Network Working Group D. Stopp
Request for Comments: 3918 Ixia
Category: Informational B. Hickman
Spirent Communications
October 2004

Methodology for IP Multicast Benchmarking

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Abstract

The purpose of this document is to describe methodology specific to the benchmarking of multicast IP forwarding devices. It builds upon the tenets set forth in RFC 2544, RFC 2432 and other IETF Benchmarking Methodology Working Group (BMWG) efforts. This document seeks to extend these efforts to the multicast paradigm.

The BMWG produces two major classes of documents: Benchmarking Terminology documents and Benchmarking Methodology documents. The Terminology documents present the benchmarks and other related terms. The Methodology documents define the procedures required to collect the benchmarks cited in the corresponding Terminology documents.

Table of Contents

1. Introduction

This document defines tests for measuring and reporting the throughput, forwarding, latency and Internet Group Management Protocol (IGMP) group membership characteristics of devices that support IP multicast protocols. The results of these tests will provide the user with meaningful data on multicast performance.

A previous document, "Terminology for IP Multicast Benchmarking" [Du98], defined many of the terms that are used in this document. The terminology document should be consulted before attempting to make use of this document.

This methodology will focus on one source to many destinations, although many of the tests described may be extended to use multiple source to multiple destination topologies.

Subsequent documents may address IPv6 multicast and related multicast routing protocol performance. Additional insight on IP and multicast networking can be found in [Hu95], [Ka98] and [Mt98].

2. Key Words to Reflect Requirements

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 set up

The set of methodologies presented in this document are for single ingress, multiple egress multicast scenarios as exemplified by Figures 1 and 2. Methodologies for multiple ingress and multiple egress multicast scenarios are beyond the scope of this document.

Figure 1 shows a typical setup for an IP multicast test, with one source to multiple destinations.

```
                     +------------+         +--------------+
                     |            |         |  destination |
   +--------+        |     Egress(-)------->|    test      |
   | source |        |            |         |   port(E1)   |
   |  test  |------>(|)Ingress    |         +--------------+
   |  port  |        |            |         +--------------+
   +--------+        |     Egress(-)------->|  destination |
                     |            |         |    test      |
                     |            |         |   port(E2)   |
                     |    DUT     |         +--------------+
                     |            |               . . .
                     |            |         +--------------+
                     |            |         |  destination |
                     |     Egress(-)------->|    test      |
                     |            |         |   port(En)   |
                     +------------+         +--------------+
```

Figure 1

If the multicast metrics are to be taken across multiple devices forming a System Under Test (SUT), then test frames are offered to a single ingress interface on a device of the SUT, subsequently forwarded across the SUT topology, and finally forwarded to the test apparatus' frame-receiving components by the test egress interface(s) of devices in the SUT. Figure 2 offers an example SUT test topology. If a SUT is tested, the test topology and all relevant configuration details MUST be disclosed with the corresponding test results.

```
               *-----------------------------------------*
               |                                         |
   +--------+  |                     +----------------+  |  +--------+
   |        |  |   +------------+    |DUT B Egress E0(-)-|->|        |
   |        |  |   |DUT A       |--->|                |  |  |        |
   | source |  |   |            |    |      Egress E1(-)-|->|  dest. |
   |  test  |--|->(-)Ingress, I |    +----------------+  |  |  test  |
   |  port  |  |   |            |    +----------------+  |  |  port  |
   |        |  |   |            |--->|DUT C Egress E2(-)-|->|        |
   |        |  |   +------------+    |                |  |  |        |
   |        |  |                     |      Egress En(-)-|->|        |
   +--------+  |                     +----------------+  |  +--------+
               |                                         |
               *------------------SUT--------------------*
```

Figure 2

Generally, the destination test ports first join the desired number of multicast groups by sending IGMP Group Report messages to the DUT/SUT. To verify that all destination test ports successfully joined the appropriate groups, the source test port MUST transmit IP multicast frames destined for these groups. After test completion, the destination test ports MAY send IGMP Leave Group messages to clear the IGMP table of the DUT/SUT.

In addition, test equipment MUST validate the correct and proper forwarding actions of the devices they test in order to ensure the receipt of the frames that are involved in the test.

3.1. Test Considerations

The methodology assumes a uniform medium topology. Issues regarding mixed transmission media, such as speed mismatch, headers differences, etc., are not specifically addressed. Flow control, QoS and other non-essential traffic or traffic-affecting mechanisms affecting the variable under test MUST be disabled. Modifications to the collection procedures might need to be made to accommodate the transmission media actually tested. These accommodations MUST be presented with the test results.

An actual flow of test traffic MAY be required to prime related mechanisms, (e.g., process RPF events, build device caches, etc.) to optimally forward subsequent traffic. Therefore, prior to running any tests that require forwarding of multicast or unicast packets, the test apparatus MUST generate test traffic utilizing the same addressing characteristics to the DUT/SUT that will subsequently be used to measure the DUT/SUT response. The test monitor should ensure the correct forwarding of traffic by the DUT/SUT. The priming action need only be repeated to keep the associated information current.

