Internet Engineering Task Force (IETF)
Request for Comments: 6201
Updates: 1242, 2544
Category: Informational
ISSN: 2070-1721

R. Asati
C. Pignataro
F. Calabria
Cisco
C. Olvera
Consulintel
March 2011

Device Reset Characterization

Abstract

An operational forwarding device may need to be restarted (automatically or manually) for a variety of reasons, an event called 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 resume the forwarding operation.

This document specifies a methodology for characterizing reset (and reset time) during benchmarking of forwarding devices and provides clarity and consistency in reset test procedures beyond what is specified in RFC 2544. Therefore, it updates RFC 2544. This document also defines the benchmarking term "reset time" and, only in this, updates RFC 1242.

Status of This Memo

Table of Contents

1. Introduction 3
   1.1. Scope 3
   1.2. Reset Time 4
   1.3. Reset Time Measurement Methods 5
   1.4. Reporting Format 6
2. Key Words to Reflect Requirements 7
3. Test Requirements 7
4. Reset Tests 8
   4.1. Hardware Reset Tests 9
        4.1.1. Routing Processor (RP) / Routing Engine Reset 9
        4.1.2. Line Card (LC) Removal and Insertion (REQUIRED) 11
   4.2. Software Reset Tests 12
        4.2.1. Operating System (OS) Reset (REQUIRED) 12
        4.2.2. Process Reset (OPTIONAL) 13
   4.3. Power Interruption Test 14
        4.3.1. Power Interruption (REQUIRED) 14
5. Security Considerations 15
6. Acknowledgments 16
7. References 16
   7.1. Normative References 16
   7.2. Informative References 16

1. Introduction

An operational forwarding device (or one of its components) may need to be restarted for a variety of reasons, an event called 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 resume the forwarding operation. In other words, the duration of the recovery time following the reset (see Section 1.2, "Reset Time") is what is in question.

However, the answer to this question is no longer simple and straightforward 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.

1.1. Scope

This document specifies a methodology for characterizing reset (and reset time) during benchmarking of forwarding devices and provides clarity and consistency in reset procedures beyond what is specified in [RFC2544]. Software upgrades involve additional benchmarking complexities and are outside the scope of this document.

These procedures may be used by other benchmarking documents such as [RFC2544], [RFC5180], [RFC5695], etc., and it is expected that other protocol-specific benchmarking documents will reference this document for reset recovery time characterization. Specific Routing Information Base (RIB) and Forwarding Information Base (FIB) scaling considerations are outside the scope of this document and can be quite complex to characterize. However, other documents can characterize specific dynamic protocols’ scaling and interactions as well as leverage and augment the reset tests defined in this document.

This document updates Section 26.6 of [RFC2544] and defines the benchmarking term "reset time", updating [RFC1242].

This document focuses only on the reset criterion of benchmarking and presumes that it would be beneficial to [RFC5180], [RFC5695], and other IETF Benchmarking Methodology Working Group (BMWG) efforts.

1.2. Reset Time

Definition

Reset time is the total time that a device is determined to be out of operation and includes the time to perform the reset and the time to recover from it.

Discussion

During a period of time after a reset or power up, network devices may not accept and forward frames. The duration of this period of forwarding unavailability can be useful in evaluating devices. In addition, some network devices require some form of reset when specific setup variables are modified. If the reset period were long, it might discourage network managers from modifying these variables on production networks.

The events characterized in this document are entire reset events. That is, the recovery period measured includes the time to perform the reset and the time to recover from it. Some reset events will be atomic (such as pressing a reset button) while others (such as power cycling) may comprise multiple actions with a recognized interval between them. In both cases, the duration considered is from the start of the event until full recovery of forwarding after the completion of the reset events.

Measurement Units

Time, in milliseconds, providing sufficient resolution to distinguish between different trials and different implementations. See Section 1.4.

Issues

There are various types of resets: hardware resets, software resets, and power interruptions. See Section 4.

See Also

This definition updates [RFC1242].

1.3. Reset Time Measurement Methods

The reset time is the time during which traffic forwarding is temporarily interrupted following a reset event. Strictly speaking, this is the time over which one or more frames are lost. This definition is similar to that of "Loss of Connectivity Period" defined in [IGPConv], Section 4.

