DetNet MPLS Data Plane Encapsulation
draft-ietf-detnet-dp-sol-mpls-02

Abstract

   This document specifies the Deterministic Networking data plane when
   operating over an 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 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 11, 2019.

Table of Contents

1. Introduction

   Deterministic Networking (DetNet) is a service that can be offered by
   a network to DetNet flows. DetNet provides these flows with a 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].

   The DetNet Architecture decomposes the DetNet related data plane
   functions into two sub-layers: a service sub-layer and a forwarding
   sub-layer. The service sub-layer is used to provide DetNet service
   protection and reordering. The forwarding sub-layer is used to
   provides congestion protection (low loss, assured latency, and
   limited reordering) leveraging MPLS Traffic Engineering mechanisms.

   This document specifies the DetNet data plane operation and the on-
   wire encapsulation of DetNet flows over an MPLS-based Packet Switched
   Network (PSN). The specified encapsulation provides the building
   blocks to enable the DetNet service and forwarding sub-layer
   functions and supports flow identification as described in the DetNet
   Architecture. As part of the service sub-layer functions, this
   document describes DetNet node data plane operation. It also
   describes the function and operation of the Packet Replication (PRF)
   Packet Elimination (PEF) and Packet Ordering (POF) functions with an
   MPLS data plane. It also describes an MPLS-based DetNet forwarding
   sub-layer that eliminates (or reduces) contention loss and provides
   bounded latency for DetNet flows.

   MPLS encapsulated DetNet flows can be carried over network
   technologies that can provide the DetNet required level of service.
   This document defines examples of such, specifically carrying DetNet
   MPLS flows over IEEE 802.1 TSN sub-networks, and over DetNet IP PSN.

   The intent is for this document to support different traffic types
   being mapped over DetNet MPLS, but this is out side the scope of this

document. An example of such can be found in
[I-D.ietf-detnet-dp-sol-ip]. This document also allows for, but does
not define, associated controller plane and Operations,
Administration, and Maintenance (OAM) functions.

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 reader is
assumed to be familiar with that document and its terminology.

The following terminology is introduced in this document:

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.

d-CW   A DetNet Control Word (d-CW) is used for sequencing and
              identifying duplicate packets of a DetNet flow at the
              DetNet service sub-layer.

2.2. Abbreviations

The following abbreviations are used in this document:

AC   Attachment Circuit.

CE   Customer Edge equipment.

CoS   Class of Service.

CW   Control Word.

DetNet   Deterministic Networking.

DF   DetNet Flow.

DN-IWF   DetNet Inter-Working Function.

L2   Layer 2.

L2VPN   Layer 2 Virtual Private Network.

L3   Layer 3.

LSR   Label Switching Router.

MPLS   Multiprotocol Label Switching.

MPLS-TE   Multiprotocol Label Switching - Traffic Engineering.

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.

PEF   Packet Elimination Function.

PRF   Packet Replication Function.

PREOF   Packet Replication, Elimination and Ordering Functions.

POF   Packet Ordering Function.

PSN   Packet Switched Network.

PW   PseudoWire.

QoS   Quality of Service.

S-PE   Switching Provider Edge.

T-PE   Terminating Provider Edge.

TSN   Time-Sensitive Network.

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.

4. DetNet MPLS Data Plane Overview

4.1. Layers of DetNet Data Plane

   This document describes how DetNet flows are carried over MPLS
   networks. The DetNet Architecture, [I-D.ietf-detnet-architecture],
   decomposes the DetNet data plane into two sub-layers: a service sub-
   layer and a forwarding sub-layer. The basic approach defined in this
   document supports the DetNet service sub-layer based on existing
   pseudowire (PW) encapsulations and mechanisms, and supports the
   DetNet forwarding sub-layer based on existing MPLS Traffic
   Engineering encapsulations and mechanisms. Background on PWs can be
   found in [RFC3985] and [RFC3031]. Background on MPLS Traffic
   Engineering can be found in [RFC3272] and [RFC3209].

```
                    DetNet          MPLS
                      .
                      .
                 +------------+
                 |  Service   | d-CW, S-Label
                 +------------+
                 | Forwarding | F-Label(s)
                 +------------+
                      .
                      .
```
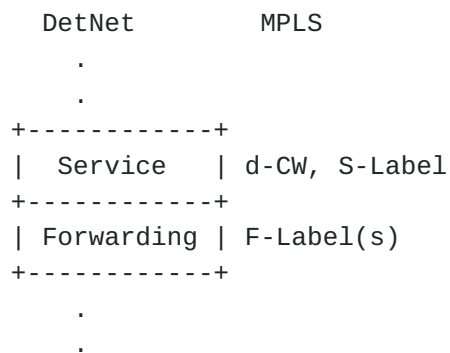
            Figure 1: DetNet Adaptation to MPLS Data Plane

   The DetNet MPLS data plane approach defined in this document is shown
   in Figure 1. The service sub-layer is supported by a DetNet control
   word (d-CW) which conforms to the Generic PW MPLS Control Word
   (PWMCW) defined in [RFC4385]. A d-CW identifying service label
   (S-Label) is also used.

