interrupts on the system.
  - Configure NIC interrupts to only use the cores that are not allocated to any other process (VNF/vSwitch).
- NUMA configuration: Any unused sockets in a multi-socket system should be disabled.
- CPU pinning: The vSwitch and the VNF should each be affinitized to separate logical cores using a combination of maxcpus, isolcpus and taskset.

- BIOS configuration: BIOS should be configured for performance where an explicit option exists, sleep states should be disabled, any virtualization optimization technologies should be enabled, and hyperthreading should also be enabled.

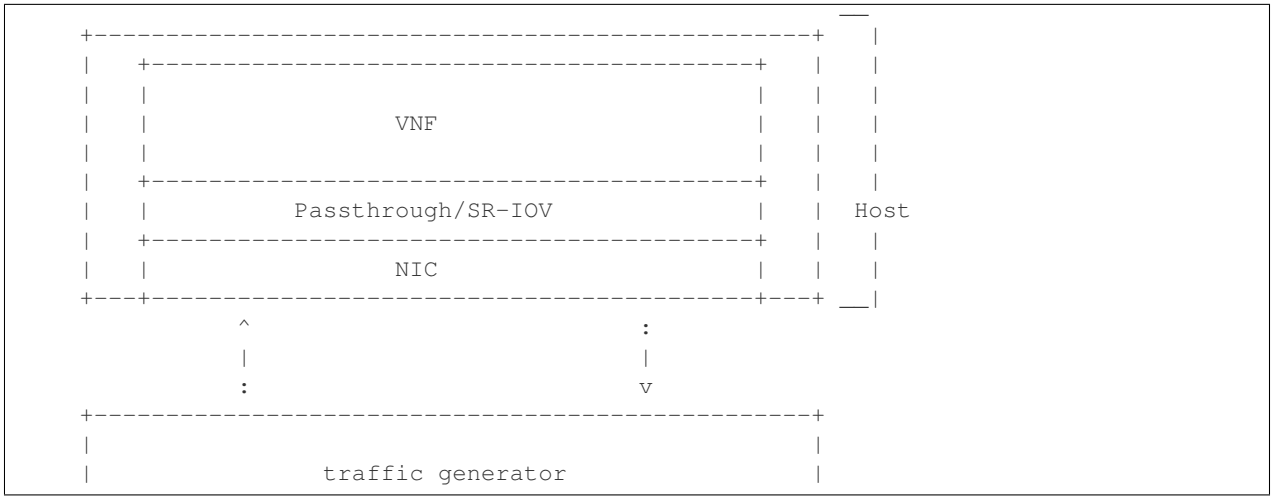
System Validation

System validation is broken down into two sub-categories: Platform validation and VNF validation. The validation test itself involves verifying the forwarding capability and stability for the sub-system under test. The rationale behind system validation is two fold. Firstly to give a tester confidence in the stability of the platform or VNF that is being tested; and secondly to provide base performance comparison points to understand the overhead introduced by the virtual switch.

- Benchmark platform forwarding capability: This is an OPTIONAL test used to verify the platform and measure the base performance (maximum forwarding rate in fps and latency) that can be achieved by the platform without a vSwitch or a VNF. The following diagram outlines the set-up for benchmarking Platform forwarding capability:



- Benchmark VNF forwarding capability: This test is used to verify the VNF and measure the base performance (maximum forwarding rate in fps and latency) that can be achieved by the VNF without a vSwitch. The performance metrics collected by this test will serve as a key comparison point for NIC passthrough technologies and vSwitches. VNF in this context refers to the hypervisor and the VM. The following diagram outlines the set-up for benchmarking VNF forwarding capability:



Methodology to benchmark Platform/VNF forwarding capability

The recommended methodology for the platform/VNF validation and benchmark is: - Run RFC2889 Maximum Forwarding Rate test, this test will produce maximum forwarding rate and latency results that will serve as the expected values. These expected values can be used in subsequent steps or compared with in subsequent validation tests. - Transmit bidirectional traffic at line rate/max forwarding rate (whichever is higher) for at least 72 hours, measure throughput (fps) and latency. - Note: Traffic should be bidirectional. - Establish a baseline forwarding rate for what the platform can achieve. - Additional validation: After the test has completed for 72 hours run bidirectional traffic at the maximum forwarding rate once more to see if the system is still functional and measure throughput (fps) and latency. Compare the measure the new obtained values with the expected values.

NOTE 1: How the Platform is configured for its forwarding capability test (BIOS settings, GRUB configuration, runlevel...) is how the platform should be configured for every test after this

NOTE 2: How the VNF is configured for its forwarding capability test (# of vCPUs, vNICs, Memory, affinitization. . .) is how it should be configured for every test that uses a VNF after this.

4.3.5 2.2.4 RFCs for testing virtual switch performance

The starting point for defining the suite of tests for benchmarking the performance of a virtual switch is to take existing RFCs and standards that were designed to test their physical counterparts and adapting them for testing virtual switches. The rationale behind this is to establish a fair comparison between the performance of virtual and physical switches. This section outlines the RFCs that are used by this specification.