It is the intent of this memo to provide the methodology for basic characterizations regarding the forwarding of multicast packets by a device or simple system of devices. These characterizations may be useful in illustrating the impact of device architectural features (e.g., message passing versus shared memory; handling multicast traffic as an exception by the general purpose processor versus the by a primary data path, etc.) in the forwarding of multicast traffic.

It has been noted that the formation of the multicast distribution tree may be a significant component of multicast performance. While this component may be present in some of the measurements or scenarios presented in this memo, this memo does not seek to explicitly benchmark the formation of the multicast distribution tree. The benchmarking of the multicast distribution tree formation is left as future, more targeted work specific to a given tree formation vehicle.

3.1.1. IGMP Support

All of the ingress and egress interfaces MUST support a version of IGMP. The IGMP version on the ingress interface MUST be the same version of IGMP that is being tested on the egress interfaces.

Each of the ingress and egress interfaces SHOULD be able to respond to IGMP queries during the test.

Each of the ingress and egress interfaces SHOULD also send LEAVE (running IGMP version 2 or later) [Ca02] [Fe97] after each test.

3.1.2. Group Addresses

There is no restriction to the use of multicast addresses [De89] to compose the test traffic other than those assignments imposed by IANA. The IANA assignments for multicast addresses [IANA1] MUST be regarded for operational consistency. Address selection does not need to be restricted to Administratively Scoped IP Multicast addresses [Me98].

3.1.3. Frame Sizes

Each test SHOULD be run with different multicast frame sizes. For Ethernet, the recommended sizes are 64, 128, 256, 512, 1024, 1280, and 1518 byte frames.

Other link layer technologies MAY be used. The minimum and maximum frame lengths of the link layer technology in use SHOULD be tested.

When testing with different frame sizes, the DUT/SUT configuration MUST remain the same.

3.1.4. TTL

The data plane test traffic should have a TTL value large enough to traverse the DUT/SUT.

The TTL in IGMP control plane messages MUST be in compliance with the version of IGMP in use.

3.1.5. Trial Duration

The duration of the test portion of each trial SHOULD be at least 30 seconds. This parameter MUST be included as part of the results reporting for each methodology.

4. Forwarding and Throughput

This section contains the description of the tests that are related to the characterization of the frame forwarding of a DUT/SUT in a multicast environment. Some metrics extend the concept of throughput presented in RFC 1242. Forwarding Rate is cited in RFC 2285 [Ma98].

4.1. Mixed Class Throughput

Objective:

To determine the throughput of a DUT/SUT when both unicast class frames and multicast class frames are offered simultaneously to a fixed number of interfaces as defined in RFC 2432.

Procedure:

Multicast and unicast traffic are mixed together in the same aggregated traffic stream in order to simulate a heterogeneous networking environment.

The following events MUST occur before offering test traffic:

- o All destination test ports configured to receive multicast traffic MUST join all configured multicast groups;
- o The DUT/SUT MUST learn the appropriate unicast and multicast addresses; and

- o Group membership and unicast address learning MUST be verified through some externally observable method.

The intended load [Ma98] SHOULD be configured as alternating multicast class frames and unicast class frames to a single ingress interface. The unicast class frames MUST be configured to transmit in an unweighted round-robin fashion to all of the destination ports.

For example, with six multicast groups and 3 destination ports with one unicast addresses per port, the source test port will offer frames in the following order:

m1 u1 m2 u2 m3 u3 m4 u1 m5 u2 m6 u3 m1 ...

Where:

m<Number> = Multicast Frame<Group>
u<Number> = Unicast Frame<Target Port>

Mixed class throughput measurement is defined in RFC 2432 [Du98]. A search algorithm MUST be utilized to determine the Mixed Class Throughput. The ratio of unicast to multicast frames MUST remain the same when varying the intended load.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version
- o Total number of multicast groups
- o Traffic distribution for unicast and multicast traffic classes
- o The ratio of multicast to unicast class traffic

The following results MUST be reflected in the test report:

- o Mixed Class Throughput as defined in RFC 2432 [Du98], including: Throughput per unicast and multicast traffic classes.

The Mixed Class Throughput results for each test SHOULD be reported in the form of a table with a row for each of the tested frame sizes per the recommendations in section 3.1.3. Each row SHOULD specify the intended load, number of multicast frames offered, number of unicast frames offered and measured throughput per class.

4.2. Scaled Group Forwarding Matrix

Objective:

To determine Forwarding Rate as a function of tested multicast groups for a fixed number of tested DUT/SUT ports.

Procedure:

This is an iterative procedure. The destination test port(s) MUST join an initial number of multicast groups on the first iteration. All destination test ports configured to receive multicast traffic MUST join all configured multicast groups. The recommended number of groups to join on the first iteration is 10 groups. Multicast traffic is subsequently transmitted to all groups joined on this iteration and the forwarding rate is measured.