   A node operating on a DetNet flow in the Detnet service sub-layer,
   i.e. a node processing a DetNet packet which has the S-Label as top
   of stack uses the local context associated with that S-Label, for
   example a received F-Label, to determine what local DetNet
   operation(s) are applied to that packet. An S-Label may be unique
   when taken from the platform label space [RFC3031], which would
   enable correct DetNet flow identification regardless of which input
   interface or LSP the packet arrives on.

   The DetNet MPLS data plane builds on MPLS Traffic Engineering
   encapsulations and mechanisms to provide a forwarding sub-layer that
   is responsible for providing resource allocation and explicit routes.
   The forwarding sub-layer is supported by one or more forwarding
   labels (F-Labels).

4.2. DetNet MPLS Data Plane Scenarios

```
   DetNet MPLS       Relay        Transit        Relay       DetNet MPLS
   End System        Node          Node          Node        End System
     (T-PE)          (S-PE)        (LSR)         (S-PE)         (T-PE)
   +----------+                                             +----------+
   |  Appl.   |<------------ End to End Service ---------->|  Appl.   |
   +----------+   +---------+                   +---------+   +----------+
   | Service  |<--| Service |-- DetNet flow --| Service |-->| Service  |
   +----------+   +---------+  +----------+   +---------+   +----------+
   |Forwarding|   |Fwd| |Fwd|  |Forwarding|   |Fwd| |Fwd|   |Forwarding|
   +--------.--+   +-.-+ +-.-+  +----.---.-+   +-.-+ +-.-+   +---.------+
         : Link :    /  ,-----.  \   : Link :    /  ,-----.  \
        +........+   +-[  Sub  ]-+   +......+    +-[  Sub  ]-+
                       [Network]                  [Network]
                        `-----'                    `-----'
        |<- LSP -->| |<-------- LSP -----------| |<--- LSP -->|

        |<---------------- DetNet MPLS -------------------->|
```
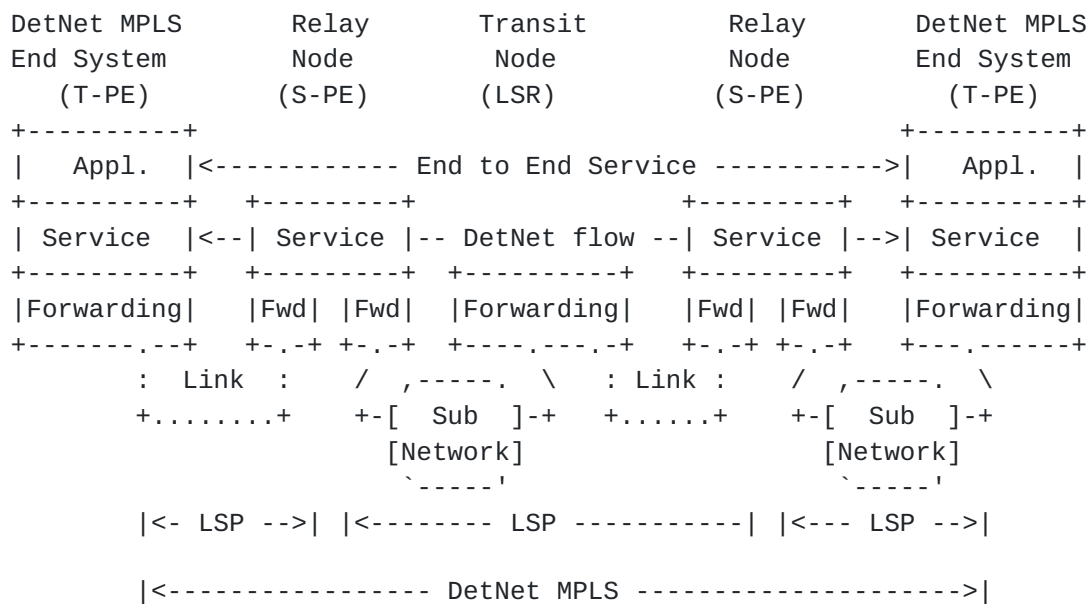
                  Figure 2: A DetNet MPLS Network

   Figure 2 illustrates a hypothetical DetNet MPLS-only network composed
   of DetNet aware MPLS enabled end systems, operating over a DetNet
   aware MPLS network. In this figure, relay nodes sit at MPLS LSP
   boundaries and transit nodes are LSRs.

   DetNet end system and relay nodes are DetNet service sub-layer aware,
   understand the particular needs of DetNet flows and provide both
   DetNet service and forwarding sub-layer functions. They add, remove
   and process d-CWs, S-Labels and F-labels as needed. MPLS enabled end
   system and relay nodes can enhance the reliability of delivery by
   enabling the replication of packets where multiple copies, possibly
   over multiple paths, are forwarded through the DetNet domain. They
   can also eliminate surplus previously replicated copies of DetNet
   packets. DetNet MPLS nodes provide functionality similar to T-PEs
   when they sit at the edge of an MPLS domain, and functionality
   similar to S-PEs when they are in the middle of an MPLS domain, see
   [RFC6073]. End system and relay nodes also include DetNet forwarding
   sub-layer functions, support for notably explicit routes, and
   resources allocation to eliminate (or reduce) congestion loss and
   jitter.

