

Benchmarking Terminology for LAN Switching Devices
< [draft-ietf-bmwg-lanswitch-02.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 ``ltd-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 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) and prepare the way for a discussion on benchmarking methodology for local area switches.

LAN switches are one of the principal sources of new bandwidth in the local area. 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 local area network switches. Although it might be found

useful to apply some of the terms defined here to a broader range of network interconnect devices, this draft primarily deals with devices which switch frames at the Medium Access Control (MAC) layer. It defines terms in relation to 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:

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 throughput.

Unidirectional traffic is also appropriate 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 traffic can be used to load the ports of a switching device in different ways. For example unidirectional traffic can be sent to two or more input ports from an external source and switched by the device under test to a single output port (n-to-1) or such traffic can be sent to a single input port and switched by the device under test to two or more output ports (1-to-n). 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.

The measurement of the minimum inter-frame gap serves to detect violations of the IEEE 802.3 standard.

Issues:

half duplex / full duplex

Measurement units:

n/a

See Also:

bidirectional traffic (2.3)

meshed traffic (2.4)

2.3. Bidirectional traffic

Definition:

Two or more streams of frames forwarded in opposite directions between at least 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 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)

meshed traffic (2.4)

[2.4.](#) Meshed traffic

Definition:

Multiple streams of frames switched simultaneously between all of a designated number of ports of a switching device such that 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 follows from the discussions in sections [14](#) and [16](#) of RFC [1944](#) on **bidirectional traffic and multi-port testing** and readily extends to configurations with multiple switching devices linked together over backbone connections.

As with bidirectional multi-port traffic, meshed 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 streams when traffic is meshed over 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 that bidirectional multiport traffic can load backbone connections linking together two switching devices more than meshed traffic.

Meshed traffic on half duplex ports is inherently bursty since ports must interrupt transmission whenever they receive frames. Bursty meshed traffic which is characteristic of real network traffic 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 medium access controller.

When offering bursty meshed traffic to a device under test a number of variables have to be considered. These 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.

Measurement units:

n/a

Issues:

half duplex / full duplex

See Also:

unidirectional traffic (2.2)

bidirectional traffic (2.3)

burst (2.5)

burst size (2.6)

inter-burst gap (2.7)

[2.5](#) Burst

Definition:

A sequence of frames transmitted with the minimum inter-frame gap allowed by

the medium.

Discussion:

This definition follows from discussions in [section 3.16 of RFC 1242](#) and [section 21 of RFC 1944](#) which describes cases where 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 burst length. When traffic is bidirectional or meshed bursts on half duplex media are finite since ports interrupt transmission intermittently to receive frames.

On real networks burst size will normally 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 meshed streams of traffic are inherently bursty since ports share their time between receiving and transmitting frames. External sources offering bursty traffic for a given frame size and burst size must adjust the inter-burst gap to achieve a specified rate of transmission.

Measurement units:

nanoseconds

microseconds

milliseconds
seconds

Issues:

See Also:
burst size (2.6)

[2.8](#) Port load

Definition:

The number of frames per second that a switched port transmits and receives.

Discussion:

Port 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 medium for a given frame size. In the case of bidirectional or meshed traffic port load is the sum of the frames transmitted and received on a port per second. The load on an Ethernet port which is transmitting and receiving a total of 7440 64-byte frames per second equals 50% given that the maximum rate of transmission on an Ethernet is 14880 64-byte frames per second.

There is room for varying the balance between incoming and outgoing traffic when loading ports with bidirectional and meshed traffic. When offering meshed traffic to a device the equal distribution of load over all ports will help avoid unwanted or inadvertent port overloading in throughput tests.

Measurement units:

bits per second
frames per second with the frame size specified
as a percentage of the maximum frame rate allowed by the medium for a given frame size.

Issues:

See Also:
bidirectional traffic (2.3)
meshed traffic (2.4)
overload (2.9)

[2.9](#) Overload

Definition:

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

Discussion:

Overloading can serve to exercise input and output buffers, buffer allocation algorithms and congestion control mechanisms. Port overloading with unidirectional traffic requires a minimum of two input and one output ports when the medium rate of all ports is the same. The number of input ports will vary according to the media rates of the output

port or ports under test.