The number of multicast groups joined by each destination test port is then incremented, or scaled, by an additional number of multicast groups. The recommended granularity of additional groups to join per iteration is 10, although the tester MAY choose a finer granularity. Multicast traffic is subsequently transmitted to all groups joined during this iteration and the forwarding rate is measured.

The total number of multicast groups joined MUST not exceed the multicast group capacity of the DUT/SUT. The Group Capacity (Section 7.1) results MUST be known prior to running this test.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version

The following results MUST be reflected in the test report:

- o The total number of multicast groups joined for that iteration

- o Forwarding rate determined for that iteration

The Scaled Group Forwarding results for each test SHOULD be reported in the form of a table with a row representing each iteration of the test. Each row or iteration SHOULD specify the total number of groups joined for that iteration, offered load, total number of frames transmitted, total number of frames received and the aggregate forwarding rate determined for that iteration.

4.3. Aggregated Multicast Throughput

Objective:

To determine the maximum rate at which none of the offered frames to be forwarded through N destination interfaces of the same multicast groups are dropped.

Procedure:

Offer multicast traffic at an initial maximum offered load to a fixed set of interfaces with a fixed number of groups at a fixed frame length for a fixed duration of time. All destination test ports MUST join all specified multicast groups.

If any frame loss is detected, the offered load is decreased and the sender will transmit again. An iterative search algorithm MUST be utilized to determine the maximum offered frame rate with a zero frame loss.

Each iteration will involve varying the offered load of the multicast traffic, while keeping the set of interfaces, number of multicast groups, frame length and test duration fixed, until the maximum rate at which none of the offered frames are dropped is determined.

Parameters to be measured MUST include the maximum offered load at which no frame loss occurred. Other offered loads MAY be measured for diagnostic purposes.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version
- o Total number of multicast groups

The following results MUST be reflected in the test report:

- o Aggregated Multicast Throughput as defined in RFC 2432 [Du98]

The Aggregated Multicast Throughput results SHOULD be reported in the format of a table with a row for each of the tested frame sizes per the recommendations in section 3.1.3. Each row or iteration SHOULD specify offered load, total number of offered frames and the measured Aggregated Multicast Throughput.

4.4. Encapsulation/Decapsulation (Tunneling) Throughput

This sub-section provides the description of tests related to the determination of throughput measurements when a DUT/SUT or a set of DUTs are acting as tunnel endpoints.

For this specific testing scenario, encapsulation or tunneling refers to a packet that contains an unsupported protocol feature in a format that is supported by the DUT/SUT.

4.4.1. Encapsulation Throughput

Objective:

To determine the maximum rate at which frames offered to one ingress interface of a DUT/SUT are encapsulated and correctly forwarded on one or more egress interfaces of the DUT/SUT without loss.

Procedure:

```
  Source               DUT/SUT                 Destination
 Test Port                                     Test Port(s)
+---------+          +-----------+             +---------+
|         |          |           |             |         |
|         |          |     Egress|--(Tunnel)-->|         |
|         |          |           |             |         |
|         |--------->|Ingress    |             |         |
|         |          |           |             |         |
|         |          |     Egress|--(Tunnel)-->|         |
|         |          |           |             |         |
+---------+          +-----------+             +---------+
```

Figure 3

Figure 3 shows the setup for testing the encapsulation throughput of the DUT/SUT. One or more tunnels are created between each egress interface of the DUT/SUT and a destination test port. Non-Encapsulated multicast traffic will then be offered by the source test port, encapsulated by the DUT/SUT and forwarded to the destination test port(s).

The DUT/SUT SHOULD be configured such that the traffic across each egress interface will consist of either:

a) A single tunnel encapsulating one or more multicast address groups OR
b) Multiple tunnels, each encapsulating one or more multicast address groups.

The number of multicast groups per tunnel MUST be the same when the DUT/SUT is configured in a multiple tunnel configuration. In addition, it is RECOMMENDED to test with the same number of tunnels on each egress interface. All destination test ports MUST join all multicast group addresses offered by the source test port. Each egress interface MUST be configured with the same MTU.

Note: when offering large frames sizes, the encapsulation process may require the DUT/SUT to fragment the IP datagrams prior to being forwarded on the egress interface. It is RECOMMENDED to limit the offered frame size such that no fragmentation is required by the DUT/SUT.

A search algorithm MUST be utilized to determine the encapsulation throughput as defined in [Du98].

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- Number of tested egress interfaces on the DUT/SUT
- Test duration
- IGMP version
- Total number of multicast groups
- MTU size of DUT/SUT interfaces
- Originating un-encapsulated frame size
- Number of tunnels per egress interface
- Number of multicast groups per tunnel
- Encapsulation algorithm or format used to tunnel the packets

The following results MUST be reflected in the test report:

- o Measured Encapsulated Throughput as defined in RFC 2432 [Du98]
- o Encapsulated frame size

The Encapsulated Throughput results SHOULD be reported in the form of a table and specific to this test there SHOULD be rows for each originating un-encapsulated frame size. Each row or iteration SHOULD specify the offered load, encapsulation method, encapsulated frame size, total number of offered frames, and the encapsulation throughput.

