

**Benchmarking Terminology for LAN Switching Devices**  
< [draft-ietf-bmwg-lanswitch-04.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) 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 forwarding performance, latency, address handling and filtering.

## **2. Term definitions**

[RFC 1242](#) "Benchmarking Terminology for Network Interconnect Devices" defines many of the terms that are used in this document. The terminology document should be consulted before attempting to make use of this document. [RFC 1944](#) "Benchmarking Methodology for Network Interconnect Devices" defines 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. The new terms are defined in groups.

### **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, its 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.

## **3. Index of terms defined**

### **3.1 Devices**

#### **3.1.1 Device under test (DUT)**

#### **3.1.2 System under test (SUT)**

### **3.2 Traffic patterns**

#### **3.2.1 Unidirectional traffic**

- [\*\*3.2.2\*\*](#) Bidirectional traffic
- [\*\*3.2.3\*\*](#) Partially meshed traffic
- [\*\*3.2.4\*\*](#) Fully meshed traffic

### [\*\*3.3\*\*](#) Bursts

- [\*\*3.3.1\*\*](#) Burst
- [\*\*3.3.2\*\*](#) Burst size
- [\*\*3.3.3\*\*](#) Inter-burst gap (IBG)

### [\*\*3.4\*\*](#) Loads

- [\*\*3.4.1\*\*](#) Intended load
- [\*\*3.4.2\*\*](#) Offered load
- [\*\*3.4.3\*\*](#) Maximum offered load
- [\*\*3.4.4\*\*](#) Overloading

### [\*\*3.5\*\*](#) Forwarding rates

- [\*\*3.5.1\*\*](#) Forwarding rate at maximum load
- [\*\*3.5.2\*\*](#) Maximum forwarding rate

### [\*\*3.6\*\*](#) Congestion control

- [\*\*3.6.1\*\*](#) Backpressure
- [\*\*3.6.2\*\*](#) Forward pressure
- [\*\*3.6.3\*\*](#) Head of line blocking

### [\*\*3.7\*\*](#) Address handling

- [\*\*3.7.1\*\*](#) Address caching capacity
- [\*\*3.7.2\*\*](#) Address learning rate
- [\*\*3.7.3\*\*](#) Flood count

### [\*\*3.8\*\*](#) Errored frame filtering

- [\*\*3.8.1\*\*](#) Errored frames

### [\*\*3.9\*\*](#) Broadcasts

- [\*\*3.9.1\*\*](#) Broadcast forwarding rate at maximum load
- [\*\*3.9.2\*\*](#) Broadcast latency

## [\*\*3.1\*\*](#) Devices

This group applies to all types of networking devcies.

### [\*\*3.1.1\*\*](#) Device under test (DUT)

Definition:

The network forwarding device to which stimuli is offered and response measured.

Discussion:

A single stand-alone or modular unit generally equipped with its own power supply.

Measurement units:

n/a

Issues:

See Also:

System under test (SUT) (3.1.2)

### **3.1.2 System Under Test (SUT).**

Definition:

The collective set of network devices to which stimuli is offered as a single entity and response measured.

Discussion:

A system under test may be comprised of a variety of networking devices. Some devices may be active in the forwarding decision-making process, such as routers or switches; other devices may be passive such as CSU/DSUs. Regardless of constituent components, the system is treated as a singular entity to which stimuli is offered and response measured.

Measurement units:

n/a

Issues:

See Also:

Device under test (DUT) (3.1.1)

## **3.2 Traffic patterns**

This group applies to the distribution of frames forwarded by any DUT/SUT.

### **3.2.1 Unidirectional traffic**

Definition:

Single or multiple streams of frames forwarded in one direction only from a set of input ports to a set of 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 DUT/SUT in different ways. For example unidirectional traffic can be sent to two

or more input ports from an external source and forwarded by the DUT/SUT to a single output port (n-to-1) or such traffic can be sent to a single input port and forwarded by the DUT/SUT 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 a DUT/SUT is 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.

It should be noted that 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 (3.2.2)

partially meshed traffic (3.2.3)

fully meshed traffic (3.2.4)

### **[3.2.2](#) Bidirectional traffic**

Definition:

Two or more streams of frames forwarded in opposite directions between two or more ports of a DUT/SUT.

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 (3.2.1)

partially meshed traffic (3.2.3)

fully meshed traffic (3.2.4)

### **[3.2.3](#) Partially meshed traffic**

Definition:

Streams of frames forwarded between a set of input ports and a set of output ports of a DUT/SUT with a one to many, many to one or many to many mapping of input ports to output ports.

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. Meshed traffic can be unidirectional or bidirectional.

Measurement units:

n/a

Issues:

half duplex / full duplex

See Also:

unidirectional traffic (3.2.1)

bidirectional traffic (3.2.2)

fully meshed traffic (3.2.4)

bursts (3.3)

### [3.2.4](#) Fully meshed traffic

Definition:

Streams of frames switched simultaneously between all of a designated number of ports of a 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:

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.