RFC 1242 Benchmarking Terminology for Network Interconnection

Devices RFC 1242 defines the terminology that is used in describing performance benchmarking tests and their results. Definitions and discussions covered include: Back-to-back, bridge, bridge/router, constant load, data link frame size, frame loss rate, inter frame gap, latency, and many more.

RFC 2544 Benchmarking Methodology for Network Interconnect Devices

RFC 2544 outlines a benchmarking methodology for network Interconnect Devices. The methodology results in performance metrics such as latency, frame loss percentage, and maximum data throughput.

In this document network “throughput” (measured in millions of frames per second) is based on RFC 2544, unless otherwise noted. Frame size refers to Ethernet frames ranging from smallest frames of 64 bytes to largest frames of 4K bytes.

Types of tests are:

1. Throughput test defines the maximum number of frames per second that can be transmitted without any error.
2. Latency test measures the time required for a frame to travel from the originating device through the network to the destination device. Please note that RFC2544 Latency measurement will be superseded with a measurement of average latency over all successfully transferred packets or frames.
3. Frame loss test measures the network’s response in overload conditions - a critical indicator of the network’s ability to support real-time applications in which a large amount of frame loss will rapidly degrade service quality.

4. Burst test assesses the buffering capability of a virtual switch. It measures the maximum number of frames received at full line rate before a frame is lost. In carrier Ethernet networks, this measurement validates the excess information rate (EIR) as defined in many SLAs.
5. System recovery to characterize speed of recovery from an overload condition.
6. Reset to characterize speed of recovery from device or software reset. This type of test has been updated by RFC6201 as such, the methodology defined by this specification will be that of RFC 6201.

Although not included in the defined RFC 2544 standard, another crucial measurement in Ethernet networking is packet delay variation. The definition set out by this specification comes from RFC5481.

RFC 2285 Benchmarking Terminology for LAN Switching Devices

RFC 2285 defines the terminology that is used to describe the terminology for benchmarking a LAN switching device. It extends RFC 1242 and defines: DUTs, SUTs, Traffic orientation and distribution, bursts, loads, forwarding rates, etc.

RFC 2889 Benchmarking Methodology for LAN Switching

RFC 2889 outlines a benchmarking methodology for LAN switching, it extends RFC 2544. The outlined methodology gathers performance metrics for forwarding, congestion control, latency, address handling and finally filtering.

RFC 3918 Methodology for IP Multicast Benchmarking

RFC 3918 outlines a methodology for IP Multicast benchmarking.

RFC 4737 Packet Reordering Metrics

RFC 4737 describes metrics for identifying and counting re-ordered packets within a stream, and metrics to measure the extent each packet has been re-ordered.

RFC 5481 Packet Delay Variation Applicability Statement

RFC 5481 defined two common, but different forms of delay variation metrics, and compares the metrics over a range of networking circumstances and tasks. The most suitable form for vSwitch benchmarking is the “PDV” form.

RFC 6201 Device Reset Characterization

RFC 6201 extends the methodology for characterizing the speed of recovery of the DUT from device or software reset described in RFC 2544.

4.3.6 2.2.5 Details of the Test Report

There are a number of parameters related to the system, DUT and tests that can affect the repeatability of a test results and should be recorded. In order to minimise the variation in the results of a test, it is recommended that the test report includes the following information:

- Hardware details including:
  - Platform details.

  - Processor details.
  - Memory information (see below)
  - Number of enabled cores.
  - Number of cores used for the test.
  - Number of physical NICs, as well as their details (manufacturer, versions, type and the PCI slot they are plugged into).
  - NIC interrupt configuration.
  - BIOS version, release date and any configurations that were modified.
- Software details including:
  - OS version (for host and VNF)
  - Kernel version (for host and VNF)
  - GRUB boot parameters (for host and VNF).
  - Hypervisor details (Type and version).
  - Selected vSwitch, version number or commit id used.
  - vSwitch launch command line if it has been parameterised.
  - Memory allocation to the vSwitch – which NUMA node it is using, and how many memory channels.
  - Where the vswitch is built from source: compiler details including versions and the flags that were used to compile the vSwitch.
  - DPDK or any other SW dependency version number or commit id used.
  - Memory allocation to a VM - if it's from Hugpages/elsewhere.
  - VM storage type: snapshot/independent persistent/independent non-persistent.
  - Number of VMs.
  - Number of Virtual NICs (vNICs), versions, type and driver.
  - Number of virtual CPUs and their core affinity on the host.
  - Number vNIC interrupt configuration.
  - Thread affinitization for the applications (including the vSwitch itself) on the host.
  - Details of Resource isolation, such as CPUs designated for Host/Kernel (isolcpu) and CPUs designated for specific processes (taskset).
- Memory Details
  - Total memory
  - Type of memory
  - Used memory
  - Active memory
  - Inactive memory
  - Free memory
  - Buffer memory
  - Swap cache

  - Total swap
  - Used swap
  - Free swap
- Test duration.
- Number of flows.
- Traffic Information:
  - Traffic type - UDP, TCP, IMIX / Other.
  - Packet Sizes.
- Deployment Scenario.

Note: Tests that require additional parameters to be recorded will explicitly specify this.