4.4.2. Decapsulation Throughput

Objective:

To determine the maximum rate at which frames offered to one ingress interface of a DUT/SUT are decapsulated and correctly forwarded by the DUT/SUT on one or more egress interfaces without loss.

Procedure:

```
  Source                  DUT/SUT            Destination
 Test Port                                   Test Port(s)
+---------+             +-----------+        +---------+
|         |             |           |        |         |
|         |             |     Egress|------->|         |
|         |             |           |        |         |
|         |--(Tunnel)-->|Ingress    |        |         |
|         |             |           |        |         |
|         |             |     Egress|------->|         |
|         |             |           |        |         |
+---------+             +-----------+        +---------+
```

Figure 4

Figure 4 shows the setup for testing the decapsulation throughput of the DUT/SUT. One or more tunnels are created between the source test port and the DUT/SUT. Encapsulated multicast traffic will then be offered by the source test port, decapsulated by the DUT/SUT and forwarded to the destination test port(s).

The DUT/SUT SHOULD be configured such that the traffic across the ingress interface will consist of either:

a) A single tunnel encapsulating one or more multicast address groups OR
b) Multiple tunnels, each encapsulating one or more multicast address groups.

The number of multicast groups per tunnel MUST be the same when the DUT/SUT is configured in a multiple tunnel configuration. All destination test ports MUST join all multicast group addresses offered by the source test port. Each egress interface MUST be configured with the same MTU.

A search algorithm MUST be utilized to determine the decapsulation throughput as defined in [Du98].

When making performance comparisons between the encapsulation and decapsulation process of the DUT/SUT, the offered frame sizes SHOULD reflect the encapsulated frame sizes reported in the encapsulation test (See section 4.4.1) in place of those noted in section 3.1.3.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- Number of tested egress interfaces on the DUT/SUT
- Test duration
- IGMP version
- Total number of multicast groups
- Originating encapsulation algorithm or format used to tunnel the packets
- Originating encapsulated frame size
- Number of tunnels
- Number of multicast groups per tunnel

The following results MUST be reflected in the test report:

- Measured Decapsulated Throughput as defined in RFC 2432 [Du98]
- Decapsulated frame size

The Decapsulated Throughput results SHOULD be reported in the format of a table and specific to this test there SHOULD be rows for each originating encapsulated frame size. Each row or iteration SHOULD specify the offered load, decapsulated frame size, total number of offered frames and the decapsulation throughput.

4.4.3. Re-encapsulation Throughput

Objective:

To determine the maximum rate at which frames of one encapsulated format offered to one ingress interface of a DUT/SUT are converted to another encapsulated format and correctly forwarded by the DUT/SUT on one or more egress interfaces without loss.

Procedure:

```
  Source                     DUT/SUT              Destination
 Test Port                                        Test Port(s)
+---------+               +---------+             +---------+
|         |               |         |             |         |
|         |               |   Egress|-(Tunnel)->  |         |
|         |               |         |             |         |
|         |-(Tunnel)->    |Ingress  |             |         |
|         |               |         |             |         |
|         |               |   Egress|-(Tunnel)->  |         |
|         |               |         |             |         |
+---------+               +---------+             +---------+

                          Figure 5
```

Figure 5 shows the setup for testing the Re-encapsulation throughput of the DUT/SUT. The source test port will offer encapsulated traffic of one type to the DUT/SUT, which has been configured to re-encapsulate the offered frames using a different encapsulation format. The DUT/SUT will then forward the re-encapsulated frames to the destination test port(s).

The DUT/SUT SHOULD be configured such that the traffic across the ingress and each egress interface will consist of either:

a) A single tunnel encapsulating one or more multicast address groups OR
b) Multiple tunnels, each encapsulating one or more multicast address groups.

The number of multicast groups per tunnel MUST be the same when the DUT/SUT is configured in a multiple tunnel configuration. In addition, the DUT/SUT SHOULD be configured such that the number of tunnels on the ingress and each egress interface are the same. All destination test ports MUST join all multicast group addresses offered by the source test port. Each egress interface MUST be configured with the same MTU.

Note that when offering large frames sizes, the encapsulation process may require the DUT/SUT to fragment the IP datagrams prior to being forwarded on the egress interface. It is RECOMMENDED to limit the offered frame sizes, such that no fragmentation is required by the DUT/SUT.

A search algorithm MUST be utilized to determine the re-encapsulation throughput as defined in [Du98].

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version
- o Total number of multicast groups
- o MTU size of DUT/SUT interfaces
- o Originating encapsulation algorithm or format used to tunnel the packets
- o Re-encapsulation algorithm or format used to tunnel the packets
- o Originating encapsulated frame size
- o Number of tunnels per interface
- o Number of multicast groups per tunnel

The following results MUST be reflected in the test report:

- o Measured Re-encapsulated Throughput as defined in RFC 2432 [Du98]
- o Re-encapsulated frame size

The Re-encapsulated Throughput results SHOULD be reported in the format of a table and specific to this test there SHOULD be rows for each originating encapsulated frame size. Each row or iteration SHOULD specify the offered load, Re-encapsulated frame size, total number of offered frames, and the Re-encapsulated Throughput.