Port overloading with bidirectional and meshed traffic requires the sum of the traffic transmitted and received on each port to exceed the maximum rate of transmission allowed by the medium. To distribute port overload equally, the external source of traffic must transmit at the same rate situated between more than 50% and a 100% of the maximum medium rate to each of the ports under test.

Measurement units:

N-octet frames per second

Issues:

See Also:

bidirectional traffic (2.3)

meshed traffic (2.4)

port load (2.8)

2.10 Intended rate

Definition:

The number of frames per second that an external source attempts to send to a port of a device under test.

Discussion:

An external source may not transmit frames to a device under test at the intended rate due to collisions on CSMA/CD links or the action of congestion control mechanisms. This makes it useful to distinguish between intended rate and the rate at which the source can be observed to send frames to a device under test.

An external source should have sufficient internal resources to transmit frames at the intended rate and in the case of Ethernet must implement the truncated binary exponential back-off algorithm when executing bidirectional and meshed performance tests to ensure that it 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 counted as transmitted frames in performance benchmarks.

Measurement units:

bits per second

N-octets per second

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

Issues:

token ring

See also:

offered rate (2.11)

2.11 Offered rate

Definition:

The number of frames per second that an external source can be observed to send to a port of a device under test.

Discussion:

Offered rate may differ from intended rate due to collisions on half duplex media or congestion control mechanisms.

The frame count on a port of a device under test may exceed the rate at which an external device offers frames due to the presence of spanning tree BPDUs (Bridge Protocol Data Units) on 802.1D-compliant switches or SNMP frames. If such frames cannot be inhibited, they MUST be left out of frame counts in performance benchmarks.

Measurement units:

bits per second

N-octets per second

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

Issues:

token ring

See also:

intended rate (2.10)

[2.12](#) Maximum load

Definition:

The load which results on a port when traffic is transmitted or addressed to it at the maximum rate allowed by the medium.

Discussion:

Maximum port load may be less than the maximum rate allowed by the medium when the offered rate of the external sources sending traffic to the device or system under test is less than the intended rate.

Measurement units:

bits per second

frames per second with the frame size specified

as a percentage of the maximum frame rate allowed by the medium for a given frame size.

Issues:

See Also:

bidirectional traffic (2.3)

meshed traffic (2.4)

port load (2.8)

intended rate (2.10)

offered rate (2.11)

forwarding rate (2.13)

forwarding rate at maximum load (2.14)

[2.13](#) Forwarding rate

Definition:

The number of frames per second that a device is observed to deliver to the correct output port in response to a known intended rate.

Discussion:

Forwarding rate does not take frame loss into account and must only be sampled on the output side of the ports under test. It can be measured on devices offered unidirectional, bidirectional or meshed traffic. The forwarding rates of switching devices which exhibit no frame loss may be reduced through the action of congestion control mechanisms.

Measurement units:

N-octet frames per second

Issues:

See Also:

port load (2.8)
intended rate (2.10)
offered rate (2.11)
forwarding rate at maximum load (2.14)

[2.14](#) Forwarding rate at maximum load

Definition:

The number of frames per second that a device is observed to successfully deliver to the correct output port at maximum load.

Discussion:

Forwarding rate at maximum load may be less than the maximum rate at which a device might be observed to successfully forward traffic.

Measurement units:

N-octet frames per second

Issues:

See Also:

maximum load (2.12)
forwarding rate (2.13)

[2.15](#) Flooding

Definition:

Frames received on ports which do not correspond to the 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. Both known and unknown unicast frames can be

flooded.

Measurement units:

N-octet valid frames per second

Issues:

Spanning tree BPDUs.

See Also:

2.16 Backpressure

Definition:

Techniques whereby switching devices avoid frame loss by impeding external sources of traffic from transmitting frames to congested ports.

Discussion:

Some switches send jam signals, for example preamble bits, back to traffic sources when their transmit and/or receive buffers start to overfill. Switches implementing full duplex Ethernet links may use IEEE 802.3x Flow Control for the same purpose. Such devices may incur no frame loss when external sources attempt to offer traffic to congested or overloaded ports. Jamming and flow control normally slow all traffic transmitted to congested input ports including traffic intended for uncongested output ports.

Measurement units:

frame loss on congested port or ports

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

Issues:

jamming not explicitly described in standards

See Also:

forward pressure (2.17)

2.17 Forward pressure

Definition:

An illegal technique whereby a device retransmits buffered frames without waiting for the interval calculated by the normal operation of the back-off algorithm.