4.4 2.3. Test identification

4.4.1 2.3.1 Throughput tests

The following tests aim to determine the maximum forwarding rate that can be achieved with a virtual switch. The list is not exhaustive but should indicate the type of tests that should be required. It is expected that more will be added.

Test ID: LTD.Throughput.RFC2544.PacketLossRatio

Title: RFC 2544 X% packet loss ratio Throughput and Latency Test

Prerequisite Test: N/A

Priority:

Description:

This test determines the DUT's maximum forwarding rate with X% traffic loss for a constant load (fixed length frames at a fixed interval time). The default loss percentages to be tested are: - X = 0% - X = 10^-7%

Note: Other values can be tested if required by the user.

The selected frame sizes are those previously defined under *Default Test Parameters*. The test can also be used to determine the average latency of the traffic.

Under the RFC2544 test methodology, the test duration will include a number of trials; each trial should run for a minimum period of 60 seconds. A binary search methodology must be applied for each trial to obtain the final result.

Expected Result: At the end of each trial, the presence or absence of loss determines the modification of offered load for the next trial, converging on a maximum rate, or RFC2544 Throughput with X% loss. The Throughput load is re-used in related RFC2544 tests and other tests.

Metrics Collected:

The following are the metrics collected for this test:

- The maximum forwarding rate in Frames Per Second (FPS) and Mbps of the DUT for each frame size with X% packet loss.

- The average latency of the traffic flow when passing through the DUT (if testing for latency, note that this average is different from the test specified in Section 26.3 of RFC2544).
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Test ID: LTD.Throughput.RFC2544.PacketLossRatioFrameModification

Title: RFC 2544 X% packet loss Throughput and Latency Test with packet modification

Prerequisite Test: N/A

Priority:

Description:

This test determines the DUT's maximum forwarding rate with X% traffic loss for a constant load (fixed length frames at a fixed interval time). The default loss percentages to be tested are: - X = 0% - X = 10^-7%

Note: Other values can be tested if required by the user.

The selected frame sizes are those previously defined under *Default Test Parameters*. The test can also be used to determine the average latency of the traffic.

Under the RFC2544 test methodology, the test duration will include a number of trials; each trial should run for a minimum period of 60 seconds. A binary search methodology must be applied for each trial to obtain the final result.

During this test, the DUT must perform the following operations on the traffic flow:

- Perform packet parsing on the DUT's ingress port.
- Perform any relevant address look-ups on the DUT's ingress ports.
- Modify the packet header before forwarding the packet to the DUT's egress port. Packet modifications include:
  - Modifying the Ethernet source or destination MAC address.
  - Modifying/adding a VLAN tag. (**Recommended**).
  - Modifying/adding a MPLS tag.
  - Modifying the source or destination ip address.
  - Modifying the TOS/DSCP field.
  - Modifying the source or destination ports for UDP/TCP/SCTP.
  - Modifying the TTL.

Expected Result: The Packet parsing/modifications require some additional degree of processing resource, therefore the RFC2544 Throughput is expected to be somewhat lower than the Throughput level measured without additional steps. The reduction is expected to be greatest on tests with the smallest packet sizes (greatest header processing rates).

Metrics Collected:

The following are the metrics collected for this test:

- The maximum forwarding rate in Frames Per Second (FPS) and Mbps of the DUT for each frame size with X% packet loss and packet modification operations being performed by the DUT.

- The average latency of the traffic flow when passing through the DUT (if testing for latency, note that this average is different from the test specified in Section 26.3 of RFC2544).
- The RFC5481 PDV form of delay variation on the traffic flow, using the 99th percentile.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Test ID: LTD.Throughput.RFC2544.Profile

Title: RFC 2544 Throughput and Latency Profile

Prerequisite Test: N/A

Priority:

Description:

This test reveals how throughput and latency degrades as the offered rate varies in the region of the DUT's maximum forwarding rate as determined by LTD.Throughput.RFC2544.PacketLossRatio (0% Packet Loss). For example it can be used to determine if the degradation of throughput and latency as the offered rate increases is slow and graceful or sudden and severe.

The selected frame sizes are those previously defined under *Default Test Parameters*.

The offered traffic rate is described as a percentage delta with respect to the DUT's RFC 2544 Throughput as determined by LTD.Throughput.RFC2544.PacketLoss Ratio (0% Packet Loss case). A delta of 0% is equivalent to an offered traffic rate equal to the RFC 2544 Throughput; A delta of +50% indicates an offered rate half-way between the Throughput and line-rate, whereas a delta of -50% indicates an offered rate of half the maximum rate. Therefore the range of the delta figure is natuarlly bounded at -100% (zero offered traffic) and +100% (traffic offered at line rate).

The following deltas to the maximum forwarding rate should be applied:

- -50%, -10%, 0%, +10% & +50%

Expected Result: For each packet size a profile should be produced of how throughput and latency vary with offered rate.

