

Benchmarking Methodology WG
Internet Draft
Updates: [2544](#) (if approved)
Intended status: Informational
Expires: November 2010

Rajiv Asati
Cisco
Carlos Pignataro
Cisco
Fernando Calabria
Cisco
Cesar Olvera
Consulintel
May 1, 2010

Device Reset Characterization
draft-ietf-bmwg-reset-00

Abstract

An operational forwarding device may need to be re-started (automatically or manually) for a variety of reasons, an event that we call a "reset" in this document. Since there may be an interruption in the forwarding operation during a reset, it is useful to know how long a device takes to begin forwarding packets again.

This document specifies a methodology for characterizing reset during benchmarking of forwarding devices, and provides clarity and consistency in reset test procedures beyond what's specified in [RFC2544](#). It therefore updates [RFC2544](#).

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

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>

This Internet-Draft will expire on November 1, 2010.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction.....	4
1.1.	Scope.....	5
2.	Key Words to Reflect Requirements.....	5
3.	Reset Test.....	5
3.1.	Hardware Reset.....	6
3.1.1.	Routing Processor (RP) / Routing Engine reset.....	6
3.1.1.1.	RP Failure for a single-RP device (mandatory)...	6
3.1.1.2.	RP Failure for a multiple-RP device (optional)..	8
3.1.2.	Line Card (LC) Removal and Insertion (mandatory)....	10
3.2.	Software Reset.....	12
3.2.1.	Operating System (OS) reset (mandatory).....	13
3.2.2.	Process reset (optional).....	14
3.3.	Power interruption.....	16
3.3.1.	Power Interruption (mandatory).....	17
4.	Security Considerations.....	19
5.	IANA Considerations.....	19
6.	Acknowledgments.....	19
7.	References.....	20
7.1.	Normative References.....	20
7.2.	Informative References.....	20
	Authors' Addresses.....	21

1. Introduction

An operational forwarding device (or one of its components) may need to be re-started for a variety of reasons, an event that we call a "reset" in this draft. Since there may be an interruption in the forwarding operation during a reset, it is useful to know how long a device takes to begin forwarding packets again.

However, the answer to this question is no longer simple and straight-forward as the modern forwarding devices employ many hardware advancements (distributed forwarding, etc.) and software advancements (graceful restart, etc.) that influence the recovery time after the reset.

In order to provide consistent and fairness while benchmarking a set of different DUTs, the Network tester / Operator MUST (a) use identical control and data plane information during testing, (b) document & report any factors that may influence the overall time after reset / convergence.

Some of these factors follow:

1. Type of reset - Hardware (line-card crash, etc.) vs. Software (protocol reset, process crash, etc.) or even complete power failures
2. Manual vs. Automatic reset
3. Scheduled vs. non-scheduled reset
4. Local vs. Remote reset
5. Scale - Number of line cards present vs. in-use
6. Scale - Number of physical and logical interfaces
7. Scale - Number of routing protocol instances
8. Scale - Number of Routing Table entries
9. Scale - Number of Route Processors available
10. Performance - Redundancy strategy deployed for route processors and line cards
11. Performance - Interface encapsulation as well as achievable NDR (non-dropping rate)

12. Any other internal or external factor that may influence recovery time after a hardware or software reset

This document specifies a methodology for characterizing reset during benchmarking of forwarding devices, and provides clarity and consistency in reset procedures beyond what's specified in [\[RFC2544\]](#). These procedures may be used by other benchmarking documents such as [\[RFC2544\]](#), [\[RFC5180\]](#), [\[RFC5695\]](#), etc.

This document updates [Section 26.6 of \[RFC2544\]](#).

1.1. Scope

This document focuses on only the reset criterion of benchmarking, and presumes that it would be beneficial to [\[RFC2544\]](#), [\[RFC5180\]](#), [\[RFC5695\]](#), and other BMWG benchmarking efforts.

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 [BCP 14](#), [RFC 2119](#) [\[RFC2119\]](#). [RFC 2119](#) defines the use of these key words to help make the intent of standards track documents as clear as possible. While this document uses these keywords, this document is not a standards track document.

