DetNet Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2021

G. Mirsky
ZTE Corp.
M. Chen
Huawei

July 8, 2020

Operations, Administration and Maintenance (OAM) for Deterministic Networks (DetNet) with MPLS Data Plane
draft-ietf-detnet-mpls-oam-01

Abstract

This document lists functional requirements for Operations, Administration, and Maintenance (OAM) toolset in Deterministic Networks (DetNet) and, using these requirements; defines format and use principals of the DetNet service Associated Channel over a DetNet network with the MPLS data plane..

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 January 9, 2021.

Copyright Notice

Copyright (c) 2020 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

[RFC8655] introduces and explains Deterministic Networks (DetNet) architecture and how the Packet Replication and Elimination function (PREF) can be used to ensure low packet drop ratio in DetNet domain.

Operations, Administration and Maintenance (OAM) protocols are used to detect, localize defects in the network, and monitor network performance. Some OAM functions, e.g., failure detection, work in the network proactively, while others, e.g., defect localization, usually performed on-demand. These tasks achieved by a combination of active and hybrid, as defined in [RFC7799], OAM methods.

This document lists the functional requirements toward OAM for DetNet domain. The list can further be used for gap analysis of available OAM tools to identify possible enhancements of existing or whether new OAM tools are required to support proactive and on-demand path monitoring and service validation. Also, this document defines format and use principals of the DetNet service Associated Channel over a DetNet network with the MPLS data plane [I-D.ietf-detnet-mpls].

2. Conventions used in this document

2.1. Terminology and Acronyms

The term "DetNet OAM" used in this document interchangeably with longer version "set of OAM protocols, methods and tools for Deterministic Networks".

CW Control Word

DetNet Deterministic Networks

d-ACH DetNet Associated Channel Header

d-CW DetNet Control Word

DNH DetNet Header

GAL Generic Associated Channel Label

G-ACh Generic Associated Channel

OAM: Operations, Administration and Maintenance

PREF Packet Replication and Elimination Function

POF Packet Ordering Function

PW Pseudowire

RDI Remote Defect Indication

E2E End-to-end

CFM Connectivity Fault Management

BFD Bidirectional Forwarding Detection

TSN Time-Sensitive Network

F-Label A Detnet "forwarding" label that identifies the LSP used to forward a DetNet flow across an MPLS PSN, e.g., a hop-by-hop label used between label switching routers (LSR).

S-Label A DetNet "service" label that is used between DetNet nodes that implement also the DetNet service sub-layer functions. An S-Label is also used to identify a DetNet flow at DetNet service sub-layer.

Underlay Network or Underlay Layer: The network that provides connectivity between the DetNet nodes. MPLS network providing LSP connectivity between DetNet nodes is an example of the underlay layer.

DetNet Node - a node that is an actor in the DetNet domain. DetNet domain edge node and node that performs PREF within the domain are examples of DetNet node.

2.2. Keywords

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Requirements

This section lists requirements for OAM in DetNet domain with MPLS data plane:

1. It MUST be possible to initiate DetNet OAM session from any DetNet node towards another DetNet node(s) within given domain.

2. It SHOULD be possible to initialize DetNet OAM session from a centralized controller.

3. DetNet OAM MUST support proactive and on-demand OAM monitoring and measurement methods.

4. DetNet OAM packets MUST be in-band, i.e., follow precisely the same path as DetNet data plane traffic.

5. DetNet OAM MUST support unidirectional OAM methods, continuity check, connectivity verification, and performance measurement.

6. DetNet OAM MUST support bi-directional OAM methods. Such OAM methods MAY combine in-band monitoring or measurement in the forward direction and out-of-bound notification in the reverse direction, i.e., from egress to ingress end point of the OAM test session.

7. DetNet OAM MUST support proactive monitoring of a DetNet node availability in the given DetNet domain.

8. DetNet OAM MUST support Path Maximum Transmission Unit discovery.

9. DetNet OAM MUST support Remote Defect Indication (RDI) notification to the DetNet node performing continuity checking.

10. DetNet OAM MUST support performance measurement methods.

11. DetNet OAM MAY support hybrid performance measurement methods.

12. DetNet OAM MUST support unidirectional performance measurement methods. Calculated performance metrics MUST include but are not limited to throughput, packet loss, delay and delay variation metrics. [RFC6374] provides excellent details on performance measurement and performance metrics.

13. DetNet OAM MUST support defect notification mechanism, like Alarm Indication Signal. Any DetNet node in the given DetNet domain MAY originate a defect notification addressed to any subset of nodes within the domain.

14. DetNet OAM MUST support methods to enable survivability of the DetNet domain. These recovery methods MAY use protection switching and restoration.

15. DetNet OAM MUST support the discovery of Packet Replication, Elimination, and Order preservation sub-functions locations in the domain.

16. DetNet OAM MUST support testing of Packet Replication, Elimination, and Order preservation sub-functions in the domain.

17. DetNet OAM MUST support monitoring any sub-set of paths traversed through the DetNet domain by the DetNet flow.

4. Active OAM for DetNet Networks with MPLS Data Plane

OAM protocols and mechanisms act within the data plane of the particular networking layer. And thus it is critical that the data plane encapsulation supports OAM mechanisms in such a way to comply with the above-listed requirements. One of such examples that require special consideration is requirement #5:

> DetNet OAM packets MUST be in-band, i.e., follow precisely the same path as DetNet data plane traffic both for unidirectional and bi-directional DetNet paths.

The Det Net data plane encapsulation in transport network with MPLS encapsulation specified in [I-D.ietf-detnet-mpls]. For the MPLS underlay network, DetNet flows to be encapsulated analogous to pseudowires (PW) over MPLS packet switched network, as described in

[RFC3985], [RFC4385]. Generic PW MPLS Control Word (CW), defined in [RFC4385], for DetNet displayed in Figure 1.

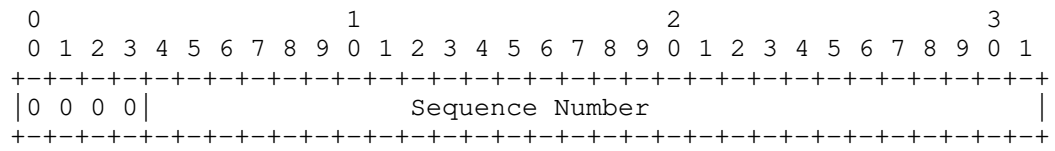

Figure 1: DetNet Control Word Format

PREF in the DetNet domain composed by a combination of nodes that perform replication and elimination sub-functions. The elimination sub-function always uses the S-Label and packet sequencing information, e.g., the value in the Sequence Number field of DetNet CW (d-CW). The replication sub-function uses the S-Label information only. For data packets Figure 2 presents an example of PREF in DetNet domain.

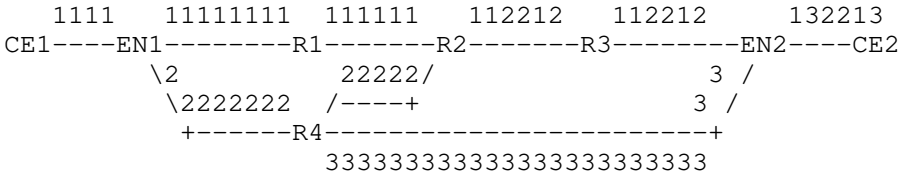

Figure 2: DetNet Data Plane Based on PW

4.1. DetNet Active OAM Encapsulation

DetNet OAM, like PW OAM, uses PW Associated Channel Header defined in [RFC4385]. Figure 3 displays the encapsulation of a DetNet MPLS [I-D.ietf-detnet-mpls] active OAM packet.

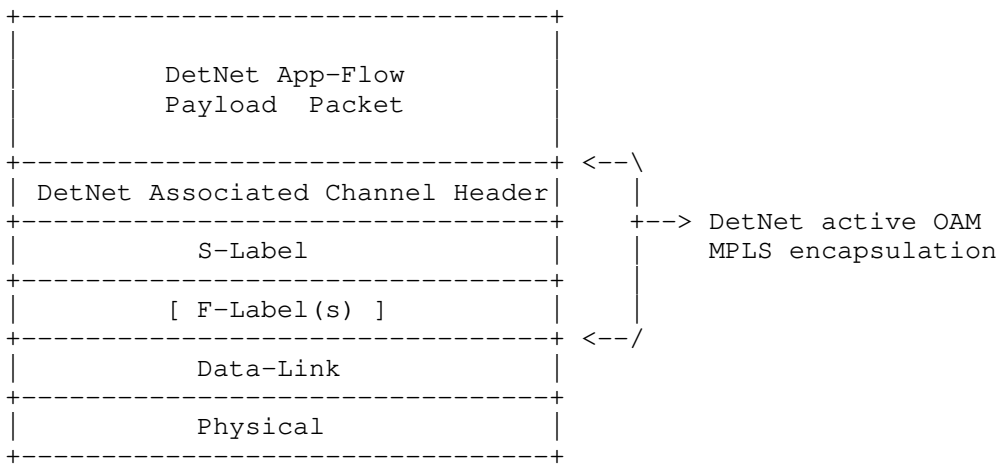

Figure 3: DetNet Active OAM Packet Encapsulation in MPLS Data Plane

Figure 4 displays encapsulation of a test packet of an active DetNet OAM protocol in case of MPLS-over-UDP/IP [I-D.ietf-detnet-mpls-over-udp-ip].

