Updates for the Back-to-back Frame Benchmark

draft-morton-bmwg-b2b-frame-03

Al Morton

Background

- RFC 2544 specifies the method to measure the Back-to-Back Frame Benchmark
- Defined in RFC 1242 – Longest burst of frames a DUT can process without loss
  - "Tests of this parameter are intended to determine the extent of data buffering in the device.“
- 2544: Very concise Objective, Procedure, and Reporting
- OPNFV Tests (2017) indicated areas for refinement

Benchmark Calculations

For each Frame size, calculate the following summary statistics for Back-to-back Frame values over the N tests:

- **Average (Benchmark)**
- **Minimum**
- **Maximum**
- **Standard Deviation**

Further, calculate the Implied DUT Buffer Time and the Corrected DUT Buffer Time in seconds, as follows:

Implied DUT Buffer Time =

Average num of Back-to-back Frames / Max Theoretical Frame Rate

Corrected DUT Buffer Time =

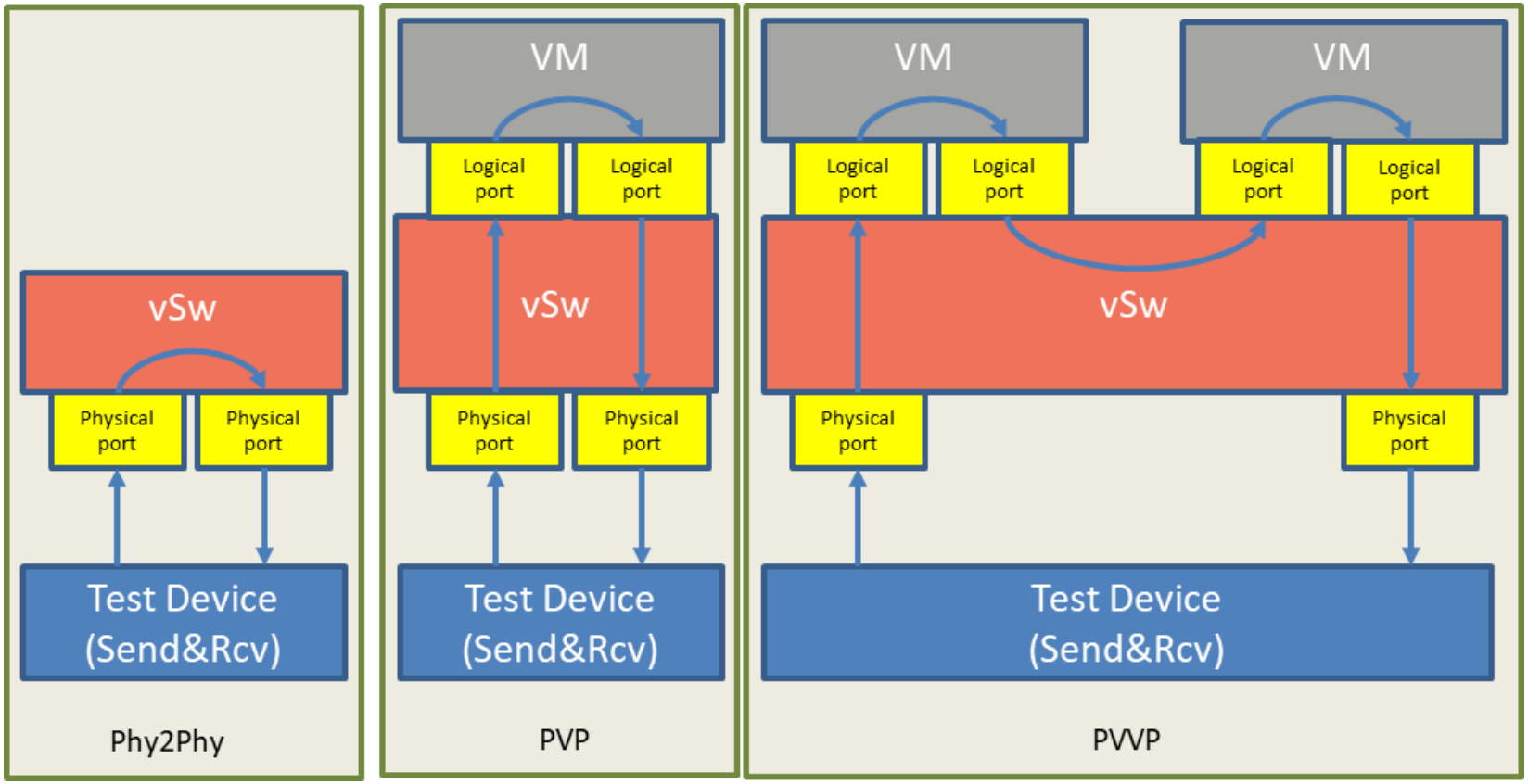
$$\text{Implied DUT Buffer Time} * \frac{\text{Measured Throughput}}{\text{Max Theoretical Frame Rate}}$$

