Network Working Group
INTERNET-DRAFT
Expiration Date: May 1997
Nov 1996

R. Craig
Cisco Systems

Terminology for Cell/Call Benchmarking

<draft-ietf-bmwg-call-00.txt>

Status of this Memo

Abstract

The purpose of this draft is to add terminology specific to the cell and call-based switch environment to that defined by the Benchmarking Methodology Working Group (BMWG) of the Internet Engineering Task Force (IETF) in RFC1242.

While primarily directed towards wide area switches, portions of the document may be useful for benchmarking other devices such as ADSU's.

1. Introduction

In light of the increasing use of cell-based and/or circuit-switched transport layers in building networks, it would be useful to develop a set of benchmarks with which to compare technologies, implementation strategies, and products.

1.1 Terminology Brought Forward

The terminology defined in RFC 1242 applies equally well to this memo. There is also a certain amount of overlap with terms defined in draft-ietf-bmwg-lanswitch-00.txt.

2. Definition Format (from RFC1242)

Term to be defined.

Definition:
The specific definition for the term.

Discussion:
A brief discussion of the term, its application and any restrictions on measurement procedures.

Measurement units:
Units used to record measurements of this term, if applicable.

3. Term Definitions

3.1 Virtual Circuit

This group applies to those switches that are connection-oriented.

3.1.1 Call setup time

Definition: the length of time for the virtual circuit to be established.

Discussion: as measured from the initiation of the signalling to circuit establishment.

Measurement units: fractional seconds

Issues:

See also:

3.1.2 Call setup rate (sustained)

Definition: the maximum sustained rate of successful connection establishment.

Discussion: without loss of existing calls.

Measurement units: calls per second

Issues:

See also:

3.1.3 Call maintenance overhead

Definition: the amount of work required to maintain the calls that have been established.

Discussion: a method to obtain the desired result would be to benchmark with PVC's in place, then with SVC's. The difference in results would be the overhead.

Measurement units:

Issues:

See also:

3.1.4 Call teardown time

Definition: the length of time for the virtual circuit to be torn down.

Discussion: measured from the start of the signalling to the freeing of the resources associated with that call (end to end, if applicable).

Measurement units: fractional seconds

Issues:

See also:

3.1.5 Call teardown rate (sustained)

Definition: the maximum rate at which calls can be successfully torn down.

Discussion: without loss of existing calls, and without failure to tear down any calls that have been signalled to be destroyed.

Measurement units: teardowns per second

Issues:

See also:

3.1.6 Impact of Signalling on Forwarding

Definition: cells per second versus calls per second

Discussion: some devices use the same engine for cell forwarding and call maintenance. In this case, interaction between the two

will be inevitable. More interesting, however, would be the case where the two processing functions are clearly separate, yet still interact.

Measurement units: cells per second versus calls per second

Issues:

See also:

3.2 Cell/Packet Interaction

This group applies to cell-based switches, connection-oriented or not.

3.2.1 Packet disassembly/reassembly time (peak)

Definition: the length of time to disassemble a layer 3 packet into layer 2 cells, or reassemble cells into a packet.

Discussion: with no packet or cell loss or corruption.

Measurement units: the appropriate fraction of a second

Issues:

See also:

3.2.2 Packet disassembly/reassembly rate (sustained)

Definition: the maximum sustained rate at which packets can be disassembled/reassembled into/from cells.

Discussion: without loss or corruption.

Measurement units: packets per second

Issues:

See also:

3.2.3 Full packet drop rate (on cell loss)

Definition: the rate at which cell loss triggering full packet drop can be detected/sustained.

Discussion: When a packet is disassembled into cells, typically many cells result. When these cells are transmitted, they are

subject to loss or corruption. The device should recognize at the cell/packet boundary that a cell or cells belonging to a given packet has been lost and should drop that packet, immediately freeing those resources. A couple of things are of interest here: whether the switch is able to detect very small amounts of cell loss and correctly drop the associated packets and whether large amounts of cell loss perturb this ability in any way.

Measurement units: (dropped) packets per second

Issues:

See also:

3.2.4 End to end data integrity

Definition: the percentage of packets (post-reassembly) that actually contain undetected data link layer corruption.