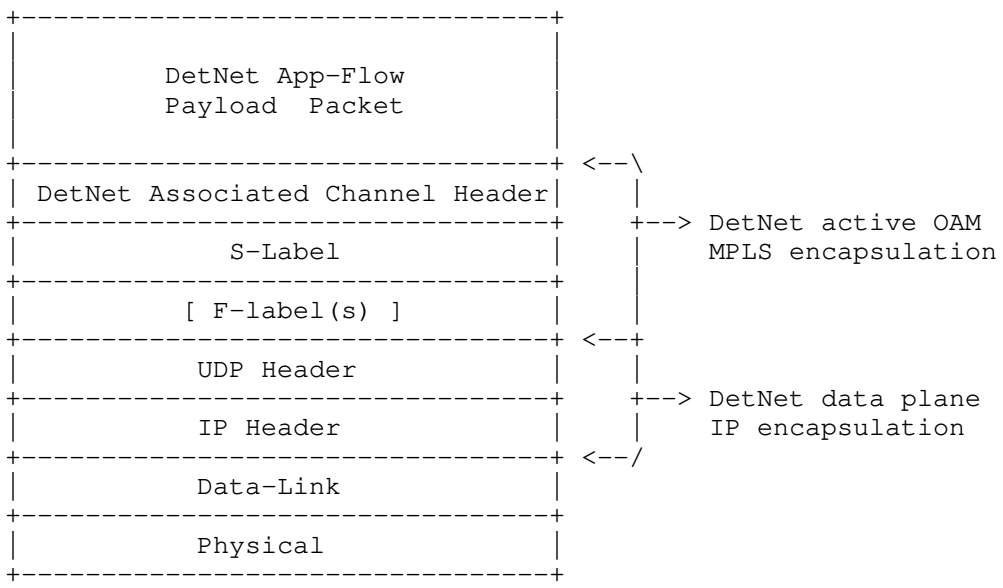

Figure 4: DetNet Active OAM Packet Encapsulation in MPLS-over-UDP/IP

Figure 5 displays the format of the DetNet Associated Channel Header (d-ACH).

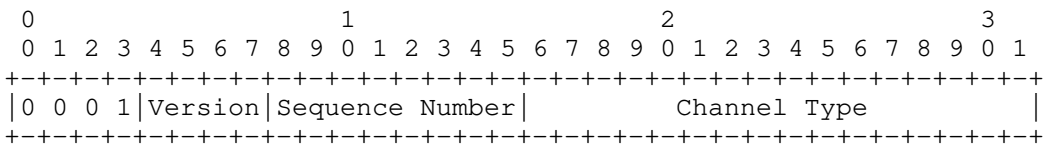

Figure 5: DetNet Associated Channel Header Format

The meanings of the fields in the d-ACH are:

Bits 0..3 MUST be 0b0001. This value of the first nibble allows the packet to be distinguished from an IP packet [RFC4928] and a DetNet data packet [I-D.ietf-detnet-mpls].

Version: this is the version number of the d-ACH. This specification defines version 0.

Sequence Number: this is unsigned eight bits-long field. The originating DetNet node MUST set the value of the Sequence Number field to a non-zero before packet being transmitted. The originating node MUST monotonically increase the value of the Sequence Number field for the every next active OAM packet.

Channel Type: the value of DetNet Associated Channel Type is one of values defined in the IANA PW Associated Channel Type registry.

The DetNet flow, according to [I-D.ietf-detnet-mpls], is identified by the S-label that MUST be at the bottom of the stack. Active OAM packet MUST have d-ACH immediately following the S-label.

4.2. DetNet Replication, Elimination, and Ordering Sub-functions Interaction with Active OAM

At the DetNet service layer, special functions MAY be applied to the particular DetNet flow - PREF to potentially lower packet loss, improve the probability of on-time packet delivery and Packet Ordering Function (POF) to ensure in-order packet delivery. As data and the active OAM packets have the same Flow ID, S-label, sub-functions that rely on sequencing information in the DetNet service layer MUST process 28 MSBs of the d-ACH as the source of the sequencing information for the OAM packet.

5. Use of Hybrid OAM in DetNet

Hybrid OAM methods are used in performance monitoring and defined in [RFC7799] as:

> Hybrid Methods are Methods of Measurement that use a combination of Active Methods and Passive Methods.

A hybrid measurement method may produce metrics as close to passive, but it still alters something in a data packet even if that is the value of a designated field in the packet encapsulation. One example of such a hybrid measurement method is the Alternate Marking method described in [RFC8321]. Reserving the field for the Alternate Marking method in the DetNet Header will enhance available to an operator set of DetNet OAM tools.

6. OAM Interworking Models

Interworking of two OAM domains that utilize different networking technology can be realized either by a peering or a tunneling model. In a peering model, OAM domains are within the corresponding network domain. When using the peering model, state changes that are detected by a Fault Management OAM protocol can be mapped from one OAM domain into another or a notification, e.g., an alarm, can be sent to a central controller. In the tunneling model of OAM interworking, usually, only one active OAM protocol is used. Its test packets are tunneled through another domain along with the data flow, thus ensuring the fate sharing among test and data packets.

6.1. OAM of DetNet MPLS Interworking with OAM of TSN

Active DetNet OAM is required to provide the E2E fault management and performance monitoring for a DetNet flow. Interworking of DetNet active OAM with MPLS data plane with the IEEE 802.1 Time-Sensitive Networking (TSN) domain based on [I-D.ietf-detnet-mpls-over-tsn].

In the case of the peering model is used in the fault management OAM, then the node that borders both TSN and DetNet MPLS domains MUST support [RFC7023]. [RFC7023] specified the mapping of defect states between Ethernet Attachment Circuits (ACs) and associated Ethernet PWs that are part of an end-to-end (E2E) emulated Ethernet service. Requirements and mechanisms described in [RFC7023] are equally applicable to using the peering model to achieve E2E FM OAM over DetNet MPLS and TSN domains. The Connectivity Fault Management (CFM) protocol [IEEE.CFM] or in [ITU.Y1731] can provide fast detection of a failure in the TSN segment of the DetNet service. In the DetNet MPLS domain BFD (Bidirectional Forwarding Detection), specified in [RFC5880] and [RFC5885], can be used. To provide E2E failure

detection, the TSN segment might be presented as a concatenated with the DetNet MPLS and the Section 6.8.17 [RFC5880] MAY be used to inform the upstream DetNet MPLS node of a failure of the TSN segment. Performance monitoring can be supported by [RFC6374] in the DetNet MPLS and [ITU.Y1731] in the TSN domains, respectively. Performance objectives for each domain should refer to metrics that additive or be defined for each domain separately.

The following considerations are to be realized when using the tunneling model of OAM interworking between DetNet MPLS and TSN domains:

- Active OAM test packet MUST be mapped to the same TSN Stream ID as the monitored DetNet flow.

- Active OAM test packets MUST be treated in the TSN domain based on its S-label and CoS marking (TC field value).

Note that the tunneling model of the OAM interworking requires that the remote peer of the E2E OAM domain supports the active OAM protocol selected on the ingress endpoint. For example, if BFD is used for proactive path continuity monitoring in the DetNet MPLS domain, a TSN endpoint of the DetNet service has also support BFD as defined in [RFC5885].

6.2. OAM of DetNet MPLS Interworking with OAM of DetNet IP

Interworking between active OAM segments in DetNet MPLS and DetNet IP domains can also be realized using either the peering or the tunneling model, as discussed in Section 6.1. Using the same protocol, e.g., BFD, over both segments, simplifies the mapping of errors in the peering model. To provide the performance monitoring over a DetNet IP domain STAMP [RFC8762] and its extensions [I-D.ietf-ippm-stamp-option-tlv] can be used.

7. IANA Considerations

TBA

8. Security Considerations

This document lists the OAM requirements for a DetNet domain and does not raise any security concerns or issues in addition to ones common to networking. Additionally, security considerations discussed in DetNet specifications: [RFC8655], [I-D.ietf-detnet-security], [I-D.ietf-detnet-mpls] are applicable to this document. Security concerns and issues related to MPLS OAM tools like LSP Ping [RFC8029], BFD over PW [RFC5885] also apply to this specification.

9. Acknowledgment

Authors extend their appreciation to Pascal Thubert for his insightful comments and productive discussion that helped to improve the document.

10. References

10.1. Normative References

[I-D.ietf-detnet-mpls]
Varga, B., Farkas, J., Berger, L., Malis, A., Bryant, S., and J. Korhonen, "DetNet Data Plane: MPLS", draft-ietf-detnet-mpls-08 (work in progress), July 2020.

