

## **Benchmarking Terminology for LAN Switching Devices**

< [draft-ietf-bmwg-lanswitch-01.txt](#) >

### Status of this Document

This document is an Internet-Draft. 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.''

To learn the current status of any Internet-Draft, please check the ``1id-abstracts.txt'' listing contained in the Internet-Drafts Shadow Directories on ds.internic.net (US East Coast), nic.nordu.net (Europe), ftp.isi.edu (US West Coast), or munnari.oz.au (Pacific Rim).

Distribution of this document is unlimited. Please send comments to [bmwg@harvard.edu](mailto:bmwg@harvard.edu) or to the editor.

### Abstract

The purpose of this draft is to define and discuss benchmarking terminology for local area switching devices. It is meant to extend the terminology already defined for network interconnect devices in RFCs 1242 and 1944 by the Benchmarking Methodology Working Group (BMWG) of the Internet Engineering Task Force (IETF).

LAN switches are one of the principal sources of new bandwidth in the local area and are handling a significantly increasing proportion of network traffic. The multiplicity of products brought to market makes it desirable to define a set of terms to be used when evaluating the performance characteristics of local area switching devices. Well-defined terminology will help in providing the user community with complete, reliable and comparable data on LAN switches.

### **1. Introduction**

The purpose of this draft is to discuss and define terminology for the benchmarking of LAN switching devices. This draft covers local area devices

which switch frames at the Media Access Control (MAC) layer. It discusses throughput, latency, address handling and filtering.

## **2. Term definitions**

A previous document, "Benchmarking Terminology for Network Interconnect Devices" ([RFC 1242](#)), 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. A more recent document, "Benchmarking Methodology for Network Interconnect Devices" ([RFC 1944](#)), defined a number of test procedures which are directly applicable to switches. Since it discusses a number of terms relevant to benchmarking switches it should also be consulted. A number of new terms applicable to benchmarking switches are defined below using the format for definitions set out in [Section 2 of RFC 1242](#). RFCs 1242 and 1944 already contain discussions of some of these terms.

### **2. 1. Reminder of [RFC 1242](#) definition format**

Term to be defined. (e.g., Latency)

Definition:

The specific definition for the term.

Discussion:

A brief discussion about the term, it's application and any restrictions on measurement procedures.

Measurement units:

The units used to report measurements of this term, if applicable.

Issues:

List of issues or conditions that effect this term.

See Also:

List of other terms that are relevant to the discussion of this term.

### **2.2. Unidirectional traffic**

Definition:

Unidirectional traffic is made up of a single or multiple streams of frames forwarded in one direction only from one or more ports of a switching device designated as input ports to one or more other ports of the device designated as output ports.

Discussion:

This definition conforms to the discussion in [section 16 of RFC 1944](#) on multi-port testing which describes how unidirectional traffic can be offered to ports of a device to measure maximum rate of throughput.

Unidirectional traffic SHOULD be offered to devices for:

- the measurement of the minimum inter-frame gap
- the creation of many-to-one or one-to-many port overload
- the detection of head of line blocking
- the measurement of throughput when congestion control mechanisms are active

Unidirectional streams of traffic can be used to create different patterns of traffic. For example unidirectional streams can be offered to two input ports so as to overload a single output port (2-to-1) or they can be offered to a single input port and switched by the device under test to two or more output ports (1-to-2). Such patterns can be combined to test for head of line blocking or to measure throughput when congestion control mechanisms are active.

When devices are equipped with ports running at different media rates the number of input streams required to load or overload an output port or ports will vary.

Issues:

half duplex / full duplex

Measurement units:

n/a

See Also:

bidirectional traffic (2.3)

multidirectional traffic (2.4)

### **2.3. Bidirectional traffic**

Definition:

Bidirectional traffic is made up of two or more streams of frames forwarded in opposite directions between at two or more ports of a switching device.

Discussion:

This definition conforms to the discussions in sections [14](#) and [16](#) of RFC [1944](#) on **bidirectional traffic and multi-port testing**.

Bidirectional traffic MUST be offered when measuring the maximum rate of throughput on full duplex ports of a switching device.

