Network Working Group
INTERNET-DRAFT
Expires in: September 2006

Scott Poretsky
Reef Point Systems

March 2006

Methodology for Benchmarking Network-layer Traffic Control Mechanisms

<draft-ietf-bmwg-dsmmeth-01.txt>

Status of this Memo

Copyright Notice

Abstract

This document describes the methodology for the benchmarking of devices that implement traffic control based on IP precedence or diff-serv code point criteria. The methodology is to be applied to measurements made on the data plane to evaluate the performance of the traffic control mechanisms. The methodology permits the specific traffic control mechanisms and configuration commands to vary between DUTs. The methodology uses much of the Terminology defined in [Pp06].

Table of Contents

1. Introduction

This document describes the methodology for the benchmarking of devices that implement traffic control based on IP precedence or diff-serv code point criteria. The methodology is to be applied to measurements made on the data plane to evaluate the performance of the traffic control mechanisms. The methodology permits the specific traffic control mechanisms and configuration commands to vary between DUTs. The methodology uses much of the Terminology defined in [Pp06].

2. Existing definitions

For the sake of clarity and continuity this RFC adopts the template for definitions set out in Section 2 of RFC 1242. Definitions are indexed and grouped together in sections for ease of reference. Reference [Pp06] for benchmarking terminology.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [Br97]. RFC 2119 defines the use of these key words to help make the intent of standards track documents as clear as possible. While this document uses these keywords, this document is not a standards track document.

3. Test Setup

3.1 Test Topologies

Figure 1 shows the Test Topology for benchmarking performance when Forwarding Congestion does not exist on the egress link. This topology is to be used when benchmarking the Undifferentiated Response and the Traffic Control without Forwarding Congestion Figure 2 shows the Test Topology for benchmarking performance when Forwarding Congestion does not exist on the egress link. This topology is to be used when benchmarking the Traffic Control with Forwarding Congestion. The Forwarding Congestion is produced by offering load to two ingress interfaces on the DUT destined for the same single egress interface. The aggregate of the ingress offered load MUST exceed the Forwarding Capacity of the egress link to produce Forwarding Congestion.

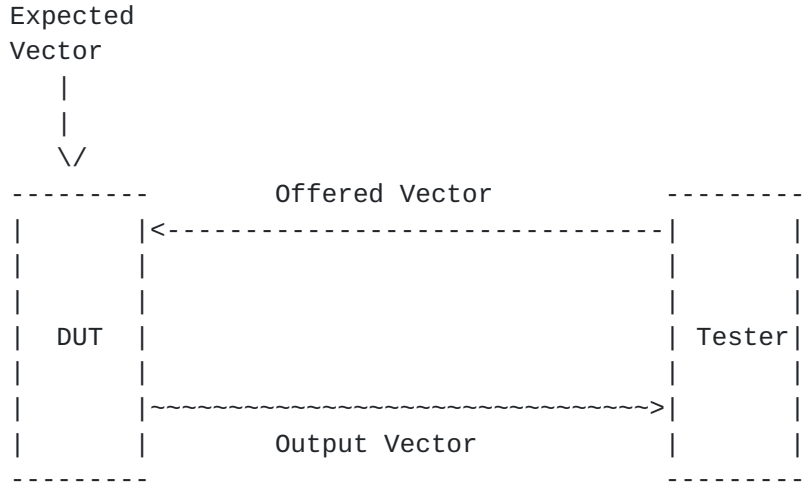
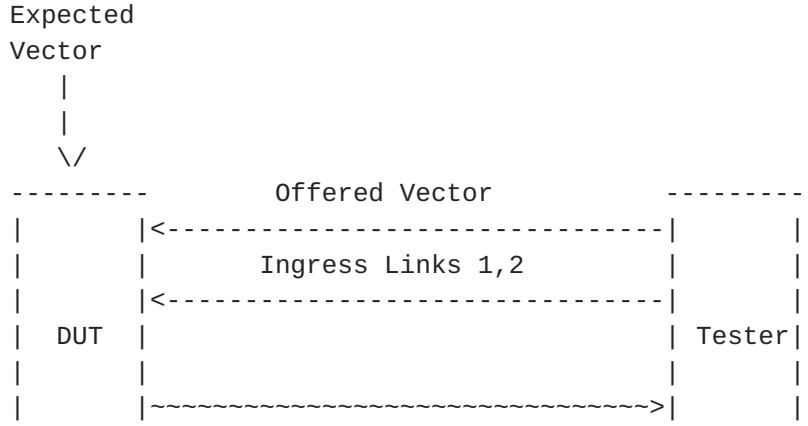


