

DetNet
Internet-Draft
Intended status: Standards Track
Expires: January 1, 2018

J. Korhonen, Ed.

L. Andersson
Y. Jiang
N. Finn
Huawei
B. Varga
J. Farkas
Ericsson
CJ. Bernardos
UC3M
T. Mizrahi
Marvell
L. Berger
LabN
June 30, 2017

DetNet Data Plane Encapsulation
draft-dt-detnet-dp-sol-01

Abstract

This document specifies Deterministic Networking data plane encapsulation solutions. The described data plane solutions can be applied over either IP or MPLS Packet Switched Networks.

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 <http://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 1, 2018.

Copyright Notice

Copyright (c) 2017 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 (<http://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	3
2. Terminology	4
2.1. Terms used in this document	4
2.2. Abbreviations	5
3. Requirements language	6
4. DetNet data plane overview	6
4.1. DetNet data plane encapsulation requirements	8
5. DetNet data plane solution	9
5.1. DetNet specific packet fields	9
5.2. DetNet encapsulation	9
5.2.1. PseudoWire-based data plane encapsulation	9
5.2.2. Native IPv6-based data plane encapsulation	11
5.3. DetNet flow identification for duplicate detection	12
5.3.1. PseudoWire encapsulation	13
5.3.2. Native IPv6 encapsulation	13
6. PREF specific considerations	13
6.1. PseudoWire-based data plane	13
6.1.1. Forwarder clarifications	13
6.1.2. Edge node processing clarifications	14
6.1.3. Relay node processing clarifications	16
6.2. Native IPv6-based data plane	17
7. Other DetNet data plane considerations	17
7.1. Class of Service	17
7.2. Quality of Service	18
7.3. Cross-DetNet flow resource aggregation	19
7.4. Bidirectional traffic	20
7.5. Layer 2 addressing and QoS Considerations	21
7.6. Interworking between PW- and IPv6-based encapsulations	21
8. Time synchronization	21
9. Management and control considerations	23
9.1. PW Label and IPv6 Flow Label assignment and distribution	23

9.2. Packet replication and elimination	23
9.3. Explicit paths	23
9.4. Congestion protection and latency control	23
9.5. Flow aggregation control	24
10. Security considerations	24
11. IANA considerations	24
12. Acknowledgements	24
13. References	25
13.1. Normative references	25
13.2. Informative references	27
Appendix A. Example of DetNet data plane operation	28
Appendix B. Example of pinned paths using IPv6	29
Authors' Addresses	29

1. Introduction

Deterministic Networking (DetNet) is a service that can be offered by a network to DetNet flows. DetNet provides these flows extremely low packet loss rates and assured maximum end-to-end delivery latency. General background and concepts of DetNet can be found in [I-D.ietf-detnet-architecture].

This document specifies the DetNet data plane. It defines how DetNet traffic is encapsulated at the network layer, and how DetNet-aware nodes can identify DetNet flows. Two data plane definitions are given.

- o PW-based: One solution is based on PseudoWires (PW) [RFC3985] and makes use of multi-segment pseudowires (MS-PW) [RFC6073] to map DetNet Relay and Edge Nodes [I-D.ietf-detnet-architecture] [I-D.ietf-detnet-dp-alt] to PW architecture. The PW-based data plane can be run over an MPLS [RFC4448] [RFC6658] Packet Switched Network (PSN).
- o Native-IP: The other solution is based on IP header fields, namely on the IPv6 Flow Label and a new DetNet Control Word extension header option. It is targeted for native IPv6 networks.

It is worth noting that while PWs are designed to work over IP PSNs this document describes a native-IP solution that operates without PWs. The primary reason for this is the benefit gained by enabling the use of a normal application stack, where transport protocols such as TCP or UDP are directly encapsulated in IP.

This document specifies the encapsulation for DetNet flows, including a DetNet Control Word (CW). Furthermore, it describes how DetNet flows are identified, how DetNet Relay and Edge nodes work, and how the Packet Replication and Elimination function (PREF) is implemented

with these two data plane solutions. This document does not define the associated control plane functions, or Operations, Administration, and Maintenance (OAM). It also does not specify traffic handling capabilities required to deliver congestion protection and latency control to DetNet flows as this is defined to be provided by the underlying MPLS or IP network.

2. Terminology

2.1. Terms used in this document

This document uses the terminology established in the DetNet architecture [I-D.ietf-detnet-architecture] and the DetNet Data Plane Solution Alternatives [I-D.ietf-detnet-dp-alt].

The following terms are also used in this document:

DA-T-PE	MPLS based DetNet edge node: a DetNet-aware PseudoWire Terminating Provider Edge (T-PE).
DA-S-PE	MPLS based DetNet relay node: a DetNet-aware PseudoWire Switching Provider Edge (S-PE).
T-Label	A label used to identify the LSP used to transport a DetNet flow across an MPLS PSN, e.g., a hop-by-hop label used between label switching routers (LSR).
S-Label	A DetNet node to DetNet node "service" label that is used between DA-*-PE devices.
PW Label	A PseudoWire label that is used to identify DetNet flow related PW Instances within a PE node.
Flow Label	IPv6 header field that is used to identify a DetNet flow (together with the source IP address field).
local-ID	An edge and relay node internal construct that uniquely identifies a DetNet flow. It may be used to select proper forwarding and/or DetNet specific service function.
PREF	A Packet Replication and Elimination Function (PREF) does the replication and elimination processing of DetNet flow packets in edge or relay nodes. The replication function is essentially the existing 1+1 protection mechanism. The elimination function reuses and extends the existing duplicate detection mechanism

to operate over multiple (separate) DetNet member flows of a DetNet compound flow.

2.2. Abbreviations

The following abbreviations used in this document:

AC	Attachment Circuit.
CE	Customer Edge equipment.
CoS	Class of Service.
CW	Control Word.
d-CW	DetNet Control Word.
DetNet	Deterministic Networking.
DF	DetNet Flow.
L2VPN	Layer 2 Virtual Private Network.
LSR	Label Switching Router.
MPLS	Multiprotocol Label Switching.
MPLS-TP	Multiprotocol Label Switching - Transport Profile.
MS-PW	Multi-Segment PseudoWire (MS-PW).
NSP	Native Service Processing.
OAM	Operations, Administration, and Maintenance.
PE	Provider Edge.
PREF	Packet Replication and Elimination Function.
PSN	Packet Switched Network.
PW	PseudoWire.
QoS	Quality of Service.
TSN	Time-Sensitive Network.

3. Requirements language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

4. DetNet data plane overview

This document describes how to use IP and/or MPLS to support a data plane method of flow identification and packet forwarding over layer-3. Two different cases are covered: (i) the inter-connect scenario, in which IEEE802.1 TSN is routed over a layer-3 network (i.e., to enlarge the layer-2 domain), and (ii) native connectivity between DetNet-aware end systems. Figure 1 illustrates an exemplary scenario.