Issues:

truncated binary exponential back-off algorithm

Measurement units:

n/a

See Also:

unidirectional traffic (2.2)

multidirectional traffic (2.4)

### **2.4. Multidirectional traffic**

Definition:

Multidirectional traffic is made up of streams of frames that are switched

simultaneously between multiple ports of a switching device. When such streams are fully meshed each of the ports under test will both send frames to and receive frames from all of the other ports under test.

Discussion:

This definition extends the discussions in sections [14](#) and [16](#) of [RFC 1944](#) on bidirectional traffic and multi-port testing.

As with bidirectional multi-port tests, multidirectional traffic exercises both the transmission and reception sides of the ports of a switching device. Since ports are not divided into two groups every port forwards frames to and receives frames from every other port. The total number of individual unidirectional streams offered in a multidirectional test for  $n$  switched ports equals  $n \times (n - 1)$ . This compares with  $n \times (n / 2)$  such streams in a bidirectional multi-port test. It should be noted however that bidirectional multiport tests create a greater load than multidirectional tests on backbone connections linking together two switching devices because none of the transmitted frames are forwarded locally. Backbone tests SHOULD use bidirectional multi-port traffic.

Multidirectional traffic on half duplex ports is inherently bursty since ports must interrupt transmission intermittently to receive frames. When offering such bursty traffic to a device under test a number of variables have to be considered. They include frame size, the number of frames within bursts as well as the interval between bursts. The terms burst, burst size and inter-burst gap are defined in sections [2.5](#), [2.6](#) and [2.7](#) below.

Bursty multidirectional traffic is characteristic of real network traffic. It simultaneously exercises many of the component parts of a switching device such as input and output buffers, buffer allocation mechanisms, aggregate switching capacity, processing speed and behavior of the media access controller.

Measurement units:

n/a

Issues:

half duplex / full duplex

See Also:

unidirectional traffic (2.2)

bidirectional traffic (2.3)

## [2.5](#) Burst

Definition:

A group of frames transmitted with the minimum inter-frame gap allowed by the media. This definition allows for single frame bursts and infinite bursts.

Discussion:

This definition follows from the discussion in [section 21 of RFC 1944](#). It is useful to consider isolated frames as single frame bursts.

Measurement units:

n/a

Issues:

See Also:

burst size (2.6)

## **2.6 Burst size**

Definition:

The number of frames in a burst.

Discussion:

Burst size can range from one to infinity. In unidirectional streams there is no theoretical limit to the burst length. Bursts in bidirectional and multidirectional streams of traffic on half duplex media are finite since ports interrupt transmission intermittently to receive frames.

On real networks burst size can increase with window size. This makes it desirable to test devices with small as well as large burst sizes.

Measurement units:

number of N-octet frames

Issues:

See Also: burst (2.5)

## **2.7 Inter-burst gap (IBG)**

Definition:

The interval between two bursts.

Discussion:

This definition conforms to the discussion in [section 20 of RFC 1944](#) on bursty traffic.

Bidirectional and multidirectional streams of traffic are inherently bursty since ports share their time between receiving and transmitting frames. The rate of transmission of an external source of traffic is a function of the number of frames per burst, frame length and the inter-burst gap. External sources offering bursty multidirectional traffic for a given frame size and burst size MUST adjust the inter-burst gap to achieve a specified rate of transmission.

When a burst contains a single frame inter-burst gap and inter-frame gap are equal.

When a burst is infinite the interburst gap equals the minimum inter-frame gap.

Measurement units:

nanoseconds

microseconds

milliseconds

seconds

Issues:

See Also: burst size (2.6), load (2.8)

## **2.8 Load, nominal and real**

Definition:

The amount of traffic per second that a port transmits and receives.

Discussion:

Load can be expressed in a number of ways: bits per second, frames per second with the frame size specified or as a percentage of the maximum frame rate allowed by the media for a given frame size. A unidirectional stream of **7440 64-byte Ethernet frames per second is equivalent to a 50% load** given that the maximum rate of transmission on an Ethernet is 14880 64-byte frames per second.