Metrics Collected:

The following are the metrics collected for this test:

- The forwarding rate in Frames Per Second (FPS) and Mbps of the DUT for each delta to the maximum forwarding rate and for each frame size.
- The average latency for each delta to the maximum forwarding rate and for each frame size.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.
- Any failures experienced (for example if the vSwitch crashes, stops processing packets, restarts or becomes unresponsive to commands) when the offered load is above Maximum Throughput MUST be recorded and reported with the results.

Test ID: LTD.Throughput.RFC2544.SystemRecoveryTime

Title: RFC 2544 System Recovery Time Test

Prerequisite Test LTD.Throughput.RFC2544.PacketLossRatio

Priority:

Description:

The aim of this test is to determine the length of time it takes the DUT to recover from an overload condition for a constant load (fixed length frames at a fixed interval time). The selected frame sizes are those previously defined under *Default Test Parameters*, traffic should be sent to the DUT under normal conditions. During the duration of the test and while the traffic flows are passing though the DUT, at least one situation leading to an overload condition for the DUT should occur. The time from the end of the overload condition to when the DUT returns to normal operations should be measured to determine recovery time. Prior to overloading the DUT, one should record the average latency for 10,000 packets forwarded through the DUT.

The overload condition SHOULD be to transmit traffic at a very high frame rate to the DUT (150% of the maximum 0% packet loss rate as determined by LTD.Throughput.RFC2544.PacketLossRatio or line-rate whichever is lower), for at least 60 seconds, then reduce the frame rate to 75% of the maximum 0% packet loss rate. A number of time-stamps should be recorded: - Record the time-stamp at which the frame rate was reduced and record a second time-stamp at the time of the last frame lost. The recovery time is the difference between the two timestamps. - Record the average latency for 10,000 frames after the last frame loss and continue to record average latency measurements for every 10,000 frames, when latency returns to within 10% of pre-overload levels record the time-stamp.

Expected Result:

Metrics collected

The following are the metrics collected for this test:

- The length of time it takes the DUT to recover from an overload condition.
- The length of time it takes the DUT to recover the average latency to pre-overload conditions.

Deployment scenario:

- Physical → virtual switch → physical.

Test ID: LTD.Throughput.RFC2544.BackToBackFrames

Title: RFC2544 Back To Back Frames Test

Prerequisite Test: N

Priority:

Description:

The aim of this test is to characterize the ability of the DUT to process back-to-back frames. For each frame size previously defined under *Default Test Parameters*, a burst of traffic is sent to the DUT with the minimum inter-frame gap between each frame. If the number of received frames equals the number of frames that were transmitted, the burst size should be increased and traffic is sent to the DUT again. The value measured is the back-to-back value, that is the maximum burst size the DUT can handle without any frame loss.

Expected Result:

Tests of back-to-back frames with physical devices have produced unstable results in some cases. All tests should be repeated in multiple test sessions and results stability should be examined.

Metrics collected

The following are the metrics collected for this test:

- The back-to-back value, which is the the number of frames in the longest burst that the DUT will handle without the loss of any frames.

- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Deployment scenario:

- Physical → virtual switch → physical.

Test ID: LTD.Throughput.RFC2889.MaxForwardingRateSoak

Title: RFC 2889 X% packet loss Max Forwarding Rate Soak Test

Prerequisite Test LTD.Throughput.RFC2544.PacketLossRatio

Priority:

Description:

The aim of this test is to understand the Max Forwarding Rate stability over an extended test duration in order to uncover any outliers. To allow for an extended test duration, the test should ideally run for 24 hours or, if this is not possible, for at least 6 hours. For this test, each frame size must be sent at the highest Throughput rate with X% packet loss, as determined in the prerequisite test. The default loss percentages to be tested are: - X = 0% - X = 10^-7%

Note: Other values can be tested if required by the user.

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- Max Forwarding Rate stability of the DUT.
  - This means reporting the number of packets lost per time interval and reporting any time intervals with packet loss. The RFC2889 Forwarding Rate shall be measured in each interval. An interval of 60s is suggested.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.
- The RFC5481 PDV form of delay variation on the traffic flow, using the 99th percentile.

Test ID: LTD.Throughput.RFC2889.MaxForwardingRateSoakFrameModification

Title: RFC 2889 Max Forwarding Rate Soak Test with Frame Modification

Prerequisite Test: LTD.Throughput.RFC2544.PacketLossRatioFrameModification (0% Packet Loss)

Priority:

Description:

The aim of this test is to understand the Max Forwarding Rate stability over an extended test duration in order to uncover any outliers. To allow for an extended test duration, the test should ideally run for 24 hours or, if this is not possible, for at least 6 hour. For this test, each frame size must be sent at the highest Throughput rate with 0% packet loss, as determined in the prerequisite test.

During this test, the DUT must perform the following operations on the traffic flow:

- Perform packet parsing on the DUT's ingress port.
- Perform any relevant address look-ups on the DUT's ingress ports.

- Modify the packet header before forwarding the packet to the DUT's egress port. Packet modifications include:
  - Modifying the Ethernet source or destination MAC address.
  - Modifying/adding a VLAN tag (**Recommended**).
  - Modifying/adding a MPLS tag.
  - Modifying the source or destination ip address.
  - Modifying the TOS/DSCP field.
  - Modifying the source or destination ports for UDP/TCP/SCTP.
  - Modifying the TTL.