[I-D.ietf-detnet-mpls-over-tsn]
Varga, B., Farkas, J., Malis, A., and S. Bryant, "DetNet Data Plane: MPLS over IEEE 802.1 Time Sensitive Networking (TSN)", draft-ietf-detnet-mpls-over-tsn-03 (work in progress), June 2020.

[I-D.ietf-detnet-mpls-over-udp-ip]
Varga, B., Farkas, J., Berger, L., Malis, A., and S. Bryant, "DetNet Data Plane: MPLS over UDP/IP", draft-ietf-detnet-mpls-over-udp-ip-06 (work in progress), May 2020.

[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>.

[RFC7023] Mohan, D., Ed., Bitar, N., Ed., Sajassi, A., Ed., DeLord, S., Niger, P., and R. Qiu, "MPLS and Ethernet Operations, Administration, and Maintenance (OAM) Interworking", RFC 7023, DOI 10.17487/RFC7023, October 2013, <https://www.rfc-editor.org/info/rfc7023>.

[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>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <https://www.rfc-editor.org/info/rfc8655>.

10.2. Informational References

[I-D.ietf-detnet-security]
Mizrahi, T. and E. Grossman, "Deterministic Networking (DetNet) Security Considerations", draft-ietf-detnet-security-10 (work in progress), May 2020.

[I-D.ietf-ippm-stamp-option-tlv]
Mirsky, G., Min, X., Nydell, H., Foote, R., Masputra, A., and E. Ruffini, "Simple Two-way Active Measurement Protocol Optional Extensions", draft-ietf-ippm-stamp-option-tlv-06 (work in progress), June 2020.

[IEEE.CFM]
IEEE, "Connectivity Fault Management clause of IEEE 802.1Q", IEEE 802.1Q, 2013.

[ITU.Y1731]
ITU-T, "OAM functions and mechanisms for Ethernet based Networks", ITU-T Recommendation G.8013/Y.1731, November 2013.

[RFC3985] Bryant, S., Ed. and P. Pate, Ed., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, DOI 10.17487/RFC3985, March 2005, <https://www.rfc-editor.org/info/rfc3985>.

[RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, DOI 10.17487/RFC4385, February 2006, <https://www.rfc-editor.org/info/rfc4385>.

[RFC4928] Swallow, G., Bryant, S., and L. Andersson, "Avoiding Equal Cost Multipath Treatment in MPLS Networks", BCP 128, RFC 4928, DOI 10.17487/RFC4928, June 2007, <https://www.rfc-editor.org/info/rfc4928>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/info/rfc5880>.

[RFC5885] Nadeau, T., Ed. and C. Pignataro, Ed., "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)", RFC 5885, DOI 10.17487/RFC5885, June 2010, <https://www.rfc-editor.org/info/rfc5885>.

[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>.

[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>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[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>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 14, 2021

A. Malis
Independent
X. Geng
M. Chen
Huawei
F. Qin
China Mobile
B. Varga
Ericsson

July 13, 2020

Deterministic Networking (DetNet) Controller Plane Framework
draft-malis-detnet-controller-plane-framework-04

Abstract

This document provides a framework overview for the Deterministic Networking (DetNet) controller plane. It discusses concepts and requirements that will be basis for Detnet controller plane solution documents.

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 January 14, 2021.

Copyright Notice

Copyright (c) 2020 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

Deterministic Networking (DetNet) provides the capability to carry specified unicast and/or multicast data flows for real-time applications with extremely low data loss rates and bounded latency within a network domain. As discussed in the Deterministic

Networking Architecture [RFC8655], techniques used to provide this capability include reserving data plane resources for individual (or aggregated) DetNet flows in some or all of the intermediate nodes along the path of the flow, providing explicit routes for DetNet flows that do not immediately change with the network topology, and distributing data from DetNet flow packets over time and/or space to ensure delivery of each packet's data in spite of the loss of a path.

The DetNet data plane is defined in a set of documents that are anchored by the DetNet Data Plane Framework [I-D.ietf-detnet-data-plane-framework] and the associated DetNet MPLS [I-D.ietf-detnet-mpls] and DetNet IP [I-D.ietf-detnet-ip] data plane specifications, with additional details and subnet mappings provided in [I-D.ietf-detnet-ip-over-mpls], [I-D.ietf-detnet-mpls-over-udp-ip], [I-D.ietf-detnet-mpls-over-tsn], [I-D.ietf-detnet-ip-over-tsn], and interconnection of TSN networks [I-D.ietf-detnet-tsn-vpn-over-mpls].

While the Detnet Architecture and Data Plane Framework documents are primarily concerned with data plane operations, they do contain some references and requirements for functions that would be required in order to automate DetNet service provisioning and monitoring via a DetNet controller plane. The purpose of this document is to gather these references and requirements into a single document and discuss how various possible DetNet controller plane architectures could be used to satisfy these requirements, while not providing the actual protocol details for a DetNet controller plane solution. Such controller plane protocol solutions will be the subject of subsequent documents.

Note that in the DetNet overall architecture, the controller plane includes what are more traditionally considered separate control and management planes. Traditionally, the management plane is primarily involved with node and network provisioning, operational OAM for performance monitoring, and troubleshooting network behaviors and outages, while the control plane is primarily responsible for the instantiation and maintenance of flows, MPLS label allocation and distribution, and active in-band or out-of-band signaling to support these functions. In the DetNet architecture, all of this functionality is combined into a single Controller Plane. See Section 4.4.2 of [RFC8655] and the aggregation of Control and Management planes in [RFC7426] for further details.

1.1. Terminology

This document uses the terminology established in the DetNet Architecture [RFC8655], and the reader is assumed to be familiar with that document and its terminology.

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. DetNet Controller Plane Requirements

Other DetNet documents, including [RFC8655] and [I-D.ietf-detnet-data-plane-framework], contain requirements for the Controller Plane. For convenience, these requirements have been compiled here. These requirements have been organized to show those primarily related to the control plane, those primarily relate to the management plane, and those applicable to both planes.

2.1. DetNet Control Plane Requirements

The primary requirements of the DetNet Control Plane are that it must be able to:

- Support the dynamic creation, modification, and deletion of DetNet flows. This may include some or all of explicit path determination, link bandwidth reservations, restricting flows to specific links (e.g., IEEE 802.1 Time-Sensitive Networking (TSN) links), node buffer and other resource reservations, specification of required queuing disciplines along the path, ability to manage bidirectional flows, etc., as needed for a flow.

- Support DetNet flow aggregation and de-aggregation via the ability to dynamically create and delete flow aggregates (FAs), and be able to modify existing FAs by adding or deleting participating flows.

- Allow flow instantiation requests to originate in an end application (via an Application Programming Interface (API), via static provisioning, or via a dynamic control plane, such as a centralized SDN controller or distributed signaling protocols. See Section 3 for further discussion of these options.

- In the case of the DetNet MPLS data plane, manage DetNet Service Label (S-Label), Forwarding Label (F-Label), and Aggregation Label (A-Label) [I-D.ietf-detnet-mpls] allocation and distribution.

- Also in the case of the DetNet MPLS data plane, support the DetNet service sub-layer, which provides DetNet service functions such as protection and reordering through the use of packet replication, duplicate elimination, and packet ordering functions (PREOF).

- Support queue control techniques defined in Section 4.5 of [RFC8655] and [I-D.finn-detnet-bounded-latency] that require time synchronization among network nodes.

- Advertise static and dynamic node and link resources such as capabilities and adjacencies to other network nodes (for dynamic signaling approaches) or to network controllers (for centralized approaches).

- Scale to handle the number of DetNet flows expected in a domain (which may require per-flow signaling or provisioning).

- Provision flow identification information at each of the nodes along the path. Flow identification may differ depending on the location in the network and the DetNet functionality (e.g. transit node vs. relay node).

2.2. DetNet Management Plane Requirements

The primary requirements of the DetNet Management Plane are that it must be able to:

- Monitor the performance of DetNet flows and nodes to ensure that they are meeting required objectives, both proactively and on-demand.

- Support DetNet flow continuity check and connectivity verification functions.

- Support testing and monitoring of packet replication, duplicate elimination, and packet ordering functionality in the DetNet domain.

2.3. Requirements For Both Planes

The following requirements apply to both the DetNet Controller and Management Planes:

- Operate in a converged network domain that contains both DetNet and non-DetNet flows.

- Adapt to DetNet domain topology changes such as links or nodes failures (fault recovery/restoration), additions and removals.

3. DetNet Control Plane Architecture

As noted in the Introduction, the DetNet control plane is responsible for the instantiation and maintenance of flows, allocation and distribution of flow related information (e.g., MPLS label), and active in-band or out-of-band information distribution to support these functions.

The following sections define three possible classes of DetNet control plane architectures: a fully distributed control plane utilizing dynamic signaling protocols, a fully centralized SDN-like control plane, and a control plane combining these two. They discuss the various information exchanges between entities in the network in each of these architectures and the advantages and disadvantages of each option.

In each of the following sections, examples are used to illustrate possible mechanisms that could be used in each of the architectures. These are not meant to be exhaustive or to preclude any other possible mechanism that could be used in place of those used in the examples.

