Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2019

M. Chen
A. Malis
Huawei

October 22, 2018

DetNet Packet Loss and Delay Performance Measurement
draft-chen-detnet-loss-delay-00

Abstract

Deterministic Networking (DetNet) is defined to provide end-to-end bounded latency and extremely low packet loss rates for critical flows. It's important to measure and monitor the packet loss rates and end-to-end delay and delay variation of a DetNet flow path, which allows evaluation of whether the Service Level Agreements (SLA) of the provided DetNet services are satisfied. These metrics are also useful in network/traffic planning, trouble shooting, and network performance evaluation.

This document defines protocol mechanisms to support passive Performance Measurement (PM) for DetNet service.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

Table of Contents

1. Introduction

Deterministic Networking (DetNet) [I-D.ietf-detnet-architecture] can provide end-to-end bounded latency and extremely low packet loss rates for critical flows. This achieved by dedicating network resources (e.g., link bandwidth and queue buffering) to DetNet flows, and by replicating packets along multiple paths.

It's important to measure and monitor the packet loss rate and one-way delay and delay variation of a DetNet flow path in order to evaluate whether the Service Level Agreements (SLA) of the provided DetNet services are satisfied. These metrics are also useful in

network/traffic planning, troubleshooting, and network performance evaluation.

As defined in [RFC7799], performance measurement can be classified into Active, Passive and Hybrid measurement. Active measurement is performed by injecting Operations, Administration, and Maintenance (OAM) packets to the network to estimate the performance of the network by measuring the performance of the OAM packets. However, adding extra traffic can affect the quantities to be measured to some degree. On the other hand, passive measurement monitors the actual service traffic, rather than injecting OAM packets to estimate the traffic performance. Therefore, Passive performance measurement will not affect the behavior of the real DetNet service, and also provide more accurate measurement results. Accordingly, this document defines protocol mechanisms to support Passive PM for DetNet services.

DetNet defines two encapsulations, an MPLS-based encapsulation [I-D.ietf-detnet-dp-sol-mpls], and an IP-based encapsulation [I-D.ietf-detnet-dp-sol-ip]. For the MPLS based encapsulation, a service layer is introduced, which is supported by a DetNet Service Label (S-Label) and a DetNet Control Word (d-CW). The S-Label is used to identify a DetNet flow. The d-CW contains a sequence number that is designed for supporting the Packet Replication, Elimination, and Ordering Functions (PREOF). When perform packet loss and delay measurements, the sequence number can be used for packet accounting and packet count/timestamp correlation as well.

[RFC6374] defines Loss Measurement (LM) and Delay Measurement (DM) messages to communicate packet counts, timestamps, and other relevant information between Measurement Points (MPs). This document defines three new TLVs to the [RFC6374] LM message and DM messages. Which are used for communicating the correlation information (e.g., sequence number, measurement interval, and service label) that enables the packet loss and packet delay calculation. The detailed definitions of these new TLVs are described in Section 3.

