

I-D Title(s): IPv6 Benchmarking Methodology

IETF minutes:

https://datatracker.ietf.org/public/proceeding_interim.cgi?meeting_num=68

Reviewer Name: Scott Bradner

Date: 4/9/07

Review Summary:

Overall:

- * Does/Do the draft(s) provide clear identification of the scope of work? E.g., is the class of device, system, or service being characterized clearly articulated.

yes

- * If a terminology memo, are the measurement areas clearly defined or otherwise cited? Is the working set of supporting terminology sufficient and correct? To your knowledge, are the areas of the memo that may conflict with other bodies of work? Are there any measurements or terminology that are superfluous? Are any missing?

n/a

- * If a methodology memo, does the methodology AND its corresponding terminology adequately define a benchmarking solution for its application area? Do the methodologies present sufficient detail for the experimental control of the benchmarks?

yes

- * If neither a terminology or methodology, does the offered memo offer complementary information important to the use or application of the related benchmarking solution?

n/a

- * Do you feel there are undocumented limitations or caveats to the benchmarking solution being proposed? If so, please describe.

no

- * Does the memo attempt to define acceptance criteria for any of the benchmark areas?

no

Technical Content: (Accuracy, Completeness of coverage)

Are definitions accurate? Is the terminology offered relevant?

n/a

To your knowledge, are there technical areas that are erroneous?
Are there questionable technical areas that need to be re-examined
or otherwise scrutinized.

no

Does the solution adequately address IPv6?

yes

Do you feel the memo(s) being offered are technically mature enough
for advancement to informational RFC?

yes

Clarity and Utility:

If you had a need, would you utilize the benchmarking solutions
advocated by this and its related memos? If not, why?

see suggestions at the end

Conformance to BMWG principles: (see BMWG charter)

<http://www.ietf.cnri.reston.va.us/html.charters/bmwg-charter.html>

Do you have confidence that the benchmarks, as explicitly
defined, will yield consistent results if repeated on the
same device (DUT/SUT), multiple times for a given test condition.
If not, cite benchmark(s) and issue(s).

yes

Do you have confidence that the benchmarks, if executed for a
given test condition, utilizing the documented methodology
on multiple test infrastructure (e.g., test equipment), would
yield correct and consistent results on the same DUT/SUT?
(Said differently, are the benchmark's methodology written
with enough exacting detail, that benchmark implementation
differences do not yield a difference in the measured quantities?)
If not, cite benchmark(s) and issue(s).

yes

Do you feel that the benchmarks form a basis of comparison between implementations of quantity being characterized? (I.e., are the benchmarks suitable for comparing solutions from different vendors.)

yes

If not, cite benchmarks and issues.

For those benchmarks cited above, do you feel that the benchmarks, as specified, have universal applicability for the given behavior being characterized? (i.e., benchmarks might not form a basis for cross-vendor comparison, can be used universally in a different role.)

yes

Editorial Comments:

(includes any deficiencies noted w.r.t. I-D Nits, spelling, & grammar)

section 4 - in "note" - change "The dynamic option is preferred if the test tool interacts with the DUT" to "The dynamic option is preferred wherein the test tool interacts with the DUT" - **DONE**

section 5.1 - change "Two types of media are commonly deployed and SHOULD be tested" to "Two types of media are commonly deployed and each SHOULD be tested if the network element supports that type of media:" - **DONE**

section 5.1.1 - append "if supported" to "The 4096, 8192, 9216 bytes long jumbo frame sizes SHOULD be used when benchmarking Gigabit Ethernet interfaces." - **DONE**

section 5.2.1 - add '[note to IANA, replace "xxxxx" with assigned prefix]' after "IANA reserved the IPv6 address block xxxxx/48 for use with IPv6 benchmark testing." - **DONE**

change "These addresses MUST not be assumed to be routable on the Internet" to "These addresses MUST be assumed to be not routable on the Internet" - **DONE**

last pp in section 5.2.1 - it would be good to say why - **Done**
Change last sentence to provide argument. New version: "Other prefix lengths can also be used if desired, however the indicated range reflects major prefix boundaries expected to be present in IPv6 routing tables and they should be representative to establish baseline performance metrics."

section 5.3 3rd pp - this is true as long as hop-by-hop header processing impacts more than an interface processor - I would expect that a future generation of v6 routers will implement this processing in firmware (since its not hard to do in general and because it protects the router against DoS attacks based on adding the header - **We agree with Scott that processing of Hop-by-Hop can be moved into dedicated hardware however the implementation is not simple and**

unlikely to be complete. There are multiple option available with this EH so it might become unfeasible to have dedicated HW that cover all the options.

