Network Working Group C. Popoviciu
Internet-Draft A. Hamza
Expires: July 5, 2007 G. Van de Velde
Cisco Systems
D. Dugatkin
IXIA
January 2007

IPv6 Benchmarking Methodology
<draft-ietf-bmwg-ipv6-meth-01.txt>

Abstract

The Benchmarking Methodologies defined in RFC2544 [2] are IP version independent however, they do not address some of the specificities of IPv6. This document provides additional benchmarking guidelines which in conjunction with RFC2544 lead to a more complete and realistic evaluation of the IPv6 performance of network elements.

Table of Contents

1. Introduction

The benchmarking methodologies defined by RFC2544 [2] are proving to be very useful in consistently evaluating IPv4 forwarding performance of network elements. Adherence to these testing and result analysis procedures facilitates objective comparison of IPv4 performance data measured on various products and by various individuals. While the principles behind the methodologies introduced in RFC2544 are largely IP version independent, the protocol continued to evolve, particularly in its version 6 (IPv6).

This document provides benchmarking methodology recommendations that address IPv6 specific aspects such as evaluating the forwarding performance of traffic containing extension headers as defined in RFC2460 [5]. These recommendations are defined within the RFC2544 framework and are meant to complement the test and result analysis procedures as described in RFC2544 and not to replace them.

The terms used in this document remain consistent with those defined in "Benchmarking Terminology for Network Interconnect Devices" [3]. This terminology SHOULD be consulted before using or applying the recommendations of this document.

Any methodology aspects not covered in this document SHOULD be assumed to be treated based on the RFC2544 recommendations.

2. Existing Definitions

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 [1]. 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. Tests and Results Evaluation

The recommended performance evaluation tests are described in Section 6 of this document. Not all of these tests are applicable to all network element types. Nevertheless, for each evaluated device it is recommended to perform as many of the applicable tests described in Section 6 as possible.

Test execution and the results analysis MUST be performed while observing generally accepted testing practices regarding

repeatability, variance and statistical significance of small numbers of trials.

4. Test Environment Set Up

The test environment setup options recommended for the IPv6 performance evaluation are the same as the ones described in Section 6 of RFC2544, in both single-port and multi-port scenarios. Single-port testing is used in measuring per interface forwarding performance while multi-port testing is used to measure the scalability of forwarding performance across the entire platform.

Throughout the test, the DUT can be monitored for relevant resource (Processor, Memory, etc.) utilization. This data could be beneficial in understanding traffic processing by each DUT and the resources that must be allocated for IPv6. It could reveal if the IPv6 traffic is processed in hardware, by applicable devices, under all test conditions or it is punted in the software switched path. If such data is considered of interest, it MUST be collected out of band and independent of any management data that might be recommended to be collected through the interfaces forwarding the test traffic.

Note: During testing, either static or dynamic options for neighbor discovery can be used. The static option can be used as long as it is supported by the test tool. The dynamic option is preferred if the test tool interacts with the DUT for the duration of the test to maintain the respective neighbor caches in an active state. The test scenarios assume the test traffic simulated end points and the IPv6 source and destination addresses are not directly attached to the DUT, but are seen as one hop beyond, to avoid Neighbor Solicitation (NS) and Neighbor Advertisement (NA) storms due to the Neighbor Unreachability Detection (NUD) mechanism [6].

5. Test Traffic

The traffic used for all tests described in this document SHOULD meet the requirements described in this section. These requirements are designed to reflect the characteristics of IPv6 unicast traffic in all its aspects. Using the recommended IPv6 traffic profile leads to a complete evaluation of the network element performance.

5.1. Frame Formats and Sizes

Two types of media are commonly deployed and SHOULD be tested: Ethernet and SONET. This section identifies the frame sizes that SHOULD be used for each media type. Refer to recommendations in

RFC2544 for all other media types.

Similar to IPv4, small frame sizes help characterize the per-frame processing overhead of the DUT. Note that the minimum IPv6 packet size (40 bytes) is larger than that of an IPv4 packet (20 bytes). Tests should compensate for this difference.