   DetNet transit nodes reside wholly within a DetNet domain, and also
   provide DetNet forwarding sub-layer functions in accordance with the
   performance required by a DetNet flow carried over an LSP. Unlike
   other DetNet node types, transit nodes provide no service sub-layer

processing. In a DetNet MPLS network, transit nodes may be DetNet
service aware or may be DetNet unaware MPLS Label Switching Routers
(LSRs). In this latter case, such LSRs would be unaware of the
special requirements of the DetNet service sub-layer, but would still
provide traffic engineering services and the QoS need to ensure that
the (TE) LSPs meet the service requirements of the carried DetNet
flows.

The LSPs may be provided by any MPLS controller method. For example
they may be provisioned via a management plane, RSVP-TE, MPLS-TP, or
MPLS Segment Routing (when extended to support resource allocation).

Figure 3 illustrates how an end to end MPLS-based DetNet service is
provided in a more detail. In this figure, the end systems, CE1 and
CE2, are able to send and receive MPLS encapsulated DetNet flows, and
R1, R2 and R3 are relay nodes as they sit in the middle of a DetNet
network. The 'X' in the end systems, and relay nodes represents
potential DetNet compound flow packet replication and elimination
points. In this example, service protection is supported over four
DetNet member flows and TE LSPs. For a unidirectional flow, R1
supports PRF, R2 supports PREOF and R3 supports PEF and POF. Note
that the relay nodes may change the underlying forwarding sub-layer,
for example tunneling MPLS over IEEE 802.1 TSN Section 8, or simply
over interconnect network links.
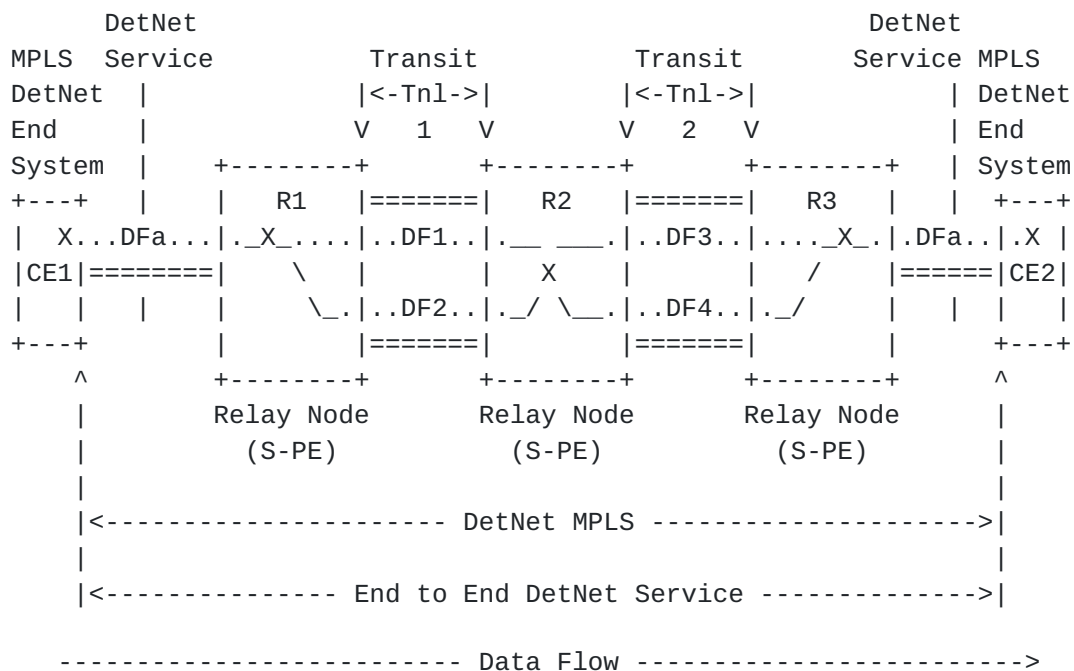
```
        DetNet                                          DetNet
   MPLS  Service           Transit          Transit      Service MPLS
   DetNet   |           |<-Tnl->|         |<-Tnl->|          | DetNet
   End      |           V   1   V         V   2   V          | End
   System   |   +--------+       +--------+       +--------+  | System
   +---+    |   |   R1   |=======|   R2   |=======|   R3   |  |   +---+
   | X...DFa...|._X_....|..DF1..|.__ ___.|..DF3..|...._X_.|.DFa..|.X |
   |CE1|=======|    \   |       |   X    |       |    /   |=====|CE2|
   |   |   |    |     \_.|..DF2..|._/ \__.|..DF4..|._/    |   |  |   |
   +---+    |        |=======|         |=======|         |      +---+
      ^      +--------+       +--------+       +--------+      ^
      |        Relay Node       Relay Node       Relay Node    |
      |          (S-PE)           (S-PE)           (S-PE)      |
      |                                                        |
      |<--------------------- DetNet MPLS --------------------->|
      |                                                        |
      |<-------------- End to End DetNet Service -------------->|

      ------------------------ Data Flow ------------------------>
```

```
   X   = Optional service protection (none, PRF, PREOF, PEF/POF)
   DFx = DetNet member flow x over a TE LSP
```