Improvements in 03

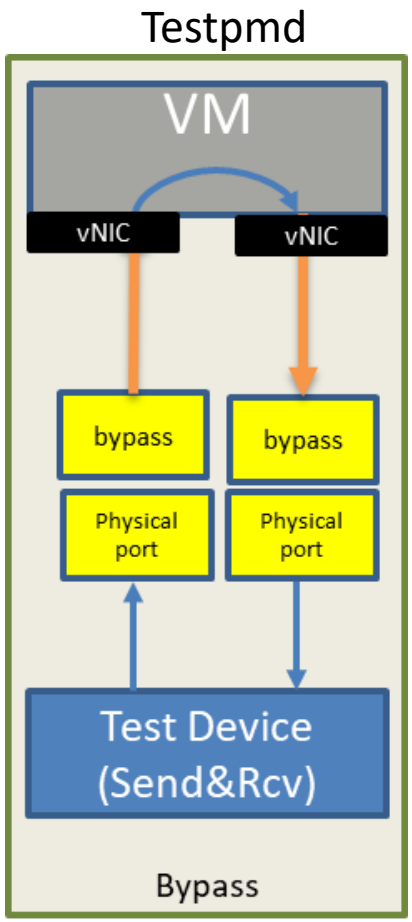
- Search Algorithms
  - Resolved Several Open Discussion Points
- Reference to the OPNFV VSPERF results presented at IETF-102
  - Buffers are key to absorbing Transient Interrupts *without Loss*
- Reference to ETSI GS NFV-TST 009
  - Binary Search and Binary Search with Loss Verification
  - MUST use one of the Algorithms above

Previous Tests with OVS vSwitch

No Transients | Observed Transients in Scenarios w/VM



Tests with SR IOV



Iter	Date	Time	Rate	rx_fps	lost	loss %ge	Duration	
1	10/4/2018	00:39.8	100	19681502	57919685	14.827	17	
2	10/4/2018	00:56.6	100	19621327	61427257	15.725	17	
3	10/4/2018	01:06.7	50	19495074	10706	0.005	10	
4	10/4/2018	01:16.9	50	19490325	0	0	10	
5	10/4/2018	01:28.4	75	19639161	67184396	22.931	12	
6	10/4/2018	01:41.3	75	19641298	43835848	14.961	13	
7	10/4/2018	01:51.8	62.5	19699001	38454729	15.75	10	
8	10/4/2018	02:02.1	62.5	19615272	46021999	18.851	11	
9	10/4/2018	02:12.2	56.25	19671347	22516545	10.247	10	
10	10/4/2018	02:22.4	56.25	19697345	22528836	10.252	10	
11	10/4/2018	02:32.5	53.125	19611975	10401037	5.012	10	
12	10/4/2018	02:42.6	53.125	19682392	10302749	4.964	10	
13	10/4/2018	02:52.8	51.562	19697436	4318670	2.144	10	
14	10/4/2018	03:02.9	51.562	19707651	4212109	2.091	10	
15	10/4/2018	03:13.0	50.781	19701824	1147274	0.578	11	
16	10/4/2018	03:23.1	50.781	19620899	1206844	0.608	10	
17	10/4/2018	03:33.3	50.391	19636419	0	0	10	THPT

Greatly Transient Frequency (this is the only one, above)