3.1. Distributed Control Plane and Signaling Protocols

In a fully distributed configuration model, User-to-Network Interface (UNI) information is transmitted over a (to-be-defined) DetNet UNI protocol from the user side to the network side, and then UNI and network configuration information propagate in the network via distributed control plane signaling protocols. Such a DetNet UNI protocol are not visible in case of DetNet capable End-systems.

Taking an RSVP-TE traffic-engineered MPLS network, where End systems are not part of the DetNet domain, as a theoretical example:

1. An IGP collects topology information and DetNet capabilities of the network nodes

2. The control plane of the ingress edge node receives a flow establishment request from the UNI and calculates one or more valid path(s);

3. Using RSVP-TE [RFC3209], the ingress edge node sends a PATH message with an explicit route. After receiving the PATH message, the egress edge node sends a RESV message with the distributed label and resource reservation request.

IGP in the above example would require extensions to incorporate DetNet capabilities. Similarly, current reservation-oriented

distributed control plane protocols, e.g. RSVP-TE, can only reserve bandwidth along the path, while the configuration of a fine-grained schedule, e.g., Enhancements for Scheduled Traffic [IEEE.802.1QBV_2015], is not supported. If RSVP-TE were to be used for serving a DetNet flow, it would require extensions in order to support queue and scheduler reservations in addition to bandwidth reservation.

As discussed in Section 4.9 of [RFC8655], scalability is a primary concern for DetNet, given the large number of expected flows in a DetNet domain. This could potentially be much larger than, for example, the number of full-mesh MPLS traffic tunnels in a network using MPLS traffic engineering, which would typically be N*(N-1) tunnels, where N is the number of edge routers in the domain.

Even when flow aggregation is used, DetNet domains can be expected to support a very large number of flows that will need particular queuing disciplines and/or resource allocation, depending on the requirements for each flow. This could require a large amount of dynamic signaling, such as an RSVP-TE session to establish and maintain each flow. Other RSVP-TE scalability concerns are further discussed in [RFC5439].

All of the above tends to argue against a purely distributed control plane for DetNet domains.

3.2. SDN/Fully Centralized Control Plane

In the fully SDN/centralized configuration model, flow/UNI information is transmitted from a Centralized User Configuration or from applications via an API or northbound interface to a Centralized Controller, which is the sole source of routing and forwarding information for the domain. Configurations of nodes for DetNet flows are performed by the controller using a protocol such as NETCONF [RFC6241]/YANG [RFC6020] or PCE-CC [RFC8283].

Taking again an MPLS network, where End systems are not part of the DetNet domain, as a theoretical example:

1. A Centralized Controller collects topology information and DetNet capabilities of the network nodes via NETCONF/YANG;

2. The Controller receives a flow establishment request from a UNI and calculates one or more valid path(s) through the network;

3. The Controller chooses the optimal path and configures the devices along that path for flow transmission via PCE-CC.

Protocols in the above example may require extensions to incorporate DetNet specific parameters.

3.3. Combined Control Plane (partly centralized, partly distributed)

In the combined model, a Controller and control plane protocols work together to provide DetNet services, and there are a number of possible combinations.

Using an RSVP-TE traffic-engineered MPLS network with centralized PCE (Path Computation Engine), where End systems are not part of the DetNet domain, as a theoretical example:

1. A Centralized Controller collects topology information and DetNet capabilities of the network nodes via an IGP and/or BGP-LS [RFC7752];

2. The Controller receives a flow establishment request from a Network Management System and calculates one or more valid path(s) through the network;

3. Based on the calculation result, the Controller distributes flow path information to the ingress edge node and other information (e.g. replication/duplicate elimination) to the relevant nodes.

4. Using RSVP-TE, the ingress edge node sends a PATH message with an explicit route. After receiving the PATH message, the egress edge node sends a RESV message with the distributed label and resource reservation request.

Similarly to Distributed Control Plane and SDN/Fully Centralized Control Plane scenarios extensions of protocols are required to incorporate DetNet specific parameters.

There are many other variations that could be included in a combined control plane. This document cannot discuss all the possible control plane mechanisms that could be used in combined configuration models. Every solution has its own mechanisms and corresponding parameters that are required for it to work.

4. DetNet Control Plane Additional Details and Issues

This section discusses some additional DetNet control plane details and issues.

4.1. Explicit Paths

Explicit paths are required in DetNet to provide a stable forwarding service and guarantee that DetNet service is not impacted when the network topology changes. The following features are necessary to have explicit paths in DetNet:

- Path computation: DetNet explicit paths need to meet the SLA (Service Level Agreement) requirements and/or resource guarantees from the application/client, which include bandwidth, maximum end-to-end delay, maximum end-to-end delay variation, maximum loss ratio, etc. In a distributed system with IGP-TE, CSPF (Constrained Shortest Path First) can be used to compute a set of feasible paths for a DetNet service. In a system with a network controller, a PCE (Path Computation Engine) can compute paths satisfying the requirements of DetNet based on the network information collected from the DetNet domain.

- Path establishment: Once the path has been computed, the options discussed in Section 3 can be used to establish the path. Also see Section 4.4.1 for some additional considerations depending on the details of the network infrastructure.

- Strict or loose paths: An explicit path is strict when every intermediate hop is specified so that its route can't change. An explicit path is loose when any IGP route is allowed along the path. Generally, end-to-end SLA guarantees require a strict explicit path in DetNet. However, when the IGP route is known to be able to meet the SLA requirements, loose explicit paths are also acceptable.

4.2. Resource Reservation

Network congestion could cause uncontrolled delay and/or packet loss. DetNet flows are supposed to be protected from congestion, so sufficient resource reservation for DetNet service is necessary. Resources in the network are complex and hard to quantize, and may include such entities as packet processing resources, packet buffering, port and link bandwidth, and so on. The resources a particular flow requires are determined by the flow's characteristics and SLA.

- Resource Allocation: Port bandwidth is one of the basic attributes of a network device which is easy to obtain or calculate. In current traffic engineering implementations, network resource allocation is synonymous with bandwidth allocation. A DetNet flow is characterized with a traffic specification as defined in [I-D.ietf-detnet-flow-information-model], including attributes

such as Interval, Maximum Packets Per Interval, and Maximum Payload Size. The traffic specification describes the worst case, rather than the average case, for the traffic, to ensure that sufficient bandwidth and buffering resources are reserved to satisfy the traffic specification. However, in case of DetNet, resource allocation is more than simple bandwidth reservation. For example, allocation of buffers and required queuing disciplines during forwarding may be required as well. Furthermore, resources must be ensured to execute DetNet service sub-layer functions on the node, such as protection and reordering through the use of packet replication, duplicate elimination, and packet ordering functions (PREOF).

- o Device configuration with or without flow discrimination: The resource allocation can be guaranteed by device configuration. For example, an output port bandwidth reservation can be configured as a parameter of queue management and the port scheduling algorithm. When DetNet flows are aggregated, a group of DetNet flows share the allocated resource in the network device. When the DetNet flows are treated independently, the device should maintains a mapping relationship between a DetNet flow and its corresponding resources.

4.3. PREOF Support

DetNet path redundancy is supported via packet replication, duplicate elimination, and packet ordering functions (PREOF). A DetNet flow is replicated and goes through multiple networks paths to avoid packet loss caused by device or link failures. In general, current control plane mechanisms that can be used to establish an explicit path, whether distributed or centralized, support point-to-point (P2P) and point-to-multipoint (P2MP) path establishment. PREOF requires the ability to compute and establish a set of multiple paths (e.g., multiple LSP segments in an MPLS network) from the point(s) of packet replication to the point(s) of packet merging and ordering. Mapping of DetNet (member) flows to explicit path segments has to be ensured as well. Protocol extensions will be required to support these new features. Terminology will also be required to refer to this coordinated set of path segments (such as an "LSP graph" in case of DetNet MPLS data plane).

4.4. Data Plane specific considerations

4.4.1. DetNet in an MPLS Domain

For the purposes of this document, "traditional MPLS" is defined as MPLS without the use of segment routing (see Section 4.4.3 for a discussion of MPLS with segment routing) or MPLS-TP [RFC5960].

In traditional MPLS domains, a dynamic control plane using distributed signaling protocols is typically used for the distribution of MPLS labels used for forwarding MPLS packets. The dynamic signaling protocols most commonly used for label distribution are LDP [RFC5036], RSVP-TE, and BGP [RFC8277] (which enables BGP/MPLS-based Layer 3 VPNs [RFC4384] and Layer 2 VPNs [RFC7432]).

Any of these protocols could be used to distribute DetNet Service Labels (S-Labels) and Aggregation Labels (A-Labels) [I-D.ietf-detnet-mpls]. As discussed in [I-D.ietf-detnet-data-plane-framework], S-Labels are similar to other MPLS service labels, such as pseudowire, L3 VPN, and L2 VPN labels, and could be distributed in a similar manner, such as through the use of targeted LDP or BGP. If these were to be used for DetNet, they would require extensions to support DetNet-specific features such as PREOF, aggregation (A-Labels), node resource allocation, and queue placement.