Figure 1. Test Topology for Benchmarking
Without Forwarding Congestion



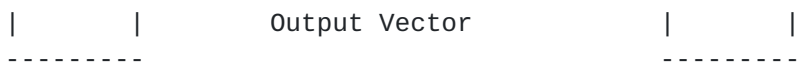


Figure 2. Test Topology for Benchmarking
With Forwarding Congestion

3.2 Test Considerations

3.2.1 Routing Configuration

Routing Protocols SHOULD NOT be used. All routing decisions SHOULD be made based upon pre-configured static routes.

3.2.2 Interface Types

All test cases in this methodology document may be executed with any interface type. All interfaces MUST be the same media and Throughput [5,6] for each test case.

3.2.3 Offered Vector

The Offered Vector MUST be configured on the Tester as follows:

a. The Offered Load MUST be the Forwarding Capacity of the device at a fixed packet size.

b. The Forwarding Capacity MUST be measured at the egress interface of the DUT

c. Each test case MUST be executed using a single, selectable packet size. Packet Size is measured in bytes and includes the IP header and payload. If IPsec packets are used then the packet size also includes it. Packet Size must be equal to or less than the interface MTU so that there is no fragmentation.

d. It is RECOMMENDED that the number of flows used be 1000, 10000, and/or 100000. A flow MUST be identified by its DSCP, IP Source Address, and IP Destination Address.

e. It is RECOMMENDED that the number of DSCPs used be 1, 2, 3, 4, 6, 8, 16, and/or 64. When the number of DSCPs is 1 then the Undifferentiated Response is benchmarked. The actual values of the DSCPs used is selectable.

3.2.4 Test Duration

It is RECOMMENDED that the Test Duration for each test case includes a minimum of 10 minutes of Offered Load and Output Vector measurement

3.2.5 Expected Vector

The Expected Vector is configured on the DUT. The Traffic Control mechanisms and specific configuration commands may vary between DUTs. Test Cases may be repeated with variation to the Expected Vector to produce a more benchmark results.

3.3 Reporting Format

For each test case, it is recommended that the following reporting format be completed:

PARAMETERS	UNITS
Offered Vector	
Offered Load	pps
Number of DSCPs	{1..64}
Codepoint Set	{0..63, 0..63, ... , x}
Number of Flows	{1000, 10000, 100000}
Number of Flows per DSCP	Number of Flows/Number of DSCPs
Packet Size	bytes
Undifferentiated Response (Number of DSCPs = 1)	
Forwarding Capacity	pps
Packet Loss	packets
Forwarding Delay	
Minimum	msec
Maximum	msec
Average	msec
Jitter	
Average	msec
Peak-to-Peak	msec
Out-of-Order Packets	packets
Duplicate Packets	packets
Expected Vector {for DSCP=n} (as configured on DUT)	
Forwarding Capacity	pps
Packet Loss	packets
Forwarding Delay	
Minimum	msec
Maximum	msec
Average	msec
Output Vector {for DSCP=n}	
Forwarding Capacity	pps
Packet Loss	packets
Forwarding Delay	
Minimum	msec

```
        Maximum          msec
        Average          msec
    Jitter
        Average          msec
        Peak-to-Peak     msec
    Out-of-Order Packets     packets
    Duplicate Packets        packets
```

4. Test Cases

4.1 Undifferentiated Response

Purpose:
To establish the baseline performance of the DUT.

Procedure:
1. Configure DUT with Expected Vector.
2. Configure the Tester for the Offered Vector.
   Number of DSCPs MUST equal 1 and the
   RECOMMENDED DSCP value is 0 (Best Effort).
   Use 1000 Flows identified by IP SA/DA. All flows
   have the same DSCP value.
3. Using the Test Topology in Figure 1, source the
   Offered Load from the Tester to the DUT.
4. Measure and record the Output Vector.
5. Maintain offered load for 10 minutes minimum
   to observe possible variations in measurements.
6. Repeat steps 2 through 5 with 10000 and 100000
   Flows.

Expected Results:
Forwarding Vector equals the Offered Load. There
is no packet loss and no out-of-order packets.

4.2 Traffic Control Baseline Performance

Purpose:
To benchmark the Output Vectors for a Codepoint Set
without Forwarding Congestion.

