

Description

Until EOS release 4.32.0F, EOS allows users to statically configure link min-delay and maxdelay used for IS-IS FlexAlgo. This feature adds support for dynamic measurement of link delay using the TWAMP Light protocol described in RFC 8186 and provides it to IS-IS FlexAlgo dynamically.

This document describes how to configure and monitor this feature.

There are 2 modes in TWAMP Light:

- 1. Active/Sender where the device initiates sending probe packets and receives reflected packets with a timestamp to compute link-delay
- 2. Reflector where the device receives a probe packet, adds timestamps, and sends it back

These two modes can operate independently of each other, and the same interface may be in Active mode as well as in Reflector mode at the same time. It is up to the user to configure different UDP ports for these two modes to avoid a conflict.

IS-IS advertises the unidirectional link delay (average delay), unidirectional min-max delay and unidirectional delay variation measured by TWAMP Light, as described in RFC 8570. Only unidirectional min-max delay gets advertised in ASLA (Application Specific Link Attributes Sub-TLV).

This feature only supports IPv4 probe packets but the measured link-delay values are utilized for both IPv4 and IPv6 topologies in IS-IS multi-topology deployments.

Platform compatibility

All 7500R3<X>, 7800R3<X>, DCS-7280<X>R3<X> except on systems where at least one FAP has "Revision" as "Q2c2T<X>".
 Run the below command to check the FAP "Revision".

Arista#show	platform fap		
Switches cu	rrently in the	system	
Name	Туре	Device	Revision
 Fap0	Jericho2	Jericho2c	Q2c2tA1



Feature History

Release	Update
4.32.1	Initial introduction
4.33.0	Dynamic min delay with per interface profile support.

Configuration

In order to use min-delay and max-delay in IS-IS FlexAlgo the user has to configure a Twamp sender profile and choose this profile in the traffic-engineering configuration.

In addition to the basic IS-IS related configurations, the user has to:

- Configure a flexible algorithm that will use the min-delay metrics.
- Bind this flexible algorithm to IS-IS.
- Enable IS-IS and traffic engineering on the interfaces.

Twamp configuration

To initiate TWAMP reflection, the user has to enter "monitor twamp" / "twamp-light" configuration mode.

Example:

```
switch(config)#monitor twamp
switch(config-monitor-twamp)#twamp-light
switch(config-twamp-light)#
```

To configure the system to actively send TWAMP Light probes, a sender profile must be created. This is done by using the "sender profile <NAME>" command.

Example:

```
switch(config-twamp-light)#sender profile profile1
switch(config-twamp-light-sender-profile-profile1)#
```

A sender profile has several parameters that can be set:



- measurement interval the interval window within which the sender sends probe packets (in seconds)
- measurement samples the number of packets sent during each interval window
- significance and offset variables used to calculate normalized min-delay (in microseconds)

The above values are optional and if not set, the default values are::

measurement interval: 60 seconds

measurement samples: 60significance: 10 microseconds

offset: 2 microseconds

Example:

```
switch(config-twamp-light-sender-profile-profile1)#measurement interval 10 seconds
switch(config-twamp-light-sender-profile-profile1)#measurement samples 5
switch(config-twamp-light-sender-profile-
profile1)#significance 10 microseconds offset 5 microseconds
```

Note: Maximum effective sampling rate is 1 packet per second. Additionally, the significance value has to be larger than the offset value.

If D_m is the observed min-delay then the normalized min-delay D_n is computed as follows: $D_n = \text{ceiling}(D_m, \text{significance}, \text{offset})$.

The ceiling function will compute the ceiling values by adding the offset value to multiples of the significance value starting from a multiple of 0. For example, if significance is specified as 10 and offset as 3 then the ceiling values computed will be 3, 13, 23, 33 etc. The measured delay value will be rounded up to the nearest ceiling value greater than the measured delay as follows:

- If measured delay value is 2 then it will be rounded up to 3
- If measured delay value is 4 then it will rounded up to 13
- If measured delay value is 25 then it will be rounded up to 33

The user can also configure which UDP ports will be used for sending and reflecting TWAMP probes. The following three port values can be globally configured:

- listen port used by the reflector to listen for incoming packets
- source port used by the sender as the source port in probe packets
- destination port used by the sender as the destination port in probe packets



Configuring the ports is optional and if not set, the default values are:

listen port: 51201source port: 51200destination port: 51201

Example:

```
switch(config-twamp-light)#reflector defaults
switch(config-twamp-light-reflector-defaults)#listen port <PORT>
switch(config-twamp-light)#sender defaults
switch(config-twamp-light-sender-defaults)#destination port <PORT>
switch(config-twamp-light-sender-defaults)#source port <PORT>
```

Traffic engineering configuration

The following commands are used to configure the IGP (IS-IS) to use the dynamic unidirectional delay of the point-to-point links, as measured using the TWAMP Light sender profile.