Discussion of SR IOV Test

- Conjecture about the VM processes attracting transients/losses seems to be wrong!
- However, the SR-IOV test config:
  - Testpmd configured with large Tx and Rx Buffers!
  - removes the ovs-dpdk vswitch processes
  - removes the logical interfaces between the VM and ovs-dpdk
  - replaces os.path.join(RTE_TARGET, 'kmod/igb_uio.ko') with ‘vfio-pci’ in 02_vswitch.conf

Next Steps

- Proposal/ASK:
  - WG Adoption of this draft
- Please Read and send your Review to the list

BACKUP

Improvements in 02

- Discussion with Yoshiaki Itou
- Clarified text describing *what* quantities are calculated:
  - Explain “Implied DUT buffer time”
  - Describe what phenomenon the correction factor accounts for
- Potential Benefit of using the Correction factor.
- Clarified Scope: does not apply to RFC 8239

Questions for BMWG Discussion

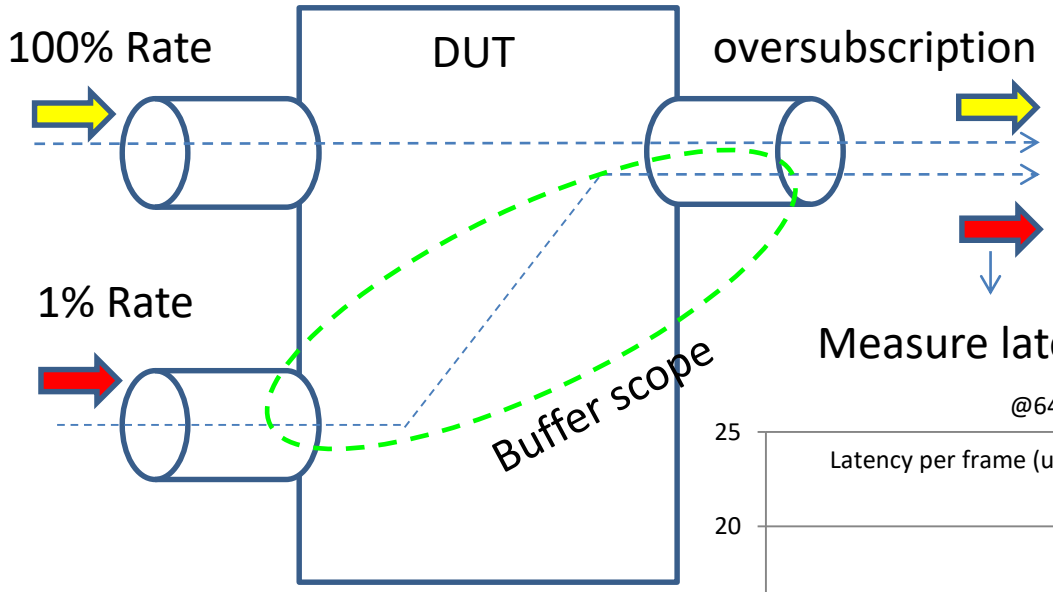
- @@@@ Should a particular search algorithm be included?
  - Yes
- @@@@ Should the search include trial repetition whenever frame loss is observed, to avoid the effects of background loss (un-related to buffer overflow)?
  - Yes, see OPNFV Hackfest Plugfest results (2018)