The frame sizes listed for IPv6 include the extension headers used in testing (see section 4.3). By definition, extension headers are part of the IPv6 packet payload. Depending on the total length of the extension headers, their use might not be possible at the smallest frame sizes.

5.1.1. Frame Sizes to be used on Ethernet

Ethernet in all its types has become the most commonly deployed interface in today's networks. The following frame sizes SHOULD be used for benchmarking over this media type: 64, 128, 256, 512, 1024, 1280, 1518 bytes. The 4096, 8192, 9216 bytes long jumbo frame sizes SHOULD be used when benchmarking Gigabit Ethernet interfaces. The maximum frame rates for each frame size and the various Ethernet interface types are provided in Appendix A.

5.1.2. Frame Sizes to be used on SONET

Packet over SONET (PoS) interfaces are commonly used for edge uplinks and high bandwidth core links. Evaluating the forwarding performance of PoS interfaces supported by the DUT is recommended. The following frame sizes SHOULD be used for this media type: 64, 128, 256, 512, 1024, 1280, 1518, 2048, 4096 bytes. The maximum frame rates for each frame size and the various PoS interface types are provided in Appendix A.

5.2. Protocol Addresses Selection

There are two aspects of IPv6 benchmarking testing where IP address selection considerations MUST be analyzed: The selection of IP addresses for the DUT and the selection of addresses for the test traffic.

5.2.1. DUT Protocol Addresses

IANA reserved the IPv6 address block xxxxx/48 for use with IPv6 benchmark testing. These addresses MUST not be assumed to be routable on the Internet and MUST not be used as Internet source or destination addresses.

Similar to RFC2544, Appendix C, addresses from the first half of this

range SHOULD be used for the ports viewed as input and addresses from the other half of the range for the output ports.

The prefix length of the IPv6 addresses configured on the DUT interfaces MUST fall into either one of the following:

o Prefix length is /126 which would simulate a point-to-point link for a core router.
o Prefix length is smaller or equal to /64.

No prefix lengths SHOULD be selected in the range between 64 and 128 except the 126 value mentioned above.

Note that /126 prefixes might not be always the best choice for addressing point-to-point links such as back-to-back Ethernet unless the autoprovisioning mechanism is disabled. Also, not all network elements support this type of addresses.

While with IPv6, the DUT interfaces can be configured with multiple global unicast addresses, the methodology described in this document does not require testing such a scenario. It is not expected that such an evaluation would bring additional data with respect to the performance of the network element.

The Interface ID portion of the global unicast IPv6 DUT addresses SHOULD be set to ::1. There are no requirements in the selection of the Interface ID portion of the link local IPv6 addresses.

It is recommended that multiple iterations of the benchmark tests be conducted using the following lengths: 32, 48, 64, 126 and 128 for the advertised traffic destination prefix. Other prefix lengths can also be used if desired, however the indicated range should be sufficient to establish baseline performance metrics.

5.2.2. Test Traffic Protocol Addresses

The IPv6 source and destination addresses for the test streams SHOULD belong to the IPv6 range assigned by IANA as discussed in section 4.2.1. The source addresses SHOULD belong to one half of the range and the destination addresses to the other, reflecting the DUT interface IPv6 address selection.

Tests SHOULD first be executed with a single stream leveraging a single source-destination address pair. The tests SHOULD then be repeated with traffic using a random destination address in the corresponding range. If the network element prefix lookup capabilities are evaluated, the tests SHOULD focus on the IPv6 relevant prefix boundaries: 0-64, 126 and 128.

Special care needs to be taken about the Neighbor Unreachability

Detection (NUD) [6] process. The IPv6 prefix reachable time in the router advertisement SHOULD be set to 30 seconds and allow 50% jitter. The IPv6 source and destination addresses SHOULD not appear to be directly connected to the DUT to avoid Neighbor Solicitation (NS) and Neighbor Advertisement (NA) storms due to multiple test traffic flows.