A sender profile can be activated for traffic engineering either globally or per interface.

Globally:

```
switch(config)#router traffic-engineering
switch(config-te)#twamp-light sender profile <PROFILE NAME>
```

Per interface:

```
switch(config)#interface Ethernet 3
switch(config-if-Et3)#traffic-engineering twamp-light sender profile <PROFILE NAME>
```

PROFILE NAME has to be one of the sender profiles configured under the "monitor twamp" / "twamp-light" mode.

If a sender profile name is not configured on an interface, the global profile configured under router traffic engineering will be used by TWAMP-Light to measure the dynamic unidirectional delay metrics of the link.



To activate TWAMP measurements and to use the measured values for traffic engineering for an interface the following command is configured on the interface:

```
switch(config-if-Et3)#traffic-engineering min-delay dynamic twamp-light fallback 

<VALUE> microseconds/milliseconds
```

Note that a fallback value also has to be configured on all the IS-IS point-to-point interfaces that use dynamic delay measurement. IS-IS will advertise this fallback value if TWAMP Light fails to compute the dynamic unidirectional min-delay metric of a link.

Show commands

Twamp show commands

Show command that displays calculated one way and two way min-delay. By default it displays information for all peer IP addresses configured on the remote end of the IS-IS point-to-point links. The output can be filtered for a specified IP address.

```
SYNTAX: show monitor twamp-light [ip IP]
SAMPLE USAGE:
switch#show monitor twamp-light
Unit: microseconds
                             One Way
IP Address Interface
                                           Two Way
                        Minimum Delay Minimum Delay
1.0.0.2
          Ethernet17/1
                                 2742
                                               5482
1.0.0.3 Ethernet18/1
                            2772
                                              5544
switch#show monitor twamp-light ip 1.0.0.2
Unit: microseconds
IP Address Interface
                             One Way
                                           Two Way
                        Minimum Delay Minimum Delay
1.0.0.2
          Ethernet17/1
                                 2742
                                               5482
```

Show command that displays detailed information about calculated results. By default it displays information for all peer IP addresses configured on the remote end of the IS-IS point-to-point links. The output can be filtered for a specified IP address.



Show command that displays packet counters. By default it displays sender, reflector and drop packet counters. Drop counters can use the 'detail' keyword which will display counters for each drop reason.



```
Ethernet17/1 0

switch# show monitor twamp-light counters drop detail

Interface: Ethernet17/1

Invalid TWAMP packet: 0

Packet too small: 0

Wrong dst port: 0

Internal error: 0

Wrong EtherType: 0

Unknown IP address: 0

Incorrect timestamps: 0
```

Show command that displays probe packet information that includes the packet's RX time, its sequence number and the delay calculated based on this packet. The output can be filtered by either an IPv4 address or an Interface. By default it displays statistics for all received probe packets but it can be limited to a specified count of most recently received probe packets with the 'count' keyword.

```
SYNTAX: show monitor twamp-light (( ip IP | interface INTF)) raw [count NUM]
SAMPLE USAGE:
switch# show monitor twamp-light ip 1.0.0.2 raw count 5
Address: 1.0.0.2
Last RX interface: Ethernet17/1
Two-way delays (in microseconds):
Rx time
                                Segnum
                                              Delay
2024-05-15 01:44:45.814785
                                     19
                                             10120.5
2024-05-15 01:44:44.824825
                                             21474.2
                                    18
2024-05-15 01:44:43.828435
                                            25934.5
                                    17
2024-05-15 01:44:42.828052
                                    16
                                             26654.2
2024-05-15 01:44:41.811446
                                             11387.4
                                    15
```

Additionally, there are two clear commands that can clear the statistics maintained by the system.

Clear packet counters. By default it clears all the gathered packet counters. The command can take an optional IP address to only clear statistics for the specified peer.

```
SYNTAX: clear monitor twamp-light counters [reflector | sender] [ip IP]
```



Clear TWAMP Light calculated results. By default this command clears all the calculated results. The command can take an optional IP address to only clear statistics for the specified peer.

```
SYNTAX: clear monitor twamp-light [ip IP]
```

Traffic engineering show commands

No new show commands are introduced for this feature. Existing show commands outputs have been updated to display dynamic delay related information.