In the case of bidirectional or multidirectional traffic port load is the sum of the frames transmitted and received on a port per second.

There is room for varying the balance between incoming and outgoing traffic when loading ports with bidirectional and multidirectional traffic. In the case of bidirectional traffic a 100% load can be created by offering a n% load on one port and a (100 - n)% load on the opposite port.

Multidirectional traffic will be equally distributed over all ports under test when all ports are offered 50% of the target load.

It has to kept in mind that an external source may not deliver frames to a device under test at the desired rate due to collisions on CSMA/CD links or the action of congestion control mechanisms. Because of this it is often necessary to distinguish between the desired or target load (nominal load) and the actual load (real load) offered to the device under test.

External sources of Ethernet traffic MUST implement the truncated binary exponential back-off algorithm when executing bidirectional and multidirectional performance tests to ensure that the external source of traffic is accessing the medium legally.

Frames which are not successfully transmitted by the external source of traffic to the device under test MUST NOT be not counted as transmitted frames in performance benchmarks.

Measurement units:

bits per second

N-octets per second

$(N\text{-octets per second} / \text{media\_maximum-octets per second}) \times 100$

Issues:

token ring

## **2.9 Overload**

Definition:

Loading a port or ports in excess of the maximum rate of transmission allowed by the media.

Discussion:

Overloading can serve to exercise input and/or output buffers, buffer allocation algorithms and congestion control mechanisms. Unidirectional overloads require a minimum of two input and one output ports when all ports run at the same nominal speed.

Bidirectional and multidirectional overloading occurs when the sum of the traffic transmitted and received on each port exceeds the maximum media rate. The external source of traffic MUST transmit at the same rate situated between more than 50% and a 100% of the maximum media rate to each of the ports under test in order to equally distribute an overload over all ports under test.

Measurement units:

N-octet frames per second

Issues: nominal/real load

See Also:

## **2.10 Speed**

Definition:

The number of frames that a device is capable of delivering to the correct destination port in a given time interval. The maximum speed of a switching device is the highest number of frames it can deliver during a one second interval to the correct destination port.

Discussion:

Switching devices which exhibit no frame loss may be found to deliver frames to their proper destination ports at differing rates. This may be due to the action of congestion control mechanisms at high loads or the relative aggressiveness of the truncated binary back-off algorithm. Speed MUST only be sampled on the output side of the ports under test. This is because an input port may receive frames at higher rates when the device under test drops frames.

Measurement units:

N-octet frames per second

Issues:

See Also:

## **2.11 Flooded frame**

Definition:

A unicast frame which is received on ports which do not correspond to the frame's destination MAC address information.

Discussion:

When recording throughput statistics it is important to check that frames have been forwarded to their proper destinations. Flooded frames MUST NOT be

counted as received frames.

Measurement units:

N-octet valid frames per second

Issues:

Spanning tree BPDUs.

See Also:

### **2.11 Backpressure**

Definition:

A jamming technique used by some switching devices to avoid frame loss when one or more of its ports are saturated.

Discussion:

Some switches are designed to send jamming signals back to traffic sources when ports begin to saturate. Such devices may incur no frame loss when ports are offered target loads in excess of 100% by external traffic sources. Jamming however affects traffic destined to congested as well as uncongested ports so it is important to measure the maximum speed at which a device can forward frames to both congested and uncongested ports when backpressure mechanisms are active.

Measurement units:

N--octet frames per second between the jamming port and an uncongested destination port

Issues:

not explicitly described in standards

See Also:

forward pressure (2.12)

### **2.12 Forward pressure**

Definition:

A technique which modifies the truncated binary exponential backoff algorithm to avoid frame loss when congestion on one or more of the ports under test occurs.

Discussion:

Some switches avoid buffer overload by retransmitting buffered frames without waiting for the interval calculated by the normal operation of the backoff algorithm. It is useful to measure how aggressive a switch's backoff algorithm is in both congested and uncongested states. Forward pressure reduces the number of collisions when congestion on a port builds up.

Measurement units:

intervals in microseconds between transmission retries during 16 successive



collisions.