6th pp - "bottom size"?? -Done

The new version: "a common value SHOULD be selected for the smallest frame size"

section 8 - I'd remove the sentence beginning "Most network infrastructures are allocated a /48" because this is no longer the standard assignment - just say that the requested size will meet the requirements for testing in large routers and large emulated networks - Done

New version: "The requested size meets the requirements for testing large network elements and large emulated networks."

Reviewer Name: Bill Cervený

Date: March 14, 2007

Please organize your comments in the following categories below.

Review Summary:

This document accurately captures the current state-of-the-art for performance issues that should be tested in the IPv6-enabled devices.

Overall:

- * Does/Do the draft(s) provide clear identification of the scope of work? E.g., is the class of device, system, or service being characterized clearly articulated.

Yes.

- * If a terminology memo, are the measurement areas clearly defined or otherwise cited? Is the working set of supporting terminology sufficient and correct? To your knowledge, are the areas of the memo that may conflict with other bodies of work? Are there any measurements or terminology that are superfluous? Are any missing?
- * If a methodology memo, does the methodology AND its corresponding terminology adequately define a benchmarking solution for its application area? Do the methodologies present sufficient detail for the experimental control of the benchmarks?

Yes, excluding the areas discussed in the editorial comments. In particular, some clarification in the area of filtering could be helpful. - Addressed, see comments below

- * If neither a terminology or methodology, does the offered memo offer complementary information important to the use or application of the related benchmarking solution?

* Do you feel there are undocumented limitations or caveats to the benchmarking solution being proposed? If so, please describe.

No

* Does the memo attempt to define acceptance criteria for any of the benchmark areas?

Technical Content: (Accuracy, Completeness of coverage)

Are definitions accurate? Is the terminology offered relevant?

Yes and yes

To your knowledge, are there technical areas that are erroneous? Are there questionable technical areas that need to be re-examined or otherwise scrutinized.

Not to my knowledge

Does the solution adequately address IPv6?

Yes

Do you feel the memo(s) being offered are technically mature enough for advancement to informational RFC?

Yes. This said, the amount of operational experience with IPv6 has been very limited, compared to that of IPv4. It is likely that the document will need to be updated to reflect new findings as more is understood with IPv6 operationally. On the other hand, this document will provide valuable guidance to early adopters. **- We agree with the comment, so far, the feedback received from those who used the recommendations of the document indicate that it is perceived to be complete.**

Clarity and Utility:

If you had a need, would you utilize the benchmarking solutions advocated by this and its related memos? If not, why?

Yes, they are consistent with what I would consider if I had a need to perform general IPv6 benchmark and performance testing.

Conformance to BMWG principles: (see BMWG charter)

<http://www.ietf.cnri.reston.va.us/html.charters/bmwg-charter.html>

Do you have confidence that the benchmarks, as explicitly defined, will yield consistent results if repeated on the same device (DUT/SUT), multiple times for a given test condition. If not, cite benchmark(s) and issue(s).

Yes

Do you have confidence that the benchmarks, if executed for a given test condition, utilizing the documented methodology on multiple test infrastructure (e.g., test equipment), would yield correct and consistent results on the same DUT/SUT? (Said differently, are the benchmark's methodology written with enough exacting detail, that benchmark implementation differences do not yield a difference in the measured quantities?) If not, cite benchmark(s) and issue(s).

Yes

Do you feel that the benchmarks form a basis of comparison between implementations of quantity being characterized? (I.e., are the benchmarks suitable for comparing solutions from different vendors.)

Yes

If not, cite benchmarks and issues.

For those benchmarks cited above, do you feel that the benchmarks, as specified, have universal applicability for the given behavior being characterized? (i.e., benchmarks might not form a basis for cross-vendor comparison, can be used universally in a different role.)

Yes

Editorial Comments:

(includes any deficiencies noted w.r.t. I-D Nits, spelling, & grammar)

Review of draft-ietf-bmwg-ipv6-meth-01.txt

Section 1.

s/are proving to be very useful/are proving to be useful/ - Done

s/procedures as described in RFC2544 and not to replace them/procedures described in RFC2544 and not replace them/ - Done

Section 4.

s/test traffic simulated end points/test traffic simulates end points/
- Done

Sentence beginning "The test scenarios assume ..." Consider rewording to something like: "To avoid neighbor solicitation (NS) and neighbor advertisement (NA) storms due to the neighbor unreachability detection (NUD) mechanism, the test scenarios assume the test traffic simulates end points and the IPv6 source and destination addresses are one hop beyond the DUT." - Done