The show command that displays the traffic engineering parameters of an interface has been updated to display if the min-delay of the interface is dynamic or static. For dynamic min-delay it shows additional information if the value is measured one or fallback.

```
SYNTAX: show traffic-engineering interfaces [ Interface Identity ]

SAMPLE USAGE:
switch# show traffic-engineering interfaces
Interface Ethernet5:
Traffic engineering metric: 7
Administrative groups: 0
Minimum delay (dynamic): 50 microseconds (measured)

switch# show traffic-engineering interfaces Ethernet4
Interface Ethernet4:
Traffic engineering metric: 6
Administrative groups: 0
Minimum delay (dynamic): 500 microseconds (fallback)
```

The command that displays the IS-IS database for traffic engineering will show the dynamic delay metrics of the neighbors.

```
SYNTAX: show isis database traffic-engineering [ level/VRF ] [LSP id ]

SAMPLE USAGE:
switch# show isis database traffic-engineering
Legend:
H - hostname conflict
U - node unreachable

IS-IS Instance: 1 VRF: default
IS-IS Level 2 Link State Database
```



```
Seg Num Cksum Life Length IS Received LSPID
                                                                                  Fla
   LSPID
gs
                                                    197 L2 1111.1111.1001.00-00
   smd418.00-00
                                142
                                      4625
                                            1178
                                                                                  <>
     LSP generation remaining wait time: 0 ms
     Time remaining until refresh: 878 s
     NLPID: 0xCC(IPv4)
     Hostname: smd418
     TE IPv4 router ID: 1.4.5.6
     Area addresses: 49.0001
     Interface address: 1.0.1.1
     Interface address: 1.0.0.1
     IS Neighbor
                          : 1111.1111.1002.00 Metric: 10
       IPv4 Neighbor Address: 1.0.0.2
       IPv4 Interface Address: 1.0.0.1
       Maximum link BW: 100.00 Gbps
       Maximum reservable link BW: 0.00 bps
       Unreserved BW:
           TE class 0: 0.00 bps
                                       TE class 1: 0.00 bps
                                                                TE class 2: 0.00 bps
           TE class 3: 0.00 bps
                                       TE class 4: 0.00 bps
                                                                TE class 5: 0.00 bps
           TE class 6: 0.00 bps
                                        TE class 7: 0.00 bps
       Min/Max unidirectional link delay: 542/642 us
       Average unidirectional link delay: 612 us
       Unidirectional link delay variation: 3342 us
       Application Specific Link Attributes:
         Standard applications: Flex-Algo
           Min/Max unidirectional link delay: 542/642 us
     Reachability
                           : 1.0.1.0/24 Metric: 10 Type: 1 Up
     Reachability
                           : 1.0.0.0/24 Metric: 10 Type: 1 Up
     Router Capabilities: Router Id: 1.0.0.1 Flags: []
       Area leader priority: 250 algorithm: 0
```

Syslog messages

There are no new syslog messages.

Troubleshooting

1. Verify if min-delay has been calculated:

```
switch(config)#show monitor twamp-light
Unit: microseconds
```



```
IP Address Interface One Way Two Way
                      Minimum Delay Minimum Delay
                     _____
1.0.0.2
         Ethernet29/1
                               2542
                                            5092
switch(config)#show monitor twamp-light detail
IP address: 1.0.0.2
Interface: Ethernet29/1
Description:
Sender profile name: defaultSender
Effective sampling rate: 1 packet every 1.0 second
Packets sent: 558
Packets received: 21
Unit: microseconds
Measurement Min Max Avg Variance
_____ ____
One-way delay 2542 24382 6872 46609812
Two-way delay 5092 48762 13742 93219622
```

2. Verify that the sender profile exists and all default ports are set correctly:

```
switch(config-twamp-light)#show active
monitor twamp

twamp-light

reflector defaults

listen port 51202

!

sender defaults

destination port 51202

source port 51201
!

sender profile defaultSender

measurement interval 20 seconds

measurement samples 20
```

3. Verify that the packets are being sent

```
switch(config)#show monitor twamp-light ip 1.0.0.2 raw count 5
Address: 1.0.0.2
Last RX interface: Ethernet29/1
Two-way delays (in microseconds):
```