5. Forwarding Latency

This section presents methodologies relating to the characterization of the forwarding latency of a DUT/SUT in a multicast environment. It extends the concept of latency characterization presented in RFC 2544.

The offered load accompanying the latency-measured packet can affect the DUT/SUT packet buffering, which may subsequently impact measured packet latency. This SHOULD be a consideration when selecting the intended load for the described methodologies below.

RFC 1242 and RFC 2544 draw a distinction between device types: "store and forward" and "bit-forwarding." Each type impacts how latency is collected and subsequently presented. See the related RFCs for more information.

5.1. Multicast Latency

Objective:

To produce a set of multicast latency measurements from a single, multicast ingress interface of a DUT/SUT through multiple, egress multicast interfaces of that same DUT/SUT as provided for by the metric "Multicast Latency" in RFC 2432 [Du98].

The procedures below draw from the collection methodology for latency in RFC 2544 [Br96]. The methodology addresses two topological scenarios: one for a single device (DUT) characterization; a second scenario is presented or multiple device (SUT) characterization.

Procedure:

If the test trial is to characterize latency across a single Device Under Test (DUT), an example test topology might take the form of Figure 1 in section 3. That is, a single DUT with one ingress interface receiving the multicast test traffic from frame-transmitting component of the test apparatus and n egress interfaces on the same DUT forwarding the multicast test traffic back to the frame-receiving component of the test apparatus. Note that n reflects the number of TESTED egress interfaces on the DUT actually expected to forward the test traffic (as opposed to configured but untested, non-forwarding interfaces, for example).

If the multicast latencies are to be taken across multiple devices forming a System Under Test (SUT), an example test topology might take the form of Figure 2 in section 3.

The trial duration SHOULD be 120 seconds to be consistent with RFC 2544 [Br96]. The nature of the latency measurement, "store and forward" or "bit forwarding", MUST be associated with the related test trial(s) and disclosed in the results report.

A test traffic stream is presented to the DUT. It is RECOMMENDED to offer traffic at the measured aggregated multicast throughput rate (Section 4.3). At the mid-point of the trial's duration, the test apparatus MUST inject a uniquely identifiable ("tagged") frame into the test traffic frames being presented. This tagged frame will be the basis for the latency measurements. By "uniquely identifiable", it is meant that the test apparatus MUST be able to discern the "tagged" frame from the other frames comprising the test traffic set. A frame generation timestamp, Timestamp A, reflecting the completion of the transmission of the tagged frame by the test apparatus, MUST be determined.

The test apparatus will monitor frames from the DUT's tested egress interface(s) for the expected tagged frame(s) and MUST record the time of the successful detection of a tagged frame from a tested egress interface with a timestamp, Timestamp B. A set of Timestamp B values MUST be collected for all tested egress interfaces of the DUT/SUT. See RFC 1242 [Br91] for additional discussion regarding store and forward devices and bit forwarding devices.

A trial MUST be considered INVALID should any of the following conditions occur in the collection of the trial data:

- o Unexpected differences between Intended Load and Offered Load or unexpected differences between Offered Load and the resulting Forwarding Rate(s) on the DUT/SUT egress ports.
- o Forwarded test frames improperly formed or frame header fields improperly manipulated.
- o Failure to forward required tagged frame(s) on all expected egress interfaces.
- o Reception of tagged frames by the test apparatus more than 5 seconds after the cessation of test traffic by the source test port.

The set of latency measurements, M, composed from each latency measurement taken from every ingress/tested egress interface pairing MUST be determined from a valid test trial:

```
M = { (Timestamp B(E0) - Timestamp A),
      (Timestamp B(E1) - Timestamp A), ...
      (Timestamp B(En) - Timestamp A) }
```

where (E0 ... En) represents the range of all tested egress interfaces and Timestamp B represents a tagged frame detection event for a given DUT/SUT tested egress interface.

A more continuous profile MAY be built from a series of individual measurements.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version
- o Offered load
- o Total number of multicast groups

The following results MUST be reflected in the test report:

- o The set of all latencies with respective time units related to the tested ingress and each tested egress DUT/SUT interface.

The time units of the presented latency MUST be uniform and with sufficient precision for the medium or media being tested.

The results MAY be offered in a tabular format and should preserve the relationship of latency to ingress/egress interface for each multicast group to assist in trending across multiple trials.

5.2. Min/Max Multicast Latency

Objective:

To determine the difference between the maximum latency measurement and the minimum latency measurement from a collected set of latencies produced by the Multicast Latency benchmark.

Procedure:

Collect a set of multicast latency measurements over a single test duration, as prescribed in section 5.1. This will produce a set of multicast latencies, M, where M is composed of individual forwarding latencies between DUT frame ingress and DUT frame egress port pairs. E.g.:

M = {L(I,E1),L(I,E2), ..., L(I,En)}

where L is the latency between a tested ingress interface, I, of the DUT, and Ex a specific, tested multicast egress interface of the DUT. E1 through En are unique egress interfaces on the DUT.