5.3. Traffic with Extension Headers

Extension headers are an intrinsic part of the IPv6 architecture [5]. They are used with various types of practical traffic such as: Fragmented traffic, mobile IP based traffic, authenticated and encrypted traffic. For these reasons, all tests described in this document SHOULD be performed with both traffic that has no extension headers and traffic that has a set of extension headers. They can be selected from the following list [5] which reflects the recommended order of multiple extension headers in a packet:

- o Hop-by-hop header
- o Destination options header
- o Routing header
- o Fragment header
- o Authentication header
- o Encapsulating Security Payload header
- o Destination options header
- o Mobility header

Considering the fact that extension headers are an intrinsic part of the protocol and that they fulfill different roles, benchmarking of traffic containing each extension header SHOULD be executed individually.

The special processing rules for the Hop-by-hop extension header require a specific benchmarking approach. Unlike the other extension headers, this header must be processed by each node that forwards the traffic. Tests with traffic containing this extension headers type will not measure the forwarding performance of the DUT but rather its extension headers processing ability which is dependent on the information contained in the extension headers. The concern is that this traffic, at high rates, could have a negative impact on the operational resources of the router and could be used as a security threat. When benchmarking with traffic that contains the Hop-by-hop extension header, the goal is not to measure throughput [2] as in the case of the other extension headers but rather to evaluate the impact of such traffic on the router. In this case, traffic with the Hop-by-hop extension headers should be sent at 1%, 10% and 50% of the interface total bandwidth. Device resources must be monitored at each traffic rate to determine the impact.

The tests with traffic containing each individual extension header MUST be complemented with tests that contain a chain of two or more extension headers (the chain MUST not contain the Hop-by-hop extension header). The chain should also exclude the ESP extension header since traffic with an encrypted payload can not be used in tests with modifiers such as filters based on upper layer information (see Section 5). Since the DUT is not analyzing the content of the extension headers, any combination of extension headers can be used in testing. The extension headers chain recommended to be used in testing is:

- o Routing header - 24-32 bytes
- o Destination options header - 8 bytes
- o Fragment header - 8 bytes

This is a real life extension headers chain that would be found in an IPv6 packet between two mobile nodes exchanged over an optimized path that requires fragmentation. The listed extension headers lengths represent test tool defaults. The total length of the extension headers chain SHOULD be larger than 32 bytes.

Extension headers add extra bytes to the payload size of the IP packets which MUST be factored in when used in testing at low frame sizes. Their presence will modify the minimum packet size used in testing. For direct comparison between the data obtained with traffic that has extension headers and with traffic that doesn't have them, at low frame size, a common bottom size SHOULD be selected for both types of traffic.

For the most cases, the network elements ignore the extension headers when forwarding IPv6 traffic. For these reasons it is most likely that the extension headers related performance impact will be observed only when testing the DUT with traffic filters that contain matching conditions for the upper layer protocol information. In those cases, the DUT MUST traverse the chain of extension headers, a process that could impact performance.

5.4. Traffic set up

All tests recommended in this document SHOULD be performed with bi-directional traffic. For asymmetric situations, tests MAY be performed with unidirectional traffic for a more granular characterization of the DUT performance. In these cases, the bidirectional traffic testing would reveal only the lowest performance between the two directions.

All other traffic profile characteristics described in RFC2544 SHOULD be applied in this testing as well. IPv6 multicast benchmarking is outside the scope of this document.

6. Modifiers

RFC2544 underlines the importance of evaluating the performance of network elements under certain operational conditions. The conditions defined in RFC2544 Section 11 are common to IPv4 and IPv6 with the exception of broadcast frames. IPv6 does not use layer 2 or layer 3 broadcasts. This section provides additional conditions that are specific to IPv6. The suite of tests recommended in this document SHOULD be first executed in the absence of these conditions and then repeated under each of these conditions separately.

6.1. Management and Routing Traffic

The procedures defined in RFC2544 sections 11.2 and 11.3 are applicable for IPv6 management and routing update Frames as well.

6.2. Filters

The filters defined in Section 11.4 of RFC2544 apply to IPv6 benchmarking as well. The filter definitions however must be expanded to include upper layer protocol information matching in order to analyze the handling of traffic with extension headers which are an important architectural component of IPv6.