Figure 3: MPLS-Based Native DetNet

As previously mentioned, this document specifies how MPLS is used to
support DetNet flows using an MPLS data plane as well as how such can
be mapped to IEEE 802.1 TSN and IP DetNet PSNs. An equally import
scenario is when IP is supported over DetNet MPLS and this is covered
in [I-D.ietf-detnet-dp-sol-ip]. Another important scenario is where
an Ethernet Layer 2 service is supported over DetNet MPLS and this is
covered in [TBD-TSN-OVER-DETNET].

4.2.1. IP Over DetNet MPLS Data Plane Scenarios

[Author's note: this section to be moved to IP sol draft]

```
     IP  DetNet         Relay         Transit         Relay         IP DetNet
     End System         Node           Node            Node         End System
                        (T-PE)         (LSR)          (T-PE)
     +----------+                                                  +----------+
     |   Appl.  |<------------ End to End Service ----------->|  Appl.  |
     +----------+   ...........+                     +-----.....   +----------+
     | Service  |<--: Service |-- DetNet flow --| Service :-->| Service  |
     +----------+   +---------+  +----------+   +---------+   +----------+
     |Forwarding|   |Fwd| |Fwd|  |Forwarding|   |Fwd| |Fwd|   |Forwarding|
     +-------.--+   +-.-+ +-.-+  +----.---.-+   +-.-+ +-.-+   +---.------+
             :  Link :    /  ,-----.  \   : Link :    /  ,-----.  \
             +........+    +-[  Sub  ]-+   +......+    +-[  Sub  ]-+
                            [Network]                  [Network]
                            `-----'                    `-----'


          |<- DN IP->|  |<---- DetNet MPLS ---->|  |< -DN IP ->|
```

                  Figure 4: DetNet IP Over MPLS Network

Figure 4 illustrates DetNet enabled End Systems (hosts), connected to
DetNet (DN) enabled IP networks, operating over a DetNet aware MPLS
network. In this figure, Relay nodes sit at the boundary of the MPLS
domain since the non-MPLS domain is DetNet aware. This figure is
very similar to Figure 2. The primary difference is that the Relay
nodes are at the edges of the MPLS domain and therefore function as
T-PEs, and that service sub-layer functions are not provided over the
DetNet IP network. There is no difference in transit node function.

Figure 5 illustrates how relay nodes can provide service protection
over the MPLS domain. In this case, CE1 and CE2 are IP DetNet end
systems which are interconnected via a MPLS domain such as previously
shown in Figure 3. Note that R1 and R3 sit at the edges of an MPLS
domain and therefore are similar to T-PEs, while R2 sits in the
middle of the domain and is therefore similar to an S-PE.

```
          DetNet                                   DetNet
   IP     Service        Transit        Transit    Service  IP
  DetNet                 |<-Tnl->|      |<-Tnl->|           DetNet
  End       |            V  1  V      V   2  V      |       End
  System |     +--------+       +--------+       +--------+  |  System
+---+    |     |  R1    |=======|  R2    |=======|  R3    |  |    +---+
|   |----|-----|._X_....|..DF1..|.__ ___.|..DF3..|...._X_.|--|----|   |
|CE1|    |     |   \    |       |    X   |       |   /    |  |    |CE2|
|   |    |     |    \_. |..DF2..|._/ \__.|..DF4..|._/     |  |    |   |
+---+    |     |        |=======|        |=======|        |  |    +---+
    ^        +--------+        +--------+        +--------+      ^
    |          Relay Node        Relay Node        Relay Node    |
    |            (T-PE)            (S-PE)            (T-PE)       |
    |                                                            |
    |<-DN IP-> <-------- DetNet MPLS --------------> <-DN IP->|
    |                                                            |
    |<-------------- End to End DetNet Service --------------->|

     ------------------------ Data Flow ------------------------>

     X   = Service protection (PRF, PREOF, PEF/POF)
     DFx = DetNet member flow x over a TE LSP
```