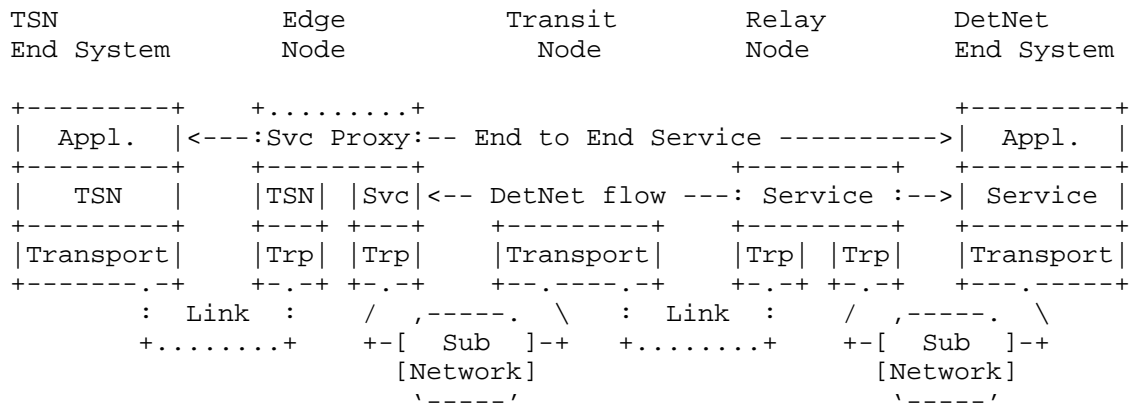


Figure 1: A simple DetNet enabled network architecture

Figure 2 illustrates how DetNet can provide services for IEEE 802.1 TSN end systems over a DetNet enabled network. The edge nodes insert and remove required DetNet data plane encapsulation. The 'X' in the edge and relay nodes represents a potential DetNet flow packet replication and elimination point. This conceptually parallels L2VPN services, and could leverage existing related solutions as discussed below.

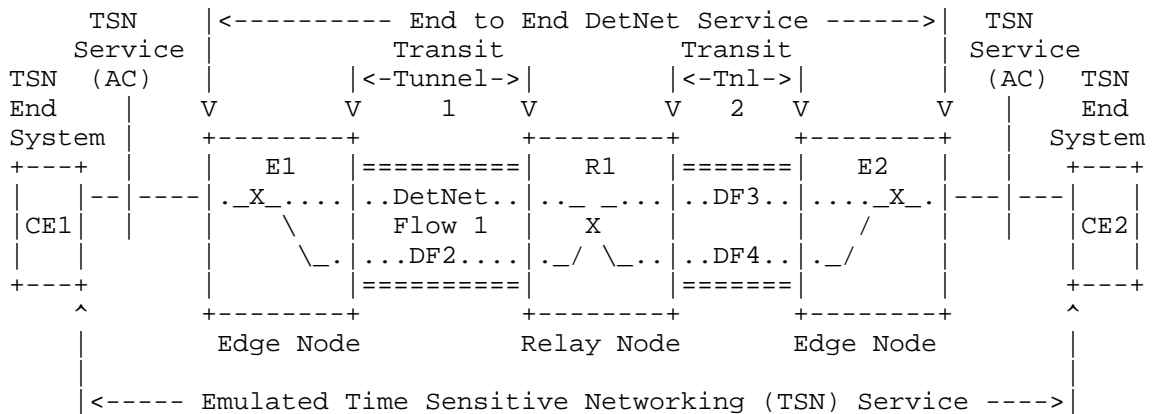


Figure 2: IEEE 802.1TSN over DetNet

Figure 3 illustrates how end to end PW-based DetNet service can be provided. In this case, the end systems are able to send and receive DetNet flows. For example, an end system sends data encapsulated in PseudoWire (PW) and in MPLS. Like earlier the 'X' in the end systems, edge and relay nodes represents potential DetNet flow packet replication and elimination points. Here the relay nodes may change the underlying transport, for example tunneling IP over MPLS, or simply interconnect network segments.

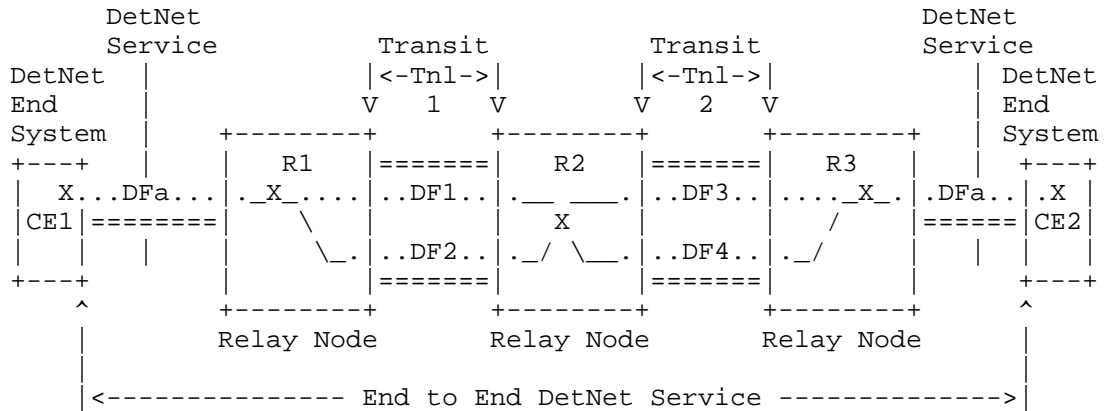


Figure 3: PW-Based Native DetNet

Figure 4 illustrates how end to end IP-based DetNet service can be provided. In this case, the end systems are able to send and receive DetNet flows. [Editor's note: TBD]

NOTE: This figures is TBD

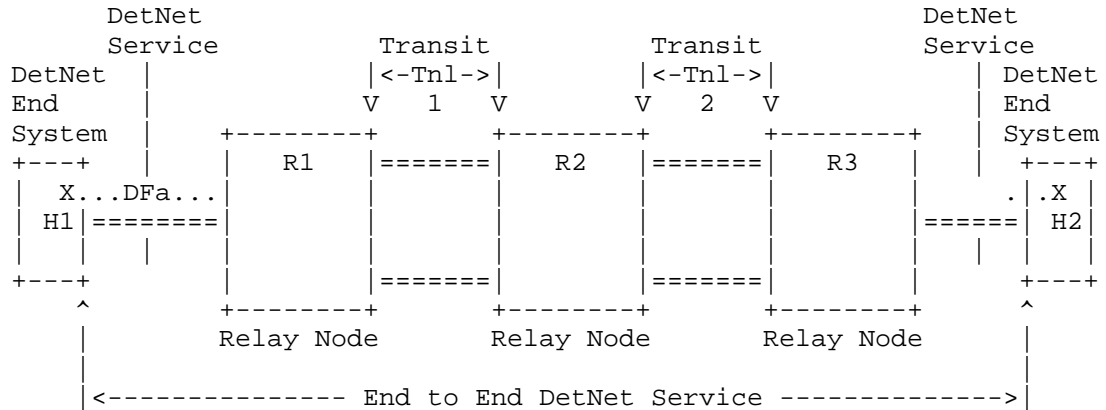


Figure 4: IP-Based Native DetNet

4.1. DetNet data plane encapsulation requirements

Two major groups of scenarios can be distinguished which require flow identification during transport:

1. DetNet function related scenarios:

- * Congestion protection and latency control: usage of allocated resources (queuing, policing, shaping).
- * Explicit routes: select/apply the flow specific path.
- * Service protection: recognize DetNet compound and member flows for replication and elimination.