Procedure:
1. Configure DUT with Expected Vector for each DSCP in
   the Codepoint Set.
2. Configure the Tester for the Offered Vector.
   Number of DSCPs MUST 2 or more. Any DSCP values can
   be used. Use 1000 Flows identified by IP SA/DA
   and DSCP value.
3. Using the Test Topology in Figure 1, source the
   Offered Load from the Tester to the DUT.
4. Measure and record the Output Vector for each DSCP
   in the Codepoint Set.
5. Maintain offered load for 10 minutes minimum
   to observe possible variations in measurements.
6. Repeat steps 2 through 5 with 10000 and 100000
   Flows.
7. Increment number of DSCPs used and repeat steps

1 through 6.

Expected Results:
Forwarding Vector equals the Offered Load. There is no packet loss and no out-of-order packets. Output vectors match the Expected Vectors for each DSCP in the Codepoint Set.

4.3 Traffic Control Performance with Forwarding Congestion

Purpose:
To benchmark the Output Vectors for a Codepoint Set with Forwarding Congestion.

Procedure:

1. Configure DUT with Expected Vector for each DSCP in the Codepoint Set.
2. Configure the Tester for the Offered Vector. Number of DSCPs MUST 2 or more. Any DSCP values can be used. Use 1000 Flows identified by IP SA/DA and DSCP value. The Offered Load MUST exceed the Forwarding Capacity of a single egress link by 25% using 2 ingress links.
3. Using the Test Topology in Figure 2, source the Offered Load from the Tester to the DUT. The aggregate of the ingress offered load MUST exceed the Forwarding Capacity of the egress link to produce Forwarding Congestion.
4. Measure and record the Output Vector for each DSCP in the Codepoint Set.
5. Maintain offered load for 10 minutes minimum to observe possible variations in measurements.
6. Repeat steps 2 through 5 with 10000 and 100000 Flows.
7. Increment offered load by 25% to 200% maximum.
8. Increment number of DSCPs used and repeat steps 1 through 6.

Expected Results:
Forwarding Vector equals the Offered Load. There is no packet loss and no out-of-order packets. Output vectors match the Expected Vectors for each DSCP in the Codepoint Set.

5. IANA Considerations

This document requires no IANA considerations.

6. Security Considerations

Documents of this type do not directly affect the security of the Internet or of corporate networks as long as benchmarking is not performed on devices or systems connected to production networks.

Packets with unintended and/or unauthorized DSCP or IP precedence values may present security issues. Determining the security consequences of such packets is out of scope for this document.

7. Acknowledgments

8. References

8.1 Normative References

[Br91] Bradner, S., "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, July 1991.

[Br97] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997

[Br98] Braden, B., Clark, D., Crowcroft, J., Davie, B., Deering, S., Estrin, D., Floyd, S., Jacobson, V., Minshall, G., Partridge, C., Peterson, L., Ramakrishnan, K., Shenker, S., Wroclawski, J. and L. Zhang, "Recommendations on Queue Management and Congestion Avoidance in the Internet", RFC 2309, April 1998.

[Ma98] Mandeville, R., "Benchmarking Terminology for LAN Switching Devices", RFC 2285, July 1998.

[Ni98] Nichols, K., Blake, S., Baker, F., Black, D., "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[Pp06] Perser, J., Poretsky, S., Erramilli, S., and Khurana, S., "Terminology for Benchmarking Network-layer Traffic Control Mechanisms", draft-ieft-bwmg-dsmterm-12, work in progress, 2006.

8.2 Informative References

[Bl98] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., Weiss, W., "An Architecture for Differentiated Services", RFC 2475, December 1998.

[Br99] Bradner, S., McQuaid, J. "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999

[Fl93] Floyd, S., and Jacobson, V., "Random Early Detection gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, V.1 N.4, August 1993, p. 397-413. URL "ftp://ftp.ee.lbl.gov/papers/early.pdf".

[Ja99] Jacobson, V., Nichols, K., Poduri, K., "An Expedited Forwarding PHB", RFC 2598, June 1999

[Ma91] Mankin, A., Ramakrishnan, K., "Gateway Congestion Control Survey", RFC 1254, August 1991

[Sc96] Schulzrinne, H., Casner, S., Frederick, R., Jacobson, V.,

"RTP: A Transport Protocol for Real-Time Applications",
RFC 1889, January 1996

9. Authors' Addresses

Scott Poretsky
Reef Point Systems
8 New England Executive Park
Burlington, MA 01803
USA

Phone: + 1 508 439 9008
EMail: sporetsky@reefpoint.com

copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.