Section 5

Change "Also, not all network elements support this type of addresses"

to something like "Also, not all network elements support addresses of this prefix length." - Done

s/Interface ID portion of the global/interface ID portion of global/ - Done

s/tests be conducted using the following lengths/tests be conducted using the following prefix lengths/ - Done

Consider changing sentence beginning "Other prefix lengths can also be used..." to "Other prefix lengths can be used. However, the indicated range should be sufficient to establish baseline performance metrics." - Done

Regarding prefix lengths and IANA recommendation -- To be consistent with recommendations in document, the IANA recommendations should be for a prefix /31 or shorter. If a shorter IANA prefix is obtained, either all references recommending prefixes shorter than the IANA block should be adjusted and/or there should be text indicating that some tests may use blocks other than the IANA block. - Done

This is a very good observation. Based on discussions within the WG it was agreed that a /48 will be sufficient for the benchmarking needs. Based on this observation, we removed the recommendation of including /32 prefixes in the routing tables during test.

There is no section 4.2.1 referencing IANA recommendations - Done, replaced with 5.2.1

Uncapitalize neighbor discovery, neighbor advertisement, neighbor solicitation, neighbor unreachability, hop-by-hop, source addresses, destination addresses, protocol, address plans, preamble, inter frame gap. - Done

s/various types of practical traffic such as: Fragmented/various types of practical traffic such as fragmented/ - Done

Consider a scenario where specific types of extension headers either are blocked or not used on the network of the customer paying for the test -- is it still considered necessary to test with these headers? - **This would be a valid scenario and the tester can make the decision to eliminate these tests based on user interest. However, these tests will be considered incomplete and could not be used for comparisons against other platforms. We will not make changes to the recommendations but leave it to users to decide their test strategy.**

s/Considering the fact that/Considering that/ - Done

s/containing this extension headers type/containing this extension header type/ - Done

s/extension headers processing capability which/extension header processing capability, which/ - Done

s/The tests with traffic containing each individual extension header MUST be complemented with tests that contain/The tests with traffic containing each individual extension header MUST be complemented with tests containing/ - Done

ESP not defined earlier, as far as I can tell. - Done
Inserted reference to RFC2406

s/The extension headers chain/The extension header chain/ - Done

s/real life extension headers chain/real life extension header chain/ - Done

s/the extension headers chain SHOULD/the extension header chain SHOULD/ - Done

s/For the most cases/For most cases/ - Done

s/it is most likely/it is likely/ - Done

Section 6

SA and DA defined twice. - **We did not find two definitions**

Perhaps RFC2544 explains this better, but shouldn't the addresses in the filter examples have prefix lengths, such as 2001:DB8::1/40?
- **This is a good observation. The examples however list host IPv6 addresses which imply /128 masks.**

In the text "The protocol field is defined as ...", shouldn't it be indicated that this is not intended to be an all-inclusive list of protocols?? - Done

Added: "The upper layer protocols listed above are recommended selection, however they do not represent an all-inclusive list of upper layer protocols which could be used in defining filters."

Section 7

s/Extension headers specific/Extension header specific/ - Done

s/For these reasons, this test is not recommended anymore for IPv6 benchmarking/For these reasons, this test is no longer recommended for IPv6 benchmarking/ - Done

Section 8

As mentioned earlier an IANA /48 wouldn't be sufficient, as the tests in this document are currently described. - Done, **see comment above**

Appendix A

s/4 bytes header/4-byte header/ - Done

s/a 2 or 4 bytes FCS field and a 1 byte/a 2- or 4-byte FCS field and a 1-byte/ - Done

Reviewer Name: Rajiv Asati

Date: March 20, 2007

Please organize your comments in the following categories below.

Review Summary:

This document provides an IPv6 specific update to RFC2544 which is very important in supporting the integration of IPv6. The document accurately covers all relevant aspects of benchmarking IPv6.

Overall:

- * Does/Do the draft(s) provide clear identification of the scope of work? E.g., is the class of device, system, or service being characterized clearly articulated.

Yes.

- * If a terminology memo, are the measurement areas clearly defined or otherwise cited? Is the working set of supporting terminology sufficient and correct? To your knowledge, are the areas of the memo that may conflict with other bodies of work? Are there any measurements or terminology that are superfluous? Are any missing?

Yes, the terminology is clearly defined.

The document does not conflict with any other ongoing work.

The document contains no superfluous sections and it is complete.

- * If a methodology memo, does the methodology AND its corresponding terminology adequately define a benchmarking solution for its application area? Do the methodologies present sufficient detail for the experimental control of the benchmarks?

Yes.

- * If neither a terminology or methodology, does the offered memo offer complementary information important to the use or application of the related benchmarking solution?