However, as discussed in Section 3.1, distributed signaling protocols may have difficulty meeting DetNet's scalability requirements. MPLS also allows SDN-like centralized label management and distribution as an alternative to distributed signaling protocols, using protocols such as PCEP and OpenFlow [OPENFLOW].

PCEP, particularly when used as a part of PCE-CC, is a possible candidate protocol to use for centralized management of traditional MPLS-based DetNet domains. However, PCE path calculation algorithms would need to be extended to include the location determination for PREOF nodes in a path, and the means to signal the necessary resource reservation and PREOF function placement information to network nodes. See ((?I-D.ietf-pce-pcep-extension-for-pce-controller)) for further discussion of PCE-CC and PCEP for centralized control of an MPLS domain.

4.4.2. DetNet in an IP Domain

For the purposes of this document, "traditional IP" is defined as IP without the use of segment routing (see Section 4.4.3 for a discussion of IP with segment routing). In a later revision of this document, this section will discuss possible protocol extensions to existing IP routing protocols such as OSPF, IS-IS, and BGP. It should be noted that a DetNet IP data plane [I-D.ietf-detnet-ip] is simpler than a DetNet MPLS data plane [I-D.ietf-detnet-mpls], and doesn't support PREOF, so only one path per flow or flow aggregate is required.

4.4.3. DetNet in a Segment Routing Domain

Segment Routing [RFC8402] is a scalable approach to building network domains that provides explicit routing via source routing encoded in packet headers and it is combined with centralized network control to compute paths through the network. Forwarding paths are distributed with associated policy to network edge nodes for use in packet headers. As such, segment routing can be considered as a new data plane for both MPLS and IP. It reduces the amount of network signaling associated with distributed signaling protocols such as RSVP-TE, and also reduces the amount of state in core nodes compared with that required for traditional MPLS and IP routing, as the state is now in the packets rather than in the routers. This could be useful for DetNet, where a very large number of flows through a network domain are expected, which would otherwise require the instantiation of state for each flow traversing each node in the network. However, further analysis is needed on the expected gain, as DetNet flows may require various type of DetNet specific resources as well.

In a later revision of this document, this section will discuss the impact of DetNet on the Segment Routing Control and Management planes. Note that the DetNet MPLS and IP data planes described in [I-D.ietf-detnet-mpls] and [I-D.ietf-detnet-ip] were constructed to be compatible with both types of segment routing, SR-MPLS [RFC8660] and SRv6 [I-D.ietf-6man-segment-routing-header]. However, as of this writing, traffic engineering and resource reservation for segment routing are currently unsolved problems.

Editor's note: this section may be collapsed to previous sections and listing MPLS segment routing in the MPLS section as one of the possible explicit routing techniques for MPLS, and do the same for IP.

5. Management Plane Overview

The Management Plane includes the ability to statically provision network nodes and to use OAM to monitor DetNet performance and detect outages or other issues at the DetNet layer.

5.1. Provisioning

Static provisioning in a Detnet network nodes will be performed via the use of appropriate YANG models, including [I-D.ietf-detnet-yang] and [I-D.ietf-detnet-topology-yang].

5.2. DetNet Operations, Administration and Maintenance (OAM)

The overall framework and requirements for DetNet OAM are discussed in [I-D.mirsky-detnet-oam]. This document currently includes additional OAM details that may eventually be merged into that document.

5.2.1. OAM for Performance Monitoring (PM)

5.2.1.1. Active PM

Active PM is performed by injecting OAM packets into the network to estimate the performance of the network by measuring the performance of the OAM packets. Adding extra traffic can affect the delay and throughput performance of the network, and for this reason active PM is not recommended for use in operational DetNet domains. However, it is a useful test tool when commissioning a new network or during troubleshooting.

5.2.1.2. Passive PM

Passive PM monitors the actual service traffic in a network domain in order to measure its performance without having a detrimental affect on the network. As compared to Active PM, Passive PM is much preferred for use in DetNet domains.

5.2.2. OAM for Connectivity and Fault/Defect Management (CFM)

[I-D.mirsky-detnet-oam] contains requirements for connectivity and fault/defect detection and management in a DetNet domain.

6. Gap Analysis

In a later revision of this document, this section will contain a gap analysis of existing IETF control and management plane protocols not already discussed elsewhere in this document for their ability (or inability) to satisfy the requirements in Section 2, and discuss possible protocol extensions to existing protocols to fill the gaps, if any.

7. IANA Considerations

This document has no actions for IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

8. Security Considerations

Editor's note: This section needs more details.

The overall security considerations of DetNet are discussed in [RFC8655] and [I-D.ietf-detnet-security]. For DetNet networks that make use of Segment Routing (whether SR-MPLS or SRv6), the security considerations in [RFC8402] also apply.

DetNet networks that make use of a centralized controller plane may be threatened by the loss of connectivity (whether accidental or malicious) between the central controller and the network nodes, and/or the spoofing of control messages from the controller to the network nodes. This is important since such networks depend on centralized controllers to calculate flow paths and instantiate flow state in the network nodes. For networks that use both DetNet and Segment Routing with a centralized controller, this would also include the calculation of SID lists and their installation in edge/border routers.

In both cases, such threats may be mitigated through redundant controllers, the use of authentication between the controller(s) and the network nodes, and other mechanisms for protection against DOS attacks. A mechanism for supporting one or more alternative central controllers and the ability to fail over to such an alternative controller will be required.

9. Acknowledgments

Thanks to Jim Guichard, Donald Eastlake, and Stewart Bryant for their review comments.

10. References

10.1. Normative References

[I-D.ietf-detnet-data-plane-framework]
Varga, B., Farkas, J., Berger, L., Malis, A., and S. Bryant, "DetNet Data Plane Framework", draft-ietf-detnet-data-plane-framework-06 (work in progress), May 2020.

[I-D.ietf-detnet-flow-information-model]
Varga, B., Farkas, J., Cummings, R., Jiang, Y., and D. Fedyk, "DetNet Flow Information Model", draft-ietf-detnet-flow-information-model-10 (work in progress), May 2020.

[I-D.ietf-detnet-ip]
Varga, B., Farkas, J., Berger, L., Fedyk, D., and S. Bryant, "DetNet Data Plane: IP", draft-ietf-detnet-ip-07 (work in progress), July 2020.

[I-D.ietf-detnet-mpls]
Varga, B., Farkas, J., Berger, L., Malis, A., Bryant, S., and J. Korhonen, "DetNet Data Plane: MPLS", draft-ietf-detnet-mpls-09 (work in progress), July 2020.

[I-D.ietf-detnet-security]
Mizrahi, T. and E. Grossman, "Deterministic Networking (DetNet) Security Considerations", draft-ietf-detnet-security-10 (work in progress), May 2020.

[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>.

[RFC7426] Haleplidis, E., Ed., Pentikousis, K., Ed., Denazis, S., Hadi Salim, J., Meyer, D., and O. Koufopavlou, "Software-Defined Networking (SDN): Layers and Architecture Terminology", RFC 7426, DOI 10.17487/RFC7426, January 2015, <https://www.rfc-editor.org/info/rfc7426>.

[RFC8174] .

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <https://www.rfc-editor.org/info/rfc8655>.

10.2. Informative References

[I-D.finn-detnet-bounded-latency]
Finn, N., Boudec, J., Mohammadpour, E., Zhang, J., Varga, B., and J. Farkas, "DetNet Bounded Latency", draft-finn-detnet-bounded-latency-04 (work in progress), June 2019.

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Dukes, D., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-26 (work in progress), October 2019.

[I-D.ietf-detnet-ip-over-mpls]
Varga, B., Berger, L., Fedyk, D., Bryant, S., and J. Korhonen, "DetNet Data Plane: IP over MPLS", draft-ietf-detnet-ip-over-mpls-06 (work in progress), May 2020.

[I-D.ietf-detnet-ip-over-tsn]
Varga, B., Farkas, J., Malis, A., and S. Bryant, "DetNet Data Plane: IP over IEEE 802.1 Time Sensitive Networking (TSN)", draft-ietf-detnet-ip-over-tsn-03 (work in progress), June 2020.

[I-D.ietf-detnet-mpls-over-tsn]
Varga, B., Farkas, J., Malis, A., and S. Bryant, "DetNet Data Plane: MPLS over IEEE 802.1 Time Sensitive Networking (TSN)", draft-ietf-detnet-mpls-over-tsn-03 (work in progress), June 2020.

[I-D.ietf-detnet-mpls-over-udp-ip]
Varga, B., Farkas, J., Berger, L., Malis, A., and S. Bryant, "DetNet Data Plane: MPLS over UDP/IP", draft-ietf-detnet-mpls-over-udp-ip-06 (work in progress), May 2020.

[I-D.ietf-detnet-topology-yang]
Geng, X., Chen, M., Li, Z., and R. Rahman, "Deterministic Networking (DetNet) Topology YANG Model", draft-ietf-detnet-topology-yang-00 (work in progress), January 2019.

[I-D.ietf-detnet-tsn-vpn-over-mpls]
Varga, B., Farkas, J., Malis, A., Bryant, S., and D. Fedyk, "DetNet Data Plane: IEEE 802.1 Time Sensitive Networking over MPLS", draft-ietf-detnet-tsn-vpn-over-mpls-03 (work in progress), June 2020.