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- Max Forwarding Rate stability of the DUT.
  - This means reporting the number of packets lost per time interval and reporting any time intervals with packet loss. The RFC2889 Forwarding Rate shall be measured in each interval. An interval of 60s is suggested.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.
- The RFC5481 PDV form of delay variation on the traffic flow, using the 99th percentile.

Test ID: LTD.Throughput.RFC6201.ResetTime

Title: RFC 6201 Reset Time Test

Prerequisite Test: N/A

Priority:

Description:

The aim of this test is to determine the length of time it takes the DUT to recover from a reset.

Two reset methods are defined - planned and unplanned. A planned reset requires stopping and restarting the virtual switch by the usual 'graceful' method defined by it's documentation. An unplanned reset requires simulating a fatal internal fault in the virtual switch - for example by using kill -SIGKILL on a Linux environment.

Both reset methods SHOULD be exercised.

For each frame size previously defined under *Default Test Parameters*, traffic should be sent to the DUT under normal conditions. During the duration of the test and while the traffic flows are passing through the DUT, the DUT should be reset and the Reset time measured. The Reset time is the total time that a device is determined to be out of operation and includes the time to perform the reset and the time to recover from it (cf. RFC6201).

RFC6201 defines two methods to measure the Reset time:

- Frame-Loss Method: which requires the monitoring of the number of lost frames and calculates the Reset time based on the number of frames lost and the offered rate according to the following formula:

```
                        Frames_lost (packets)
Reset_time = ------------------------------------
                  Offered_rate (packets per second)
```

- Timestamp Method: which measures the time from which the last frame is forwarded from the DUT to the time the first frame is forwarded after the reset. This involves time-stamping all transmitted frames and recording the timestamp of the last frame that was received prior to the reset and also measuring the timestamp of the first frame that is received after the reset. The Reset time is the difference between these two timestamps.

According to RFC6201 the choice of method depends on the test tool's capability; the Frame-Loss method SHOULD be used if the test tool supports: - Counting the number of lost frames per stream. - Transmitting test frame despite the physical link status.

whereas the Timestamp method SHOULD be used if the test tool supports: - Timestamping each frame. - Monitoring received frame's timestamp. - Transmitting frames only if the physical link status is up.

Expected Result:

Metrics collected

The following are the metrics collected for this test: - Average Reset Time over the number of trials performed.

Results of this test should include the following information: - The reset method used. - Throughput in Fps and Mbps. - Average Frame Loss over the number of trials performed. - Average Reset Time in milliseconds over the number of trials performed. - Number of trials performed. - Protocol: IPv4, IPv6, MPLS, etc. - Frame Size in Octets - Port Media: Ethernet, Gigabit Ethernet (GbE), etc. - Port Speed: 10 Gbps, 40 Gbps etc. - Interface Encapsulation: Ethernet, Ethernet VLAN, etc.

Deployment scenario:

- Physical → virtual switch → physical.

Test ID: LTD.Throughput.RFC2889.MaxForwardingRate

Title: RFC2889 Forwarding Rate Test

Prerequisite Test: LTD.Throughput.RFC2544.PacketLossRatio

Priority:

Description:

This test measures the DUT's Max Forwarding Rate when the Offered Load is varied between the throughput and the Maximum Offered Load for fixed length frames at a fixed time interval. The selected frame sizes are those previously defined under *Default Test Parameters*. The throughput is the maximum offered load with 0% frame loss (measured by the prerequisite test), and the Maximum Offered Load (as defined by RFC2285) is *"the highest number of frames per second that an external source can transmit to a DUT/SUT for forwarding to a specified output interface or interfaces"*.

Traffic should be sent to the DUT at a particular rate (TX rate) starting with TX rate equal to the throughput rate. The rate of successfully received frames at the destination counted (in FPS). If the RX rate is equal to the TX rate, the TX rate should be increased by a fixed step size and the RX rate measured again until the Max Forwarding Rate is found.

The trial duration for each iteration should last for the period of time needed for the system to reach steady state for the frame size being tested. Under RFC2889 (Sec. 5.6.3.1) test methodology, the test duration should run for a minimum period of 30 seconds, regardless whether the system reaches steady state before the minimum duration ends.

Expected Result: According to RFC2889 The Max Forwarding Rate is the highest forwarding rate of a DUT taken from an iterative set of forwarding rate measurements. The iterative set of forwarding rate measurements are made by setting the intended load transmitted from an external source and measuring the offered load (i.e what the DUT is capable of forwarding). If the Throughput == the Maximum Offered Load, it follows that Max Forwarding Rate is equal to the Maximum Offered Load.

Metrics Collected:

The following are the metrics collected for this test:

- The Max Forwarding Rate for the DUT for each packet size.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Deployment scenario:

- Physical → virtual switch → physical. Note: Full mesh tests with multiple ingress and egress ports are a key aspect of RFC 2889 benchmarks, and scenarios with both 2 and 4 ports should be tested. In any case, the number of ports used must be reported.