- * Do you feel there are undocumented limitations or caveats to the benchmarking solution being proposed? If so, please describe.

No

- * Does the memo attempt to define acceptance criteria for any of the benchmark areas?

Technical Content: (Accuracy, Completeness of coverage)

Are definitions accurate? Is the terminology offered relevant?

Yes, the definitions are accurate and the terminology is relevant.

To your knowledge, are there technical areas that are erroneous?
Are there questionable technical areas that need to be re-examined

or otherwise scrutinized.

Not to my knowledge

Does the solution adequately address IPv6?

Yes

Do you feel the memo(s) being offered are technically mature enough for advancement to informational RFC?

Yes. IPv6 is a mature technology and this document is addressing its architectural characteristics but operational experience, which is being developed, will provide very valuable feedback.

Clarity and Utility:

If you had a need, would you utilize the benchmarking solutions advocated by this and its related memos? If not, why?

Yes and I am aware of ongoing benchmarking efforts that use these recommendations.

Conformance to BMWG principles: (see BMWG charter)

<http://www.ietf.cnri.reston.va.us/html.charters/bmwg-charter.html>

Do you have confidence that the benchmarks, as explicitly defined, will yield consistent results if repeated on the same device (DUT/SUT), multiple times for a given test condition. If not, cite benchmark(s) and issue(s).

Yes

Do you have confidence that the benchmarks, if executed for a given test condition, utilizing the documented methodology on multiple test infrastructure (e.g., test equipment), would yield correct and consistent results on the same DUT/SUT? (Said differently, are the benchmark's methodology written with enough exacting detail, that benchmark implementation differences do not yield a difference in the measured quantities?) If not, cite benchmark(s) and issue(s).

Yes

Do you feel that the benchmarks form a basis of comparison between implementations of quantity being characterized? (I.e., are the benchmarks suitable for comparing solutions from different vendors.)

Yes

If not, cite benchmarks and issues.

For those benchmarks cited above, do you feel that the benchmarks, as specified, have universal applicability for the given behavior being characterized? (i.e., benchmarks might not form

a basis for cross-vendor comparison, can be used universally in a different role.)

Yes

Editorial Comments:

Other changes suggested:

*****David Newman**

- Appendices A.1 and A.2 refer to "maximum throughput." There is no such metric. RFC 1242 defines throughput as a single rate, not a range with minimum and maximum values. - Done

Removed maximum.

- Appendix A.1 erroneously rounds up some maximum rates.

For example, the maximum rate for 64-byte frames on gigabit Ethernet is given as 1,488,096 fps. In fact, the formula given in this same section yields the correct rate of 1,488,095.24 fps.

Even if a rate were x.74 fps, the number should either be rounded down if expressed in integer form or (better yet, in my opinion) presented as a floating-point number.

- Ethernet's maximum rates are subject to some tolerance due to clock slop. It would be useful to note that the rates given are theoretical maximums, and that actual rates may vary by +/- 100 ppm (with conversions for the appropriate pps for each flavor of Ethernet). - Done

Introduced "theoretical" in the Appendix title and in text. Also, introduced the following note: "Note: Ethernet's maximum frame rates are subject to variances due to clock slip. The listed rates are theoretical maximums and actual tests should account for a +/- 100 ppm tolerance."

*****Frame Recommendation Changes**

The frame size selection discussion concluded with Jim McQuaid's proposal to possibly have a document dedicated to the topic.

Maximum Frame size

- **Dan Romascanu**

"I agree about jumbo frames which are not standardized at this point in time. I would observe at the same time that RFC 2544 is out of synch by two generations in what Ethernet standard (non-jumbo) maximum frame length is concerned. The figure of 1518 mentioned all over RFC 2544 was replaced long time back by 1522 in order to accommodate the IEEE 802.1Q VLAN header, and 1522 was more recently replaced by 2000, as per IEEE 802.3as."

- **Scott Bradner**

"true - but I'm not sure how much of a performance difference it would make (other than time-on-the-wire) since the max payload size has not changed."

- **David Newman**

"There's no standard for jumbo frames. Some boxes support 9000-byte frames, while it's 9216 on others. And it only gets worse over Sonet/SDH. Different cards from one popular vendor have MTUs ranging from 4474 to 9192. While I'd agree that it's useful to test with jumbo frames, I don't believe it's possible yet to define a general case."

Changed from:

"The 4096, 8192, 9216 bytes long jumbo frame sizes SHOULD be used when benchmarking Gigabit Ethernet interfaces if these frame sizes are supported."