It should be noted that bidirectional meshed traffic on half duplex ports is inherently bursty since ports must interrupt transmission whenever they receive frames. This kind of bursty meshed traffic is characteristic of real network traffic and can be advantageously used to diagnose a DUT/SUT by exercising many of its component parts simultaneously. Additional inspection may be warranted to correlate the frame forwarding capacity of a DUT/SUT when offered meshed traffic and the behavior of individual elements such as input and output buffers, buffer allocation mechanisms, aggregate switching capacity, processing speed or medium access control.

When offering bursty meshed traffic to a DUT/SUT a number of variables have to be considered. These include frame size, the number of frames within bursts, the interval between bursts as well as the

distribution of load between incoming and outgoing traffic. Terms related to bursts are defined in [section 3.3](#) below.

Measurement units:

n/a

Issues:

half duplex / full duplex

See Also:

unidirectional traffic (3.2.1)

bidirectional traffic (3.2.2)

partially meshed traffic (3.2.4)

bursts (3.3)

### **[3.3](#) Bursts**

This group applies to the intervals defining traffic forwarded by DUT/SUT.

#### **[3.3.1](#) 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 (3.3.2)

inter-burst gap (IBG) (3.2.3)

#### **[3.2.2](#) 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 (3.3.1)  
inter-burst gap (IBG) (3.2.3)

### **[3.2.3](#) 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 (3.3.1)  
burst size (3.2.2)

## **[3.4](#) Loads**

This group applies to the rates that traffic is offered to any DUT/SUT.

### **[3.4.1](#) Intended load**

Definition:  
The number of frames per second that an external source attempts to transmit to a DUT/SUT for forwarding to a specified output port or ports.

Discussion:  
Collisions on CSMA/CD links or the action of congestion control mechanisms can effect the rate at which an external source of traffic transmits frames to a DUT/SUT. This makes it useful to distinguish



the load that an external source attempts to apply to a DUT/SUT and the load it is observed or measured to apply.

In the case of Ethernet an external source of traffic must implement the truncated binary exponential back-off algorithm to ensure that it is accessing the medium legally.

Measurement units:

bits per second

N-octets per second

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

Issues:

See Also:

offered load (3.4.2)

### **[3.4.2 Offered load](#)**

Definition:

The number of frames per second that an external source can be observed or measured to transmit to a DUT/SUT for forwarding to a specified output port or ports.

Discussion:

The load which an external device can be observed to apply to a DUT/SUT may be less than the load the external device attempts to apply due to collisions or the action of congestion control mechanisms.

Frames which are not successfully transmitted by an external source of traffic to a DUT/SUT MUST NOT be counted as transmitted frames when measuring the forwarding rate of a DUT/SUT.

The frame count on a port of a DUT/SUT 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. Such frames should be treated as modifiers as described in [section 11 of RFC 1944](#).

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 load (3.4.1)

### **[3.4.3 Maximum offered load](#)**

Definition:

The highest number of frames per second that an external source can

transmit to a DUT/SUT for forwarding to a specified output port or ports.

**Discussion:**

The maximum load that an external device can apply to a DUT/SUT may not equal the maximum load allowed by the medium. This will be the case when an external source lacks the resources to transmit frames at the minimum legal inter-frame gap or when it has sufficient resources to transmit frames below the minimum legal inter-frame gap. Moreover, maximum load may vary with respect to parameters other than a medium's maximum theoretical utilization. For example, on those media employing tokens, maximum load may vary as a function of Token Rotation Time, Token Holding Time, or the ability to chain multiple frames to a single token.

The maximum load that an external device applies to a DUT/SUT MUST be specified when measuring forwarding rates.

**Measurement units:**

bits per second

N-octets per second

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

**Issues:**

**See Also:**

offered load (3.4.2)

### **3.4.4 Overloading**

**Definition:**

Attempting to load a DUT/SUT in excess of the maximum rate of transmission allowed by the medium.

**Discussion:**

Overloading can serve to exercise buffers and buffer allocation algorithms as well as congestion control mechanisms.

The number of input port or ports required to overload one or more output ports of a DUT/SUT will vary according to the media rates of the ports involved. An external source can also overload a port by transmitting frames below the minimum inter-frame gap. This can serve to determine whether a device respects the minimum inter-frame gap. Overloading can be achieved with unidirectional, bidirectional and meshed traffic.

**Measurement units:**

N-octets per second

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

N-octet frames per second

**Issues:**

See Also:  
offered load (3.4.2)

### **3.5 Forwarding rates**

This group applies to the rates at which traffic is forwarded by any DUT/SUT in response to a stimulus.

#### **3.5.1 Forwarding rate**

Definition:

The number of frames per second that a device can be observed to successfully transmit to the correct destination port in response to a specified offered load.

Discussion:

Unlike throughput defined in [section 3.17 of RFC 1242](#), forwarding rate makes no explicit reference to frame loss. Forwarding rate, which must only be sampled on the output side of the ports under test, can be measured for unidirectional, bidirectional or meshed traffic and should be sampled in fixed time intervals of one second. If longer or shorter intervals are used they should be cited when reporting a device's forwarding rate.