2. OAM function related scenarios:

- * troubleshooting (e.g., identify misbehaving flows, etc.)
- * recognize flow(s) for analytics (e.g., increase counters, etc.)
- * correlate events with flows (e.g., volume above threshold, etc.)
- * etc.

Each node (edge, relay and transit) use a local-ID of the DetNet-(compound)-flow in order to accomplish its role during transport.

Recognizing the DetNet flow is more relaxed for edge and relay nodes, as they are fully aware of both the DetNet service and transport layers. The primary DetNet role of intermediate transport nodes is limited to ensuring congestion protection and latency control for the above listed DetNet functions.

The DetNet data plane allows for the aggregation of DetNet flows, e.g., via MPLS hierarchical LSPs, to improved scaling. When DetNet flows are aggregated, transit nodes may have limited ability to provide service on per-flow DetNet identifiers. Therefore, identifying each individual DetNet flow on a transit node may not be achieved in some network scenarios, but DetNet service can still be assured in these scenarios through resource allocation and control.

On each node dealing with DetNet flows, a local-ID is assumed to determine what local operation a packet goes through. Therefore, local-IDs MUST be unique on each edge and relay nodes. Local-ID MUST be unambiguously bound to the DetNet flow.

5. DetNet data plane solution

5.1. DetNet specific packet fields

The DetNet data plane encapsulation should include two DetNet specific information element in each packet of a DetNet flow: (1) flow identification and (2) sequence number.

The DetNet data plane encapsulation may consists further elements used for overlay tunneling, to distinguish between DetNet member flows of the same DetNet compound flow or to support OAM functions.

5.2. DetNet encapsulation

This document specifies two encapsulations for the DetNet data plane: (1) PseudoWire (PW) for MPLS PSN and (2) native IPv6 based encapsulation for IP PSN.

5.2.1. PseudoWire-based data plane encapsulation

Figure 5 illustrates a DetNet PW encapsulation over an MPLS PSN. The PW-based encapsulation of the DetNet flows fits perfectly for the Layer-2 interconnect deployment cases (see Figure 2). Furthermore, end to end DetNet service i.e., native DetNet deployment (see Figure 3) is also possible if DetNet-aware end systems are capable of initiating and termination MPLS encapsulated PWs. It is also possible use the same encapsulation format with a Packet PW over MPLS [RFC6658]. Transport of IP encapsulated DetNet flows, see Section 5.2.2, over DetNet PWs is also possible. Interworking

between PW- and IPv6-based encapsulations is discussed further in Section 7.6.

The PW-based DetNet data plane encapsulation consists of:

- o DetNet control word (d-CW) containing sequencing information for packet replication and duplicate elimination purposes. There is a separate sequence number space for each DetNet flow.
- o PseudoWire Label (PW Label) that is a standard PW label identifying a DetNet flow and a PW Instance within a (DA-)T-PE or (DA-)S-PE device.
- o An optional S-Label that represents DetNet Service LSP used between (DA-)T-PE or (DA-)S-PE nodes. One possible use of an S-Label is to identify the different DetNet member flows used to provide protection to a DetNet composite flow, perhaps even when both LSPs appear on the same link for some reason.
- o MPLS transport LSP label(s) (T-label) which may be a hop-by-hop label used between LSRs.

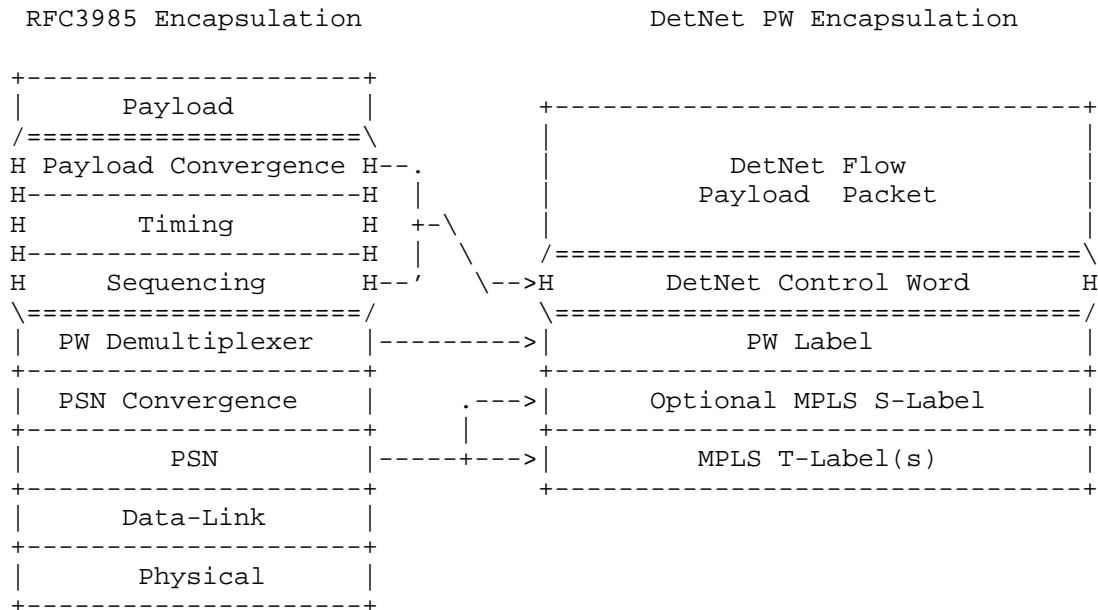


Figure 5: Encapsulation of a DetNet flow in a PW with MPLS(-TP) PSN

The DetNet control word (d-CW) is identical to the control word defined for Ethernet over MPLS networks in [RFC4448]. The DetNet control word is illustrated in Figure 6.

