Network Working Group
INTERNET-DRAFT
Expires in: January 2001

Hardev Soor
Debra Stopp
IXIA

Ralph Daniels
Netcom Systems
July 2000

Methodology for IP Multicast Benchmarking
<draft-ietf-bmwg-mcastm-04.txt>

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet- Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Abstract

The purpose of this draft 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.

1 Introduction

This document defines a specific set of tests that vendors can use to measure and report the performance characteristics and forwarding capabilities of network devices that support IP multicast protocols. The results of these tests will provide the user comparable data from different vendors with which to evaluate these devices.

A previous document, " Terminology for IP Multicast Benchmarking" (RFC 2432), 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 IP multicast communication.

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 RFC 2119.

3 Test set up

Figure 1 shows a typical setup for an IP multicast test, with one source to multiple destinations, although this MAY be extended to multiple source to multiple destinations.

```
                                                 +----------------+
                          +------------+         |                |
  +--------+              |            |--------->| destination(1) |
  |        |              |            |         |                |
  | source |-------->|            |         +----------------+
  |        |              |            |         +----------------+
  +--------+              |   D U T    |--------->|                |
                          |            |         | destination(2) |
                          |            |         |                |
                          |            |         +----------------+
                          |            |              . . .
                          |            |         +----------------+
                          |            |         |                |
                          |            |--------->| destination(n) |
```

```
                             |            |           |                |
                             |            |           +----------------+
                             |            |
                             +------------+
```

Figure 1

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

In addition, all transmitted frames MUST contain a recognizable pattern that can be filtered on in order to ensure the receipt of only the frames that are involved in the test.

3.1   Test Considerations

3.2   IGMP Support

Each of the destination ports should support and be able to test all IGMP versions 1, 2 and 3. The minimum requirement, however, is IGMP version 2.

Each destination port should be able to respond to IGMP queries during the test.

Each destination port should also send LEAVE (running IGMP version 2) after each test.

3.3   Group Addresses

The Class D Group address SHOULD be changed between tests. Many DUTs have memory or cache that is not cleared properly and can bias the results.

The following group addresses are recommended by use in a test:

```
224.0.1.27-224.0.1.255
224.0.5.128-224.0.5.255
224.0.6.128-224.0.6.255
```

If the number of group addresses accommodated by these ranges do not satisfy the requirements of the test, then these ranges may be overlapped. The total number of configured group addresses must be less than or equal to the IGMP table size of the DUT/SUT.

3.4   Frame Sizes

Each test SHOULD be run with different Multicast Frame Sizes. The recommended frame sizes are 64, 128, 256, 512, 1024, 1280, and 1518 byte frames.

3.5 TTL

The source frames should have a TTL value large enough to accommodate the DUT/SUT.

3.6 Layer 2 Support

Each of the destination ports should support GARP/GMRP protocols to join groups on Layer 2 DUTs/SUTs.

4 Forwarding and Throughput

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

4.1 Mixed Class Throughput

Objective

To determine the maximum throughput rate at which none of the offered frames, comprised from a unicast Class and a multicast Class, to be forwarded are dropped by the device across a fixed number of ports as defined in RFC 2432.

Procedure

Multicast and unicast traffic are mixed together in the same aggregated traffic stream in order to simulate the non-homogenous networking environment. While the multicast traffic is transmitted from one source to multiple destinations, the unicast traffic MAY be evenly distributed across the DUT/SUT architecture. In addition, the DUT/SUT SHOULD learn the appropriate unicast IP addresses, either by sending ARP frames from each unicast address, sending a RIP packet or by assigning static entries into the DUT/SUT address table.

The mixture of multicast and unicast traffic MUST be set up in one of two ways:

a) As a percent of the total traffic flow resulting in a ratio. For example, 20 percent multicast traffic to 80 percent unicast traffic.

b) In evenly distributed bursts of multicast and unicast traffic, resulting in a 50-50 ratio of multicast to unicast traffic.

The transmission of the frames MUST be set up so that they form a deterministic distribution while still maintaining the specified forwarding rates. See Appendix A for a discussion on non-homogenous vs. homogenous packet distribution.

Similar to the Frame loss rate test in RFC 2544, the first trial SHOULD be run for the frame rate that corresponds to 100% of the maximum rate for the frame size on the input media. Repeat the procedure for the rate that corresponds to 90% of the maximum rate used and then for 80% of this rate. This sequence SHOULD be continued (at reducing 10% intervals) until there are two successive trials in which no frames are lost. The maximum granularity of the trials MUST be 10% of the maximum rate, a finer granularity is encouraged.

Result

Parameters to be measured SHOULD include the frame loss and percent loss for each class of traffic per destination port. The ratio of unicast traffic to multicast traffic MUST be reported.