                  Figure 5: DetNet IP Over DetNet MPLS Network

```
  IP                 Edge                    Edge        IP
  End System         Node                    Node        End System
                     (T-PE)       (LSR)      (T-PE)
+----------+   +...,.-----+               +-----,...+   +----------+
|  Appl.   |<--:Svc Proxy|-- E2E Service --|Svc Proxy:-->|  Appl.  |
+----------+   +.....+---+               +---+.....+   +----------+
|   IP     |<--:IP : |Svc|-- IP/DN Flow ---|Svc| :IP :-->|   IP    |
+----------+   +---+ +---+  +----------+   +---+ +---+   +----------+
|Forwarding|   |Fwd| |Fwd|  |Forwarding|   |Fwd| |Fwd|   |Forwarding|
+--------.--+   +-.-+ +-.-+  +----.---.-+   +-.-+ +-.-+   +---.------+
      : Link :     /  ,-----.  \    : Link :    /  ,-----.  \
      +........+    +-[  Sub  ]-+   +......+   +-[  Sub  ]-+
                     [Network]                 [Network]
                      `-----'                   `-----'

      |<--- IP --->| |<----- DetNet MPLS ----->| |<--- IP --->|
```
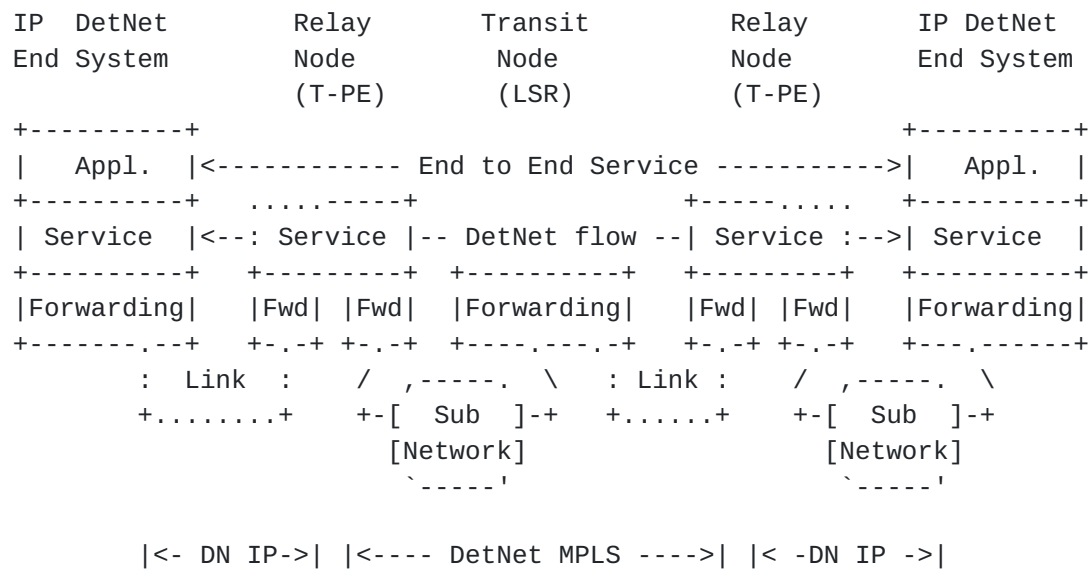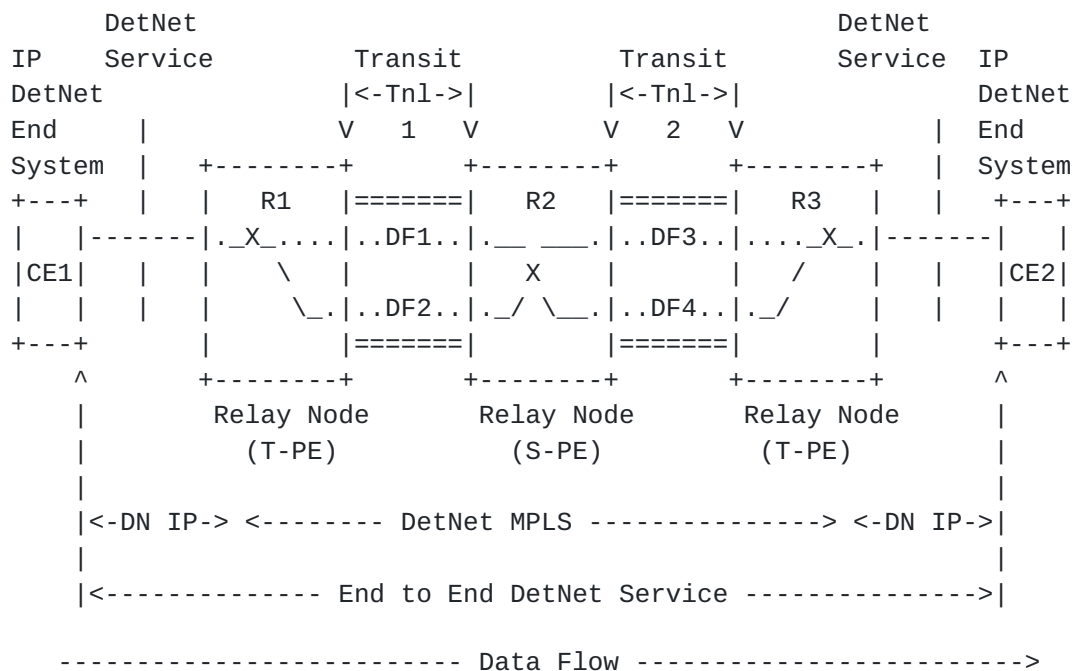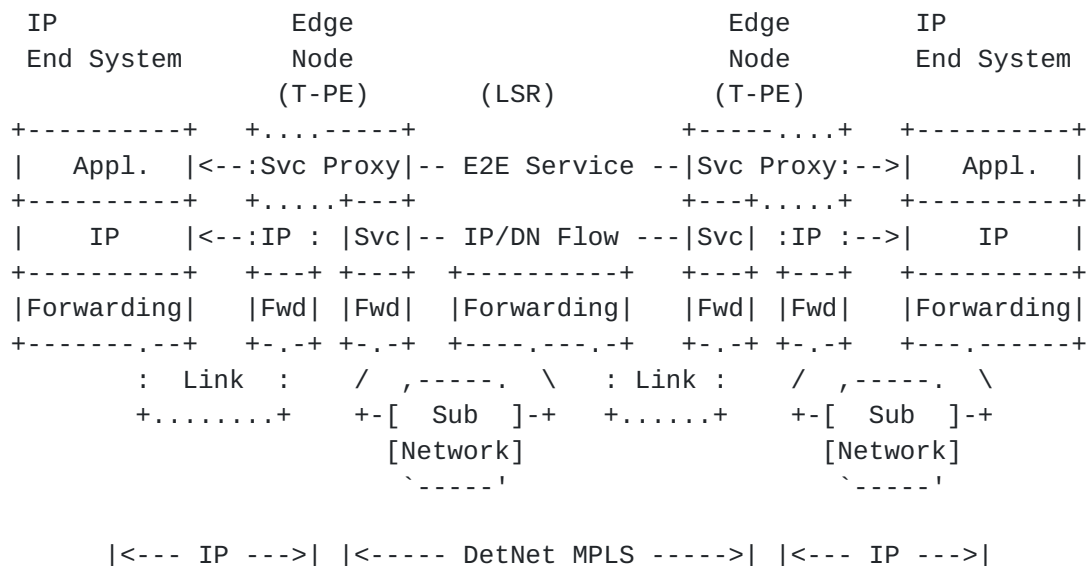
             Figure 6: Non-DetNet Aware IP Over DetNet MPLS Network