Procedure*

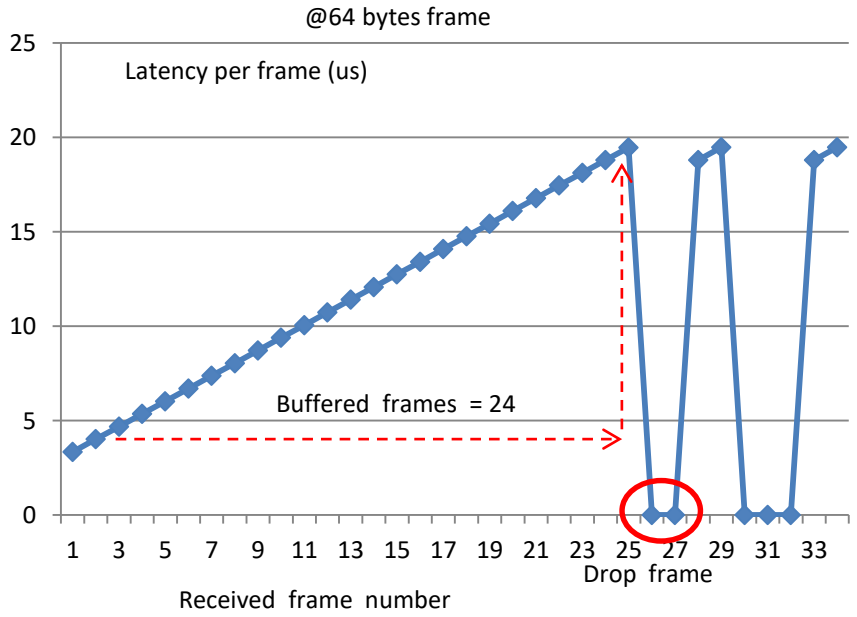
- List of Frame Sizes selected from TPUT tests
- Tests are composed of Repeated Trials
- For Each Frame size, a Trial requires sending a burst length counting forwarded frames
- Trials seek the longest burst length with 0 loss
- Test outcome is the burst length
- Tests are Repeated (N times), burst lengths are Averaged
- Average length is the Benchmark

Yoshiaki Itou
RFC8239 Data Center Benchmarking Methodology
3.Buffering Testing

2018.7.3



Measure latency per frame

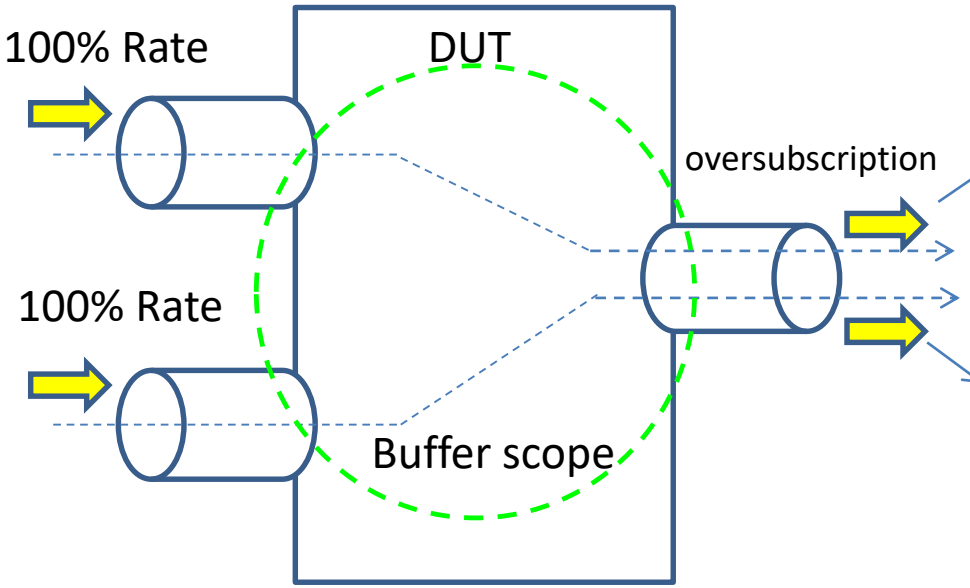


To cause congestion at the same time
Traffic generator has following configuration.
Port Send Mode = Synchronous
Flow Control = disable

