Benchmarking Methodology for EVPN VPWS
draft-kishjac-bmwg-evpnvpwstest-03

Abstract

This document defines methodologies for benchmarking EVPN-VPWS performance. EVPN-VPWS is defined in RFC 8214, and is being deployed in Service Provider networks. Specifically, this document defines the methodologies for benchmarking EVPN-VPWS scale convergence, failover, core isolation, high availability, and longevity.

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). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on June 14, 2020.

Copyright Notice

Copyright (c) 2019 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 (https://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 .................................................. 2
   1.1.  Requirements Language .................................... 2
   1.2.  Terminologies ............................................. 3
2.  Test Topology .................................................. 4
3.  Test Cases ..................................................... 6
   3.1.  Local Failure ............................................. 6
   3.2.  Fail over Test in Remote PE ................................ 7
   3.3.  Core Failure .............................................. 7
   3.4.  Link Flap .................................................. 8
4.  Activate/deactivate AC’s ....................................... 9
   4.1.  Deactivate/Activate M number of attachment circuits ...... 9
5.  Scale Convergence ............................................... 9
   5.1.  To measure the packet loss during the core link failure .. 9
6.  High Availability ............................................... 10
   6.1.  To Record the whether there is traffic loss due to routing engine failover for redundancy test. .......... 10
7.  SOAK Test ..................................................... 11
   7.1.  To Measure the stability of the DUT with scale and traffic. ........................................... 11
8.  Acknowledgments ................................................. 11
9.  IANA Considerations ........................................... 12
10. Security Considerations ......................................... 12
11. References ..................................................... 12
    11.1. Normative References ..................................... 12
    11.2. Informative References ................................... 12
Appendix A. Appendix ............................................ 13
Authors’ Addresses ............................................ 13

1. Introduction

EVPN-VPWS is defined in RFC 8214, discusses how VPWS can be combined with EVPNs to provide a new/combined solution. This draft defines methodologies that can be used to benchmark RFC 8214 solutions. Further, this draft provides methodologies for benchmarking the performance of EVPN VPWS Scale, Scale Convergence, Core isolation, longevity, high availability.

1.1. Requirements Language

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 RFC 2119 [RFC2119].
1.2. Terminologies

All-Active Redundancy Mode: When all PEs attached to an Ethernet segment are allowed to forward known unicast traffic to/from that Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in All-Active redundancy mode.

AA All Active mode

AC Attachment Circuits

CE Customer Router/Devices/Switch.

DF Designated Forwarder

DUT Device under test.

Ethernet Segment (ES): When a customer site (device or network) is connected to one or more PEs via a set of Ethernet links, then that set of links is referred to as an ‘Ethernet segment’.

EVI: An EVPN instance spanning the Provider Edge (PE) devices participating in that EVPN.

Ethernet Segment Identifier (ESI): A unique non-zero identifier that identifies an Ethernet segment is called an ‘Ethernet Segment Identifier’.

Ethernet Tag: An Ethernet tag identifies a particular broadcast domain, e.g., a VLAN. An EVPN instance consists of one or more broadcast domains.

Interface Physical interface of a router/switch.

IRB Integrated routing and bridging interface

MAC Media Access Control addresses on a PE.

MHPE2 Multi homed Provider Edge router 2.

MHPE1 Multi homed Provider Edge router 1.

SHPE3 Single homed Provider Edge Router 3.

PE: Provider Edge device.

P Provider Router.
RR Route Reflector.

RT Traffic Generator.

Sub Interface Each physical Interfaces is subdivided into Logical units.

SA Single Active

Single-Active Redundancy Mode: When only a single PE, among all the PEs attached to an Ethernet segment, is allowed to forward traffic to/from that Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in Single-Active redundancy mode.

2. Test Topology

EVPN-VPWS Services running on SHPE3, MHPE1 and MHPE2 in Single Active Mode:
Traffic Generator acts as a sender/receiver of layer 2 traffic with multiple vlan.

There are five routers in the topology. SHPE3, RR/P, MHPE1 and MHPE2 emulating a service provider network. CE is a customer device connected to MHPE1 and MHPE2, it is configured with bridge domains in different vlans. The traffic generator is connected to CE and SHPE3. The MHPE1 acts as DUT. The traffic generator will act as sender and receiver. The measurement will be taken in DUT.
All routers except CE is configured with OSPF/IS-IS, LDP, MPLS, BGP with EVPN address family.

All routers except CE must have Interior Border Gateway protocol configured, RR acting as route reflector.