In addition, the transmit and receive rates in frames per second for each source and destination port for both unicast and multicast traffic, together with the number of frames transmitted and received per port per class type traffic SHOULD be reported.

4.2   Scaled Group Forwarding Matrix

Definition

A table that demonstrates Forwarding Rate as a function of tested multicast groups for a fixed number of tested DUT/SUT ports.

Procedure

Multicast traffic is sent at a fixed percent of maximum offered load with a fixed number of receive ports of the tester at a fixed frame length.

The receive ports SHOULD continue joining incrementally by 10 multicast groups until a user defined maximum is reached.

The receive ports will continue joining in the incremental fashion until a user defined maximum is reached.

Results

Parameters to be measured SHOULD include the frame loss and percent loss per destination port for each multicast group address.

In addition, the transmit and receive rates in frames per second for each source and destination port for all multicast groups, together with the number of frames transmitted and received per port per multicast groups SHOULD be reported.

4.3 Aggregated Multicast Throughput

Definition

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

Procedure

Multicast traffic is sent at a fixed percent of maximum offered load with a fixed number of groups at a fixed frame length for a fixed duration of time.

The initial number of receive ports of the tester will join the group(s) and the sender will transmit to the same groups after a certain delay (a few seconds).

Then the an incremental number of receive ports will join the same groups and then the Multicast traffic is sent as stated.

The receive ports will continue to be added and multicast traffic sent until a user defined maximum number of ports is reached.

Results

Parameters to be measured SHOULD include the frame loss and percent loss per destination port for each multicast group address.

In addition, the transmit and receive rates in frames per second for each source and destination port for all multicast groups, together with the number of frames transmitted and received per port per multicast groups SHOULD be reported.

4.4   Encapsulation (Tunneling) Throughput

This sub-section provides the description of tests that help in obtaining throughput measurements when a DUT/SUT or a set of DUTs are acting as tunnel endpoints. The following Figure 2 presents the scenario for the tests.

```
   Client A      DUT/SUT A      Network      DUT/SUT B      Client B


              ----------                   ----------
              |        |      ------       |        |
 -----(a)  (b)|        |(c)  (      )   (d)|        |(e) (f)-----
 ||||| -----> |        |---->(      )----->|        |-----> |||||
 -----        |        |      ------       |        |       -----
              |        |                   |        |
              ----------                   ----------
```

Figure 2

A tunnel is created between DUT/SUT A (the encapsulator) and DUT/SUT B (the decapsulator). Client A is acting as a source and Client B is the destination. Client B joins a multicast group (for example, 224.0.1.1) and it sends an IGMP Join message to DUT/SUT B to join that group. Client A now wants to transmit some traffic to Client B. It will send the multicast traffic to DUT/SUT A which encapsulates the multicast frames, sends it to DUT/SUT B which will decapsulate the same frames and forward them to Client B.

4.4.1 Encapsulation Throughput

Definition

The maximum rate at which frames offered a DUT/SUT are encapsulated and correctly forwarded by the DUT/SUT without loss.

Procedure

To test the forwarding rate of the DUT/SUT when it has to go through the process of encapsulation, a test port B is injected at the other end of DUT/SUT A (Figure B) that will receive the encapsulated frames and measure the throughput. Also, a test port A is used to generate multicast frames that will be passed through the tunnel.

The following is the test setup:

```
Test port A     DUT/SUT A              Test port B

               ---------- (c')      (d')---------
               |        |--------------->|       |
-------(a)  (b)|        |                |       |
||||||| -----> |        |      ------   ---------
-------        |        |(c)  ( N/W  )
               |        |---->(      )
               ----------      ------
```

Figure 3

In Figure 2, a tunnel is created with the local IP address of DUT/SUT A as the beginning of the tunnel (point c) and the IP address of DUT/SUT B as the end of the tunnel (point d). DUT/SUT B is assumed to have the tunneling protocol enabled so that the frames can be decapsulated. When the test port B is inserted in between the DUT/SUT A and DUT/SUT B (Figure 3), the endpoint of tunnel has to be re-configured to be directed to the test port B's IP address. For example, in Figure 3, point c' would be assigned as the beginning of the tunnel and point d' as the end of the tunnel. The test port B is acting as the end of the tunnel, and it does not have to support any tunneling protocol since the frames do not have to be decapsulated. Instead, the received encapsulated frames are used to calculate the throughput and other necessary measurements.

Result

Parameters to be measured SHOULD include the frame loss and percent loss per destination port for each multicast group address.

In addition, the transmit and receive rates in frames per second for each source and destination port for all multicast groups, together with the number of frames transmitted and received per port per multicast groups SHOULD be reported.