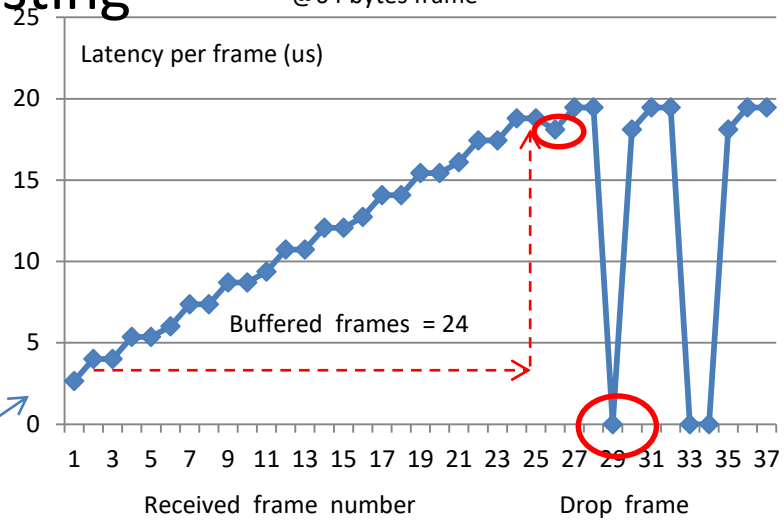
Yoshiaki Itou

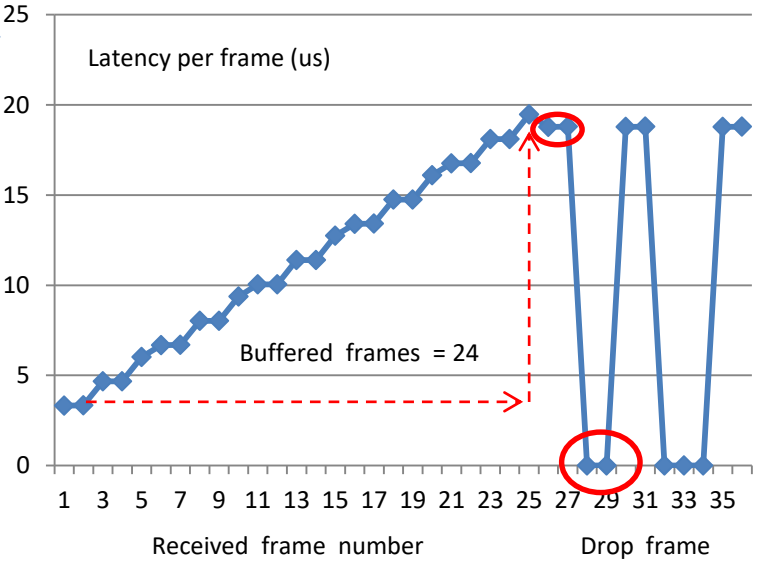
RFC8239 Data Center Benchmarking Methodology

