
Data Center Benchmarking Drafts

draft-bmwg-dcbench-terminology-05(8th revision)
draft-bmwg-dcbench-methodology-02(6th revision)

Lucien Avramov, lavramov@cisco.com

Jacob Rapp, jrapp@vmware.com

July 2016 IETF 96 – Berlin

Data Center Benchmarking Drafts Overview

Terminology

draft-ietf-bmwg-dcbench-terminology-04

Latency

Redefine how latency calculations are used
Update usage of FIFO, FILO, LIFO and LILO

Jitter

Define the application Jitter RFC 3393 and packet size requirement and histogram for DC devices

Physical Layer Calibration

Cable test calibrations and documentation

Line Rate

Consequences of PPM: 99.98%

Buffering

Define Buffering and Buffer Efficiency, Burst, Intensity of Microburst
Define Incast [many-one, many-many]

Application Throughput

Goodput definition and how to measure it

Methodology

draft-bmwg-dcbench-methodology-02

Line Rate Testing

Test all ports at 99.98% including latency, jitter histogram for min/max/avg and drops

Buffering Testing

Buffer highest efficiency
Maximum port buffer size
Maximum port pair buffer size
Maximum DUT buffer size
Microburst

MicroBurst Testing

Use all ports, at 100% intensity of microburst

Head of Line Blocking Testing

Measure two groups (8 ports) of DUT, up to all ports
Reports provides percent of traffic loss during HOLB

Incast Stateful and Stateless Traffic

measure TCP goodput while measuring UDP latency

Changes 1

Added discussion on cut-through vs. store-and-forward switches

skipping to change at *page 5, line 40*

2.2 Discussion

FILO is the most important measuring definition. Any type of switches MUST be measured with the FILO mechanism: FILO will include the latency of the switch and the latency of the frame as well as the serialization delay. It is a picture of the 'whole' latency going through the DUT. For applications, which are latency sensitive and can function with initial bytes of the frame, FIFO MAY be an additional type of measuring to supplement FILO.

LIFO mechanism can be used with store forward type of switches but not with cut-through type of switches, as it will provide negative latency values for larger packet sizes. Therefore this mechanism MUST NOT be used when comparing latencies of two different DUTs.

2.3 Measurement Units

The measuring methods to use for benchmarking purposes are as follow:

- 1) FILO MUST be used as a measuring method, as this will include the latency of the packet; and today the application commonly need to read the whole packet to process the information and take an action.
- 2) FIFO MAY be used for certain applications able to proceed data as

skipping to change at *page 5, line 40*

2.2 Discussion

FILO is the most important measuring definition. Any type of switches MUST be measured with the FILO mechanism: FILO will include the latency of the switch and the latency of the frame as well as the serialization delay. It is a picture of the 'whole' latency going through the DUT. For applications, which are latency sensitive and can function with initial bytes of the frame, FIFO MAY be an additional type of measuring to supplement FILO.

Not all DUTs are exclusively cut-through or store-and-forward. Data Center DUTs are frequently store-and-forward for smaller packet sizes and then adopting a cut-through behavior. FILO covers all scenarios.

LIFO mechanism can be used with store forward type of switches but not with cut-through type of switches, as it will provide negative latency values for larger packet sizes because LIFO removes the serialization delay. Therefore, this mechanism MUST NOT be used when comparing latencies of two different DUTs.

2.3 Measurement Units

The measuring methods to use for benchmarking purposes are as follow:

- 1) FILO MUST be used as a measuring method, as this will include the latency of the packet; and today the application commonly need to read the whole packet to process the information and take an action.
- 2) FIFO MAY be used for certain applications able to proceed data as

Changes 2

Added in note on RFC 2647 and goodput definition – (removed “bits lost” wording)

skipping to change at *page 13, line 39*

a MUST to have a non null percentage of synchronization, which MUST be specified.

7 Application Throughput: Data Center Goodput

7.1. Definition

In Data Center Networking, a balanced network is a function of maximal throughput 'and' minimal loss at any given time. This is defined by the Goodput. Goodput is the application-level throughput. It is measured in bytes / second. Goodput is the measurement of the actual application payload of the packet being sent.

7.2. Discussion

In data center benchmarking, the goodput is a value that SHOULD be measured. It provides a realistic idea of the usage of the available bandwidth. A goal in data center environments is to maximize the goodput while minimizing the loss.

7.3. Measurement Units

skipping to change at *page 13, line 46*

a MUST to have a non null percentage of synchronization, which MUST be specified.

7 Application Throughput: Data Center Goodput

7.1. Definition

In Data Center Networking, a balanced network is a function of maximal throughput 'and' minimal loss at any given time. This is defined by the Goodput. Goodput is the application-level throughput. The definition used is a variance of the definition in RFC 2647.

Goodput is the number of bits per unit of time forwarded to the correct destination interface of the DUT/SUT, minus any bits retransmitted.

7.2. Discussion

In data center benchmarking, the goodput is a value that SHOULD be measured. It provides a realistic idea of the usage of the available bandwidth. A goal in data center environments is to maximize the goodput while minimizing the loss.

7.3. Measurement Units

Changes 3

Big thanks to Al & Scott!

skipping to change at *page 15, line 25*

3.3. URL References

- [5] Yanpei Chen, Rean Griffith, Junda Liu, Randy H. Katz, Anthony D. Joseph, "Understanding TCP Incast Throughput Collapse in Datacenter Networks", <http://www.eecs.berkeley.edu/~ychen2/professional/TCPIncastWREN2009.pdf>".

3.4. Acknowledgments

The authors would like to thank Ian Cox and Tim Stevenson for their reviews and feedback.

skipping to change at *page 15, line 33*

3.3. URL References

- [5] Yanpei Chen, Rean Griffith, Junda Liu, Randy H. Katz, Anthony D. Joseph, "Understanding TCP Incast Throughput Collapse in Datacenter Networks", <http://www.eecs.berkeley.edu/~ychen2/professional/TCPIncastWREN2009.pdf>".

3.4. Acknowledgments

The authors would like to thank Alfred Morton, Scott Bradner, Ian Cox, Tim Stevenson for their reviews and feedback.

Next Steps

What's needed next?

- Currently Last Call
- “looks good to me
Scott”