3. Reset Test

This section contains the description of the tests that are related to the characterization of DUTs (Device Under Test) / SUTs (System Under Test) speed to recover from a reset. There are three types of reset considered in this document:

1. Hardware resets
2. Software resets
3. Power interruption

Different types of reset have potentially different impact on the forwarding behavior of the device. As an example, a software reset (of a routing process) might not result in forwarding interruption, whereas a hardware reset (of a line card) most likely will.

[Section 3.1](#) describes various hardware resets, whereas [Section 3.2](#) describes various software resets. Additionally, [Section 3.3](#) describes power interruption tests. These sections define and characterize these resets.

Additionally, since device specific implementations may vary for hardware and software type resets, it is desirable to classify each test case as "MUST" or "optional".

[3.1. Hardware Reset](#)

A test designed to characterize the time it takes a DUT to recover from the hardware reset.

A "hardware reset" generally involves the re-initialization of one or more physical components in the DUT, but not the entire DUT.

A hardware reset is executed by the operator for example by physical removal of a physical component, by pressing on a "reset" button for the component, or could even be triggered from the command line interface.

For routers that do not contain separate Routing Processor and Line Card modules, the hardware reset tests are not performed since they are not relevant; instead, the power interruption tests MUST be performed (see [Section 3.3](#)) in these cases.

[3.1.1. Routing Processor \(RP\) / Routing Engine reset](#)

The Routing Processor (RP) is the DUT module that is primarily concerned with Control Plane functions.

[3.1.1.1. RP Failure for a single-RP device \(mandatory\)](#)

Objective

To characterize the speed at which a DUT recovers from a Route processor hardware reset in a single RP environment.

Procedure

First, ensure that the RP is in a permanent state to which it will return to after the reset, by performing some or all of the following operational tasks: save the current DUT configuration, specify boot parameters, ensure the appropriate software files are available, or perform additional Operating System or hardware related task.

Second, ensure that the DUT is able to forward the traffic for at least 15 seconds before any test activities are performed. The traffic should use the minimum frame size possible on the media used in the testing and rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the recovery time measurement.

Modern traffic generators support the feature of transmitting packets while ignoring the link status. The operator / tester MUST ensure that this feature is enabled.

Third, perform the Route Processor (RP) hardware reset at this point. This entails for example physically removing the RP to later re-insert it, or triggering a hardware reset by other means (e.g., command line interface, physical switch, etc.)

Finally, the characterization is completed by measuring the frame loss and recovery time from the moment the RP is re-initialized or reinserted.

Reporting format

The reset results are reported in a simple statement including the frame loss and recovery times.

For each test case, it is RECOMMENDED that the following parameters be reported in these units:

Parameter	Units or Examples
Throughput	Frames per second and bits per second
Loss	Frames
Time	Seconds, with sufficient resolution

	to convey meaningful info
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, GigE (Gigabit Ethernet), POS (Packet over SONET), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, HDLC, etc.

Additionally, the DUT and test bed provisioning, configuration, and deployed methodologies that may influence the overall recovery time MUST be listed. (Refer to the additional factors listed in [Section 1](#)).

The reporting of results MUST regard repeatability considerations from [Section 4 of \[RFC2544\]](#). It is RECOMMENDED to perform multiple trials and report average results.

[3.1.1.2](#). RP Failure for a multiple-RP device (optional)

Objective

To characterize the speed at which a "secondary" Route Processor (sometimes referred to as "backup" RP) of a DUT becomes active after a "primary" (or "active") Route Processor hardware reset. This process is often referred to as "RP Switchover". The characterization in this test should be done for the default DUT behavior as well as a DUT's non-default configuration that minimizes frame loss.

Procedure

This test characterizes "RP Switchover". Many implementations allow for optimized switchover capabilities that minimize the

downtime during the actual switchover. This test consists of two sub-cases from a switchover characteristics standpoint: First, a default behavior (with no switchover-specific configurations); and second, a non-default behavior with switchover configuration to minimize frame loss. Therefore, the procedures hereby described are executed twice, and reported separately.