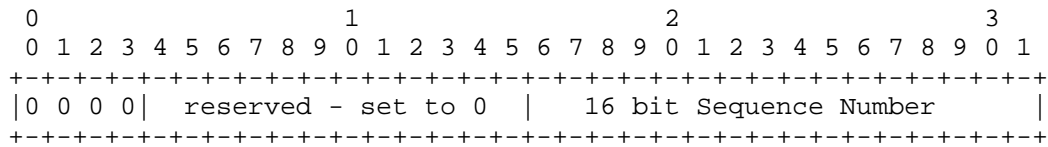


Figure 6: DetNet Control Word

5.2.2. Native IPv6-based data plane encapsulation

Figure 7 illustrates a DetNet native IPv6 encapsulation. The native IPv6 encapsulation is meant for end to end Detnet service use cases, where the end stations are DetNet-aware (see Figure 4). Technically it is possible to use the IPv6 encapsulation to tunnel any traffic over a DetNet enabled network, which would make native IPv6 encapsulation also a valid data plane choice for an interconnect use case (see Figure 2).

The native IPv6-based DetNet data plane encapsulation consists of:

- o IPv6 header as the transport protocol.
- o IPv6 header Flow Label that is used to help to identify a DetNet flow (i.e., roughly an equivalent to the PW Label). A Flow Label together with the IPv6 source address uniquely identifies a DetNet flow.
- o DetNet Control Word IPv6 Destination Option containing sequencing information for packet replication and duplicate elimination function (PREF) purposes. The DetNet Destination Option is equivalent to the DetNet Control Word.

A DetNet-aware end station (a host) or an intermediate node initiating an IPv6 packet is responsible for setting the Flow Label, adding the required DetNet Destination Option, and possibly adding a routing header such as the segment routing option (for pre-defined paths [I-D.ietf-6man-segment-routing-header]). The payload of the native IPv6 encapsulation is any payload protocol that can be identified using the Next Header field either in the IPv6 packet header or in the last IPv6 extension header.

A DetNet-aware end station (a host) or an intermediate node receiving an IPv6 packet destined to it and containing a DetNet Destination

Option does the appropriate processing of the packet. This may involve packet duplication and elimination (PREF processing), terminating a tunnel or delivering the packet to the upper layers/Applications.

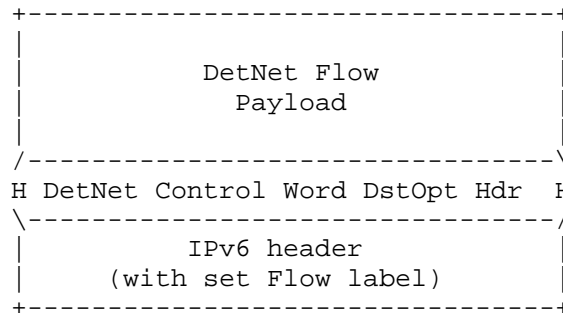


Figure 7: Encapsulation of a native IPv6 DetNet flow

A DetNet flow must carry sequencing information for packet replication and elimination function (PREF) purposes. This document specifies a new IPv6 Destination Option: the DetNet Destination Option for that purpose. The format of the option is illustrated in Figure 8.

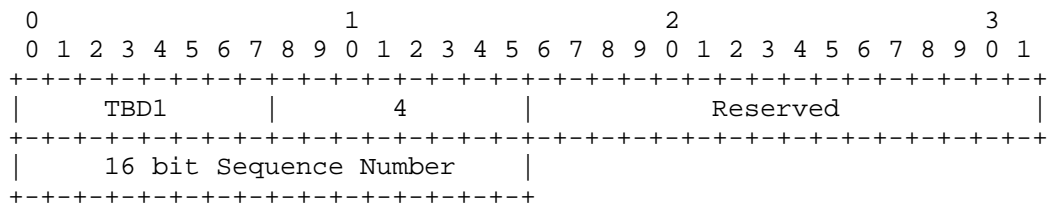


Figure 8: DetNet Destination Option

The Option Type for the DetNet Destination Option is set to TBD1. [To be removed from the final version of the document: The Option Type MUST have the two most significant bits set to 10b]

5.3. DetNet flow identification for duplicate detection

Duplicate elimination depends on flow identification. Mapping between packet fields and Local-ID may impact the implementation of duplicate elimination.

5.3.1. PseudoWire encapsulation

RFC3985 Section 5.2.1. describes PW sequencing provides a duplicate detection service among other things. This specification clarifies this definition as follows:

DetNet flows that need to undergo PREF processing MUST have the same PW Label when they arrive at the DA-*-PE node.

From the label stack processing point of view receiving the same label from multiple sources is analogous to Fast Reroute backup tunnel behavior [RFC4090]. The PW Label for a DetNet flow can be different on each PW segment.

5.3.2. Native IPv6 encapsulation

The DetNet flow identification is based on the IPv6 Flow Label and the source address combination. The two fields uniquely identify the end to end native IPv6 encapsulated DetNet flow. Obviously, the identification fails if any intermediate node modifies either the source address or the Flow Label.

6. PREF specific considerations

This section applies equally to DetNet flows transported via IPv6 and MPLS. While flow identification and some header related processing will differ between the two, the considerations covered in this section are common to both.

6.1. PseudoWire-based data plane

6.1.1. Forwarder clarifications

The DetNet specific new functionality in an edge or relay node processing is the packet replication and duplication elimination function (PREF). This function is a part of the DetNet-aware "extended" forwarder. The PREF processing is triggered by the received packet of a DetNet flow. Basically the forwarding entry has to be extended with a "PREF enabled" boolean configuration switch that is associated with the normal forwarding actions (e.g., in case of MPLS a swap, push, pop, ..). The output of the PREF elimination function is always a single packet. The output of the PREF replication function is always one or more packets (i.e., 1:M replication). The replicated packets MUST share the same DetNet control word sequence number.

The complex part of the DetNet PREF processing is tracking the history of received packets for multiple DetNet member flows. These

ingress DetNet member flows (to a node) MUST have the same local-ID if they belong to the same DetNet-(compound)-flow and share the same sequence number counter and the history information.

The edge and relay node internal procedures of the PREF are implementation specific. The order of a packet elimination or replication is out of scope in this specification. However, care should be taken that the replication function does not actually loopback packets as "replicas". Looped back packets include artificial delay when the node that originally initiated the packet receives it again. Also, looped back packets may make the network condition to look healthier than it actually is (in some cases link failures are not reflected properly because looped back packets make the situation appear better than it actually is).

6.1.2. Edge node processing clarifications

The DetNet data plane solution overloads the edge node with DetNet Edge Node functions. Edge nodes are also aware of DetNet flows and may need to operate upon those. Figure 9 illustrates the overall edge device functions. The figure shows both physical attachment circuit (AC) (e.g., Ethernet [RFC4448]) connecting to the edge node, and a packet service connecting to the edge node via an embedded router function (similarly as described e.g., in [RFC6658]). Whether traffic flow from a client AC and PSN tunnel receives DetNet specific treatment is up to a local configuration and policy.