   Figure 6 illustrates non-DetNet enabled End Systems (hosts),
   connected to DetNet (DN) enabled MPLS network. It differs from
   Figure 4 in that the hosts and edge IP networks are not DetNet aware.
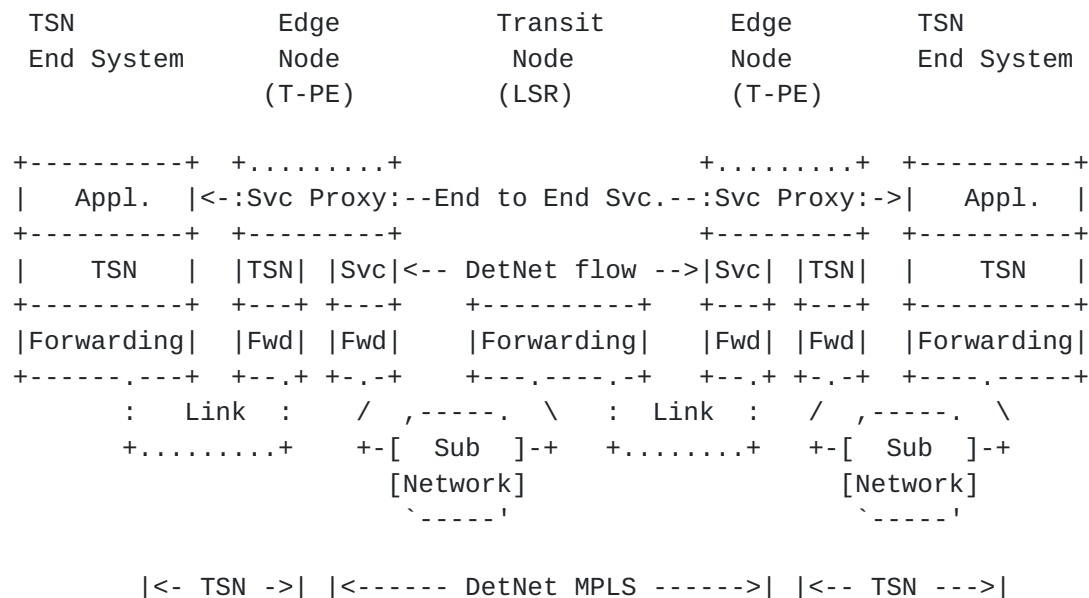   In this case, edge nodes sit at the boundary of the MPLS domain since

it is also a DetNet domain boundary. The edge nodes provide DetNet
service proxies for the end applications by initiating and
terminating DetNet service for the application's IP flows. See
[I-D.ietf-detnet-dp-sol-ip] for more information.