First, ensure that the RPs are in a permanent state such that the secondary will be activated to the same state as the active is, by performing some or all of the following operational tasks: save the current DUT configuration, specify boot parameters, ensure the appropriate software files are available, or perform additional Operating System or hardware related task.

Second, ensure that the DUT is able to forward the traffic for at least 15 seconds before any test activities are performed. The traffic should use the minimum frame size possible on the media used in the testing and rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the recovery time measurement.

Modern traffic generators support the feature of transmitting packets while ignoring the link status. The operator / tester MUST ensure that this feature is enabled.

Third, perform the primary Route Processor (RP) hardware reset at this point. This entails for example physically removing the RP, or triggering a hardware reset by other means (e.g., command line interface, physical switch, etc.) Is up to the Operator to decide if the active RP needs to be re-inserted after a grace period or not.

Finally, the characterization is completed by measuring the complete frame loss and recovery time from the moment the active RP is hardware-reset.

Reporting format

The reset results are reported twice, one for the default switchover behavior and the other for the non-default one. For each, the report consists of a simple statement including the frame loss and recovery times, as well as any specific redundancy scheme in place.

For each test case, it is RECOMMENDED that the following parameters be reported in these units:

Parameter	Units or Examples
Throughput	Frames per second and bits per second
Loss	Frames
Time	Seconds, with sufficient resolution to convey meaningful info
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, GigE (Gigabit Ethernet), POS (Packet over SONET), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, HDLC, etc.

Additionally, the DUT and test bed provisioning, configuration, and deployed methodologies that may influence the overall recovery time MUST be listed. (Refer to the additional factors listed in [Section 1](#)).

The reporting of results MUST regard repeatability considerations from [Section 4 of \[RFC2544\]](#). It is RECOMMENDED to perform multiple trials and report average results.

[3.1.2.](#) Line Card (LC) Removal and Insertion (mandatory)

The Line Card (LC) is the DUT component that is responsible with packet forwarding.

Objective

To characterize the speed at which a DUT recovers from a Line Card removal and insertion event.

Procedure

For this test, the Line Card that is being hardware-reset MUST be on the forwarding path and all destinations MUST be directly connected.

First, complete some or all of the following operational tasks: save the current DUT configuration, specify boot parameters, ensure the appropriate software files are available, or perform additional Operating System or hardware related task.

Second, ensure that the DUT is able to forward the traffic for at least 15 seconds before any test activities are performed. The traffic should use the minimum frame size possible on the media used in the testing and rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the recovery time measurement.

Modern traffic generators support the feature of transmitting packets while ignoring the link status. The operator / tester MUST ensure that this feature is enabled.

Third, perform the Line Card (LC) hardware reset at this point. This entails for example physically removing the LC to later re-insert it, or triggering a hardware reset by other means (e.g., command line interface (CLI), physical switch, etc.). However, most accurate results will be obtained using the CLI or a physical switch, and therefore these are RECOMMENDED. Otherwise, the time spend trying to physically seat the LC will get mixed into the results.

Finally, the characterization is completed by measuring the frame loss and recovery time from the moment the LC is reinitialized or reinserted.

Reporting Format

The reset results are reported in a simple statement including the frame loss and recovery times.

For each test case, it is RECOMMENDED that the following parameters be reported in these units:

Parameter	Units or Examples
Throughput	Frames per second and bits per second
Loss	Frames
Time	Seconds, with sufficient resolution to convey meaningful info
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, GigE (Gigabit Ethernet), POS (Packet over SONET), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, HDLC, etc.

Additionally, the DUT and test bed provisioning, configuration, and deployed methodologies that may influence the overall recovery time MUST be listed. (Refer to the additional factors listed in [Section 1](#)).

The reporting of results MUST regard repeatability considerations from [Section 4 of \[RFC2544\]](#). It is RECOMMENDED to perform multiple trials and report average results.

[3.2. Software Reset](#)

To characterize the speed at which a DUT recovers from the software reset.

In contrast to a "hardware reset", a "software reset" involves only the re-initialization of the execution, data structures, and partial state within the software running on the DUT module(s).