Discussion: some network devices have been known to regenerate CRC's over the re-assembled packet (i.e., the CRC is not carried end to end), resulting in undetected data link layer corruption or re-ordering of cells in a packet.

Measurement units: percentage

Issues: production of a stream of traffic containing internal checksums sufficiently strong to detect cell re-ordering (the IP checksum is not). The ISIS LSP checksum is.

See also:

3.3 Switch Fabric

This group applies to all switches.

3.3.1 Switch type

Definition: the type of switch architecture.

Discussion: Is this of any importance? We are concerned with interesting "metrics" and how they affect the performance of a device. I'm not sure switch architecture falls into this category except as an perhaps interesting bit of trivia.

Measurement units: n/a

Issues:

See also:

3.3.2 Topology Table Size

Definition: number of network elements supported.

Discussion: switches may support a limited topology due to static table sizes or processing limitations. This is true whether it's a "LAN" switch running spanning tree or a "WAN" switch running OSPF. The effect of a limited topology table on a switch in a real-world environment can be disastrous.

A similar metric (2.14 Address handling) is mentioned in "draft-ietf-bmwg-lanswitch-00.txt". Here, a more general metric is intended.

Measurement units: number

Issues: Measuring the effects of an overflow is probably meaningless, since in the multi-switch case, there is no longer any network to speak of, hence, nothing to measure.

If a device handles table overflow gracefully, this should be noted. Similarly, if a device crashes and burns on table overflow, this should be noted.

See also:

3.3.3 Topology Table Learning Rate

Definition: the rate at which the topology table can be filled or updated.

Discussion: a single switch in isolation learning MAC addresses will flood frames when the rate exceeds its learning capability. This metric is covered in "2.15 Address learning speed" of "draft-ietf-bmwg-lanswitch-00.txt". We generalize the metric here to include the topological databases of routing protocols used in switched networks (among the switches themselves) as well as the spanning tree recalculation among multiple LAN switches.

Measurement units: frames per second 1) with maximum diversity of addresses, 2) with routing instability introduced.

Issues:

See also:

3.3.4 "Bandwidth"

Definition: internal bandwidth of the switch fabric.

Discussion: open to some interpretation ;-). Should probably be stated as some combination of the slowest and fastest elements in the switching path.

Measurement units: bits per second

Issues:

See also:

3.3.5 Throughput (from RFC1242) (Cell forwarding rate)

Definition: The maximum rate at which none of the offered frames are dropped by the device.

Discussion: This metric probably overlaps work being done in the ATM Forum.

Measurement units: cells per second

Issues:

See also:

3.3.6 Non-Blocking factor

Definition: simultaneous communication amongst multiple ports.

Discussion: a switch is termed "non-blocking" if multiple ports are able to communicate across the switch fabric at the same time. If a popular destination port can accept connections from more than one source port, the number of those connections is the non-blocking factor. We are interested in the number of ports which can simultaneously transmit to a single port (N), the number of ports which can simultaneously receive from N other ports (M), and the total number of ports on the switch (P).

Measurement units: N:1, N:M:P (switch-wide measurement)

Issues:

See also:

3.4 Buffering

This group applies to all switches.

3.4.1 Buffering strategy

Definition: central pool of buffers versus distributed pools. Pools of one size versus multiple MTU sizes.

Discussion: There are tradeoffs in each approach: bus bandwidth and arbitration cycles for centrality, over-configuration of memory for distributed pools and one-size-fits-all, greater number of drops due to buffer exhaustion with MTU-tailored buffers.

The effectiveness of the given strategy is revealed by the performance of the device in overload conditions. For example, one might cause the majority of input buffers to migrate to one port which is experiencing a sustained burst of traffic, and then cause another port to burst, creating input drops due to lack of buffers while the device re-allocates its buffer pool.

Measurement units: underruns (can't feed transmitting interface quickly enough, indicative of bus bw or access problem), input/output drops (buffer exhaustion), overruns (another indicator of either buffer or CPU exhaustion)

Issues:

See also:

3.4.2 Buffering per output

Definition: the number of buffers per output port and their size.

Discussion: It must also be noted whether the buffers are local to the line card, whether they are dynamically allocated from a central pool, whether they are MTU-tailored, and so on.

Measurement units: octets

Issues:

See also: 3.4.1

3.4.3 Buffering per input

Definition: the number of buffers per input port and their size.