There are two accepted methods to measure the reset time:

1. Frame-Loss Method - This method requires test tool capability to monitor the number of lost frames. In this method, the offered stream rate (frames per second) must be known. The reset time is calculated per the equation below:

$$Reset\_time = \frac{Frames\_lost\ (packets)}{Offered\_rate\ (packets\ per\ second)}$$

2. Timestamp Method - This method requires test tool capability to timestamp each frame. In this method, the test tool timestamps each transmitted frame and monitors the received frame’s timestamp. During the test, the test tool records the timestamp (Timestamp A) of the frame that was last received prior to the reset interruption and the timestamp of the first frame after the interruption stopped (Timestamp B). The difference between Timestamp B and Timestamp A is the reset time.

The tester/operator MAY use either method for reset time measurement depending on the test tool capability. However, the Frame-Loss method SHOULD be used if the test tool is capable of (a) counting the number of lost frames per stream and (b) transmitting test frame despite the physical link status, whereas the Timestamp method SHOULD be used if the test tool is capable of (a) timestamping each frame, (b) monitoring received frame’s timestamp, and (c) transmitting frames only if the physical link status is UP. That is, specific test tool capabilities may dictate which method to use. If the test tool supports both methods based on its capabilities, the tester/operator SHOULD use the one that provides more accuracy.

1.4. Reporting Format

All reset results are reported in a simple statement including the frame loss (if measured) and reset 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 (average)	Frames
Reset Time (average)	Milliseconds
Number of trials	Integer count
Protocol	IPv4, IPv6, MPLS, etc.
Frame Size	Octets
Port Media	Ethernet, Gigabit Ethernet (GbE), Packet over SONET (POS), etc.
Port Speed	10 Gbps, 1 Gbps, 100 Mbps, etc.
Interface Encap.	Ethernet, Ethernet VLAN, PPP, High-Level Data Link Control (HDLC), etc.

For mixed protocol environments, frames SHOULD be distributed between all the different protocols. The distribution MAY approximate the network conditions of deployment. In all cases, the details of the mixed protocol distribution MUST be included in the reporting.

Additionally, the DUT (Device Under Test) or SUT (System Under Test) and test bed provisioning, port and line-card arrangement, configuration, and deployed methodologies that may influence the overall reset time MUST be listed. (Refer to the additional factors listed in Section 3).

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

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. Test Requirements

Tests SHOULD first be performed such that the forwarding state re-establishment is independent from an external source (i.e., using static address resolution, routing and forwarding configuration, and not dynamic protocols). However, tests MAY subsequently be performed using dynamic protocols that the forwarding state depends on (e.g., dynamic Interior Gateway Protocols (IGP), Address Resolution Protocol (ARP), PPP Control Protocols, etc.). The considerations in this section apply.

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

Some of these factors include the following:

1. Type of reset - hardware (line-card crash, etc.) versus software (protocol reset, process crash, etc.) or even complete power failures

2. Manual versus automatic reset

3. Scheduled versus non-scheduled reset

4. Local versus remote reset

5. Scale - Number of line cards present versus 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 throughput [RFC2544]

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

The reset time is one of the key characterization results reported after each test run. While the reset time during a reset test event may be zero, there may still be effects on traffic, such as transient delay variation or increased latency. However, that is not covered and is deemed outside the scope of this document. In this case, only "no loss" is reported.

4. Reset Tests

This section contains descriptions of the tests that are related to the characterization of the time needed for DUTs (Devices Under Test) or SUTs (Systems Under Test) to recover from a reset. There are three types of resets considered in this document:

1. Hardware resets

2. Software resets

3. Power interruption

Different types of resets potentially have a 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 4.1 describes various hardware resets, whereas Section 4.2 describes various software resets. Additionally, Section 4.3 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 "REQUIRED" or "OPTIONAL".

4.1. Hardware Reset Tests

A hardware reset test is a test designed to characterize the time it takes a DUT to recover from a 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 hardware component, by pressing a reset button for the component, or by being triggered from the command line interface (CLI).

Reset procedures that do not require the physical removal and insertion of a hardware component are RECOMMENDED. These include using the command line interface (CLI) or a physical switch or button. If such procedures cannot be performed (e.g., because of a lack of platform support or because the corresponding test case calls for them), human operation time SHOULD be minimized across different platforms and test cases as much as possible, and variation in human operator time SHOULD also be minimized across different vendors' products as much as practical by having the same person perform the operation and by practicing the operation. Additionally, the time between removal and insertion SHOULD be recorded and reported.

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 4.3) in these cases.

4.1.1. Routing Processor (RP) / Routing Engine Reset

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

4.1.1.1. RP Reset for a Single-RP Device (REQUIRED)

Objective

To characterize the time needed for a DUT to recover 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 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 tasks.

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 the rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the reset time measurement.

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, complete the characterization by recording the frame loss or timestamps (as reported by the test tool) and calculating the reset time (as defined in Section 1.3).