4.4.2   Decapsulation Throughput

Definition

The maximum rate at which frames offered a DUT/SUT are decapsulated and correctly forwarded by the DUT/SUT without loss.

Procedure

The decapsulation process returns the tunneled unicast frames back to their multicast format. This test measures the throughput of the DUT/SUT when it has to perform the process of decapsulation, therefore, a test port C is used at the end of the tunnel to receive the decapsulated frames (Figure 4).

```
Test port A  DUT/SUT A    Test port B     DUT/SUT B   Test port C

             ----------                ----------
             |        |                |        |
-----(a)  (b)|        |(c)   ----    (d)|        |(e) (f)-----
||||| -----> |        |----> |||| ----->|        |-----> |||||
-----        |        |      ----       |        |       -----
             |        |                |        |
             ----------                ----------
```

Figure 4

In Figure 4, the encapsulation process takes place in DUT/SUT A. This may effect the throughput of the DUT/SUT B. Therefore, two test ports should be used to separate the encapsulation and decapsulation processes. Client A is replaced with the test port A which will generate a multicast frame that will be encapsulated by DUT/SUT A. Another test port B is inserted between DUT/SUT A and DUT/SUT B that will receive the encapsulated frames and forward it to DUT/SUT B. Test port C will receive the decapsulated frames and measure the throughput.

Result

Parameters to be measured SHOULD include the frame loss and percent loss per destination port for each multicast group address.

In addition, the transmit and receive rates in frames per second for each source and destination port for all multicast groups, together with the number of frames transmitted and received per port per multicast groups SHOULD be reported.

4.4.3 Re-encapsulation Throughput

Definition

The maximum rate at which frames of one encapsulated format offered a DUT/SUT are converted to another encapsulated format and correctly forwarded by the DUT/SUT without loss.

Procedure

Re-encapsulation takes place in DUT/SUT B after test port C has received the decapsulated frames. These decapsulated frames will be re-inserted with a new encapsulation frame and sent to test port B which will measure the throughput. See Figure 5.

```
      Test port A   DUT/SUT A    Test port B   DUT/SUT B  Test port
C

                 ----------                ----------
                 |        |                |        |
   -----(a)  (b)|        |(c)   ----    (d)|        |(e) (f)-----
   ||||| -----> |        |----> |||| <---->|        |<----> |||||
   -----        |        |      ----       |        |       -----
                 |        |                |        |
                 ----------                ----------
```

Figure 5

Result

Parameters to be measured SHOULD include the frame loss and percent loss per destination port for each multicast group address.

In addition, the transmit and receive rates in frames per second for each source and destination port for all multicast groups, together with the number of frames transmitted and received per port per multicast groups SHOULD be reported.

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.

5.1 Multicast Latency

Definition

The set of individual latencies from a single input port on the DUT/SUT or SUT to all tested ports belonging to the destination multicast group.

Procedure

According to RFC 2544, a tagged frame is sent half way through the transmission that contains a timestamp used for calculation of latency. In the multicast situation, a tagged frame is sent to all destinations for each multicast group and latency calculated on a per multicast group basis. Note that this test MUST be run using the transmission rate that is less than the multicast throughput of the DUT/SUT. Also, the test should take into account the DUT's/SUT's need to cache the traffic in its IP cache, fastpath cache or shortcut tables since the initial part of the traffic will be utilized to build these tables.

Result

The parameter to be measured is the latency value for each multicast group address per destination port. An aggregate latency MAY also be reported.

5.2 Min/Max/Average Multicast Latency

Definition

The difference between the maximum latency measurement and the minimum latency measurement from the set of latencies produced by the Multicast Latency benchmark.

Procedure

First determine the throughput for DUT/SUT at each of the listed frame sizes determined by the forwarding and throughput tests of section 4. Send a stream of frames to a fixed number of multicast groups through the DUT at the determined throughput rate. An identifying tag SHOULD be included in all frames to ensure proper identification of the transmitted frame on the receive side, the type of tag being implementation dependent.

Latencies for each transmitted frame are calculated based on the description of latencies in RFC 2544. The average latency is the total of all accumulated latency values divided by the total number of those values. The minimum latency is the smallest latency; the maximum latency is the largest latency of all accumulated latency values.

Results

The parameters to be measured are the minium, maximum and average latency values for each multicast group address per destination port.

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

Definition

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

Procedure

Traffic is sent on the source port at the same time as the IGMP JOIN Group message is transmitted from the destination ports. The join delay is the difference in time from when the IGMP Join is sent (timestamp A) and the first frame is forwarded to a receiving member port (timestamp B).

Group Join delay = timestamp B - timestamp A

One of the keys is to transmit at the fastest rate the DUT/SUT can handle multicast frames. This is to get the best resolution and the least margin of error in the Join Delay.