Rx time	Seqnum	Delay
2024-05-22 01:42:45.694000	0	29247133.4
2024-05-22 01:42:35.475005	19	11827.7
2024-05-22 01:42:34.483524	18	21125.3
2024-05-22 01:42:33.481068	17	20094.4
2024-05-22 01:42:32.482278	16	21952.6

4. Verify that tcpdump captures packets

```
switch#tcpdump interface <INTF> verbose
```

5. Verify devices can reach each other

```
switch#ping <IP>
```

6. Verify there aren't any drops

```
switch(config)#show monitor twamp-light counters drop detail
Interface: Ethernet29/1
Invalid TWAMP packet: 0
Packet too small: 0
Wrong dst port: 0
Internal error: 0
Wrong EtherType: 0
Unknown IP address: 0
Incorrect timestamps: 0
```

7. Verify time is correctly synchronized



```
Operational State : Synchronizing

Latest ECI ToD ( in ns ) : 100000000

Latest CMIC ToD ( in ns ) : 0

Latest Phase Differences ( in ns ) : N/A

Number of ECI ToD updates received : 1
```

All FAPs should be in Stable state.

8. Verify IS-IS advertises measured unidirectional delay metrics (average, min max and delay variation) in the LSP.

```
switch#show isis database traffic-engineering 1111.1111.1001.00-00 level-2
Legend:
H - hostname conflict
U - node unreachable
IS-IS Instance: 1 VRF: default
 IS-IS Level 2 Link State Database
   LSPID
                           Seg Num Cksum Life Length IS Received LSPID
   Flags
   smd418.00-00
                                48
                                    7331 1189 182 L2 1111.1111.1001.00-
00 <>
     LSP generation remaining wait time: 0 ms
     Time remaining until refresh: 889 s
     NLPID: 0xCC(IPv4)
     Hostname: smd418
     TE IPv4 router ID: 1.4.5.6
     Area addresses: 49.0001
     Interface address: 1.0.1.1
     Interface address: 1.0.0.1
     IS Neighbor
                         : 1111.1111.1002.00 Metric: 10
       IPv4 Neighbor Address: 1.0.0.2
       IPv4 Interface Address: 1.0.0.1
       Maximum link BW: 100.00 Gbps
       Maximum reservable link BW: 0.00 bps
       Unreserved BW:
           TE class 0: 0.00 bps
                                      TE class 1: 0.00 bps
                                                              TE class 2: 0.
00 bps
           TE class 3: 0.00 bps
                                      TE class 4: 0.00 bps
                                                              TE class 5: 0.
00 bps
           TE class 6: 0.00 bps
                                       TE class 7: 0.00 bps
       Min/Max unidirectional link delay: 642/702 us
```



```
Average unidirectional link delay: 682 us

Unidirectional link delay variation: 1292 us

Reachability : 1.0.1.0/24 Metric: 10 Type: 1 Up

Reachability : 1.0.0.0/24 Metric: 10 Type: 1 Up

Router Capabilities: Router Id: 1.0.0.1 Flags: []

Area leader priority: 250 algorithm: 0
```

In case there is a need to create a customer escalation please add following information:

- TwampAgent log files, located in /var/log/agents/
- Output of the internal debug command: "show agent TwampAgent debug"
- Output of following show commands:
 - o "show monitor twamp-light"
 - "show monitor twamp-light counters"
 - o "show monitor twamp-light drop detail"
 - o "show monitor twamp-light interface <INTF> raw" for all used interfaces
- Output of the Isis show commands
 - o "show isis database traffic-engineering"
 - "show traffic-engineering interfaces < interface name> "
- Output of following platform related commands
 - "show platform fap acl tcam hw bank 12"
 - "show hardware counter drop"
 - "show platform fap eci hardware-sync"
 - "platform fap diag get IPPD_TOD_LAST_VALUE"

Please note that before gathering logs tracing needs to be enabled:

```
switch(config)#trace TwampAgent enable Twamp* all
switch(config)#trace Isis enable Twamp* all
switch(config)#trace Isis enable LinkDynamicDelay all
```

Tracing

Disclaimer: In some cases, enabling tracing can seriously impact the performance of the switch. Please use it cautiously and seek advice from an Arista representative before enabling this in any production environments.

Twamp agent traces

To enable tracing enter this command in config mode:

```
switch(config)#trace TwampAgent enable Twamp* all
```



As it traces quite a lot, remember to disable it after collecting traces.