From the collected multicast latency measurements in set M, identify MAX(M), where MAX is a function that yields the largest latency value from set M.

Identify MIN(M), when MIN is a function that yields the smallest latency value from set M.

The Max/Min value is determined from the following formula:

Result = MAX(M) - MIN(M)

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version
- o Offered load
- o Total number of multicast groups

The following results MUST be reflected in the test report:

- o The Max/Min value

The following results SHOULD be reflected in the test report:

- o The set of all latencies with respective time units related to the tested ingress and each tested egress DUT/SUT interface.

The time units of the presented latency MUST be uniform and with sufficient precision for the medium or media being tested.

The results MAY be offered in a tabular format and should preserve the relationship of latency to ingress/egress interface for each multicast group.

6. Overhead

This section presents methodology relating to the characterization of the overhead delays associated with explicit operations found in multicast environments.

6.1. Group Join Delay

Objective:

To determine the time duration it takes a DUT/SUT to start forwarding multicast frames from the time a successful IGMP group membership report has been issued to the DUT/SUT.

Procedure:

The Multicast Group Join Delay measurement may be influenced by the state of the Multicast Forwarding Database <MFDB> of the DUT/SUT. The states of the MFDB may be described as follows:

- o State 0, where the MFDB does not contain the specified multicast group address. In this state, the delay measurement includes the time the DUT/SUT requires to add the address to the MFDB and begin forwarding. Delay measured from State 0 provides information about how the DUT/SUT is able to add new addresses into MFDB.

- o State 1, where the MFDB does contain the specified multicast group address. In this state, the delay measurement includes the time the DUT/SUT requires to update the MFDB with the newly joined node<s> and begin forwarding to the new node<s> plus packet replication time. Delay measured from State 1 provides information about how well the DUT/SUT is able to update the MFDB for new nodes while transmitting packets to other nodes for the same IP multicast address. Examples include adding a new user to an event that is being promoted via multicast packets.

The methodology for the Multicast Group Join Delay measurement provides two alternate methods, based on the state of the MFDB, to measure the delay metric. The methods MAY be used independently or in conjunction to provide meaningful insight into the DUT/SUT ability to manage the MFDB.

Users MAY elect to use either method to determine the Multicast Group Join Delay; however the collection method MUST be specified as part of the reporting format.

In order to minimize the variation in delay calculations as well as minimize burden on the DUT/SUT, the test SHOULD be performed with one multicast group. In addition, all destination test ports MUST join the specified multicast group offered to the ingress interface of the DUT/SUT.

Method A:

Method A assumes that the Multicast Forwarding Database <MFDB> of the DUT/SUT does not contain or has not learned the specified multicast group address; specifically, the MFDB MUST be in State 0. In this scenario, the metric represents the time the DUT/SUT takes to add the multicast address to the MFDB and begin forwarding the multicast packet. Only one ingress and one egress MUST be used to determine this metric.

Prior to sending any IGMP Group Membership Reports used to calculate the Multicast Group Join Delay, it MUST be verified through externally observable means that the destination test port is not currently a member of the specified multicast group. In addition, it MUST be verified through externally observable means that the MFDB of the DUT/SUT does not contain the specified multicast address.

Method B:

Method B assumes that the MFDB of the DUT/SUT does contain the specified multicast group address; specifically, the MFDB MUST be in State 1. In this scenario, the metric represents the time the DUT/SUT takes to update the MFDB with the additional nodes and their corresponding interfaces and to begin forwarding the multicast packet. One or more egress ports MAY be used to determine this metric.

Prior to sending any IGMP Group Membership Reports used to calculate the Group Join Delay, it MUST be verified through externally observable means that the MFDB contains the specified multicast group address. A single un-instrumented test port MUST be used to join the specified multicast group address prior to sending any test traffic. This port will be used only for insuring that the MFDB has been populated with the specified multicast group address and can successfully forward traffic to the un-instrumented port.

Join Delay Calculation

Once verification is complete, multicast traffic for the specified multicast group address MUST be offered to the ingress interface prior to the DUT/SUT receiving any IGMP Group Membership Report messages. It is RECOMMENDED to offer traffic at the measured aggregated multicast throughput rate (Section 4.3).

After the multicast traffic has been started, the destination test port (See Figure 1) MUST send one IGMP Group Membership Report for the specified multicast group.

The join delay is the difference in time from when the IGMP Group Membership message is sent (timestamp A) and the first frame of the multicast group is forwarded to a receiving egress interface (timestamp B).