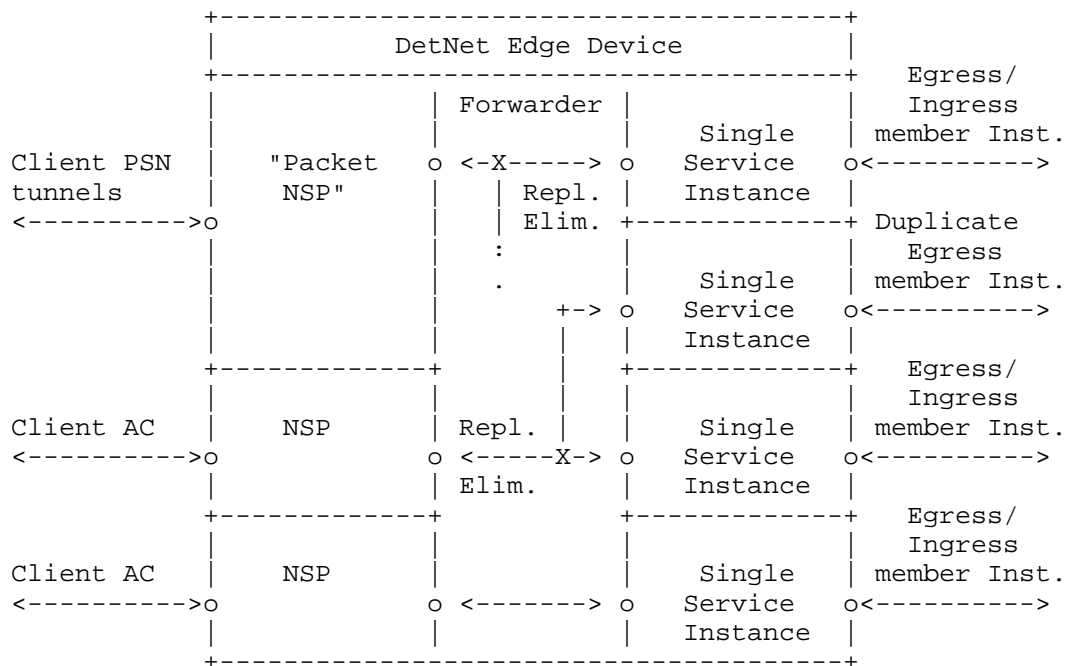


Figure 9: DetNet Edge Node processing

An edge node participates to the packet replication and duplication elimination. Required processing is done within an extended forwarder function. In the case the native service processing (NSP) is IEEE 802.1CB [IEEE8021CB] capable, the packet replication and duplicate elimination MAY entirely be done in the NSP and bypassing the DetNet flow encapsulation and logic entirely, and thus is able to operate over unmodified implementation and deployment. The NSP approach works only between edge nodes and cannot make use of relay nodes (see Section 6.1.3).

The DetNet-aware extended forwarder selects the egress DetNet member flow based on the DetNet forwarding rules. In both "normal AC" and "Packet AC" cases there may be no DetNet encapsulation header available yet as it is the case with relay nodes (see Section 6.1.3). It is the responsibility of the extended forwarder within the edge node to push the DetNet specific encapsulation (if not already present) to the packet before forwarding it to the appropriate egress DetNet member flow instance(s). The extended forwarder MAY copy the sequencing information from the native DetNet packet into the DetNet sequence number field and vice versa. If there is no existing sequencing information available in the native packet or the

forwarder chose not to copy it from the native packet, then the extended forwarder MUST maintain a sequence number counter for each DetNet flow (indexed by the DetNet flow identification).

6.1.3. Relay node processing clarifications

The DetNet data plane solution overloads a relay node with DetNet Relay node functions. Relay node is aware of DetNet flows and may operate upon those. Figure 10 illustrates the overall DetNet relay device functions.

A DetNet Relay node participates to the packet replication and duplication elimination. This processing is done within an extended forwarder function. Whether an ingress DetNet member flow receives DetNet specific processing depends on how the forwarding is programmed. For some DetNet member flows the relay node can act as a normal relay node and for some apply the DetNet specific processing (i.e., PREF). It is also possible to treat the relay node as a transit node, see Section 7.3. Again, this is entirely up to how the forwarding has been programmed.

The DetNet-aware forwarder selects the egress DetNet member flow segment based on the flow identification. The mapping of ingress DetNet member flow segment to egress DetNet member flow segment may be statically or dynamically configured. Additionally the DetNet-aware forwarder does duplicate frame elimination based on the flow identification and the sequence number combination. The packet replication is also done within the DetNet-aware forwarder. During elimination and the replication process the sequence number of the DetNet member flow MUST be preserved and copied to the egress DetNet member flow.

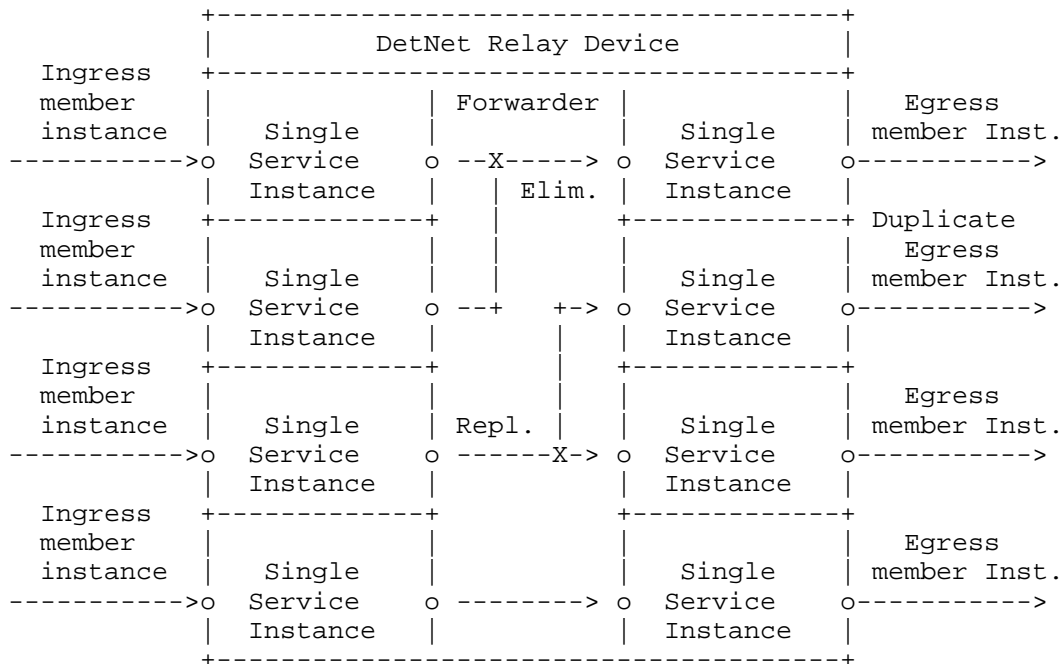


Figure 10: DetNet Relay Node processing

6.2. Native IPv6-based data plane

[Editor's note: this section is TBD.]