Test ID: LTD.Throughput.RFC2889.ForwardPressure

Title: RFC2889 Forward Pressure Test

Prerequisite Test: LTD.Throughput.RFC2889.MaxForwardingRate

Priority:

Description:

The aim of this test is to determine if the DUT transmits frames with an inter-frame gap that is less than 12 bytes. This test overloads the DUT and measures the output for forward pressure. Traffic should be transmitted to the DUT with an inter-frame gap of 11 bytes, this will overload the DUT by 1 byte per frame. The forwarding rate of the DUT should be measured.

Expected Result: The forwarding rate should not exceed the maximum forwarding rate of the DUT collected by LTD.Throughput.RFC2889.MaxForwardingRate.

Metrics collected

The following are the metrics collected for this test:

- Forwarding rate of the DUT in FPS or Mbps.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Deployment scenario:

- Physical → virtual switch → physical.

Test ID: LTD.Throughput.RFC2889.ErrorFramesFiltering

Title: RFC2889 Error Frames Filtering Test

Prerequisite Test: N/A

Priority:

Description:

The aim of this test is to determine whether the DUT will propagate any erroneous frames it receives or whether it is capable of filtering out the erroneous frames. Traffic should be sent with erroneous frames included within the flow at random intervals. Illegal frames that must be tested include: - Oversize Frames. - Undersize Frames. - CRC Errored Frames. - Dribble Bit Errored Frames - Alignment Errored Frames

The traffic flow exiting the DUT should be recorded and checked to determine if the erroneous frames where passed through the DUT.

Expected Result: Broken frames are not passed!

Metrics collected

No Metrics are collected in this test, instead it determines:

- Whether the DUT will propagate erroneous frames.
- Or whether the DUT will correctly filter out any erroneous frames from traffic flow with out removing correct frames.

Deployment scenario:

- Physical → virtual switch → physical.

Test ID: LTD.Throughput.RFC2889.BroadcastFrameForwarding

Title: RFC2889 Broadcast Frame Forwarding Test

Prerequisite Test: N

Priority:

Description:

The aim of this test is to determine the maximum forwarding rate of the DUT when forwarding broadcast traffic. For each frame previously defined under *Default Test Parameters*, the traffic should be set up as broadcast traffic. The traffic throughput of the DUT should be measured.

The test should be conducted with at least 4 physical ports on the DUT. The number of ports used MUST be recorded.

As broadcast involves forwarding a single incoming packet to several destinations, the latency of a single packet is defined as the average of the latencies for each of the broadcast destinations.

The incoming packet is transmitted on each of the other physical ports, it is not transmitted on the port on which it was received. The test MAY be conducted using different broadcasting ports to uncover any performance differences.

Expected Result:

Metrics collected:

The following are the metrics collected for this test:

- The forwarding rate of the DUT when forwarding broadcast traffic.
- The minimum, average & maximum packets latencies observed.

Deployment scenario:

- Physical → virtual switch 3x physical. In the Broadcast rate testing,

four test ports are required. One of the ports is connected to the test device, so it can send broadcast frames and listen for miss-routed frames.

4.4.2 2.3.2 Packet Latency tests

These tests will measure the store and forward latency as well as the packet delay variation for various packet types through the virtual switch. The following list is not exhaustive but should indicate the type of tests that should be required. It is expected that more will be added.

Test ID: LTD.PacketLatency.InitialPacketProcessingLatency

Title: Initial Packet Processing Latency

Prerequisite Test: N/A

Priority:

Description:

In some virtual switch architectures, the first packets of a flow will take the system longer to process than subsequent packets in the flow. This test determines the latency for these packets. The test will measure the latency of the packets as they are processed by the flow-setup-path of the DUT. There are two methods for this test, a recommended method and a nalternative method that can be used if it is possible to disable the fastpath of the virtual switch.

Recommended method: This test will send 64,000 packets to the DUT, each belonging to a different flow. Average packet latency will be determined over the 64,000 packets.

Alternative method: This test will send a single packet to the DUT after a fixed interval of time. The time interval will be equivalent to the amount of time it takes for a flow to time out in the virtual switch plus 10%. Average packet latency will be determined over 1,000,000 packets.

This test is intended only for non-learning virtual switches; For learning virtual switches use RFC2889.

For this test, only unidirectional traffic is required.

Expected Result: The average latency for the initial packet of all flows should be greater than the latency of subsequent traffic.

Metrics Collected:

The following are the metrics collected for this test:

- Average latency of the initial packets of all flows that are processed by the DUT.

Deployment scenario:

- Physical → Virtual Switch → Physical.

Test ID: LTD.PacketDelayVariation.RFC3393.Soak

Title: Packet Delay Variation Soak Test

Prerequisite Tests: LTD.Throughput.RFC2544.PacketLossRatio (0% Packet Loss)

Priority:

Description:

The aim of this test is to understand the distribution of packet delay variation for different frame sizes over an extended test duration and to determine if there are any outliers. To allow for an extended test duration, the test should ideally run for 24 hours or, if this is not possible, for at least 6 hour. For this test, each frame size must be sent at the highest possible throughput with 0% packet loss, as determined in the prerequisite test.