A software reset is initiated for example from the DUT's Command Line Interface (CLI).

3.2.1. Operating System (OS) reset (mandatory)

Objective

To characterize the speed at which a DUT recovers from an Operating System (OS) software reset.

Procedure

First, complete some or all of the following operational tasks: save the current DUT configuration, specify software boot parameters, ensure the appropriate software files are available, or perform additional Operating System task.

Second, ensure that the DUT is able to forward the traffic for at least 15 seconds before any test activities are performed. The traffic should use the minimum frame size possible on the media used in the testing and rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the recovery time measurement.

Modern traffic generators support the feature of transmitting packets while ignoring the link status. The operator / tester MUST ensure that this feature is enabled.

Third, trigger an Operating System re-initialization in the DUT, by operational means such as use of the DUT's Command Line Interface (CLI) or other management interface.

Finally, the characterization is completed by measuring the complete frame loss and recovery time from the moment the reset instruction was given until the Operating System finished the reload and re-initialization (inferred by the re-establishing of traffic).

Reporting format

The reset results are reported in a simple statement including the frame loss and recovery times.

For each test case, it is RECOMMENDED that the following parameters be reported in these units:

Parameter	Units or Examples
Throughput	Frames per second and bits per second
Loss	Frames
Time	Seconds, with sufficient resolution to convey meaningful info
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, GigE (Gigabit Ethernet), POS (Packet over SONET), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, HDLC, etc.

Additionally, the DUT and test bed provisioning, configuration, and deployed methodologies that may influence the overall recovery time MUST be listed. (Refer to the additional factors listed in [Section 1](#)).

The reporting of results MUST regard repeatability considerations from [Section 4 of \[RFC2544\]](#). It is RECOMMENDED to perform multiple trials and report average results.

[3.2.2](#). Process reset (optional)

Objective

To characterize the speed at which a DUT recovers from a software process reset.

Such speed may depend upon the number and types of process running in the DUT and which ones are tested. Different implementations of forwarding devices include various common processes. A process reset should be performed only in the processes most relevant to the tester.

Procedure

First, complete some or all of the following operational tasks: save the current DUT configuration, specify software parameters or environmental variables, or perform additional Operating System task.

Second, ensure that the DUT is able to forward the traffic for at least 15 seconds before any test activities are performed. The traffic should use the minimum frame size possible on the media used in the testing and rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the recovery time measurement.

Modern traffic generators support the feature of transmitting packets while ignoring the link status. The operator / tester MUST ensure that this feature is enabled.

Third, trigger a process reset for each process running in the DUT and considered for testing from a management interface (e.g., by means of the Command Line Interface (CLI), etc.)

Finally, the characterization for each individual process is completed by measuring the complete frame loss and recovery time from the moment the reset instruction was given until the Operating System finished the reload and re-initialization (inferred by the re-establishing of traffic).

Reporting format

The reset results are reported in a simple statement including the frame loss and recovery times for each process running in the DUT and tested. Given the implementation nature of this test, details of the actual process tested should be included along with the statement.

For each test case, it is RECOMMENDED that the following parameters be reported in these units:

Parameter	Units or Examples
Throughput	Frames per second and bits per second
Loss	Frames
Time	Seconds, with sufficient resolution to convey meaningful info
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, GigE (Gigabit Ethernet), POS (Packet over SONET), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, HDLC, etc.

Additionally, the DUT and test bed provisioning, configuration, and deployed methodologies that may influence the overall recovery time **MUST** be listed. (Refer to the additional factors listed in [Section 1](#)).

The reporting of results **MUST** regard repeatability considerations from [Section 4 of \[RFC2544\]](#). It is **RECOMMENDED** to perform multiple trials and report average results.

[3.3. Power interruption](#)

"Power interruption" refers to the complete loss of power on the DUT. It can be viewed as a special case of a hardware reset,

triggered by the loss of the power supply to the DUT or its components, and is characterized by the re-initialization of all hardware and software in the DUT.