7. Other DetNet data plane considerations

7.1. Class of Service

Class and quality of service, i.e., CoS and QoS, are terms that are often used interchangeably and confused. In the context of DetNet, CoS is used to refer to mechanisms that provide traffic forwarding treatment based on aggregate group basis and QoS is used to refer to mechanisms that provide traffic forwarding treatment based on a specific DetNet flow basis. Examples of existing network level CoS mechanisms include DiffServ which is enabled by IP header differentiated services code point (DSCP) field [RFC2474] and MPLS label traffic class field [RFC5462], and at Layer-2, by IEEE 802.1p priority code point (PCP).

CoS for DetNet flows carried in PWs and MPLS is provided using the existing MPLS Differentiated Services (DiffServ) architecture

[RFC3270]. Both E-LSP and L-LSP MPLS DiffServ modes MAY be used to support DetNet flows. The Traffic Class field (formerly the EXP field) of an MPLS label follows the definition of [RFC5462] and [RFC3270]. The Uniform, Pipe, and Short Pipe DiffServ tunneling and TTL processing models are described in [RFC3270] and [RFC3443] and MAY be used for MPLS LSPs supporting DetNet flows. MPLS ECN MAY also be used as defined in ECN [RFC5129] and updated by [RFC5462].

CoS for DetNet flows carried in IPv6 is provided using the standard differentiated services code point (DSCP) field [RFC2474] and related mechanisms. The 2-bit explicit congestion notification (ECN) [RFC3168] field MAY also be used.

One additional consideration for DetNet nodes which support CoS services is that they MUST ensure that the CoS service classes do not impact the congestion protection and latency control mechanisms used to provide DetNet QoS. This requirement is similar to requirement for MPLS LSRs to that CoS LSPs do not impact the resources allocated to TE LSPs via [RFC3473].

7.2. Quality of Service

Quality of Service (QoS) mechanisms for flow specific traffic treatment typically includes a guarantee/agreement for the service, and allocation of resources to support the service. Example QoS mechanisms include discrete resource allocation, admission control, flow identification and isolation, and sometimes path control, traffic protection, shaping, policing and remarking. Example protocols that support QoS control include Resource ReSerVation Protocol (RSVP) [RFC2205] (RSVP) and RSVP-TE [RFC3209] and [RFC3473]. The existing MPLS mechanisms defined to support CoS [RFC3270] can also be used to reserve resources for specific traffic classes.

In addition to path pinning and packet replication and elimination, described in Section 5 above, DetNet provides zero congestion loss and bounded latency and jitter. As described in [I-D.ietf-detnet-architecture], there are different mechanisms that maybe used separately or in combination to deliver a zero congestion loss service. These mechanisms are provided by the either the MPLS or IP layers, and may be combined with the mechanisms defined by the underlying network layer such as 802.1TSN.

A baseline set of QoS capabilities for DetNet flows carried in PWs and MPLS can provided by MPLS with Traffic Engineering (MPLS-TE) [RFC3209] and [RFC3473]. TE LSPs can also support explicit routes (path pinning). Current service definitions for packet TE LSPs can be found in "Specification of the Controlled Load Quality of Service", [RFC2211], "Specification of Guaranteed Quality of

Service", [RFC2212], and "Ethernet Traffic Parameters", [RFC6003]. Additional service definitions are expected in future documents to support the full range of DetNet services. In all cases, the existing label-based marking mechanisms defined for TE-LSPs and even E-LSPs are used to support the identification of flows requiring DetNet QoS.

QoS for DetNet flows carried in IPv6 MUST be provided locally by the DetNet-aware hosts and routers supporting DetNet flows. Such support will leverage the underlying network layer such as 802.1TSN. The traffic control mechanisms used to deliver QoS for IP encapsulated DetNet flows are expected to be defined in a future document. From an encapsulation perspective, and as defined in Section 5.2.2, the combination of the Flow Label together with the IP source address uniquely identifies a DetNet flow.

Packets that are marked with a DetNet Class of Service value, but that have not been the subject of a completed reservation, can disrupt the QoS offered to properly reserved DetNet flows by using resources allocated to the reserved flows. Therefore, the network nodes of a DetNet network SHOULD:

- o Defend the DetNet QoS by discarding or remarking (to a non-DetNet CoS) packets received that are not the subject of a completed reservation.
- o Not use a DetNet reserved resource, e.g. a queue or shaper reserved for DetNet flows, for any packet that does not carry a DetNet Class of Service marker.

7.3. Cross-DetNet flow resource aggregation

The ability to aggregate individual flows, and their associated resource control, into a larger aggregate is an important technique for improving scaling of control in the data, management and control planes. This document identifies the traffic identification related aspects of aggregation of DetNet flows. The resource control and management aspects of aggregation (including the queuing/shaping/policing implications) will be covered in other documents. The data plane implications of aggregation are independent for PW/MPLS and IP encapsulated DetNet flows.

DetNet flows transported via MPLS can leverage MPLS-TE's existing support for hierarchical LSPs (H-LSPs), see [RFC4206]. H-LSPs are typically used to aggregate control and resources, they may also be used to provide OAM or protection for the aggregated LSPs. Arbitrary levels of aggregation naturally fall out of the definition for hierarchy and the MPLS label stack [RFC3032]. DetNet nodes which

support aggregation (LSP hierarchy) map one or more LSPs (labels) into and from an H-LSP. Both carried LSPs and H-LSPs may or may not use the TC field, i.e., L-LSPs or E-LSPs. Such nodes will need to ensure that traffic from aggregated LSPs are placed (shaped/policed/enqueued) onto the H-LSPs in a fashion that ensures the required DetNet service is preserved.

DetNet flows transported via IP have more limited aggregation options, due to the available traffic flow identification fields of the IP solution. One available approach is to manage the resources associated with a DSCP identified traffic class and to map (remark) individually controlled DetNet flows onto that traffic class. This approach also requires that nodes support aggregation ensure that traffic from aggregated LSPs are placed (shaped/policed/enqueued) in a fashion that ensures the required DetNet service is preserved.

In both the MPLS and IP cases, additional details of the traffic control capabilities needed at a DetNet-aware node may be covered in the new service descriptions mentioned above or in separate future documents. Management and control plane mechanisms will also need to ensure that the service required on the aggregate flow (H-LSP or DSCP) are provided, which may include the discarding or remarking mentioned in the previous sections.