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- The packet delay variation value for traffic passing through the DUT.
- The RFC5481 PDV form of delay variation on the traffic flow, using the 99th percentile, for each 60s interval during the test.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

4.4.3 2.3.3 Scalability tests

The general aim of these tests is to understand the impact of large flow table size and flow lookups on throughput. The following list is not exhaustive but should indicate the type of tests that should be required. It is expected that more will be added.

Test ID: LTD.Scalability.RFC2544.0PacketLoss

Title: RFC 2544 0% loss Scalability throughput test

Prerequisite Test: LTD.Throughput.RFC2544.PacketLossRatio, IF the delta Throughput between the single-flow RFC2544 test and this test with a variable number of flows is desired.

Priority:

Description:

The aim of this test is to measure how throughput changes as the number of flows in the DUT increases. The test will measure the throughput through the fastpath, as such the flows need to be installed on the DUT before passing traffic.

For each frame size previously defined under *Default Test Parameters* and for each of the following number of flows:

- 1,000
- 2,000
- 4,000
- 8,000
- 16,000
- 32,000
- 64,000
- Max supported number of flows.

This test will be conducted under two conditions following the establishment of all flows as required by RFC 2544, regarding the flow expiration time-out:

1. The time-out never expires during each trial.

2) The time-out expires for all flows periodically. This would require a short time-out compared with flow re-appearance for a small number of flows, and may not be possible for all flow conditions.

The maximum 0% packet loss Throughput should be determined in a manner identical to LTD.Throughput.RFC2544.PacketLossRatio.

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- The maximum number of frames per second that can be forwarded at the specified number of flows and the specified frame size, with zero packet loss.

Test ID: LTD.MemoryBandwidth.RFC2544.0PacketLoss.Scalability

Title: RFC 2544 0% loss Memory Bandwidth Scalability test

Prerequisite Tests: LTD.Throughput.RFC2544.PacketLossRatio, IF the delta Throughput between an undisturbed RFC2544 test and this test with the Throughput affected by cache and memory bandwidth contention is desired.

Priority:

Description:

The aim of this test is to understand how the DUT's performance is affected by cache sharing and memory bandwidth between processes.

During the test all cores not used by the vSwitch should be running a memory intensive application. This application should read and write random data to random addresses in unused physical memory. The random nature of the data and addresses is intended to consume cache, exercise main memory access (as opposed to cache) and exercise all memory buses equally. Furthermore:

- the ratio of reads to writes should be recorded. A ratio of 1:1 SHOULD be used.
- the reads and writes MUST be of cache-line size and be cache-line aligned.
- in NUMA architectures memory access SHOULD be local to the core's node. Whether only local memory or a mix of local and remote memory is used MUST be recorded.
- the memory bandwidth (reads plus writes) used per-core MUST be recorded; the test MUST be run with a per-core memory bandwidth equal to half the maximum system memory bandwidth divided by the number of cores. The test MAY be run with other values for the per-core memory bandwidth.
- the test MAY also be run with the memory intensive application running on all cores.

Under these conditions the DUT's 0% packet loss throughput is determined as per LTD.Throughput.RFC2544.PacketLossRatio.

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- The DUT's 0% packet loss throughput in the presence of cache sharing and memory bandwidth between processes.

4.4.4 2.3.4 Activation tests

The general aim of these tests is to understand the capacity of the and speed with which the vswitch can accomodate new flows.

Test ID: LTD.Activation.RFC2889.AddressCachingCapacity

Title: RFC2889 Address Caching Capacity Test

Prerequisite Test: N/A

Priority:

Description:

Please note this test is only applicable to virtual switches that are capable of MAC learning. The aim of this test is to determine the address caching capacity of the DUT for a constant load (fixed length frames at a fixed interval time). The selected frame sizes are those previously defined under *Default Test Parameters*.

In order to run this test the aging time, that is the maximum time the DUT will keep a learned address in its flow table, and a set of initial addresses, whose value should be >= 1 and <= the max number supported by the implementation must be known. Please note that if the aging time is configurable it must be longer than the time necessary to produce frames from the external source at the specified rate. If the aging time is fixed the frame rate must be brought down to a value that the external source can produce in a time that is less than the aging time.

Learning Frames should be sent from an external source to the DUT to install a number of flows. The Learning Frames must have a fixed destination address and must vary the source address of the frames. The DUT should install flows in its flow table based on the varying source addresses. Frames should then be transmitted from an external source at a suitable frame rate to see if the DUT has properly learned all of the addresses. If there is no frame loss and no flooding, the number of addresses sent to the DUT should be increased and the test is repeated until the max number of cached addresses supported by the DUT determined.

Expected Result:

Metrics collected:

The following are the metrics collected for this test:

- Number of cached addresses supported by the DUT.
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Deployment scenario:

- Physical → virtual switch → 2 x physical (one receiving, one listening).