3.3.1. Power Interruption (mandatory)

Objective

To characterize the speed at which a DUT recovers from a complete loss of electric power or complete power interruption. This test simulates a complete power failure or outage, and should be indicative of the DUT/SUTs behavior during such event.

Procedure

First, ensure that the entire DUT is at a permanent state to which it will return to after the power interruption, by performing some or all of the following operational tasks: save the current DUT configuration, specify boot parameters, ensure the appropriate software files are available, or perform additional Operating System or hardware related task.

Second, ensure that the DUT is able to forward the traffic for at least 15 seconds before any test activities are performed. The traffic should use the minimum frame size possible on the media used in the testing and rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the recovery time measurement.

Modern traffic generators support the feature of transmitting packets while ignoring the link status. The operator / tester MUST ensure that this feature is enabled.

Third, interrupt the power (AC or DC) that feeds the corresponding DUTs power supplies at this point. This entails for example physically removing the power supplies in the DUT to later re-insert them, or simply disconnecting or switching off their power feeds (AC or DC as applicable). The actual power interruption should last at least 15 seconds.

Finally, the characterization is completed by measuring the frame loss and recovery time from the moment the power is restored or the power supplies reinserted in the DUT.

Reporting format

The reset results are reported in a simple statement including the frame loss and recovery times.

For each test case, it is RECOMMENDED that the following parameters be reported in these units:

Parameter	Units or Examples
Throughput	Frames per second and bits per second
Loss	Frames
Time	Seconds, with sufficient resolution to convey meaningful info
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, GigE (Gigabit Ethernet), POS (Packet over SONET), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, HDLC, etc.

Additionally, the DUT and test bed provisioning, configuration, and deployed methodologies that may influence the overall recovery time MUST be listed. (Refer to the additional factors listed in [Section 1](#)).

The reporting of results MUST regard repeatability considerations from [Section 4 of \[RFC2544\]](#). It is RECOMMENDED to perform multiple trials and report average results.

4. Security Considerations

Benchmarking activities, as described in this memo, are limited to technology characterization using controlled stimuli in a laboratory environment, with dedicated address space and the constraints specified in the sections above.

The benchmarking network topology will be an independent test setup and MUST NOT be connected to devices that may forward the test traffic into a production network or misroute traffic to the test management network.

Furthermore, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the DUT/SUT.

Special capabilities SHOULD NOT exist in the DUT/SUT specifically for benchmarking purposes. Any implications for network security arising from the DUT/SUT SHOULD be identical in the lab and in production networks.

There are no specific security considerations within the scope of this document.

5. IANA Considerations

There is no IANA consideration for this document.

6. Acknowledgments

The authors would like to thank Ron Bonica, who motivated us to write this document. The authors would also like to thank Al Morton, Andrew Yourtchenko, David Newman, and John E Dawson for providing thorough review, useful suggestions, and valuable input.

This document was prepared using 2-Word-v2.0.template.dot.

[7. References](#)

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2544] Bradner, S. and McQuaid, J., "Benchmarking Methodology for Network Interconnect Devices", [RFC 2544](#), March 1999.

7.2. Informative References

- [RFC5180] Popoviciu, C., et al, "IPv6 Benchmarking Methodology for Network Interconnect Devices", [RFC 5180](#), May 2008.
- [RFC5695] Akhter, A., Asati, R., and C. Pignataro, "MPLS Forwarding Benchmarking Methodology for IP Flows", [RFC 5695](#), November 2009.

Authors' Addresses

Rajiv Asati
Cisco Systems
7025-6 Kit Creek Road
RTP, NC 27709
USA

Email: rajiva@cisco.com

Carlos Pignataro
Cisco Systems
7200-12 Kit Creek Road
RTP, NC 27709
USA

Email: cpignata@cisco.com

Fernando Calabria
Cisco Systems
7200-12 Kit Creek Road
RTP, NC 27709
USA

Email: fcablabri@cisco.com

Cesar Olvera
Consulintel
Joaquin Turina, 2
Pozuelo de Alarcon, Madrid, E-28224
Spain

Phone: +34 91 151 81 99
Email: cesar.olvera@consulintel.es