Discussion: see 3.4.2

Measurement units: octets

Issues:

See also: 3.4.1

3.5 Congestion Control

This group applies to all switches.

3.5.1 Congestion avoidance

Definition: effectiveness of measures taken by the switch to avoid congestion.

Discussion: connections that are bursting above their committed rate may have cells buffered at the ingress, in order to avoid congestion in the trunks and impact on other connections, or they may simply be marked "discard-eligible" and forwarded into the network, hoping for the best.

Distinguishing between these two approaches should be relatively simple. In the first case, latency for the bursting session increases, but there is no cell loss. Other sessions are unaffected. In the second case, there may be cell loss across any of the sessions, and latency may increase across all.

Measurement units: dropped cells, latency

Issues:

See also:

3.5.2 Congestion management

Definition: effectiveness of measures taken by the switch to deal with congestion.

Discussion: in the face of sustained traffic above committed rate on multiple sessions, a switch has little choice but to begin discarding cells, since buffering cannot be infinite. This case might arise if one were wildly profligate in over-subscribing trunk bandwidth, or if one had neglected to analyze the network applications to be run over the network and they were found to be network-hostile (UDP, IPX, AT, NetBIOS, for example).

The switch has some discretion in deciding which cells to drop. Presumably, the strategy should involve something resembling

"fairness".

The basic idea is that ill-behaved connections should not starve others for resources.

Measurement units: latency, cell drops

Issues:

See also:

3.5.3 Queueing strategies

Definition: the method used for queueing frames.

Discussion: FIFO, WFQ, SFQ, tail drop, RED. Queue per interface, per rate or per connection?

Measurement units:

Issues:

See also:

3.6 Inter-switch protocols

This group applies to all switches.

3.6.1 Impact of Routing on Forwarding

Definition: interaction between routing protocol and data forwarding operations.

Discussion: No amount of routing fluctuation should have an impact on data forwarding for unaffected destinations. Similarly, no amount of data forwarding should cause the routing to become unstable.

Measurement units: route flaps per second versus cells per second, cells per second versus route stability (table fluctuation or peer loss).

Issues:

See also:

3.6.2 Impact of Congestion Control

Definition: interaction between congestion control and data forwarding operations.

Discussion: switches may share views of congestion in-band through the network. Should these feedback messages be delayed or lost, the potential exists for an incorrect picture of current network conditions, which may exacerbate congestion and lead to cell loss. Worse, it is possible to enter a stable oscillation state, where ever-increasing waves of congestion overwhelm the switches.

Measurement units:

Issues:

See also:

3.7 Quality of Service

This group applies to all switches.

3.7.1 Traffic Management

Definition: impact of misbehaving class on others, for example data forwarding on voice or video frames and vice versa.

Discussion: we wish to quantify the potential interaction amongst the various classes of service. Constant bit rate (CBR), variable bit rate (VBR) (real and non-real time?), and available bit rate (ABR) streams are established, within their respective service levels, but sufficient to subscribe the trunk to 90%. The bit rate of each is increased until it has exceeded its allocation by a degree which should cause loss or delay in the other streams.

Measurement units: cells (lost) per second, latency

Issues: some switches perform compression and silence suppression. Should these features be disabled?

See also:

3.7.2 Mapping of IP ToS/Precedence onto QoS

Definition: some method is required to map IP type of service and/or precedence values onto the switch's notion of quality of service.

Discussion:

Measurement units:

Issues:

See also:

3.8 Multicast

3.8.1 Cell replication

Definition: the device's ability to forward a cell to multiple ports simultaneously (multicast).

Discussion:

Measurement units: replication factor 1:N and cells per second measured at ingress versus cells per second measured at the egresses

Issues:

See also:

3.8.2 Impact of multicast on unicast

Definition: switch's ability to insulate unicast traffic from the effects of multicast.

Discussion: a poorly-designed replication scheme could easily swamp unicast traffic. Yet, multicast traffic often has QoS needs. How does one reconcile the competing requirements?

Measurement units: cell loss, delay

Issues:

See also:

Security Considerations

Security issues are not addressed in this memo.

Editor's Address

Robert Craig
Cisco Systems
7025 Kit Creek Road

PO Box 14987
Research Triangle Park, NC 27709
(919) 472-2886
rcraig@cisco.com