Group Join delay time = timestamp B - timestamp A

Timestamp A MUST be the time the last bit of the IGMP group membership report is sent from the destination test port; timestamp B MUST be the time the first bit of the first valid multicast frame is forwarded on the egress interface of the DUT/SUT.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- Frame size(s)
- Number of tested egress interfaces on the DUT/SUT
- IGMP version
- Total number of multicast groups
- Offered load to ingress interface
- Method used to measure the join delay metric

The following results MUST be reflected in the test report:

- The group join delay time in microseconds per egress interface(s)

The Group Join Delay results for each test MAY be reported in the form of a table, with a row for each of the tested frame sizes per the recommendations in section 3.1.3. Each row or iteration MAY specify the group join delay time per egress interface for that iteration.

6.2. Group Leave Delay

Objective:

To determine the time duration it takes a DUT/SUT to cease forwarding multicast frames after a corresponding IGMP Leave Group message has been successfully offered to the DUT/SUT.

Procedure:

In order to minimize the variation in delay calculations as well as minimize burden on the DUT/SUT, the test SHOULD be performed with one multicast group. In addition, all destination test ports MUST join the specified multicast group offered to the ingress interface of the DUT/SUT.

Prior to sending any IGMP Leave Group messages used to calculate the group leave delay, it MUST be verified through externally observable means that the destination test ports are currently members of the specified multicast group. If any of the egress interfaces do not forward validation multicast frames then the test is invalid.

Once verification is complete, multicast traffic for the specified multicast group address MUST be offered to the ingress interface prior to receipt or processing of any IGMP Leave Group messages. It is RECOMMENDED to offer traffic at the measured aggregated multicast throughput rate (Section 4.3).

After the multicast traffic has been started, each destination test port (See Figure 1) MUST send one IGMP Leave Group message for the specified multicast group.

The leave delay is the difference in time from when the IGMP Leave Group message is sent (timestamp A) and the last frame of the multicast group is forwarded to a receiving egress interface (timestamp B).

Group Leave delay time = timestamp B - timestamp A

Timestamp A MUST be the time the last bit of the IGMP Leave Group message is sent from the destination test port; timestamp B MUST be the time the last bit of the last valid multicast frame is forwarded on the egress interface of the DUT/SUT.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o IGMP version
- o Total number of multicast groups
- o Offered load to ingress interface

The following results MUST be reflected in the test report:

- o The group leave delay time in microseconds per egress interface(s)

The Group Leave Delay results for each test MAY be reported in the form of a table, with a row for each of the tested frame sizes per the recommendations in section 3.1.3. Each row or iteration MAY specify the group leave delay time per egress interface for that iteration.

7. Capacity

This section offers a procedure relating to the identification of multicast group limits of a DUT/SUT.

7.1. Multicast Group Capacity

Objective:

To determine the maximum number of multicast groups a DUT/SUT can support while maintaining the ability to forward multicast frames to all multicast groups registered to that DUT/SUT.

Procedure:

One or more destination test ports of DUT/SUT will join an initial number of multicast groups.

After a minimum delay as measured by section 6.1, the source test ports MUST transmit to each group at a specified offered load.

If at least one frame for each multicast group is forwarded properly by the DUT/SUT on each participating egress interface, the iteration is said to pass at the current capacity.

For each successful iteration, each destination test port will join an additional user-defined number of multicast groups and the test repeats. The test stops iterating when one or more of the egress interfaces fails to forward traffic on one or more of the configured multicast groups.

Once the iteration fails, the last successful iteration is the stated Maximum Group Capacity result.

Reporting Format:

The following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o IGMP version
- o Offered load

The following results MUST be reflected in the test report:

- o The total number of multicast group addresses that were successfully forwarded through the DUT/SUT

The Multicast Group Capacity results for each test SHOULD be reported in the form of a table, with a row for each of the tested frame sizes per the recommendations in section 3.1.3. Each row or iteration SHOULD specify the number of multicast groups joined per destination interface, number of frames transmitted and number of frames received for that iteration.

8. Interaction

Network forwarding devices are generally required to provide more functionality than just the forwarding of traffic. Moreover, network-forwarding devices may be asked to provide those functions in a variety of environments. This section offers procedures to assist in the characterization of DUT/SUT behavior in consideration of potentially interacting factors.

8.1. Forwarding Burdened Multicast Latency

Objective:

To produce a set of multicast latency measurements from a single multicast ingress interface of a DUT/SUT through multiple egress multicast interfaces of that same DUT/SUT as provided for by the metric "Multicast Latency" in RFC 2432 [Du98] while forwarding meshed unicast traffic.

Procedure:

The Multicast Latency metrics can be influenced by forcing the DUT/SUT to perform extra processing of packets while multicast class traffic is being forwarded for latency measurements.

The Burdened Forwarding Multicast Latency test MUST follow the described setup for the Multicast Latency test in Section 5.1. In addition, another set of test ports MUST be used to burden the DUT/SUT (burdening ports). The burdening ports will be used to transmit unicast class traffic to the DUT/SUT in a fully meshed traffic distribution as described in RFC 2285 [Ma98]. The DUT/SUT MUST learn the appropriate unicast addresses and verified through some externally observable method.