To

"Tests with jumbo frames SHOULD be executed. Frame sizes should be selected based on the values supported by the device under test due to the absence of a common standard defining the jumbo frame sizes. Examples of common jumbo frame sizes are 4096, 8192, 9216 bytes."

Minimum Frame size

- **Timmons C. Player**

"Why is 64 bytes used as the minimum frame size for SONET? I understand that 64 is the minimum for Ethernet, but the minimum for SONET is either 46 or 48 (40 byte packets), so it seems like that should be included in the recommended frame size list"

- **Ciprian Popoviciu**

"the 64 bytes frame for Ethernet leads to a 44 bytes IP packet which in the case of SONET would be a 51 bytes (with a 2 bytes FCS) or a 53 bytes (with a 4 bytes FCS) frame"

- **David Newman**

"Sonet supports 40-byte packets, the shortest length possible for IPv6. Ergo, that length should be included in any methodology, regardless of where a Sonet device ultimately gets deployed."

- **Curtis Villamizar**

"Average packet sizes in real networks have never been as low as 64 bytes so 64 bytes as a means to determine pps limitations is reasonable."

We recommended starting from 53 bytes for SONET.

*****RFC2544 back-to-back test removed**

- **Timmons C. Player**

"I'm glad to see the back-to-back test being deprecated. I've always thought that it was mostly useless for modern DUT's."

- **Curtis Villamizar**

"RFC2544 indicates that 1024 frames back to back would be called for for testing any protocol. Its a SHOULD, not a MUST. IMHO the figure of 1024 is too small and testing using bursts of small packets is also called for. The figure you'll get is not as important as passing the test without packet drops. The recommendation in the draft is based on people who do vendor benchmarking not liking non-repeatable figures as opposed to people who do ISP benchmarking not liking equipment which can't be relied upon to deliver busty traffic. I think your point is that the latter is or at least should be extinct.

ISPs (smart ones) will test for this whether its in the spec as a SHOULD or a MUST or a SHOULD NOT as it is here. You might be surprised

by some "modern DUTs" and smart ISPs prefer to be surprised in their test lab rather than on their production network. If they don't get surprised in the lab its still time well spent."

- **Scott Bradner**

"my experience (a long time ago with hardware from a long time ago) was that there was no useful repeatability in the back-to-back test (17% variance in one product between runs, in the same range in a number of other products) - do I think that the b2b test should be deprecated but I do agree with Curtis that some useful information can be found - if the max burst is very short that could be an indication of problems that could hurt but its more of a warning sign than anything quantitative and I agree that 1024 is too short a burst these days - not sure what it should be though"

We followed Scott's recommendation to remove the test on the account of repeatability concerns.

*****Throughput definition**

Scott Bradner clarified the history and reasoning behind the definition of throughput within BMWG:

> The term "throughput" is commonly referred to as an average rate that
> is achieved given a particular load. In an operating network, the
> throughput is the number of bits actually flowing, not the maximum
> achievable rate. For example, if a 10g interface has 1g of traffic on
> it, the throughput is 1g, not 10g. The maximum achievable throughput
> might be 10g (we can pretend it is, though for practical reasons it
> rarely is). It makes no sense for the benchmarking community to
> redefine a commonly used term that has somewhat different meaning to
> the rest of the community. I don't think that was ever the intention
> of RFC1242 but I can't speak for the author. Scott can clarify.

a lot of thought and discussion went into the definition of throughput the max zero-loss rate was carefully selected because it turns out to be quite meaningful

for example, about the same time that RFC 1242 was published a particular Wellfleet router was able to forward packets at a rate of over 85,000 pps (over six 10 Mbps Ethernet paths) but its zero-loss rate was less than 5,000 (as I recall) because there was a bug which caused packets to be lost when the real time clock updated the console screen every second - this was important information - it led to teh bug being found - in this case fixing that bug increased TCP throughput for paths going through the router.

bottom line - selecting the max zero-loss value as the "throughput" was exactly what the WG wanted to do.

the intention was that people using the RFC 1242 terminology should say so in their reports to avoid confusion over what they mean with their test results

Scott

- Done

No changes made to the draft

*****Minimum Frame Size - Test Tool Perspective**

We were contacted by several Test Tool vendors with the request to put a note about the fact that when selecting minimum size frames there might not be enough space left in the packets to insert signatures used to recognize packets for determining packet drops, out of order packets or for measuring delay and latency of traffic.

This is a reasonable warning since all test tools are currently using packet signatures. We decided to add the following note:

"Note: Test tools are commonly using signatures to identify test traffic packets to verify that there are no packet drops, out of order packets or to calculate various statistics such as delay and jitter. This could be the reason why the minimum frame size selectable through the test tool might not be as low as the theoretical one presented in this document."