Sample trace that signifies the delay has been correctly calculated:

```
2024-05-22 01:01:03.715739 17859 TwampLightSm 1 U64 Twamp::Light::calculateDe lay(U64, U64, U64, U64): senderRxTime: 1716364863712409278 senderTxTime: 171636486370 2925824 clientRxTime: 1716364863703354880 clientTxTime: 1716364863703783936 senderTot alTime: 9483454 clientTotalTime: 429056 => DELAY: 9054398

2024-05-22 01:01:03.715760 17859 TwampLightSm 5 Twamp::Packet::Test::Reflecti on Twamp::Light::IntfPamSm::processResults(const Ptr&, const Twamp::TwampTime&, U32, const Arnet::IpGenAddr&): Probe to 1.0.0.2 on Ethernet25/1 with seqnum: 117/117 has t wo way delay (in nanoseconds): 9054398
```

Sample trace when packet is too small:

Sample trace when packet is invalid:

```
2024-05-22 01:18:04.149089 8021 TwampLightSm 2 void Twamp::Light::IntfPamSm:
:reflectPkt(const Ptr&, U32, U8Ptr, U8Ptr, const Ptr&, const Ptr&, const Twamp::Twamp
Time&, U16, const Arnet::EthAddr&, const Arnet::EthAddr&, U16): Ethernet1 discarding
invalid packet received
```

Sample trace when packet has wrong destination port:

```
2024-05-22 01:18:51.435175 8047 TwampLightSm 2 void Twamp::Light::IntfPamSm: :handleReadableCount(): intfId:Ethernet1 ignore pkt with wrong dport (53)
```

Sample trace when there's an internal error:



```
2 void Twamp::Light::IntfPamSm:
:handleReadableCount(): intfId:Ethernet1 ignore pkt with wrong ethType

2024-05-22 01:20:18.780946 8097 TwampLightSm 2 void Twamp::Light::IntfPamSm:
:handleReadableCount(): intfId:Ethernet1 ignore pkt with wrong dport (0)

2024-05-22 01:20:59.665946 8160 TwampLightSm 2 void Twamp::Light::IntfPamSm:
:handleReadableCount(): intfId:Ethernet1 ignore pkt with wrong sport (0)
```

This may happen when there is a platform error and packet fields are incorrect. In this situation you may need to ask for additional help from the platform team.

Sample trace when packet has wrong eth type:

```
2024-05-22 01:22:24.464261 8187 TwampLightSm 2 void Twamp::Light::IntfPamSm:
:handleReadableCount(): intfId:Ethernet1 ignore pkt with wrong ethTypeOrig
```

Sample trace when packet has unknown IP:

```
2024-05-22 01:23:04.829428 8213 TwampLightSm 4 Twamp::Packet::Test::Reflection Twamp::Light::IntfPamSm::processResults(const Ptr&, const Twamp::TwampTime&, U32, const Arnet::IpGenAddr&): Ethernet1 discarding packet received from unmonitored ip: 10.0.0.1
```

Sample trace when packet has incorrect timestamps:



Sample trace to verify reflection is using PTP or NTP timestamps:

```
2024-05-22 01:24:20.248936 8250 TwampLightSm 8 Twamp::Packet::Test::Reflecti
on Twamp::Light::IntfPamSm::processResults(const Ptr&, const Twamp::TwampTime&, U32,
const Arnet::IpGenAddr&): Reflection has PTP time
```

IS-IS agent traces

To enable tracing on Isis agent for dynamic delay feature we recommend following tracing commands:

```
switch(config)#trace Isis enable Twamp* all
switch(config)#trace Isis enable LinkDynamicDelay all
```

Limitations

- Only IPv4 probing is supported. IPv6 probe packets are not supported
- Only P2P interfaces are supported show
- Only ISIS is supported as IGP
- If a peer switch sends a timestamp in NTP format, the reflector copies it back to the sender timestamp location in the same format. However, a sender only sends probe packets in PTP timestamp format.
- Only dynamic min/max unidirectional delay metrics get advertised in ASLA as per RFC 8919. Dynamic unidirectional link delay (average) and unidirectional delay variation are not advertised in ASLA.
- This feature is supported only on ethernet and port-channel interfaces, not on subinterfaces or SVIs.

Resources

- IETF RFC 8186: https://datatracker.ietf.org/doc/rfc8186/
- ISIS flexible algorithm: https://www.arista.com/en/support/toi/eos-4-27-0f/14893-is-is-flexible-algorithm