Perform a baseline measurement of Multicast Latency as described in Section 5.1. After the baseline measurement is obtained, start transmitting the unicast class traffic at a user-specified offered load on the set of burdening ports and rerun the Multicast Latency test. The offered load to the ingress port MUST be the same as was used in the baseline measurement.

Reporting Format:

Similar to Section 5.1, the following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o Test duration
- o IGMP version
- o Offered load to ingress interface
- o Total number of multicast groups
- o Offered load to burdening ports
- o Total number of burdening ports

The following results MUST be reflected in the test report:

- o The set of all latencies related to the tested ingress and each tested egress DUT/SUT interface for both the baseline and burdened response.

The time units of the presented latency MUST be uniform and with sufficient precision for the medium or media being tested.

The latency results for each test SHOULD be reported in the form of a table, with a row for each of the tested frame sizes per the recommended frame sizes in section 3.1.3, and SHOULD preserve the relationship of latency to ingress/egress interface(s) to assist in trending across multiple trials.

8.2. Forwarding Burdened Group Join Delay

Objective:

To determine the time duration it takes a DUT/SUT to start forwarding multicast frames from the time a successful IGMP Group Membership Report has been issued to the DUT/SUT while forwarding meshed unicast traffic.

Procedure:

The Forwarding Burdened Group Join Delay test MUST follow the described setup for the Group Join Delay test in Section 6.1. In addition, another set of test ports MUST be used to burden the DUT/SUT (burdening ports). The burdening ports will be used to transmit unicast class traffic to the DUT/SUT in a fully meshed traffic pattern as described in RFC 2285 [Ma98]. The DUT/SUT MUST learn the appropriate unicast addresses and verified through some externally observable method.

Perform a baseline measurement of Group Join Delay as described in Section 6.1. After the baseline measurement is obtained, start transmitting the unicast class traffic at a user-specified offered load on the set of burdening ports and rerun the Group Join Delay test. The offered load to the ingress port MUST be the same as was used in the baseline measurement.

Reporting Format:

Similar to Section 6.1, the following configuration parameters MUST be reflected in the test report:

- o Frame size(s)
- o Number of tested egress interfaces on the DUT/SUT
- o IGMP version
- o Offered load to ingress interface
- o Total number of multicast groups
- o Offered load to burdening ports
- o Total number of burdening ports
- o Method used to measure the join delay metric

The following results MUST be reflected in the test report:

- o The group join delay time in microseconds per egress interface(s) for both the baseline and burdened response.

The Group Join Delay results for each test MAY be reported in the form of a table, with a row for each of the tested frame sizes per the recommendations in section 3.1.3. Each row or iteration MAY specify the group join delay time per egress interface, number of frames transmitted and number of frames received for that iteration.

9. Security Considerations

As this document is solely for the purpose of providing metric methodology and describes neither a protocol nor a protocol's implementation, there are no security considerations associated with this document specifically. Results from these methodologies may identify a performance capability or limit of a device or system in a particular test context. However, such results might not be representative of the tested entity in an operational network.

10. Acknowledgements

The Benchmarking Methodology Working Group of the IETF and particularly Kevin Dubray, Juniper Networks, are to be thanked for the many suggestions they collectively made to help complete this document.

11. Contributions

The authors would like to acknowledge the following individuals for their help and participation of the compilation of this document: Hardev Soor, Ixia, and Ralph Daniels, Spirent Communications, both who made significant contributions to the earlier versions of this document. In addition, the authors would like to acknowledge the members of the task team who helped bring this document to fruition: Michele Bustos, Tony De La Rosa, David Newman and Jerry Perser.

12. References

12.1. Normative References

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

[Br96] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999.

[Br97] Bradner, S. "Use of Keywords in RFCs to Reflect Requirement Levels, RFC 2119, March 1997.

[Du98] Dubray, K., "Terminology for IP Multicast Benchmarking", RFC 2432, October 1998.

[IANA1] IANA multicast address assignments, http://www.iana.org/assignments/multicast-addresses

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

[Me98] Meyer, D., "Administratively Scoped IP Multicast", BCP 23, RFC 2365, July 1998.

12.2. Informative References

[Ca02] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

[De89] Deering, S., "Host Extensions for IP Multicasting", STD 5, RFC 1112, August 1989.

[Fe97] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.

[Hu95] Huitema, C., "Routing in the Internet", Prentice-Hall, 1995.

[Ka98] Kosiur, D., "IP Multicasting: the Complete Guide to Interactive Corporate Networks", John Wiley & Sons Inc., 1998.

[Mt98] Maufer, T., "Deploying IP Multicast in the Enterprise", Prentice-Hall, 1998.

13. Authors' Addresses

Debra Stopp
Ixia
26601 W. Agoura Rd.
Calabasas, CA 91302
USA

Phone: + 1 818 871 1800
EMail: debby@ixiacom.com

Brooks Hickman
Spirent Communications
26750 Agoura Rd.
Calabasas, CA 91302
USA

Phone: + 1 818 676 2412
EMail: brooks.hickman@spirentcom.com

14. Full Copyright Statement

Intellectual Property

Acknowledgement

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