[I-D.ietf-detnet-yang]
Geng, X., Chen, M., Ryoo, Y., Li, Z., Rahman, R., and D. Fedyk, "Deterministic Networking (DetNet) Configuration YANG Model", draft-ietf-detnet-yang-06 (work in progress), June 2020.

[I-D.mirsky-detnet-oam]
Mirsky, G. and M. Chen, "Operations, Administration and Maintenance (OAM) for Deterministic Networks (DetNet)", draft-mirsky-detnet-oam-03 (work in progress), May 2019.

[IEEE.802.1QBV_2015]
IEEE, "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks - Amendment 25: Enhancements for Scheduled Traffic", IEEE 802.1Qbv-2015, DOI 10.1109/IEEESTD.2016.7572858, March 2016, <http://ieeexplore.ieee.org/servlet/opac?punumber=7572858>.

[OPENFLOW]
Open Networking Foundation, "OpenFlow Switch Specification, Version 1.5.1 (Protocol version 0x06)", ONF TS-025, March 2015, <https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-switch-v1.5.1.pdf>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC4384] Meyer, D., "BGP Communities for Data Collection", BCP 114, RFC 4384, DOI 10.17487/RFC4384, February 2006, <https://www.rfc-editor.org/info/rfc4384>.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <https://www.rfc-editor.org/info/rfc5036>.

[RFC5439] Yasukawa, S., Farrel, A., and O. Komolafe, "An Analysis of Scaling Issues in MPLS-TE Core Networks", RFC 5439, DOI 10.17487/RFC5439, February 2009, <https://www.rfc-editor.org/info/rfc5439>.

[RFC5960] Frost, D., Ed., Bryant, S., Ed., and M. Bocci, Ed., "MPLS Transport Profile Data Plane Architecture", RFC 5960, DOI 10.17487/RFC5960, August 2010, <https://www.rfc-editor.org/info/rfc5960>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <https://www.rfc-editor.org/info/rfc6020>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/info/rfc8277>.

[RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <https://www.rfc-editor.org/info/rfc8283>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

Authors' Addresses

Andrew G. Malis
Independent

Email: agmalis@gmail.com

Xuesong Geng
Huawei

Email: gengxuesong@huawei.com

Mach (Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Fengwei Qin
China Mobile

Email: qinfengwei@chinamobile.com

Balazs Varga
Ericsson

Email: balazs.a.varga@ericsson.com

DetNet Working Group
Internet-Draft
Intended status: Informational
Expires: September 24, 2020

G. Mirsky
ZTE Corp.
M. Chen
Huawei
D. Black
Dell EMC

March 23, 2020

Operations, Administration and Maintenance (OAM) for Deterministic Networks (DetNet) with IP Data Plane
draft-mirsky-detnet-ip-oam-02

Abstract

This document defines the principals for using Operations, Administration, and Maintenance protocols and mechanisms in the Deterministic Networking networks with IP data plane.

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 September 24, 2020.

Copyright Notice

Copyright (c) 2020 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

[RFC8655] introduces and explains Deterministic Networks (DetNet) architecture.

Operations, Administration and Maintenance (OAM) protocols are used to detect, localize defects in the network, and monitor network performance. Some OAM functions, e.g., failure detection, work in the network proactively, while others, e.g., defect localization, usually performed on-demand. These tasks achieved by a combination of active and hybrid, as defined in [RFC7799], OAM methods.

[I-D.mirsky-detnet-mpls-oam] lists the functional requirements toward OAM for DetNet domain. The list can further be used for gap analysis of available OAM tools to identify possible enhancements of existing or whether new OAM tools are required to support proactive and on-demand path monitoring and service validation. Also, the document defines the OAM use principals for the DetNet networks with IP data plane.

2. Conventions used in this document

2.1. Terminology

The term "DetNet OAM" used in this document interchangeably with longer version "set of OAM protocols, methods and tools for Deterministic Networks".

DetNet Deterministic Networks

DiffServ Differentiated Services

OAM: Operations, Administration and Maintenance

PREF Packet Replication and Elimination Function

POF Packet Ordering Function

RDI Remote Defect Indication

ICMP Internet Control Message Protocol

Underlay Network or Underlay Layer: The network that provides connectivity between the DetNet nodes. MPLS network providing LSP connectivity between DetNet nodes is an example of the underlay layer.

DetNet Node – a node that is an actor in the DetNet domain. DetNet domain edge node and node that performs PREF within the domain are examples of DetNet node.

2.2. Keywords

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Active OAM for DetNet Networks with IP Data Plane

OAM protocols and mechanisms act within the data plane of the particular networking layer. And thus it is critical that the data plane encapsulation supports OAM mechanisms in such a way that DetNet OAM packets are in-band with a DetNet flow being monitored, i.e., DetNet OAM test packets follow precisely the same path as DetNet data plane traffic both for unidirectional and bi-directional DetNet paths.

The DetNet data plane encapsulation in a transport network with IP encapsulations specified in Section 6 of [I-D.ietf-detnet-ip]. For the IP underlay network, DetNet flows are identified by the ordered match to the provisioned information set that, among other elements, includes the IP protocol, source port number, destination port number. Active IP OAM protocols like Bidirectional Forwarding Detection (BFD) [RFC5880] or STAMP [RFC8762], use UDP transport and the well-known UDP port numbers as the destination port. Thus a DetNet node MUST be able to associate an IP DetNet flow with the particular test session to ensure that test packets experience the same treatment as the DetNet flow packets.

Most of on-demand failure detection and localization in IP networks is being done by using the Internet Control Message Protocol (ICMP) Echo Request, Echo Reply and the set of defined error messages, e.g., Destination Unreachable, with the more detailed information provided through code points. [RFC0792] and [RFC4443] define the ICMP for IPv4 and IPv6 networks, respectively. Because ICMP is another IP protocol like, for example, UDP, a DetNet node MUST able to associate an ICMP packet generated by the specified IP DetNet node and addressed to the another IP DetnNet node with an IP DetNet flow between this pair of endpoints.

3.1. Active OAM Using DetNet-in-UDP Encapsulation

Active OAM in IP DetNet can be realized using DetNet-in-UDP encapsulation [Ed.note: Do we define it in this document or start a new one?]. Using DetNet-in-UDP tunnel between IP DetNet nodes ensures that active OAM test packets are fate-sharing with the packets of the being monitored IP DetNet flow. As a result, a test packet shares the tunnel with IP DetNet flow and shares the fate, statistically speaking, of the IP DetNet flow being monitored.

3.2. Mapping Active OAM and IP DetNet flows

IP OAM protocols that use UDP transport, e.g., BFD and STAMP, can be used to detect failures or performance degradation that affects an IP DetNet flow. When the UDP destination port number used by the OAM protocol is one of the assigned by IANA, then the UDP source port can be used to achieve co-routedness of OAM, and the monitored IP DetNet flow in the multipath environments, e.g., LAG or ECMP. To maximize the accuracy of OAM results in detecting failures and monitoring performance of IP DetNet, test packets should receive the same treatment by the nodes as experienced by the IP DetNet packet. Hence, the DSCP value used for a test packet MUST be mapped to DetNet.

3.3. Active OAM Using GRE-in-UDP Encapsulation

[RFC8086] has defined the method of encapsulating GRE (Generic Routing Encapsulation) headers in UDP. GRE-in-UDP encapsulation can be used for IP DetNet OAM as it eases the task of mapping an OAM test session to a particular IP DetNet flow that is identified by N-tuple. Matching a GRE-in-UDP tunnel to the monitored IP DetNet flow enables the use of Y.1731/G.8013 [ITU-T.1731] as a comprehensive toolset of OAM. The Protocol Type field in GRE header MUST be set to 0x8902 assigned by IANA to IEEE 802.1ag Connectivity Fault Management (CFM) Protocol / ITU-T Recommendation Y.1731. Y.1731/G.8013 supports necessary for IP DetNet OAM functions, i.e., continuity check, one-way packet loss and packet delay measurement.

4. Use of Hybrid OAM in DetNet

Hybrid OAM methods are used in performance monitoring and defined in [RFC7799] as:

> Hybrid Methods are Methods of Measurement that use a combination of Active Methods and Passive Methods.

A hybrid measurement method may produce metrics as close to passive, but it still alters something in a data packet even if that is the value of a designated field in the packet encapsulation. One example of such a hybrid measurement method is the Alternate Marking method (AMM) described in [RFC8321]. One of the advantages of the use of AMM in a DetNet domain with IP data plane is that the marking is applied to a data flow, thus ensuring that a measured metrics are directly applicable to the DetNet flow.

5. IANA Considerations

This document does not have any requests for IANA allocation. This section can be deleted before the publication of the draft.

6. Security Considerations

This document describes the applicability of the existing Fault Management and Performance Monitoring IP OAM protocols, and does not raise any security concerns or issues in addition to ones common to networking or already documented for the referenced DetNet and OAM protocols.