3.Buffering Testing



Measure latency per frame
@64 bytes frame



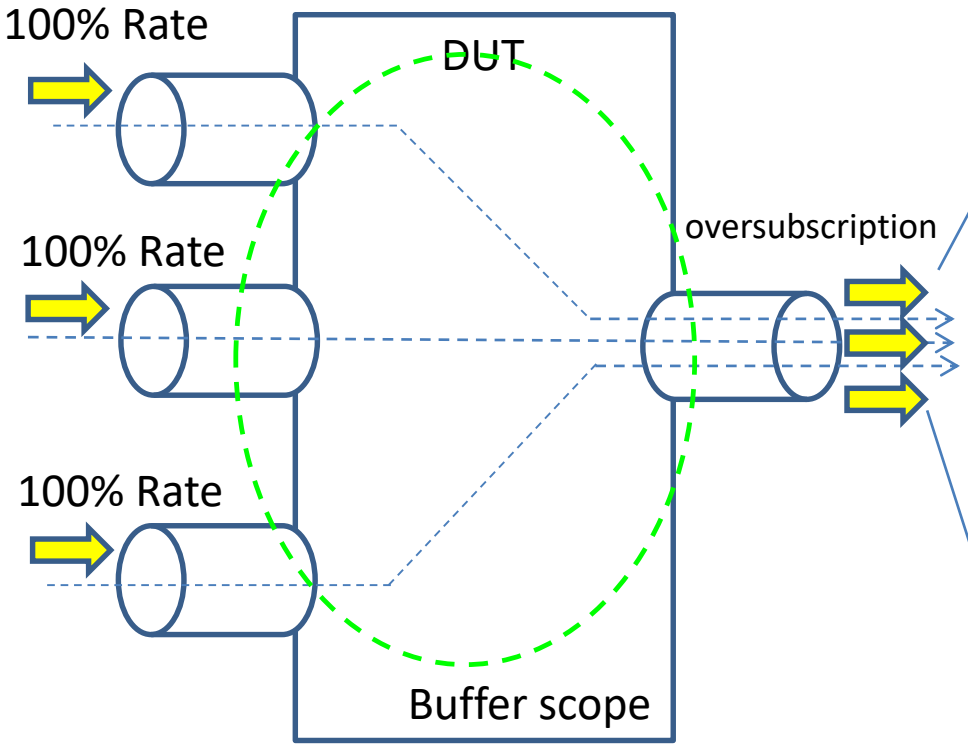


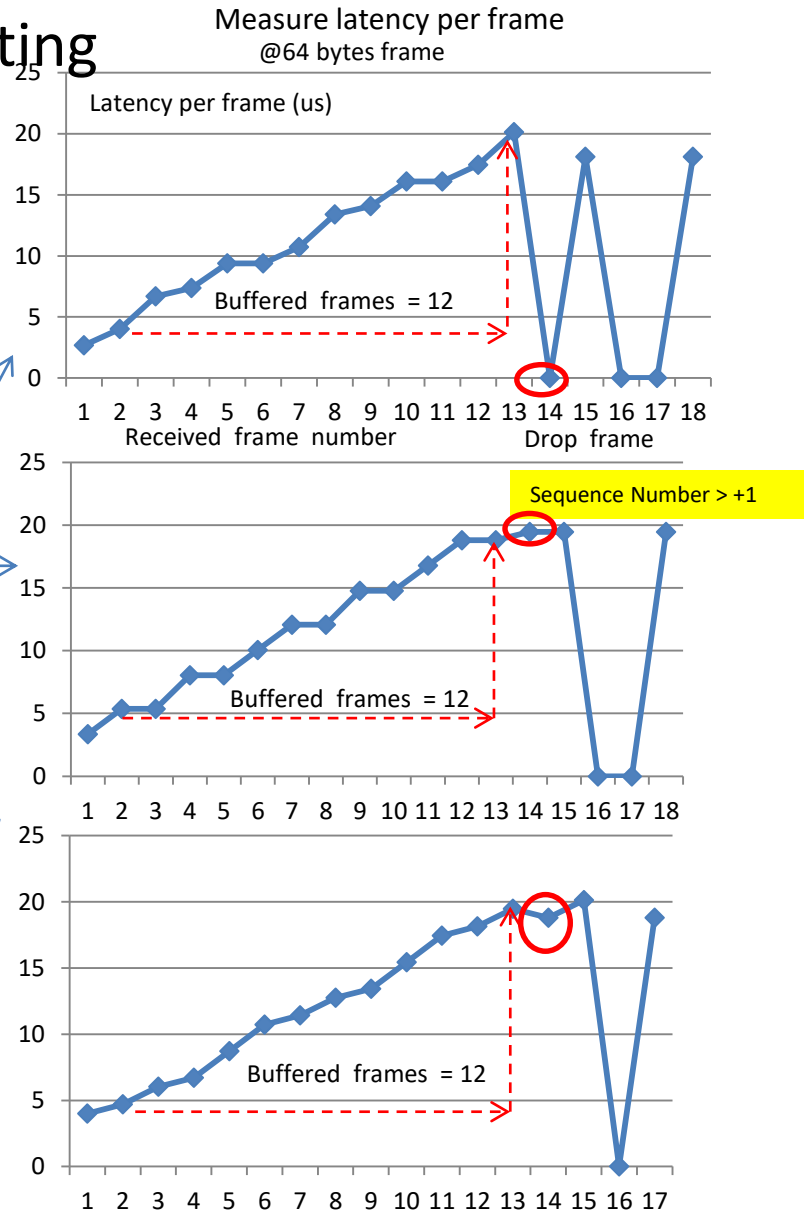
Total buffered frames = 24 + 24
= 48

Yoshiaki Itou

RFC8239 Data Center Benchmarking Methodology

3.Buffering Testing





Total buffered frames = 12 + 12 + 12
= 36

Reporting

If the tester operates using a maximum burst length in frames, then this maximum length SHOULD be reported.

Frame Size, octets	Ave B2B Length, frames	Min,Max,StdDev	Corrected Buff Time, Sec
64	26000	25500,27000,20	0.00004

Back-to-Back Frame Results

Static and configuration parameters:

Number of test repetitions, N

Minimum Step Size (during searches), in frames.