Reporting Format

The reporting format is defined in Section 1.4.

4.1.1.2. RP Switchover for a Multiple-RP Device (OPTIONAL)

Objective

To characterize the time needed for the "secondary" Route Processor (sometimes referred to as the "backup" RP) of a DUT to become 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 and, if it exists, for the 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 characteristic's standpoint: first, a default behavior (with no switchover-specific configurations) and, potentially 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 RP will be activated to the same state as the active RP 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 tasks.

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 the rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the reset time measurement.

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.). It is up to the operator to decide whether or not the primary RP needs to be re-inserted after a grace period.

Finally, complete the characterization by recording the frame loss or timestamps (as reported by the test tool) and calculating the reset time (as defined in Section 1.3).

Reporting Format

The reset results are potentially reported twice, one for the default switchover behavior (i.e., the DUT without any switchover-specific enhanced configuration) and the other for the switchover-specific behavior if it exists (i.e., the DUT configured for optimized switchover capabilities that minimize the downtime during the actual switchover).

The reporting format is defined in Section 1.4 and also includes any specific redundancy scheme in place.

4.1.2. Line Card (LC) Removal and Insertion (REQUIRED)

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

Objective

To characterize the time needed for a DUT to recover 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 tasks.

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 the rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the reset time measurement.

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., CLI, physical switch, etc.).

Finally, complete the characterization by recording the frame loss or timestamps (as reported by the test tool) and calculating the reset time (as defined in Section 1.3).

Reporting Format

The reporting format is defined in Section 1.4.

4.2. Software Reset Tests

A software reset test characterizes the time needed for a DUT to recover from a 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 CLI.

4.2.1. Operating System (OS) Reset (REQUIRED)

Objective

To characterize the time needed for a DUT to recover 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 tasks.

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 the rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the reset time measurement.

Third, trigger an operating system re-initialization in the DUT by operational means such as use of the DUT's CLI or other management interface.

Finally, complete the characterization by recording the frame loss or timestamps (as reported by the test tool) and calculating the reset time (as defined in Section 1.3).

Reporting Format

The reporting format is defined in Section 1.4.

4.2.2. Process Reset (OPTIONAL)

Objective

To characterize the time needed for a DUT to recover from a software process reset.

Such a time period may depend upon the number and types of processes 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 and most impactful to forwarding.

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 tasks.

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 the rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the reset time measurement.

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 CLI, etc.).

Finally, complete the characterization by recording the frame loss or timestamps (as reported by the test tool) and calculating the reset time (as defined in Section 1.3).

Reporting Format

The reporting format is defined in Section 1.4 and is used 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.

4.3. Power Interruption Test

"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.

4.3.1. Power Interruption (REQUIRED)

Objective

To characterize the time needed for a DUT to recover 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/SUT’s behavior during such event.

Procedure

First, ensure that the entire DUT is at a permanent state to which it will return 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 tasks.

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 the rate should be sufficient for the DUT to attain the maximum forwarding throughput. This enables a finer granularity in the reset time measurement.

Third, interrupt the power (AC or DC) that feeds the corresponding DUT's 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, complete the characterization by recording the frame loss or timestamps (as reported by the test tool) and calculating the reset time (as defined in Section 1.3).

For easier comparison with other testing, 15 seconds are removed from the reported reset time.

Reporting Format

The reporting format is defined in Section 1.4.

5. Security Considerations

Benchmarking activities, as described in this document, 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 externally 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.

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, John E. Dawson, Timmons C. Player, Jan Novak, Steve Maxwell, Ilya Varlashkin, and Sarah Banks for providing thorough review, useful suggestions, and valuable input.

7. References

7.1. Normative References

[RFC1242] Bradner, S., "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, July 1991.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999.

7.2. Informative References

[IGPConv] Poretsky, S., Imhoff, B., and K. Michielsen, "Benchmarking Methodology for Link-State IGP Data Plane Route Convergence", Work in Progress, February 2011.

[RFC5180] Popoviciu, C., Hamza, A., Van de Velde, G., and D. Dugatkin, "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
Research Triangle Park, NC 27709
USA

EMail: rajiva@cisco.com

Carlos Pignataro
Cisco Systems
7200-12 Kit Creek Road
Research Triangle Park, NC 27709
USA

EMail: cpignata@cisco.com

Fernando Calabria
Cisco Systems
7200-12 Kit Creek Road
Research Triangle Park, NC 27709
USA

EMail: fcalabri@cisco.com

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

EMail: cesar.olvera@consulintel.es