Discussion:

Some switches illegally inhibit or abort the truncated binary exponential backoff algorithm and force access to the medium to avoid frame loss.

The backoff algorithm should be fair whether the device under test is in a congested or an uncongested state.

Measurement units:

intervals in microseconds between transmission retries during 16 successive collisions.

Issues:

truncated binary exponential backoff algorithm

See also:

backpressure (2.16)

2.18 Head of line blocking

Definition:

Frame loss observed on an uncongested output port whenever frames are received from an input port which is also attempting to forward frames to a congested output port.

Discussion:

It is important to verify that a switch does not slow transmission or drop frames on ports which are not congested whenever overloading on one of its other 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 port.

Issues:

input buffers

See Also:

unidirectional traffic (2.2)

2.19 Address handling

Definition:

The number of MAC addresses per n ports, 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 n ports, per module and per chassis before a switch begins flooding frames.

Measurement units:

number of MAC addresses

Issues:

See Also:

Address learning rate (2.20)

2.20 Address learning rate

Definition:

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

Discussion:

Users may want to know how long it takes a switch to build 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 (2.19)

2.21 Illegal frames

Definition:

Frames which are over-sized, under-sized, misaligned or with an errored Frame Check Sequence.

Discussion:

Switches, unlike IEEE 802.1d compliant bridges, do not necessarily filter all types of illegal frames. Some switches, for example, which do not store frames before forwarding them to their destination ports may not filter over-sized frames (jabbers) or verify the validity of the Frame Check Sequence field. Other illegal frames are under-sized frames (runts) and misaligned frames.

Measurement units:

N-octet frames filtered or not filtered

Issues:

See Also:

2.22 Maximum broadcast forwarding rate

Definition:

The number of broadcast frames per second that a switch can deliver to all ports at maximum load.

Discussion:

There is no standard forwarding mechanism used by switches to forward broadcast frames. It is useful to determine the broadcast forwarding rate for frames switched between ports on the same card, ports on different cards

in the same chassis and ports on different chassis linked together over backbone connections.

Measurement units:

N-octet frames per second

Issues:

See Also:

broadcast latency (2.23)

[2.23](#) Broadcast latency

Definition:

The time required by a switch to forward a broadcast frame to each port located within a broadcast domain.

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. The latency measurements SHOULD be bit oriented as described in 3.8 of [RFC 1242](#). It is useful to determine broadcast latency for frames forwarded between ports on the same card, ports on different cards in the same chassis and ports on different chassis linked together over backbone connections.

Measurement units:

nanoseconds

microseconds

milliseconds

seconds

Issues:

See Also:

broadcast forwarding rate (2.20)

[3](#). Index of definitions

[2.1](#) Reminder of [RFC 1242](#) definition format

[2.2](#) Unidirectional traffic

[2.3](#) Bidirectional traffic

[2.4](#) Meshed traffic

[2.5](#) Burst

[2.6](#) Burst size

[2.7](#) Inter-burst gap (IBG)

[2.8](#) Port load

[2.9](#) Overload

[2.10](#) Intended rate

[2.11](#) Offered rate

[2.12](#) Maximum load

[2.13](#) Forwarding rate

- [2.14](#) Forwarding rate at maximum load
- [2.15](#) Flooding
- [2.16](#) Backpressure
- [2.17](#) Forward pressure
- [2.18](#) Head of line blocking
- [2.19](#) Address handling
- [2.20](#) Address learning rate
- [2.21](#) Illegal frames
- [2.22](#) Maximum broadcast forwarding rate
- [2.23](#) Broadcast latency

[4. Acknowledgments](#)

In order of appearance Jean-Christophe Bestaux of European Network Laboratories, Ajay Shah of Wandel & Goltermann, Henry Hamon of Netcom Systems, Stan Kopek of Digital Equipment Corporation, Kevin Dubray of Bay Networks, and Doug Ruby of Prominet 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

[5. Editor's Address](#)

Robert Mandeville
ENL (European Network Laboratories)
35, rue Beaubourg
[75003](#) Paris
France
mobile phone: +33 6 07 47 67 10
phone: +33 1 39 44 12 05
fax: + 33 1 39 44 12 06
email: bob.mandeville@eunet.fr