7.4. Bidirectional traffic

Some DetNet applications generate bidirectional traffic. Using MPLS definitions [RFC5654] there are associated bidirectional flows, and co-routed bidirectional flows. MPLS defines a point-to-point associated bidirectional LSP as consisting of two unidirectional point-to-point LSPs, one from A to B and the other from B to A, which are regarded as providing a single logical bidirectional transport path. This would be analogous of standard IP routing, or PWs running over two reciprocal unidirection LSPs. MPLS defines a point-to-point co-routed bidirectional LSP as an associated bidirectional LSP which satisfies the additional constraint that its two unidirectional component LSPs follow the same path (in terms of both nodes and links) in both directions. An important property of co-routed bidirectional LSPs is that their unidirectional component LSPs share fate. In both types of bidirectional LSPs, resource allocations may differ in each direction. The concepts of associated bidirectional flows and co-routed bidirectional flows can be applied to DetNet flows as well whether IPv6 or MPLS is used.

While the IPv6 and MPLS data planes must support bidirectional DetNet flows, there are no special bidirectional features with respect to the data plane other than need for the two directions take the same paths. Fate sharing and associated vs co-routed bidirectional flows

can be managed at the control level. Note, that there is no stated requirement for bidirectional DetNet flows to be supported using the same IPv6 Flow Labels or MPLS Labels in each direction. Control mechanisms will need to support such bidirectional flows for both IPv6 and MPLS, but such mechanisms are out of scope of this document. An example control plane solution for MPLS can be found in [RFC7551].

7.5. Layer 2 addressing and QoS Considerations

The Time-Sensitive Networking (TSN) Task Group of the IEEE 802.1 Working Group have defined (and are defining) a number of amendments to IEEE 802.1Q [IEEE8021Q] that provide zero congestion loss and bounded latency in bridged networks. IEEE 802.1CB [IEEE8021CB] defines packet replication and elimination functions that should prove both compatible with and useful to, DetNet networks.

As is the case for DetNet, a Layer 2 network node such as a bridge may need to identify the specific DetNet flow to which a packet belongs in order to provide the TSN/DetNet QoS for that packet. It also will likely need a CoS marking, such as the priority field of an IEEE Std 802.1Q VLAN tag, to give the packet proper service.

Although the flow identification methods described in IEEE 802.1CB [IEEE8021CB] are flexible, and in fact, include IP 5-tuple identification methods, the baseline TSN standards assume that every Ethernet frame belonging to a TSN stream (i.e. DetNet flow) carries a multicast destination MAC address that is unique to that flow within the bridged network over which it is carried. Furthermore, IEEE 802.1CB [IEEE8021CB] describes three methods by which a packet sequence number can be encoded in an Ethernet frame.

Ensuring that the proper Ethernet VLAN tag priority and destination MAC address are used on a DetNet/TSN packet may require further clarification of the customary L2/L3 transformations carried out by routers and edge label switches. Edge nodes may also have to move sequence number fields among Layer 2, PW, and IPv6 encapsulations.

7.6. Interworking between PW- and IPv6-based encapsulations

[Editor's note: add considerations for interworking between PW-based and native IPv6-based DetNet encapsuations.]

8. Time synchronization

[Editor's note: describe a bit of issues and deployment considerations related to time-synchronization within DetNet. Refer to DT discussion and the slides that summarize different approaches

and rough synchronization performance numbers. Finally, scope time-synchronization solution outside data plane.]

When DetNet is used, there is an underlying assumption that the application(s) require clock synchronization such as the Precision Time Protocol (PTP) [IEEE1588]. The relay nodes may or may not utilize clock synchronization in order to provide zero congestion loss and controlled latency delivery. In either case, there are a few possible approaches of how synchronization protocol packets are forwarded and handled by the network:

- o PTP packets can be sent either as DetNet flows or as high-priority best effort packets. Using DetNet for PTP packets requires careful consideration to prevent unwanted interactions between clock-synchronized network nodes and the packets that synchronize the clocks.
- o PTP packets are sent as a normal DetNet flow through network nodes that are not time-synchronized: in this approach PTP traffic is forwarded as a DetNet flow, and as such it is forwarded in a way that allows a low delay variation. However, since intermediate nodes do not take part in the synchronization protocol, this approach provides a relatively low degree of accuracy.
- o PTP with on-path support: in this approach PTP packets are sent as ordinary or as DetNet flows, and intermediate nodes take part in the protocol as Transparent Clocks or Boundary Clocks [IEEE1588]. The on-path PTP support by intermediate nodes provides a higher degree of accuracy than the previous approach. The actual accuracy depends on whether all intermediate nodes are PTP-capable, or only a subset of them.
- o Time-as-a-service: in this approach accurate time is provided as-a-service to the DetNet source and destination, as well as the intermediate nodes. Since traffic between the source and destination is sent over a provider network, if the provider supports time-as-a-service, then accurate time can be provided to both the source and the destination of DetNet traffic. This approach can potentially provide the highest degree of accuracy.

It is expected that the latter approach will be the most common one, as it provides the highest degree of accuracy, and creates a layer separation between the DetNet data and the synchronization service.

It should be noted that in all four approaches it is not recommended to use replication and elimination for synchronization packets; the replication/elimination approach may in some cases reduce the

synchronization accuracy, since the observed path delay will be bivalent.

9. Management and control considerations

While management plane and control planes are traditionally considered separately, from the Data Plane perspective there is no practical difference based on the origin of flow provisioning information. This document therefore does not distinguish between information provided by a control plane protocol, e.g., RSVP-TE [RFC3209] and [RFC3473], or by a network management mechanisms, e.g., RestConf [RFC8040] and YANG [RFC7950].

[Editor's note: This section is a work in progress. discuss here what kind of enhancements are needed for DetNet and specifically for PREF and DetNet zero congest loss and latency control. Need to cover both traffic control (queuing) and connection control (control plane).]

9.1. PW Label and IPv6 Flow Label assignment and distribution

The PW label distribution follows the same mechanisms specified for MS-PW [RFC6073]. The details of the control plane protocol solution required for the label distribution and the management of the label number space are out of scope of this document.

The IPv6 Flow Label distribution and the label number space are out of scope of this document. However, it should be noted that the combination of the IPv6 source address and the IPv6 Flow Label is assumed to be unique within the DetNet-enabled network. Therefore, as long as each node is able to assign unique Flow Labels for the source address(es) it is using the DetNet-enabled network wide flow identification uniqueness is guaranteed.

9.2. Packet replication and elimination

The control plane protocol solution required for managing the PREF processing is outside the scope of this document.

9.3. Explicit paths

[TBD: based on MPLS TE and SR.]

9.4. Congestion protection and latency control

[TBD]

9.5. Flow aggregation control

[TBD]

10. Security considerations

The security considerations of DetNet in general are discussed in [I-D.ietf-detnet-architecture] and [I-D.sdt-detnet-security]. Other security considerations will be added in a future version of this draft.

11. IANA considerations

TBD.

12. Acknowledgements

The author(s) ACK and NACK.