7. Acknowledgment

TBA

8. References

8.1. Normative References

[I-D.ietf-detnet-ip]
Varga, B., Farkas, J., Berger, L., Fedyk, D., Malis, A., and S. Bryant, "DetNet Data Plane: IP", draft-ietf-detnet-ip-05 (work in progress), February 2020.

[I-D.mirsky-detnet-mpls-oam]
Mirsky, G. and M. Chen, "Operations, Administration and Maintenance (OAM) for Deterministic Networks (DetNet) with MPLS Data Plane", draft-mirsky-detnet-mpls-oam-01 (work in progress), January 2020.

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, <https://www.rfc-editor.org/info/rfc792>.

[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>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[RFC8086] Yong, L., Ed., Crabbe, E., Xu, X., and T. Herbert, "GRE-in-UDP Encapsulation", RFC 8086, DOI 10.17487/RFC8086, March 2017, <https://www.rfc-editor.org/info/rfc8086>.

[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>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <https://www.rfc-editor.org/info/rfc8655>.

8.2. Informational References

[ITU-T.1731]
ITU-T, "Operations, administration and maintenance (OAM) functions and mechanisms for Ethernet-based networks", ITU-T G.8013/Y.1731, August 2015.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/info/rfc5880>.

[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>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

David Black
Dell EMC
176 South Street
Hopkinton, MA 01748
United States of America

Email: david.black@dell.com

RAW
Internet-Draft
Intended status: Standards Track
Expires: January 11, 2021

F. Theoleyre
CNRS
G. Papadopoulos
IMT Atlantique
G. Mirsky
ZTE Corp.

July 10, 2020

Operations, Administration and Maintenance (OAM) features for RAW
draft-theoleyre-raw-oam-support-03

Abstract

Some critical applications may use a wireless infrastructure. However, wireless networks exhibit a bandwidth of several orders of magnitude lower than wired networks. Besides, wireless transmissions are lossy by nature; the probability that a packet cannot be decoded correctly by the receiver may be quite high. In these conditions, guaranteeing the network infrastructure works properly is particularly challenging, since we need to address some issues specific to wireless networks. This document lists the requirements of the Operation, Administration, and Maintenance (OAM) features recommended to construct a predictable communication infrastructure on top of a collection of wireless segments. This document describes the benefits, problems, and trade-offs for using OAM in wireless networks to achieve Service Level Objectives (SLO).

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 January 11, 2021.

Copyright Notice

Copyright (c) 2020 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

Reliable and Available Wireless (RAW) is an effort that extends DetNet to approach end-to-end deterministic performances over a network that includes scheduled wireless segments. In wired networks, many approaches try to enable Quality of Service (QoS) by implementing traffic differentiation so that routers handle each type of packets differently. However, this differentiated treatment was expensive for most applications.

Deterministic Networking (DetNet) [RFC8655] has proposed to provide a bounded end-to-end latency on top of the network infrastructure, comprising both Layer 2 bridged and Layer 3 routed segments. Their work encompasses the data plane, OAM, time synchronization, management, control, and security aspects.

However, wireless networks create specific challenges. First of all, radio bandwidth is significantly lower than for wired networks. In these conditions, the volume of signaling messages has to be very limited. Even worse, wireless links are lossy: a layer 2 transmission may or may not be decoded correctly by the receiver, depending on a broad set of parameters. Thus, providing high reliability through wireless segments is particularly challenging.

Wired networks rely on the concept of _links_. All the devices attached to a link receive any transmission. The concept of a link in wireless networks is somewhat different from what many are used to in wireline networks. A receiver may or may not receive a transmission, depending on the presence of a colliding transmission, the radio channel's quality, and the external interference. Besides, a wireless transmission is broadcast by nature: any _neighboring_ device may be able to decode it. The document includes detailed information on what the implications for the OAM features are.

Last but not least, radio links present volatile characteristics. If the wireless networks use an unlicensed band, packet losses are not anymore temporally and spatially independent. Typically, links may exhibit a very bursty characteristic, where several consecutive packets may be dropped. Thus, providing availability and reliability on top of the wireless infrastructure requires specific Layer 3 mechanisms to counteract these bursty losses.

Operations, Administration, and Maintenance (OAM) Tools are of primary importance for IP networks [RFC7276]. It defines a toolset for fault detection, isolation, and performance measurement.

The primary purpose of this document is to detail the specific requirements of the OAM features recommended to construct a

predictable communication infrastructure on top of a collection of wireless segments. This document describes the benefits, problems, and trade-offs for using OAM in wireless networks to provide availability and predictability.

In this document, the term OAM will be used according to its definition specified in [RFC6291]. We expect to implement an OAM framework in RAW networks to maintain a real-time view of the network infrastructure, and its ability to respect the Service Level Objectives (SLO), such as delay and reliability, assigned to each data flow.

1.1. Terminology

- OAM entity: a data flow to be controlled;
- Maintenance End Point (MEP): OAM devices crossed when entering/ exiting the network. In RAW, it corresponds mostly to the source or destination of a data flow. OAM message can be exchanges between two MEPs;
- Maintenance Intermediate endPoint (MIP): OAM devices along the flow; OAM messages can be exchanged between a MEP and a MIP;
- Defect: a temporary change in the network (e.g., a radio link which is broken due to a mobile obstacle);
- Fault: a definite change which may affect the network performance, e.g., a node runs out of energy.

1.2. Acronyms

OAM Operations, Administration, and Maintenance

DetNet Deterministic Networking

SLO Service Level Objective

QoS Quality of Service

SNMP Simple Network Management Protocol

SDN Software-Defined Network

1.3. 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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Role of OAM in RAW

RAW networks expect to make the communications reliable and predictable on top of a wireless network infrastructure. Most critical applications will define an SLO to be required for the data flows it generates. RAW considers network plane protocol elements such as OAM to improve the RAW operation at the service and the forwarding sub-layers.

To respect strict guarantees, RAW relies on an orchestrator able to monitor and maintain the network. Typically, a Software-Defined Network (SDN) controller is in charge of scheduling the transmissions in the deployed network, based on the radio link characteristics, SLO of the flows, the number of packets to forward. Thus, resources have to be provisioned a priori to handle any defect. OAM represents the core of the pre-provisioning process and maintains the network operational by updating the schedule dynamically.

Fault-tolerance also assumes that multiple paths have to be provisioned so that an end-to-end circuit keeps on existing whatever the conditions. The Packet Replication and Elimination Function ([PREF-draft]) on a node is typically controlled by a central controller/orchestrator. OAM mechanisms can be used to monitor that PREOF is working correctly on a node and within the domain.

To be energy-efficient, reserving some dedicated out-of-band resources for OAM seems idealistic, and only in-band solutions are considered here.

RAW supports both proactive and on-demand troubleshooting.

The specific characteristics of RAW are discussed below.

2.1. Link concept and quality

In wireless networks, a _link_ does not exist. A common convention is to define a wireless link as a pair of devices that have a non-null probability of transmitting and decoding a packet. Similarly, we designate as *neighbor* any device which as a link with a specific transmitter.

Each wireless link is associated with a link quality, often measured as the Packet Delivery Ratio (PDR), i.e., the probability that the receiver can decode the packet correctly. It is worth noting that this link quality depends on many criteria, such as the level of external interference, the presence of concurrent transmissions, or the radio channel state. This link quality is even time-variant.

2.2. Broadcast Transmissions

In modern switching networks, the unicast transmission is delivered uniquely to the destination. Wireless networks are much closer to the ancient shared access wireless networks. Unicast transmission is similar to a broadcast one and can be received by any neighbor.

However, contrary to wired networks, we cannot be sure that a packet is received by *all* the devices attached to the network. It depends on the radio channel state between the transmitter(s) and the receiver(s). In particular, concurrent transmissions may be possible or not, depending on the radio conditions.

2.3. Complex Layer 2 Forwarding

Multiple neighbors may receive a transmission. Thus, anycast layer-2 forwarding helps to maximize the reliability by assigning multiple receivers to a single transmission. That way, the packet is lost only if none of the receivers decode it. Practically, it has been proven that different neighbors may exhibit very different radio conditions, and that reception independency may hold for some of them [anycast-property].

3. Operation

OAM features will enable RAW with robust operation both for forwarding and routing purposes.

3.1. Information Collection

Several solutions (e.g., Simple Network Management Protocol (SNMP), YANG-based data models) are already in charge of collecting the statistics. That way, we can encapsulate these statistics in specific monitoring packets, to send them to the controller.

3.2. Continuity Check

We need to verify that two endpoints are connected. In other words, there exists "one" way to deliver the packets between two endpoints A and B. The solution may not here defer from those of detnet.

3.3. Connectivity Verification

Additionally, to the Continuity Check, we have to verify the connectivity. This verification considers additional constraints, i.e., the absence of misconnection.

In particular, the resources have to be reserved by a given flow, and no packets from other flows steal the corresponding resources. Similarly, the destination does not receive packets from different flows through its interface.