Figure 7 illustrates how it is still possible to provided DetNet
service protection for non-DetNet aware end systems. This figures is
basically the same as Figure 5, with the exception that CE1 and CE2
are non-DetNet aware end systems and E1 and E3 are edge nodes that
replace the relay nodes R1 and R3.

```
          IP                                          IP
  Non    Service              Transit          Transit        Service Non
  DetNet                   |<-Tnl->|         |<-Tnl->|             DetNet
  End      |               V   1   V         V   2   V         | End
  System   |    +--------+      +--------+      +--------+    | System
  +---+    |    |  E1    |======|   R2   |======|   E3   |    |   +---+
  |   |--------|._X_....|..DF1..|.__ ___.|..DF3..|....X_.|------|   |
  |CE1|    |    |   \    |       |   X    |       |   /   |    |   |CE2|
  |   |    |    |    \_.|..DF2..|._/ \__.|..DF4..|._/     |    |   |   |
  +---+    |        |======|        |======|        |         +---+
           +--------+      +--------+      +--------+
           ^ Edge Node      Relay Node      Edge Node^
           | (T-PE)           (S-PE)          (T-PE)  |
           |                                          |
      <--IP-->| <-------- IP Over DetNet MPLS --------> |<--IP-->
           |                                          |
           |<------ End to End DetNet Service ------>|

     X   = Optional service protection (none, PRF, PREOF, PEF/POF)
     DFx = DetNet member flow x over a TE LSP

          Figure 7: MPLS-Based DetNet (non-MPLS End System)
```

4.2.2. IEEE 802.1 TSN Over DetNet MPLS Data Plane Scenario

[Author's note: this section to be moved to TSN over mpls sol draft -
TBD-TSN-OVER-DETNET]

```
       TSN              Edge          Transit         Edge          TSN
    End System          Node           Node           Node       End System
                       (T-PE)          (LSR)         (T-PE)


   +----------+  +.........+                      +.........+  +----------+
   |  Appl.   |<-:Svc Proxy:--End to End Svc.--:Svc Proxy:->|  Appl.   |
   +----------+  +---------+                      +---------+  +----------+
   |   TSN    |  |TSN| |Svc|<-- DetNet flow -->|Svc| |TSN|  |    TSN   |
   +----------+  +---+ +---+   +----------+   +---+ +---+  +----------+
   |Forwarding|  |Fwd| |Fwd|   |Forwarding|   |Fwd| |Fwd|  |Forwarding|
   +------.---+  +--.+ +-.-+   +---.----.-+   +--.+ +-.-+  +----.-----+
        :   Link  :   /  ,-----.  \   :   Link  :   /  ,-----.  \
        +.........+   +-[  Sub  ]-+   +........+   +-[  Sub  ]-+
                        [Network]                   [Network]
                         `-----'                     `-----'


        |<- TSN ->| |<------ DetNet MPLS ------>| |<-- TSN --->|
```

              Figure 8: A TSN over DetNet MPLS Enabled Network

   Figure 8 shows IEEE 802.1 TSN end stations operating over a TSN aware
   DetNet service running over an MPLS network. DetNet Edge Nodes sit
   at the boundary of a DetNet domain. They are responsible for mapping
   non-DetNet aware L2 traffic to DetNet services. They also support
   the imposition and disposition of the required DetNet encapsulation.
   These are functionally similar to pseudowire (PW) Terminating
   Provider Edge (T-PE) nodes which use MPLS-TE LSPs. In this example
   they understand and support IEEE 802.1 TSN and are able to map TSN
   flows into DetNet flows. The specific of this operation are
   discussed in [TBD-TSN-OVER-DETNET].

   Native TSN flow and DetNet MPLS flow differ not only by the
   additional MPLS specific encapsulation, but DetNet MPLS flows have on
   each DetNet node an associated DetNet specific data structure, what
   defines flow related characteristics and required forwarding
   functions. In this example, edge Nodes provide a service proxy
   function that "associates" the DetNet flows and native flows at the
   edge of the DetNet domain. This ensures that the DN Flow is properly
   served at the Edge node (and inside the domain).