The following people were part of the DetNet Data Plane Solution Design Team:

Jouni Korhonen

Janos Farkas

Norman Finn

Balazs Varga

Loa Andersson

Tal Mizrahi

David Mozes

Yuanlong Jiang

Carlos J. Bernardos

The DetNet chairs serving during the DetNet Data Plane Solution Design Team:

Lou Berger

Pat Thaler

13. References

13.1. 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, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC2211] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, DOI 10.17487/RFC2211, September 1997, <<http://www.rfc-editor.org/info/rfc2211>>.
- [RFC2212] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, DOI 10.17487/RFC2212, September 1997, <<http://www.rfc-editor.org/info/rfc2212>>.
- [RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <<http://www.rfc-editor.org/info/rfc2474>>.
- [RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <<http://www.rfc-editor.org/info/rfc3032>>.
- [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <<http://www.rfc-editor.org/info/rfc3168>>.
- [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, <<http://www.rfc-editor.org/info/rfc3209>>.
- [RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, DOI 10.17487/RFC3270, May 2002, <<http://www.rfc-editor.org/info/rfc3270>>.

- [RFC3443] Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks", RFC 3443, DOI 10.17487/RFC3443, January 2003, <<http://www.rfc-editor.org/info/rfc3443>>.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <<http://www.rfc-editor.org/info/rfc3473>>.
- [RFC3985] Bryant, S., Ed. and P. Pate, Ed., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, DOI 10.17487/RFC3985, March 2005, <<http://www.rfc-editor.org/info/rfc3985>>.
- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, DOI 10.17487/RFC4206, October 2005, <<http://www.rfc-editor.org/info/rfc4206>>.
- [RFC4448] Martini, L., Ed., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, DOI 10.17487/RFC4448, April 2006, <<http://www.rfc-editor.org/info/rfc4448>>.
- [RFC5129] Davie, B., Briscoe, B., and J. Tay, "Explicit Congestion Marking in MPLS", RFC 5129, DOI 10.17487/RFC5129, January 2008, <<http://www.rfc-editor.org/info/rfc5129>>.
- [RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <<http://www.rfc-editor.org/info/rfc5462>>.
- [RFC6003] Papadimitriou, D., "Ethernet Traffic Parameters", RFC 6003, DOI 10.17487/RFC6003, October 2010, <<http://www.rfc-editor.org/info/rfc6003>>.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, DOI 10.17487/RFC6073, January 2011, <<http://www.rfc-editor.org/info/rfc6073>>.

- [RFC6658] Bryant, S., Ed., Martini, L., Swallow, G., and A. Malis, "Packet Pseudowire Encapsulation over an MPLS PSN", RFC 6658, DOI 10.17487/RFC6658, July 2012, <<http://www.rfc-editor.org/info/rfc6658>>.
- [RFC7510] Xu, X., Sheth, N., Yong, L., Callon, R., and D. Black, "Encapsulating MPLS in UDP", RFC 7510, DOI 10.17487/RFC7510, April 2015, <<http://www.rfc-editor.org/info/rfc7510>>.

13.2. Informative references

- [I-D.ietf-6man-segment-routing-header]
Previdi, S., Filsfils, C., Raza, K., Leddy, J., Field, B., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Matsushima, S., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., Lebrun, D., Steinberg, D., and R. Raszuk, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-06 (work in progress), March 2017.
- [I-D.ietf-detnet-architecture]
Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-02 (work in progress), June 2017.
- [I-D.ietf-detnet-dp-alt]
Korhonen, J., Farkas, J., Mirsky, G., Thubert, P., Zhuangyan, Z., and L. Berger, "DetNet Data Plane Protocol and Solution Alternatives", draft-ietf-detnet-dp-alt-00 (work in progress), October 2016.
- [I-D.sdt-detnet-security]
Mizrahi, T., Grossman, E., Hacker, A., Das, S., "Deterministic Networking (DetNet) Security Considerations, draft-sdt-detnet-security, work in progress", 2017.
- [IEEE1588]
IEEE, "IEEE 1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems Version 2", 2008.
- [IEEE8021CB]
Finn, N., "Draft Standard for Local and metropolitan area networks - Seamless Redundancy", IEEE P802.1CB /D2.1 P802.1CB, December 2015, <<http://www.ieee802.org/1/files/private/cb-drafts/d2/802-1CB-d2-1.pdf>>.

- [IEEE8021Q] IEEE 802.1, "Standard for Local and metropolitan area networks--Bridges and Bridged Networks (IEEE Std 802.1Q-2014)", 2014, <<http://standards.ieee.org/about/get/>>.
- [RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, DOI 10.17487/RFC2205, September 1997, <<http://www.rfc-editor.org/info/rfc2205>>.
- [RFC4023] Worster, T., Rekhter, Y., and E. Rosen, Ed., "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", RFC 4023, DOI 10.17487/RFC4023, March 2005, <<http://www.rfc-editor.org/info/rfc4023>>.
- [RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <<http://www.rfc-editor.org/info/rfc4090>>.
- [RFC5654] Niven-Jenkins, B., Ed., Brungard, D., Ed., Betts, M., Ed., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, DOI 10.17487/RFC5654, September 2009, <<http://www.rfc-editor.org/info/rfc5654>>.
- [RFC7551] Zhang, F., Ed., Jing, R., and R. Gandhi, Ed., "RSVP-TE Extensions for Associated Bidirectional Label Switched Paths (LSPs)", RFC 7551, DOI 10.17487/RFC7551, May 2015, <<http://www.rfc-editor.org/info/rfc7551>>.
- [RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <<http://www.rfc-editor.org/info/rfc7950>>.
- [RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <<http://www.rfc-editor.org/info/rfc8040>>.

Appendix A. Example of DetNet data plane operation

[Editor's note: Add a simplified example of DetNet data plane and how labels etc work in the case of MPLS-based PSN and utilizing PREF. The figure is subject to change depending on the further DT decisions on the label handling..]

Appendix B. Example of pinned paths using IPv6

TBD.

Authors' Addresses

Jouni Korhonen (editor)

Email: jouni.nospam@gmail.com

Loa Andersson
Huawei

Email: loa@pi.nu

Yuanlong Jiang
Huawei

Email: jiangyuanlong@huawei.com

Norman Finn
Huawei
3101 Rio Way
Spring Valley, CA 91977
USA

Email: norman.finn@mail01.huawei.com

Balazs Varga
Ericsson
Konyves Kalman krt. 11/B
Budapest 1097
Hungary

Email: balazs.a.varga@ericsson.com

Janos Farkas
Ericsson
Konyves Kalman krt. 11/B
Budapest 1097
Hungary

Email: janos.farkas@ericsson.com

Carlos J. Bernardos
Universidad Carlos III de Madrid
Av. Universidad, 30
Leganes, Madrid 28911
Spain

Phone: +34 91624 6236
Email: cjbc@it.uc3m.es
URI: <http://www.it.uc3m.es/cjbc/>

Tal Mizrahi
Marvell
6 Hamada st.
Yokneam
Israel

Email: talmi@marvell.com

Lou Berger
LabN Consulting, L.L.C.

Email: lberger@labn.net