A small area for clarification

RFC 2544 Section 26.4 says:

> **The back-to-back value is the number of frames in the longest burst that the DUT will handle without the loss of any frames. The trial length MUST be at least 2 seconds and SHOULD be repeated at least 50 times with the average of the recorded values being reported.**

@@@@ What should be repeated?
Trial outcomes are loss or no-loss for a given burst length.
Repeating the search for longest burst makes more sense...

BACKUP

Updates in 01

- Clarified text describing *what* is measured
  - Knowledge of approximate buffer storage size (in time or bytes) may be useful to estimate whether frame losses will occur if DUT forwarding is temporarily suspended in a production deployment, due to an unexpected interruption of frame processing (an interruption of duration greater than the estimated buffer would certainly cause lost frames).
- Potential Benefit of using the Correction factor.

Recent Tests @VSPERF

- Presented at IETF-99
- [OPNFV-2017] https://wiki.opnfv.org/download/attachments/10293193/VSPERF-Dataplane-Perf-Cap-Bench.pptx?api=v2
- [VSPERF-b2b] https://wiki.opnfv.org/display/vsperf/Traffic+Generator+Testing#TrafficGeneratorTesting-AppendixB:Back2BackTestingTimeSeries(fromCI)

Several Notable Aspects of the results:

1. **Back-to-back Frame Benchmark was very consistent for some fixed frame sizes, and somewhat variable for others.**

2. **The Back-to-back Frame length reported for large frame sizes was unexpectedly long, and no explanation or measurement limit condition was indicated.**

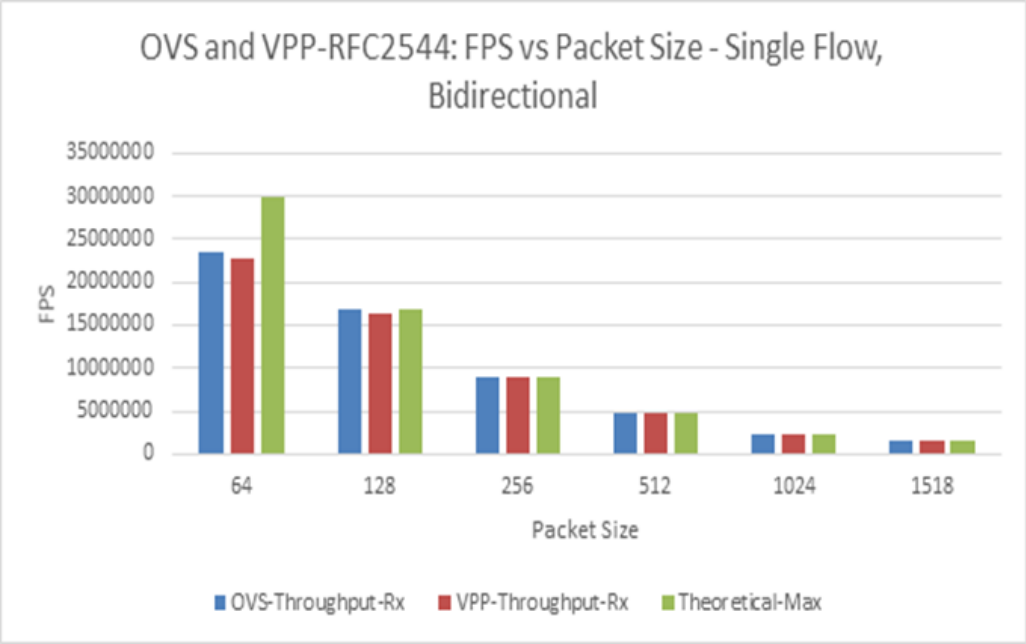
3. **Calculation of the extent of buffer time in the DUT helped explain the results with all frame sizes (some frame sizes cannot exceed the frame header processing rate of the DUT, and therefore no buffering occurs).**

4. **It was observed that the actual buffer time in the DUT could be estimated using results from the Throughput tests conducted according to Section 26.1 of [RFC2544].**

Throughput test results from [RFC2544] can be used to reduce the number of Frame Sizes tested.

Expanding the B2B Frames Method

- Pre-requisite Tests: RFC 2544 Throughput
  - Test setup and configuration must MATCH
  - Recommended Frame Sizes; MUST be zero-loss
  - Reduce Frame size set for B2B (e.g., 64 and 128)



[OPNFV-2017]

Title

(continued)