It should be noted that the forwarding rate of a DUT/SUT may be sensitive to the action of congestion control mechanisms.

Measurement units:

N-octet frames per second

Issues:

See Also:  
offered load (3.4.2)  
forwarding rate at maximum offered load (3.5.2)  
maximum forwarding rate (3.5.3)

#### **3.5.2 Forwarding rate at maximum offered load**

Definition:

The number of frames per second that a device can be observed to successfully transmit to the correct destination port in response to the maximum offered load.

Discussion:

Forwarding rate at maximum offered load may be less than the maximum rate at which a device can be observed to successfully forward traffic. This will be the case when the ability of a device to forward frames degenerates when offered traffic at maximum load. Maximum offered load must be cited when reporting forwarding rate at maximum offered load.

Measurement units:

N-octet frames per second

Issues:

See Also:

maximum offered load (3.4.3)

forwarding rate (3.5.1)

maximum forwarding rate (3.5.3)

### **3.5.3 Maximum forwarding rate**

Definition:

The highest forwarding rate of a DUT/SUT taken from an iterative set of forwarding rate measurements.

Discussion:

The forwarding rate of a device may degenerate before maximum load is reached. The load applied to a device must be cited when reporting maximum forwarding rate.

Measurement units:

N-octet frames per second

Issues:

See Also:

offered load (3.4.2)

forwarding rates (3.5.1)

forwarding rate at maximum load (3.5.2)

## **3.6 Congestion**

This group applies to the behavior of a DUT/SUT when congestion or contention is present.

### **3.6.1 Backpressure**

Definition:

Any technique used by a DUT/SUT to attempt to 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 overflow. 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.

It should be noted that jamming and other flow control methods may 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 (3.6.2)

### **3.6.2 Forward pressure**

Definition:

Methods which depart from or otherwise violate a defined standardized protocol in an attempt to increase the forwarding performance of a DUT/SUT.

Discussion:

A DUT/SUT may be found to inhibit or abort backoff algorithms in order to force access to the medium when contention occurs. It should be noted that the backoff algorithm should be fair whether the DUT/SUT is in a congested or an uncongested state. Transmission below the minimum inter-frame gap or the disregard of flow control primitives fall into this category.

Measurement units:

intervals between frames in microseconds

intervals in microseconds between transmission retries during 16 successive collisions.

Issues:

truncated binary exponential backoff algorithm

See also:

backpressure (3.6.1)

### **3.6.3 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 (3.2.1)

### **3.7 Address handling**

This group applies to the process of address resolution which enables a DUT/SUT to forward frames to the correct destination.

#### **3.7.1 Address caching capacity**

Definition:

The number of MAC addresses per n ports, per module or per device that a DUT/SUT 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 DUT/SUT. 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 DUT/SUT begins flooding frames.

Measurement units:

number of MAC addresses per n ports, per module and/or per chassis

Issues:

See Also:

Address learning rate (3.7.2)

#### **3.7.2 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 caching capacity (3.7.1)

### **3.7.3 Flood count**

**Definition:**

Frames forwarded to ports which do not correspond to the destination MAC address information when traffic is offered to a DUT/SUT for forwarding.

**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

**Issues:**

Spanning tree BPDUs.

**See Also:**

address caching capacity (3.7.1)

## **3.8 Errored frame filtering**

This group applies to frames with errors which a DUT/SUT may filter.

### **3.8.1 Errored 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/a

**Issues:**

**See Also:**

## **3.9 Broadcasts**

This group applies to MAC layer and network layer broadcast frames.

### **3.9.1 Broadcast forwarding rate**

**Definition:**

The number of broadcast frames per second that a DUT/SUT can be observed to deliver to all ports located within a broadcast domain in response to a specified offered 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. The terms maximum broadcast forwarding rate and broadcast forwarding rate at maximum load follow directly from the terms already defined for forwarding rate measurements in [section 3.5](#) above.

**Measurement units:**

N-octet frames per second

**Issues:**

**See Also:**

forwarding rate at maximum load (3.5.2)

maximum forwarding rate (3.5.3)

broadcast latency (3.9.2)

### **3.9.2 Broadcast latency**

**Definition:**

The time required by a DUT/SUT 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 (3.9.1)

#### **4. References:**

**1. [RFC 1242](#) "Benchmarking Terminology for Network Interconnect Devices"**

**2. [RFC 1944](#) "Benchmarking Methodology for Network Interconnect Devices"**

#### **5. Acknowledgments**

A special thanks goes to the IETF BenchMark WorkGroup for the many suggestions it collectively made to help complete this draft. Kevin Dubray (Bay Networks), Jean-Christophe Bestaux (ENL), Ajay Shah ( WG), Henry Hamon (Netcom Systems), Stan Kopek (3Com) and Doug Ruby (Prominet) all provided valuable input at various stages of this project.

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