MHPE1, MHPE2, SHPE3 must be configured with "N" EVPN-VPWS instances depends up on the cases.

MHPE1 and MHEPE2 must be configured with ESI per vlan or ESI on Interface.

MHPE1 and MHEPE2 are running Single Active mode of EVPN-VPWS.

CE is acting as bridge configured with vlans that is configured on MHPE1, MHPE2, SHPE3.

Depends up on the test traffic will be flowing uni directional or bi directional depends on the topology mentioned above.

The above configuration will serve as base configuration for all the test cases.

3. Test Cases

The following tests are conducted to measure the packet loss during the local link and core failure in DUT with Scaled AC’s.

3.1. Local Failure

Objective:

To Record the time taken to switch from primary to backup during local link failure.

Topology: Topology 1

Procedure:

Configure "N" AC’s in SHPE3 and MHPE1, MHPE2, working in SA mode. Ensure MHPE2 is active and DUT is backup PE. Send "X" unicast packets to CE from traffic generator to MHPE2 AC’s working in SA. Then shut the MHPE2-CE link, so that traffic from CE switches to DUT.

Measurement:
Measure the time taken to switch the traffic from active to backup, the traffic will flow from MHPE1 to SHPE3. Measure the time taken to switch the traffic.

The test is repeated for "N" times and the values are collected. The switching time is calculated by averaging the values obtained from "N" samples.

AC’s switch over from primary to backup PE in sec = (T1+T2+..Tn/N)

3.2. Fail over Test in Remote PE

Objective:

To Record the time taken by remote PE to switch traffic from primary to backup during CE link failure.

Topology: Topology 1

Procedure:

Configure "N" AC’s in SHPE3 and MHPE1,MHPE2, working in SA mode.Ensure MHPE2 is active and DUT is backup PE.Send "X" unicast packets from traffic generator to SHPE3 Ac’s.Then shut the MHPE2-CE link, this failure will be notified to remote PE and traffic switch to backup path.

Measurement:

Measure the time taken to switch the traffic from active to backup, the traffic will flow from SHPE3 to MHPE1. Measure the time taken to switch the traffic.

The test is repeated for "N" times and the values are collected. The switching time is calculated by averaging the values obtained from "N" samples.

AC’s switch over from primary to backup PE in sec = (T1+T2+..Tn/N)

3.3. Core Failure

Objective:

To Record the time taken by remote PE to switch traffic from primary to backup during core link failure.

Topology: Topology 1
Procedure:

Configure "N" AC’s in SHPE3 and MHPE1, MHPE2, working in SA mode. Ensure MHPE2 is active and DUT is backup PE. Send "X" unicast packets from traffic generator to SHPE3 Ac’s. Then shut the core link of MHPE2, this failure will be notified to remote PE and traffic switch to backup path.

Measurement:

Measure the time taken to switch the traffic from active to backup, the traffic will flow from SHPE3 to MHPE1. Measure the time taken to switch the traffic.

The test is repeated for "N" times and the values are collected. The switching time is calculated by averaging the values obtained from "N" samples.

AC’s in remote PE switches from primary to backup PE in sec due to core failure = (T1+T2+..Tn/N)

3.4. Link Flap

Objective:

To record the time taken by primary PE to regain control after the local PE-CE link flap.

Topology: Topology 1

Procedure:

Configure "N" AC’s in SHPE3 and MHPE1, MHPE2, working in SA mode. Ensure MHPE2 is standby and DUT is primary PE. Send "X" unicast packets from traffic generator connected to CE to all AC’s in MHPE1 (DUT). Then shut the link of MHPE1-CE, this failure will be notified to remote PE and traffic switch to backup path. Then bring up the link of MHPE1-CE. Now the traffic switches to DUT.

Measurement:

Measure the time taken to switch the traffic from MHPE2 to DUT, the traffic will flow from MHPE1 to SHPE3. Measure the time taken to switch the traffic.
The test is repeated for "N" times and the values are collected. The switching time is calculated by averaging the values obtained from "N" samples.

Time taken to switch back to primary (DUT) once the link is restored = (T1+T2+.Tn/N)

4. Activate/deactivate AC’s

4.1. Deactivate/Activate M number of attachment circuits.

Objective:

To measure the performance of the DUT while deactivating/activating AC’s.