Test ID: LTD.Activation.RFC2889.AddressLearningRate

Title: RFC2889 Address Learning Rate Test

Prerequisite Test: LTD.Memory.RFC2889.AddressCachingCapacity

Priority:

Description:

Please note this test is only applicable to virtual switches that are capable of MAC learning. The aim of this test is to determine the rate of address learning of the DUT for a constant load (fixed length frames at a fixed interval time). The selected frame sizes are those previously defined under *Default Test Parameters*, traffic should be sent with each IPv4/IPv6 address incremented by one. The rate at which the DUT learns a new address should be measured. The maximum caching capacity from LTD.Memory.RFC2889.AddressCachingCapacity should be taken into consideration as the maximum number of addresses for which the learning rate can be obtained.

Expected Result: It may be worthwhile to report the behaviour when operating beyond address capacity - some DUTs may be more friendly to new addresses than others.

Metrics collected:

The following are the metrics collected for this test:

- The address learning rate of the DUT.

Deployment scenario:

- Physical → virtual switch → 2 x physical (one receiving, one listening).

4.4.5 2.3.5 Coupling between control path and datapath Tests

The following tests aim to determine how tightly coupled the datapath and the control path are within a virtual switch. The following list is not exhaustive but should indicate the type of tests that should be required. It is expected that more will be added.

Test ID: LTD.CPDPCouplingFlowAddition

Title: Control Path and Datapath Coupling

Prerequisite Test:

Priority:

Description:

The aim of this test is to understand how exercising the DUT's control path affects datapath performance.

Initially a certain number of flow table entries are installed in the vSwitch. Then over the duration of an RFC2544 throughput test flow-entries are added and removed at the rates specified below. No traffic is 'hitting' these flow-entries, they are simply added and removed.

The test MUST be repeated with the following initial number of flow-entries installed: - < 10 - 1000 - 100,000 - 10,000,000 (or the maximum supported number of flow-entries)

The test MUST be repeated with the following rates of flow-entry addition and deletion per second: - 0 - 1 (i.e. 1 addition plus 1 deletion) - 100 - 10,000

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- The maximum forwarding rate in Frames Per Second (FPS) and Mbps of the DUT.
- The average latency of the traffic flow when passing through the DUT (if testing for latency, note that this average is different from the test specified in Section 26.3 of RFC2544).
- CPU and memory utilization may also be collected as part of this test, to determine the vSwitch's performance footprint on the system.

Deployment scenario:

- Physical → virtual switch → physical.

4.4.6 2.3.6 CPU and memory consumption

The following tests will profile a virtual switch's CPU and memory utilization under various loads and circumstances. The following list is not exhaustive but should indicate the type of tests that should be required. It is expected that more will be added.

Test ID: LTD.CPU.RFC2544.0PacketLoss

Title: RFC 2544 0% Loss Compute Test

Prerequisite Test:

Priority:

Description:

The aim of this test is to understand the overall performance of the system when a CPU intensive application is run on the same DUT as the Virtual Switch. For each frame size, an LTD.Throughput.RFC2544.PacketLossRatio (0% Packet Loss) test should be performed. Throughout the entire test a CPU intensive application should be run on all cores on the system not in use by the Virtual Switch. For NUMA system only cores on the same NUMA node are loaded.

It is recommended that stress-ng be used for loading the non-Virtual Switch cores but any stress tool MAY be used.

Expected Result:

Metrics Collected:

The following are the metrics collected for this test:

- CPU utilization of the cores running the Virtual Switch.
- The number of identity of the cores allocated to the Virtual Switch.
- The configuration of the stress tool (for example the command line parameters used to start it.)

4.4.7 2.3.7 Summary List of Tests

1. Throughput tests

- Test ID: LTD.Throughput.RFC2544.PacketLossRatio
- Test ID: LTD.Throughput.RFC2544.PacketLossRatioFrameModification
- Test ID: LTD.Throughput.RFC2544.Profile
- Test ID: LTD.Throughput.RFC2544.SystemRecoveryTime
- Test ID: LTD.Throughput.RFC2544.BackToBackFrames
- Test ID: LTD.Throughput.RFC2889.Soak
- Test ID: LTD.Throughput.RFC2889.SoakFrameModification
- Test ID: LTD.Throughput.RFC6201.ResetTime
- Test ID: LTD.Throughput.RFC2889.MaxForwardingRate
- Test ID: LTD.Throughput.RFC2889.ForwardPressure
- Test ID: LTD.Throughput.RFC2889.ErrorFramesFiltering
- Test ID: LTD.Throughput.RFC2889.BroadcastFrameForwarding

2. Packet Latency tests
   - Test ID: LTD.PacketLatency.InitialPacketProcessingLatency
   - Test ID: LTD.PacketDelayVariation.RFC3393.Soak
3. Scalability tests
   - Test ID: LTD.Scalability.RFC2544.0PacketLoss
   - Test ID: LTD.MemoryBandwidth.RFC2544.0PacketLoss.Scalability
4. Acivation tests
   - Test ID: LTD.Activation.RFC2889.AddressCachingCapacity
   - Test ID: LTD.Activation.RFC2889.AddressLearningRate
5. Coupling between control path and datapath Tests
   - Test ID: LTD.CPDPCouplingFlowAddition
6. CPU and memory consumption
   - Test ID: LTD.CPU.RFC2544.0PacketLoss

Revision: 0606ef6290ac1e41468f57b0ef2ff1d0571aee3e

Build date: November 26, 2015