6.2.1. Filter Format

The filter format defined in RFC2544 is applicable to IPv6 as well except that the Source Addresses (SA) and Destination Addresses (DA) are IPv6. In addition to these basic filters, the evaluation of IPv6 performance SHOULD analyze the correct filtering and handling of traffic with extension headers.

While the intent is not to evaluate a platform's capability to process the various extension header types, the goal is to measure performance impact when the network element must parse through the extension headers in order to reach upper layer information. In IPv6, routers do not have to parse through the extension headers (other than Hop-by-hop) unless, for example, the upper layer information has to be analyzed due to filters.

For these reasons, to evaluate the network element handling of IPv6 traffic with extension headers, the definition of the filters must be extended to include conditions applied to upper layer protocol information. The following filter format SHOULD be used for this type of evaluation:

```
[permit|deny]  [protocol] [SA] [DA]
```

where permit or deny indicates the action to allow or deny a packet through the interface the filter is applied to. The Protocol field is defined as:

o ipv6: any IP Version 6 traffic
o tcp: Transmission Control Protocol
o udp: User Datagram Protocol

and SA stands for the Source Address and DA for the Destination Address.

6.2.2. Filter Types

Based on the RFC2544 recommendations, two types of tests are executed when evaluating performance in the presence of modifiers. One with a single filter and one with 25 filters. The recommended filters are exemplified with the help of the IPv6 documentation prefix [9] 2001:DB8::.

Examples of single filters are:

```
Filter for TCP traffic - permit tcp 2001:DB8::1 2001:DB8::2
Filter for UDP traffic - permit udp 2001:DB8::1 2001:DB8::2
Filter for IPv6 traffic - permit ipv6 2001:DB8::1 2001:DB8::2
```

The single line filter case SHOULD verify that the network element permits all TCP/UDP/IPv6 traffic with or without any number of extension headers from IPv6 SA 2001:DB8::1 to IPv6 DA 2001:DB8::2 and deny all other traffic.

Example of 25 filters:

```
deny tcp 2001:DB8:1::1 2001:DB8:1::2
deny tcp 2001:DB8:2::1 2001:DB8:2::2
deny tcp 2001:DB8:3::1 2001:DB8:3::2
...
deny tcp 2001:DB8:C::1 2001:DB8:C::2
permit tcp 2001:DB8:99::1 2001:DB8:99::2
deny tcp 2001:DB8:D::1 2001:DB8:D::2
deny tcp 2001:DB8:E::1 2001:DB8:E::2
...
deny tcp 2001:DB8:19::1 2001:DB8:19::2
deny ipv6 any any
```

The router SHOULD deny all traffic with or without extension headers except TCP traffic with SA 2001:DB8:99::1 and DA 2001:DB8:99::2.

7. Benchmarking Tests

This document recommends the same benchmarking tests described in RFC2544 while observing the DUT setup and the traffic setup considerations described above. The following sections state the test types explicitly and highlight only the methodology differences that might exist with respect to those described in Section 26 of RFC2544.

The specificities of IPv6, particularly the definition of extension headers processing, require additional benchmarking steps. In this sense, the tests recommended by RFC2544 MUST be repeated for IPv6 traffic without extension headers and with one or multiple extension headers.

IPv6's deployment in existing IPv4 environments and the expected long co-existence of the two protocols leads network operators to place great emphasis on understanding the performance of platforms forwarding both types of traffic. While device resources are shared between the two protocols, it is important for IPv6 enabled platforms to not experience degraded IPv4 performance. Thus, IPv6 benchmarking SHOULD be performed in the context of a stand alone protocol as well as in the context of its co-existence with IPv4.

The modifiers defined are independent of extension header type so they can be applied equally to each one of the above tests.