However, you do not want to transmit the frames so fast that frames are dropped by the DUT/SUT. Traffic should be sent at the throughput rate determined by the forwarding tests of section 4.

Results

The parameter to be measured is the join delay time for each multicast group address per destination port. In addition, the number of frames transmitted and received and percent loss may be reported.

6.2   Group Leave Delay

Definition

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

Procedure

Traffic is sent on the source port at the same time as the IGMP Leave Group messages are transmitted from the destination ports. The leave delay is the difference in time from when the IGMP leave is sent (timestamp A) and the last frame is forwarded to a receiving member port (timestamp B).

Group Leave delay = timestamp B - timestamp A

One of the keys is to transmit at the fastest rate the DUT/SUT can handle multicast frames. This is to get the best resolution and least margin of error in the Leave Delay. However, you do not want to transmit the frames too fast that frames are dropped by the DUT/SUT. Traffic should be sent at the throughput rate determined by the forwarding tests of section 4.

Result

The parameter to be measured is the leave delay time for each multicast group address per destination port. In addition, the number of frames transmitted and received and percent loss may be reported.

7 Capacity

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

7.1   Multicast Group Capacity

Definition

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 ports of DUT/SUT will join an initial number of groups.

Then after a delay (enough time for all ports to join) the source port will transmit to each group at a transmission rate that the DUT/SUT can handle without dropping IP Multicast frames.

If all frames sent are forwarded by the DUT/SUT and received the test iteration is said to pass at the current capacity.

If the iteration passes at the capacity the test will add an user defined incremental value of groups to each receive port.

The iteration is to run again at the new group level and capacity tested as stated above.

Once the test fails at a capacity the capacity is stated to be the last Iteration that pass at a giving capacity.

Results

The parameter to be measured is the total number of group addresses that were successfully forwarded with no loss.

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 terms to assist in the characterization of DUT/SUT behavior in consideration of potentially interacting factors.

8.1   Forwarding Burdened Multicast Latency

The Multicast Latency metrics can be influenced by forcing the DUT/SUT to perform extra processing of packets while multicast traffic is being forwarded for latency measurements. In this test, a set of ports on the tester will be designated to be source and destination similar to the generic IP Multicast test setup. In addition to this setup, another set of ports will be selected to transmit some multicast traffic that is destined to multicast group addresses that have not been joined by these additional set of ports.

For example, if ports 1,2, 3, and 4 form the burdened response setup (setup A) which is used to obtain the latency metrics and ports 5, 6, 7, and 8 form the non-burdened response setup (setup B) which will afflict the burdened response setup, then setup B traffic will join multicast group addresses not joined by the ports in this setup. By sending such multicast traffic, the DUT/SUT will perform a lookup on the packets that will affect the processing of setup A traffic.

8.2   Forwarding Burdened Group Join Delay

The port configuration in this test is similar to the one described in section 8.1, but in this test, the multicast traffic is not sent by the ports in setup B. In this test, the setup A traffic must be influenced in such a way that will affect the DUT's/SUT's ability to process Group Join messages. Therefore, in this test, the ports in setup B will send a set of IGMP Group Join messages while the ports in setup A are also joining its own set of group addresses. Since the two sets of group addresses are independent of each other, the group join delay for setup A may be different than in the case when there were no other group addresses being joined.

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.

10 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.

[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.

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

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

[Se98] Semeria, C. and Maufer, T. "Introduction to IP Multicast Routing." http://www.3com.com/nsc/501303.html 3Com Corp., 1998.

11 Author's Addresses

Hardev Soor
IXIA
26601 W. Agoura Rd.
Calabasas, CA 91302
USA

Phone: 818 871 1800
EMail: hardev@ixiacom.com

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

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

Ralph Daniels
Netcom Systems
948 Loop Road
Clayton, NC 27520
USA

Phone: 919 550 9475
EMail: Ralph_Daniels@NetcomSystems.com

Appendix A: Determining an even distribution

It is important to understand and fully define the distribution of frames among all multicast and unicast destinations. If the distribution is not well defined or understood, the throughput and forwarding metrics are not meaningful.

In a homogeneous environment, a large single burst of multicast frames may be followed by a large burst of unicast frames. This is a very different distribution than that of a non-homogeneous environment, where the multicast and unicast frames are intermingled throughout the entire transmission.

The recommended distribution is that of the non-homogeneous environment because it more closely represents a real-world scenario. The distribution is modeled by calculating the number of multicast frames per destination port as a burst, then calculating the number of unicast frames to transmit as a percentage of the total frames transmitted. The overall effect of the distribution is small bursts of multicast frames intermingled with small bursts of unicast frames.

12
Full Copyright Statement