Topology: Topology 1

Procedure:

Configure "N" AC’s in SHPE3 and MHPE1, MHPE2, working in SA mode. Ensure MHPE2 is active and DUT is backup PE. Send "X" unicast packets from traffic generator to SHPE3 to all AC’s and send "X" unicast packets from CE to MHPE1 (DUT), let the DUT be the active and the MHPE2 must be standby. DUT will be forwarding the traffic to CE and from CE to SHPE3. Then deactivate "M" AC’s on SHPE1, DUT and MHPE2 on the fly. These AC’s must be removed from forwarding plane. Stop the traffic for these AC’s. Activate the AC’s in all PE’s. Then start the traffic, measure the time taken by "M" AC’s to forward the traffic.

Measurement:

Measure the packet loss in sec during this deactivating/activating AC’s. Repeat the test "N" times and the packet loss is calculated by averaging the values obtained from "N" samples.

packet loss in sec = (T1+T2+.Tn/N)

5. Scale Convergence

5.1. To measure the packet loss during the core link failure.

Objective:

To measure the convergence at a higher number of AC’s
Topology: Topology 1

Procedure:

Configure "N" AC's in SHPE3 and MHPE1, MHPE2, working in SA mode. DF election must be priority based not on the default RFC 7432, it should not be MOD based DF election. Send "X" unicast packets from traffic generator to SHPE3 to all Ac's and send "X" unicast packets from CE to MHPE1 (DUT), let the DUT is the active and the MHPE2 must be standby. DUT will be forwarding the traffic to CE from SHPE3 and from CE to SHPE3. Then flap the core link of the DUT.

Measurement:

Measure the packet loss in seconds once the core link is restored. Repeat the test "N" times and the packet loss is calculated by averaging the values obtained from "N" samples.

Packet loss in sec = (T1+T2+..Tn/N)

6. High Availability

6.1. To Record the whether there is traffic loss due to routing engine fail over for redundancy test.

Objective:

To record traffic loss during routing engine failover.

Topology: Topology 1

Procedure:

Configure "N" AC’s in SHPE3 and MHPE1, MHPE2, working in SA mode. Ensure MHPE2 is active and DUT is backup PE. Send "X" unicast packets from traffic generator to SHPE3 to all Ac’s and send "X" unicast packets from CE to MHPE1 (DUT), let the DUT is the active and the MHPE2 must be standby. DUT will be forwarding the traffic to CE and from CE to SHPE3. Then do a routing engine fail-over.

Measurement:

The expectation of the test is 0 traffic loss with no change in the DF role. DUT should not withdraw any routes. But in cases where the DUT is not property synchronized between master and standby, due to that packet loss are observed. In that scenario the packet loss is...
measured. The test is repeated for "N" times and the values are collected. The packet loss is calculated by averaging the values obtained by "N" samples.

Packet loss in sec = (T1+T2+..Tn/N)

7. SOAK Test

This test is carried out to measure the stability of the DUT in a scaled environment with traffic over a period of time "T’". In each interval "t1" the DUT CPU usage, memory usage are measured. The DUT is checked for any crashes during this time period.

7.1. To Measure the stability of the DUT with scale and traffic.

Objective:

To measure the stability of the DUT in a scaled environment with traffic.

Topology : Topology 1

Procedure:

Scale N AC’s in DUT, SHPE3 and MHPE2. Send F frames to DUT from CE using traffic generator with different X SA and DA for N EVI’s. Send F frames from traffic generator to SHPE3 with X different SA and DA. There is a bi-directional traffic flow with F pps in each direction. The DUT must run with traffic for 24 hours, every hour check for memory leak, crash.

Measurement :

Take the hourly reading of CPU, process memory. There should not be any leak, crashes, CPU spikes. Th CPU spike is determined as the CPU usage which shoots at 40 to 50 percent of the average usage. The average value vary from device to device. Memory leak is determined by increase usage of the memory for EVPN-VPWS process. The expectation is under steady state the memory usage for EVPN-VPWS process should not increase.

8. Acknowledgments

We would like to thank Al and Sarah for the support.
9. IANA Considerations

This memo includes no request to IANA.

10. Security Considerations

The benchmarking tests described in this document are limited to the performance characterization of controllers in a lab environment with isolated networks. 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. Further, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the controller. Special capabilities SHOULD NOT exist in the controller specifically for benchmarking purposes. Any implications for network security arising from the controller SHOULD be identical in the lab and in production networks.

11. References

11.1. Normative References


11.2. Informative References


Appendix A. Appendix

Authors’ Addresses

Sudhin Jacob (editor)
Juniper Networks
Bangalore
India

Phone: +91 8061212543
Email: sjacob@juniper.net

Kishore Tiruveedhula
Juniper Networks
10 Technology Park Dr
Westford, MA 01886
USA

Phone: +1 9785898861
Email: kishoret@juniper.net