The benchmarking tests described in this section SHOULD be performed under each of the following conditions:

Extension headers specific conditions:
- i) IPv6 traffic with no extension headers
- ii) IPv6 traffic with one extension header from the list in section 4.3
- iii) IPv6 traffic with the chain of extension headers described in section 4.3

Co-existence specific conditions:
- iv) IPv4 ONLY traffic benchmarking
- v) IPv6 ONLY traffic benchmarking
- vi) IPv4-IPv6 traffic mix with the ratio 90% vs 10%
- vii) IPv4-IPv6 traffic mix with the ratio 50% vs 50%
- viii) IPv4-IPv6 traffic mix with the ratio 10% vs 90%

Combining the test conditions listed for benchmarking IPv6 as a stand-alone protocol and the co-existence tests leads to a large coverage matrix. A minimum requirement is to cover the co-existence conditions in the case of IPv6 with no extension headers and those

where either of the traffic is 10% and 90% respectively.

The subsequent sections describe each specific tests that MUST be executed under the conditions listed above for a complete benchmarking of IPv6 forwarding performance.

7.1. Throughput

Objective: To determine the DUT throughput as defined in RFC1242.

Procedure: Same as RFC2544.

Reporting Format: Same as RFC2544.

7.2. Latency

Objective: To determine the latency as defined in RFC1242.

Procedure: Same as RFC2544.

Reporting Format: Same as RFC2544.

7.3. Frame Loss

Objective: To determine the frame loss rate, as defined in RFC1242, of a DUT throughout the entire range of input data rates and frame sizes.

Procedure: Same as RFC2544.

Reporting Format: Same as RFC2544.

7.4. Back-to-Back Frames

Objective: To characterize the ability of a DUT to process back-to-back frames as defined in RFC1242.

Based on the IPv4 experience, the Back-to-Back frames test is characterized by significant variance due to short term variations in the processing flow. For these reasons, this test is not recommended anymore for IPv6 benchmarking.

7.5. System Recovery

Objective: To characterize the speed at which a DUT recovers from an overload condition.

Procedure: Same as RFC2544.

Reporting Format: Same as RFC2544.

7.6. Reset

Objective: To characterize the speed at which a DUT recovers from a device or software reset.

Procedure: Same as RFC2544.

Reporting Format: Same as RFC2544.

8. IANA Considerations

IANA reserved prefix xxxx/48 for IPv6 benchmarking similar to 198.18.0.0/15 in RFC 3330 [8]. This prefix length provides similar flexibility as the range allocated for IPv4 benchmarking and it is taking into consideration address conservation and simplicity of usage concerns. Most network infrastructures are allocated a /48 prefix, hence this range would allow most network administrators to mimic their IPv6 Address Plans when performing IPv6 benchmarking.

9. Security Considerations

Benchmarking activities as described in this memo are limited to technology characterization using controlled stimuli in a laboratory environment, with dedicated address space and the constraints specified in the sections above.

The benchmarking network topology will be an independent test setup and MUST NOT be connected to devices that may forward the test traffic into a production network, or misroute traffic to the test management network.

Further, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the DUT/SUT.

Special capabilities SHOULD NOT exist in the DUT/SUT specifically for benchmarking purposes. Any implications for network security arising from the DUT/SUT SHOULD be identical in the lab and in production networks.

The isolated nature of the benchmarking environments and the fact that no special features or capabilities, other than those used in operational networks, are enabled on the DUT/SUT requires no security considerations specific to the benchmarking process.

10. Conclusions

The Benchmarking Methodology for Network Interconnect Devices document, RFC2544 [2], is for the most part applicable to evaluating the IPv6 performance of network elements. This document is addressing the IPv6 specific requirements that MUST be observed when applying the recommendations of RFC2544. These additional requirements stem from the architecture characteristics of IPv6. This document is not a replacement of but a complement to RFC2544.

11. Acknowledgements

Scott Bradner provided valuable guidance and recommendations for this document. The authors acknowledge the work done by Cynthia Martin and Jeff Dunn with respect to defining the terminology for IPv6 benchmarking. The authors would like to thank Bill Kine for his contribution to the initial document and to Bill Cerveny, Silvija Dry, Sven Lanckmans, Dean Lee, Athanassios Liakopoulos, Benoit Lourdelet, Al Morton, Rajiv Papejna and Pekka Savola for their very helpful feedback. Maryam Hamza inspired the authors in completing this document.

12. References

12.1. Normative References

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

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

12.2. Informative References

[3] Bradner, S., "Benchmarking terminology for network interconnection devices", RFC 1242, July 1991.