2. DetNet Control Word based PM

As discussed above, MPLS-based DetNet encapsulation introduces an S-Label and a d-CW. This document defines two new flags in the d-CW (as shown in Figure 1). The L bit is defined to indicate whether the loss measurement is enabled, and the D bit is defined to indicate whether the delay measurement is enabled.

```
          +-----------------+
          ~  IP/MPLS Tunnel ~
          +-----------------+ <--\
          |  Service Label  |    |
          +-----------------+    +-- Service Layer Header
     +----|  Control Word   |    |
     |    +-----------------+ <--/
     |    |     Payload     |
     |    +-----------------+
     |
     |    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     +--->|0 0 0 0|L|D|                 Sequence Number               |
          +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: DetNet Control Word

Where:

- o L bit: Loss measurement indicator; 1 means the loss measurement is enabled, otherwise the loss measurement is not enabled.

- o D bit: Delay measurement indicator;1 means the delay measurement is enabled, otherwise the delay measurement is not enabled. When a node receive a packet with D bit set, it will timestamp the packet and copy it for further PM processing.

2.1. Packet Loss Measurement

Assume a DetNet service path between node A and node B, where node A is the ingress node, and node B is the egress node. To measure the number of packets transmitted at node A but not received at node B within a measurement interval, there needs a way to determine which packets belong to which measurement interval. [RFC6374] uses OAM packets to demarcate different measurement intervals. However, an OAM packet-based solution cannot work when there is packet mis-ordering. This document uses the sequence number to determine to which measurement interval a packet belongs. Specifically, the measurement interval number is calculated as the modulo of the sequence number and a pre-configured constant.

- o Measurement Interval = "Sequence Number" mod "Pre-configured constant".

With this, the ingress and egress nodes can use the sequence number of a packet to calculate the measurement interval number. The packets with same interval number belong to the same measurement interval. Then the packet loss can be calculated as below:

Loss[n] = A_TxP[n] - B_RxP[n], where:

o A_TxP[n] is the number of packets transmitted at node A within the No. "n" measurement interval;

o B_RxP[n] is the number of packets received at node B within the No. "n" measurement interval;

If the calculation is performed at one side of the path, the A_TxP[n] or B_RxP[n] needs to be sent to the other side. The Loss Measurement (LM) message defined in [RFC6374] is used to communicate the counts, in order to correlate the counts from the ingress node with the counts from the egress node. The extensions to [RFC6374] to communicate the measurement interval are defined in Section 3.

If the calculation is performed at a centralized controller, then the A_TxP[n] and B_RxP[n] need to be sent to the controller. The mechanism for sending counts to a centralized controller is out side the scope of this document.

2.2. Packet Delay Measurement

To measure the delay of a packet, the D bit of the d-CW MUST be set. At the ingress node, record the time when sending the packet, with the timestamp indexed by the sequence number. At the egress node, when receiving a packet with D bit set, record the time when the packet was received, with the timestamp indexed by the sequence number. Then, with the timestamps from the ingress and egress nodes, and the sequence number, the packet delay can be calculated as below:

Delay[n] = B_RxT[n] - A_TxT[n], where:

o B_RxT[n] identifies the timestamp at node B when receiving the No. "n" packet;

o A_TxT[n] identifies the timestamp at node A when sending the No. "n" packet;

Similar to loss measurement, the Delay Measurement (DM) message defined in [RFC6374] is used to communicate the timestamps when calculation is performed at either side of a DetNet service path. In order to correlate the timestamps from the ingress node with the timestamps from the egress node, extensions to [RFC6374] to communicate the sequence number and other relevant information are needed. The detailed definitions of these extensions are described in Section 3.

The mechanism for sending timestamps to a centralized controller is out side the scope of this document.

2.3. Alternative Solutions to the "D/L" Bits

Configuration can be used to indicate whether the delay and/or loss measurements are enabled on a specific DetNet service flow. This can be done by Command Line Interface (CLI) or through the DetNet configuration model [I-D.geng-detnet-conf-yang].

Another way is to use the signalling protocol as the enabler of performance measurement. More detail will be aded in the future.

[Editor notes:

This document introduces three ways (as summarized below) to enable PM on a DetNet flow. We'd like to solicit more inputs and comments from the WG:

1. Indicated by the "D/L" bits: A straightforward way to indicate when to measure, which packets to measured. The cost is to take two bits (or at least one bit) away from the sequence number.

2. Configured by CLI or YANG: Normally, it's easy to enable/disable PM on a DetNet flow. The receiving node may take more time (e.g., by matching a local configuration item to determine) to determine whether a packet should be counted, whether a packet should be timestamped. And it is difficult to support if only partial packets of flow need to be measured. This is a common case for packet delay measurement, where sample measurement is acceptable and reasonable.

3. Signalled by control protocol: The pros and cons similar to option 2.

]

3. PM for IP-based Encapsulation

For IP-based encapsulation, since there is no service layer, the d-CW-based solution as defined in Section 2 can not be applied. The marking-based solution defined in [RFC8321] can be used. More detail will be added in future versions.

4. Extensions to RFC6374

[RFC6374] defines how to communicate the packet counts and timestamps between measurement points. In order to support passive PM, this document defines several new TLVs to carry the correlation information. The correlation information can be used to determine with which DetNet service path a packet count/timestamp correlates, and with which measurement interval a packet count correlates, and with which packet a timestamp correlates.

Three new TLVs to LM and DM messages [RFC6374] are defined in the following sub-sections.

4.1. Measurement Interval TLV

This document defines a new TLV which is referred to as Measurement Interval TLV to Loss Measurement message [RFC6374]. The Measurement Interval TLV carries the measurement interval that is used to correlate the packet counts from the ingress node with the packet counts from the egress nodes. Then the packet loss can be calculated as described in Section 2.

The format of the Measurement Interval TLV is as below:

```
   0                   1                   2                   3
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |             Type (TBD1)       |             Length            |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                     Measurement Interval                      |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: Measurement Interval TLV

Where:

- o The Type field is two octets in length, and the value is TBD1.
- o The Length field is two octets in length, with a value is 4, indicating the length of the Measurement Interval field.
- o The Measurement Interval field is 4 octets in length, and carries the measurement interval.

4.2. DetNet Control Word TLV

This document defines a new TLV which is referred to as DetNet Control Word TLV to Delay Measurement message [RFC6374]. The sequence number of the d-CW is used to correlate the timestamps from

the ingress node with the timestamp from the egress node. Then the packet delay can be calculated as described in Section 2.

The format of the DetNet Control Word TLV is as below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type (TBD2)       |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    DetNet Control Word                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: DetNet Control Word TLV

Where:

o The Type field is two octets in length, and the value is TBD2.

o The Length field is two octets in length, with a value is 4, indicating the length of the DetNet Control Word field.

o The DetNet Control Word is 4 octets in length.

4.3. Service Label TLV

This document defines a new TLV which is referred to as Service Label TLV to Loss Measurement message and Delay Measurement message [RFC6374]. The Service Label TLV carries the DetNet S-Label that is allocated by the receiving node to the DetNet service path that is being measured. Here, the receiving node can be the egress node, or an relay node. The S-Label is used to determine to which DetNet service path the packet counts/timestamps belong.

The format of the Service Label TLV is as below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type (TBD3)       |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Reserved         |          Service Label                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Service Label TLV

Where:

o The Type field is two octets in length, and the value is TBD3.

o The Length field is two octets in length, with a value is 4, indicating the length of the Reserved and the Service Label fields.

o The Service Label field is 20-bit in length.

5. IANA Considerations

IANA is requested to allocate the following TLV types from the "MPLS Loss/Delay Measurement TLV Object" sub-registry of the "Generic Associated Channel (G-ACh) Parameters" registry:

Type	Description	Reference
TBD1	Measurement Interval	This document
TBD2	DetNet Control Word	This document
TBD3	DetNet Service Label	This document

6. Security Considerations

This document enables the use of Passive monitoring to determine the SLA conformance of DetNet service flows, and does not introduce any additional Active monitoring packets to the network. As a result, this document introduces no new security considerations beyond those already described in Section 8 of [RFC6374] and Section 5 of [RFC7799].

7. Acknowledgements

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <https://www.rfc-editor.org/info/rfc6374>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

8.2. Informative References

[I-D.geng-detnet-conf-yang]
Geng, X., Chen, M., Li, Z., and R. Rahman, "DetNet Configuration YANG Model", draft-geng-detnet-conf-yang-06 (work in progress), October 2018.

[I-D.ietf-detnet-architecture]
Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-08 (work in progress), September 2018.

[I-D.ietf-detnet-dp-sol-ip]
Korhonen, J. and B. Varga, "DetNet IP Data Plane Encapsulation", draft-ietf-detnet-dp-sol-ip-01 (work in progress), October 2018.

[I-D.ietf-detnet-dp-sol-mpls]
Korhonen, J. and B. Varga, "DetNet MPLS Data Plane Encapsulation", draft-ietf-detnet-dp-sol-mpls-01 (work in progress), October 2018.

[I-D.ietf-detnet-use-cases]
Grossman, E., "Deterministic Networking Use Cases", draft-ietf-detnet-use-cases-19 (work in progress), October 2018.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321, January 2018, <https://www.rfc-editor.org/info/rfc8321>.

Authors' Addresses

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Andrew G. Malis
Huawei

Email: agmalis@gmail.com