Because of radio transmissions' broadcast nature, several receivers may be active at the same time to enable anycast Layer 2 forwarding. Thus, the connectivity verification must test any combination. We also consider priority-based mechanisms for anycast forwarding, i.e., all the receivers have different probabilities of forwarding a packet. To verify a delay SLO for a given flow, we must also consider all the possible combinations, leading to a probability distribution function for end-to-end transmissions. If this verification is implemented naively, the number of combinations to test may be exponential and too costly for wireless networks with low bandwidth.

It is worth noting that the control and data packets may not follow the same path. The connectivity verification has to be conducted in-band without impacting the data traffic. Test packets MUST share the fate with the monitored data traffic without introducing congestion in normal network conditions.

3.4. Route Tracing

ICMP tools are comprehensive tools for diagnostic. They help to identify a subset of the list of routers in the route. To ensure predictable performance, resources are reserved per flow in RAW. Thus, we need to define route tracing tools able to track the route for a specific flow.

Wireless networks are meshed by nature: we have many redundant radio links. These meshed networks are both an asset and a drawback: while several paths exist between two endpoints, and we should choose the most efficient one(s), concerning specifically the reliability, and the delay.

Thus, multipath routing can be considered to make the network fault-tolerant. Even better, we can exploit the broadcast nature of wireless networks to exploit meshed multipath routing: we may have multiple Maintenance Intermediate Endpoints (MIE) for each hop in the path. In that way, each Maintenance Intermediate Endpoint has

several possible next hops in the forwarding plane. Thus, all the possible paths between two maintenance endpoints should be retrieved, which may quickly become untractable if we apply a naive approach.

3.5. Fault Verification/detection

RAW expects to operate fault-tolerant networks. Thus, we need mechanisms able to detect faults, before they impact the network performance.

Wired networks tend to present stable performances. On the contrary, wireless networks are time-variant. We must consequently make a distinction between _normal_ evolutions and malfunction.

The network has to detect when a fault occurred, i.e., the network has deviated from its expected behavior. While the network must report an alarm, the cause may not be identified precisely. For instance, the end-to-end reliability has decreased significantly, or a buffer overflow occurs.

3.6. Fault Isolation/identification

The network has isolated and identified the cause of the fault. While detnet already expects to identify malfunctions, some problems are specific to wireless networks. We must consequently collect metrics and implement algorithms tailored for wireless networking. For instance, the quality of a specific link has decreased, requiring more retransmissions, or the level of external interference has locally increased.

4. Administration

The network has to expose a collection of metrics to support an operator making proper decisions, including:

- Packet losses: the time-window average and maximum values of the number of packet losses have to be measured. Many critical applications stop to work if a few consecutive packets are dropped;

- Received Signal Strength Indicator (RSSI) is a very common metric in wireless to denote the link quality. The radio chipset is in charge of translating a received signal strength into a normalized quality indicator;

- Delay: the time elapsed between a packet generation / enqueuing and its reception by the next hop;

- Buffer occupancy: the number of packets present in the buffer, for each of the existing flows.

These metrics should be collected:

- per virtual circuit to measure the end-to-end performance for a given flow. Each of the paths has to be isolated in multipath routing strategies;

- per radio channel to measure, e.g., the level of external interference, and to be able to apply counter-measures (e.g., blacklisting).

- per device to detect misbehaving node, when it relays the packets of several flows.

4.1. Collection of metrics

We have to minimize the number of statistics / measurements to exchange:

- energy efficiency: low-power devices have to limit the volume of monitoring information since every bit consumes energy.

- bandwidth: wireless networks exhibit a bandwidth significantly lower than wired, best-effort networks.

- per-packet cost: it is often more expensive to send several packets instead of combining them in a single link-layer frame.

Thus, localized and centralized mechanisms have to be combined together, and additional control packets have to be triggered only after a fault detection.

4.2. Worst-case metrics

RAW aims to enable real-time communications on top of a heterogeneous architecture. Wireless networks are known to be lossy, and RAW has to implement strategies to improve reliability on top of unreliable links. Hybrid Automatic Repeat reQuest (ARQ) has typically to enable retransmissions based on the end-to-end reliability and latency requirements.

To make correct decisions, the controller needs to know the distribution of packet losses for each flow, and each hop of the paths. In other words, the average end-to-end statistics are not enough. They must allow the controller to predict the worst-case.

4.3. Energy efficiency constraint

RAW targets also low-power wireless networks, where energy represents a key constraint. Thus, we have to take care of power and bandwidth consumption. The following techniques aim to reduce the cost of such maintenance:

on-path collection: some control information is inserted in the data packets if they do not fragment the packet (i.e., the MTU is not exceeded). Information Elements represent a standardized way to handle such information;

flags/fields: we have to set-up flags in the packets to monitor to be able to monitor the forwarding process accurately. A sequence number field may help to detect packet losses. Similarly, path inference tools such as [ipath] insert additional information in the headers to identify the path followed by a packet a posteriori.

5. Maintenance

RAW needs to implement a self-healing and self-optimization approach. The network must continuously retrieve the state of the network, to judge about the relevance of a reconfiguration, quantifying:

the cost of the sub-optimality: resources may not be used optimally (e.g., a better path exists);

the reconfiguration cost: the controller needs to trigger some reconfigurations. For this transient period, resources may be twice reserved, and control packets have to be transmitted.

Thus, reconfiguration may only be triggered if the gain is significant.

5.1. Replication / Elimination

When multiple paths are reserved between two maintenance endpoints, they may decide to replicate the packets to introduce redundancy, and thus to alleviate transmission errors and collisions. For instance, in Figure 1, the source node S is transmitting the packet to both parents, nodes A and B. Each maintenance endpoint will decide to trigger the replication/elimination process when a set of metrics passes through a threshold value.

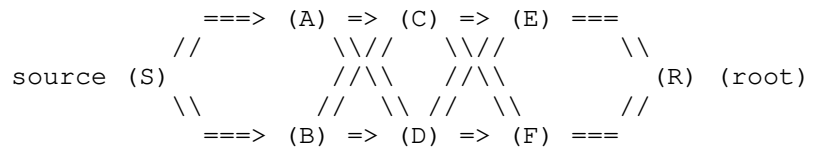

Figure 1: Packet Replication: S transmits twice the same data packet, to its DP (A) and to its AP (B).

5.2. Dynamic Resource Reservation

Wireless networks exhibit time-variant characteristics. Thus, the network has to provide additional resources along the path to fit the worst-case performance. This time-variant characteristics make the resource reservation very challenging: over-reaction waste radio and energy resources. Inversely, under-reaction jeopardize the network operations, and some SLO may be violated.

5.3. Reliable Reconfiguration

Wireless networks are known to be lossy. Thus, commands may be received or not by the node to reconfigure. Unfortunately, inconsistent states may create critical misconfigurations, where packets may be lost along a path because it has not been properly configured.

We have to propose mechanisms to guarantee that the network state is always consistent, even if some control packets are lost. Timeouts and retransmissions are not sufficient since the reconfiguration duration would be, in that case, unbounded.

6. IANA Considerations

This document has no actionable requirements for IANA. This section can be removed before the publication.

7. Security Considerations

This section will be expanded in future versions of the draft.

8. Acknowledgments

TBD

9. Informative References

[anycast-property]
Teles Hermeto, R., Gallais, A., and F. Theoleyre, "Is Link-Layer Anycast Scheduling Relevant for IEEE 802.15.4-TSCH Networks?", 2019, <https://doi.org/10.1109/LCNSymposium47956.2019.9000679>.

[ipath] Gao, Y., Dong, W., Chen, C., Bu, J., Wu, W., and X. Liu, "iPath: path inference in wireless sensor networks.", 2016, <https://doi.org/10.1109/TNET.2014.2371459>.

[PREF-draft]
Thubert, P., Eckert, T., Brodard, Z., and H. Jiang, "BIER-TE extensions for Packet Replication and Elimination Function (PREF) and OAM", 2018, <https://tools.ietf.org/html/draft-thubert-bier-replication-elimination>.

[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>.

[RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu, D., and S. Mansfield, "Guidelines for the Use of the "OAM" Acronym in the IETF", BCP 161, RFC 6291, DOI 10.17487/RFC6291, June 2011, <https://www.rfc-editor.org/info/rfc6291>.

[RFC7276] Mizrahi, T., Sprecher, N., Bellagamba, E., and Y. Weingarten, "An Overview of Operations, Administration, and Maintenance (OAM) Tools", RFC 7276, DOI 10.17487/RFC7276, June 2014, <https://www.rfc-editor.org/info/rfc7276>.

[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>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <https://www.rfc-editor.org/info/rfc8655>.

Authors' Addresses

Fabrice Theoleyre
CNRS
Building B
300 boulevard Sebastien Brant - CS 10413
Illkirch - Strasbourg 67400
FRANCE

Phone: +33 368 85 45 33
Email: theoleyre@unistra.fr
URI: http://www.theoleyre.eu

Georgios Z. Papadopoulos
IMT Atlantique
Office B00 - 102A
2 Rue de la Chataigneraie
Cesson-Sevigne - Rennes 35510
FRANCE

Phone: +33 299 12 70 04
Email: georgios.papadopoulos@imt-atlantique.fr

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com