[4] Simpson, W., "PPP in HDLC-like Framing", STD 51, RFC 1662, July 1994.

[5] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[6] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[7] Malis, A. and W. Simpson, "PPP over SONET/SDH", RFC 2615, June 1999.

[8] IANA, "Special-Use IPv4 Addresses", RFC 3330, September 2002.

[9] Huston, G., Lord, A., and P. Smith, "IPv6 Address Prefix Reserved for Documentation", RFC 3849, July 2004.

[10] Newman, D. and T. Player, "Hash and Stuffing: Overlooked Factors in Network Device Benchmarking", (draft-ietf-bmwg-hash-stuffing-08 (work in progress)", January 2007.

Appendix A. Maximum Frame Rates Reference

This appendix provides the formulas to calculate and the values for the maximum frame rates for two media types: Ethernet and SONET.

A.1. Ethernet

The maximum throughput in frames per second (fps) for various Ethernet interface types and for a frame size X can be calculated with the following formula:

```
        Line Rate (bps)
------------------------------
(8bits/byte)*(X+20)bytes/frame
```

The 20 bytes in the formula is the sum of the Preamble (8 bytes) and the Inter Frame Gap (12 bytes). The maximum throughput for various Ethernet interface types and frame sizes:

Size Bytes	10Mb/s pps	100Mb/s pps	1000Mb/s pps	10000Mb/s pps
64	14881	148810	1488096	14880952
128	8446	84449	844595	8445946
256	4529	45290	452899	4528986
512	2350	23497	234962	2349625
1024	1198	11973	119731	1197318
1280	961	9616	96153	961538
1518	813	8128	81275	812744
4096	303	3036	30369	303692
8192	152	1522	15221	152216
9216	135	1353	13534	135340

A.2. Packet over SONET

ANSI T1.105 SONET provides the formula for calculating the maximum available bandwidth for the various Packet over SONET (PoS) interface types:

STS-Nc (N = 3Y, where Y=1,2,3,etc)

[(N*87) - N/3]*(9 rows)*(8 bit/byte)*(8000 frames/sec)

Packets over SONET can use various encapsulations: PPP [7], HDLC [4] and Frame Relay. All these encapsulations use a 4 bytes header, a 2 or 4 bytes FCS field and a 1 byte Flag which are all accounted for in the overall frame size. The maximum frame rate for various interface types can be calculated with the formula (where X represents the frame size in bytes):

$$\frac{\text{Line Rate (bps)}}{\text{(8bits/byte)*(X+1)bytes/frame}}$$

The maximum throughput for various PoS interface types and frame sizes:

Size Bytes	OC-3c fps	OC-12c fps	OC-48c fps	OC-192c fps	OC-768c fps
64	288,000	1,152,000	4,608,000	18,432,000	73,728,000
128	145,116	580,465	2,321,860	9,287,442	37,149,767
256	72,840	291,362	1,165,447	4,661,790	18,647,160
512	36,491	145,965	583,860	2,335,439	9,341,754
1024	18,263	73,054	292,215	1,168,859	4,675,434
2048	9,136	36,545	146,179	584,714	2,338,858
4096	4,569	18,277	73,107	292,429	1,169,714

It is important to note that throughput test results may vary from the values presented in appendices A.1 and A.2 due to bit stuffing performed by various media types [10].

Authors' Addresses

Ciprian Popoviciu
Cisco Systems
Kit Creek Road
RTP, North Carolina 27709
USA

Phone: 919 787 8162
Email: cpopovic@cisco.com

Ahmed Hamza
Cisco Systems
3000 Innovation Drive
Kanata K2K 3E8
Canada

Phone: 613 254 3656
Email: ahamza@cisco.com

Gunter Van de Velde
Cisco Systems
De Kleetlaan 6a
Diegem 1831
Belgium

Phone: +32 2704 5473
Email: gunter@cisco.com

Diego Dugatkin
IXIA
26601 West Agoura Rd
Calabasas 91302
USA

Phone: 818 444 3124
Email: diego@ixiacom.com

Full Copyright Statement

Intellectual Property

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).