Issues:

not explicitly described in standards

See also:

backpressure (2.11)

### **[2.13](#) Head of line blocking**

Definition:

A pathological state whereby a switch drops frames forwarded to an uncongested port whenever frames are forwarded from the same source port to a congested port.

Discussion:

It is important to verify that a switch does not propagate frame loss to ports which are not congested whenever overloading on one of its ports occurs.

Measurement units:

frame loss recorded on an uncongested port when receiving frames from a port which is also forwarding frames to a congested destination port.

Issues:

Input buffers

See Also:

### **[2.14](#) Address handling**

Definition:

The number of MAC addresses per port, per module or per device which a switch can cache and successfully forward frames to without flooding or dropping frames.

Discussion:

Users building networks will want to know how many nodes they can connect to a switch. This makes it necessary to verify the number of MAC addresses that can be assigned per port, per module and per chassis before a switch begins flooding frames.

Measurement units:

number of MAC addresses

Issues:

See Also:

### **[2.15](#) Address learning rate**

Definition:

The maximum rate at which a switch can learn MAC addresses before starting to flood or drop frames.

**Discussion:**

Users may want to know how long it takes a switch to build up its address tables. This information is useful to have when considering how long it takes a network to come up when many users log on in the morning or after a network crash.

**Measurement units:**

frames per second with each successive frame sent to the switch containing a different source address.

**Issues:**

See Also: address handling (2.14)

## **2.16 Filtering illegal frames**

**Definition:**

Switches do not necessarily filter all types of illegal frames. Some switches, for example, do not store frames before forwarding them to their destination ports. These so-called cut-through switches forward frames after reading the destination and source address fields. They do not normally filter over-sized frames (jabbers) or verify the validity of the Frame Check Sequence field. Other examples of illegal frame types are under-sized frames (runts), misaligned frames and frames followed by dribble bits.

**Measurement units:**

N-octet frames filtered or not filtered

**Issues:**

See Also:

## **2.17 Broadcast rate**

**Definition:**

The number of broadcast frames forwarded by the device under test per second. The maximum broadcast rate corresponds to highest number of broadcast frames a switch can forward either locally or over a backbone connection.

**Discussion:**

There is no standard forwarding mechanism used by switches to forward broadcast frames. It is useful to determine the broadcast forwarding rate both locally and over backbone connections.

**Measurement units:**

N-octet frames per second

**Issues:**

See Also:  
broadcast latency

## **[2.18 Broadcast latency](#)**

Definition:

The time it takes a broadcast frame to go through a switching device and be forwarded to each destination port.

Discussion:

Since there is no standard way for switches to process broadcast frames, broadcast latency may not be the same on all receiving ports of a switching device. Broadcast latency SHOULD be determined on all receiving ports both locally and, if applicable, over backbone connections.

Measurement units:

The latency measurements SHOULD be bit oriented as described in 3.8 of RFC [1242](#) and reported for all connected receive ports.

Issues:

See Also:  
broadcast rate

## **[3. Acknowledgments](#)**

In order of appearance Ajay Shah of Wandel & Goltermann, Jean-Christophe Bestaux of European Network Laboratories, Stan Kopek of Digital Equipment Corporation, Henry Hamon of Netcom Systems and Kevin Dubray of Bay Networks were all instrumental in getting this draft done.

A special thanks goes to the IETF BenchMark WorkGroup for the many suggestions it collectively made to help shape this draft.

The editor

Bob Mandeville

## **[4. Editor's Address](#)**

Robert Mandeville

ENL (European Network Laboratories)

email: bob.mandeville@eunet.fr

35, rue Beaubourg

[75003](#) Paris

France

phone: +33 6 07 47 67 10

fax: + 33 1 42 78 36 71

!!!PLEASE TAKE NOTE!!!

ENL HAS MOVED TO A NEW SITE:

Robert Mandeville, ENL  
European Network Laboratories

**6 Parc Ariane, le Mercure**

Blvd des Chenes

**78284 Guyancourt**

France

NEW LAB PHONE, FAX, VOICE MAIL: +33 1 39 44 12 05

mobile phone: +33 6 07 47 67 10