   Figure 9 illustrates how DetNet can provide services for IEEE
   802.1TSN end systems, CE1 and CE2, over a DetNet enabled MPLS
   network. Similar to Figure 6, the edge nodes, E1 and E2, insert and
   remove required DetNet data plane encapsulation. The 'X' in the edge
   nodes and relay node, R1, represent a potential DetNet compound flow
   packet replication and elimination point. This conceptually
   parallels L2VPN services, and could leverage existing related
   solutions as discussed below.

```
         TSN     |<------- End to End DetNet Service ------>|  TSN
       Service  |            Transit          Transit       | Service
   TSN  (AC)    |         |<-Tnl->|        |<-Tnl->|         |  (AC)  TSN
   End   |      V         V   1  V        V   2  V          V  |  End
  System |    +--------+      +--------+      +--------+     | System
  +---+  |    |   E1   |======|   R1   |======|   E2   |     |   +---+
  |   |--|----|._X_....|..DF1..|.._ _...|..DF3..|...._X_.|---|---|   |
  |CE1|  |    |   \    |       |   X    |       |   /    |   |   |CE2|
  |   |  |    |    \_.|..DF2..|._/ \_..|..DF4..|._/      |   |   |   |
  +---+  |    |        |======|        |======|          |       +---+
    ^       +--------+      +--------+      +--------+       ^
    |         Edge Node      Relay Node      Edge Node       |
    |           (T-PE)         (S-PE)          (T-PE)        |
    |                                                        |
    |<- TSN -> <------- TSN Over DetNet MPLS -------> <- TSN ->|
    |                                                        |
    |<--- Emulated Time Sensitive Networking (TSN) Service --->|

    X   = Service protection
    DFx = DetNet member flow x over a TE LSP
```
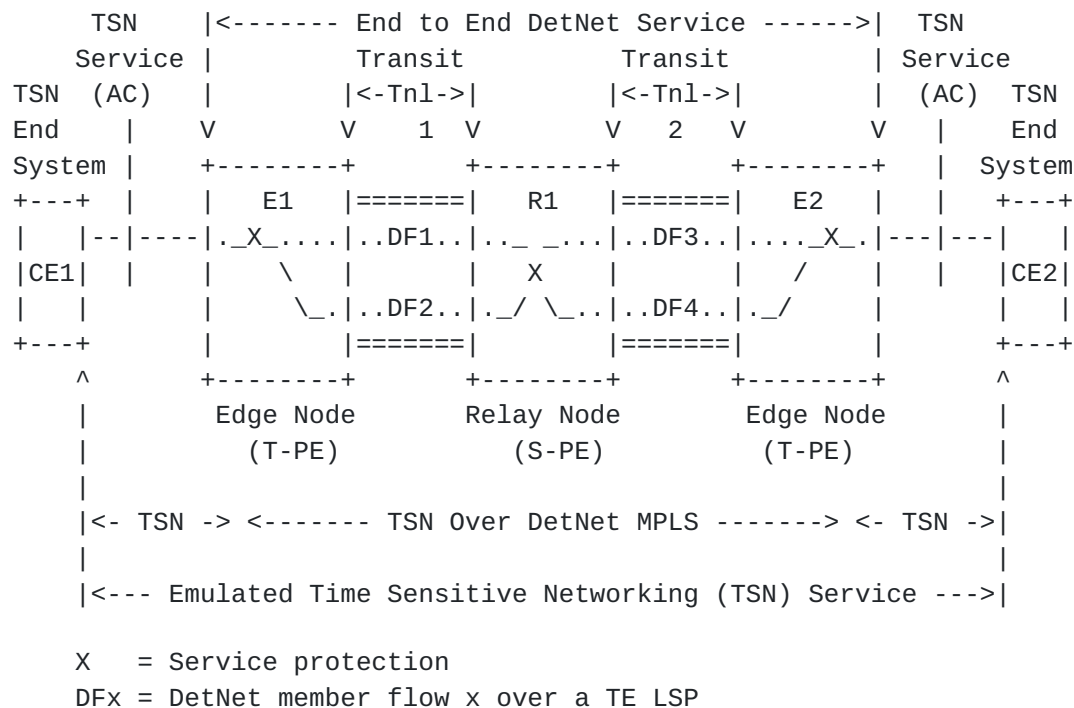
Figure 9: IEEE 802.1TSN Over DetNet

4.3. Packet Flow Example with Service Protection

An example DetNet MPLS network fragment and packet flow is
illustrated in Figure 10.

```
    1      1.1       1.1      1.2.1    1.2.1       1.2.2
  CE1----EN1--------R1-------R2-------R3--------EN2-----CE2
         \              1.2.1 /                    /
          \1.2      /-----+                   /
           +------R4-----------------------+
                   1.2.2
```
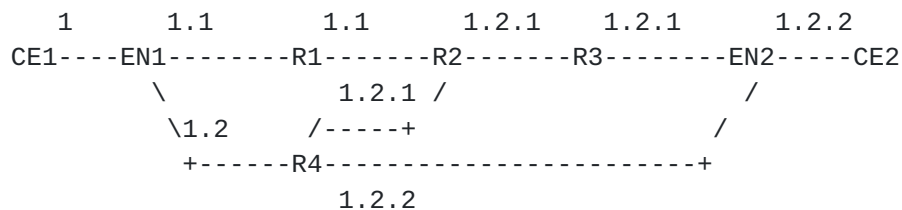
Figure 10: Example Packet Flow in DetNet Enabled MPLS Network

In Figure 10 the numbers are used to identify the instance of a
packet. Packet 1 is the original packet, and packets 1.1, and 1.2
are two first generation copies of packet 1. Packet 1.2.1 is a
second generation copy of packet 1.2 etc. Note that these numbers
never appear in the packet, and are not to be confused with sequence
numbers, labels or any other identifier that appears in the packet.
They simply indicate the generation number of the original packet so
that its passage through the network fragment can be identified to
the reader.

   Customer Equipment CE1 sends a packet into the DetNet enabled MPLS
   network. This is packet (1). Edge Node EN1 encapsulates the packet
   as a DetNet Packet and sends it to Relay node R1 (packet 1.1). EN1
   makes a copy of the packet (1.2), encapsulates it and sends this copy
   to Relay node R4.

   Note that along the MPLS path from EN1 to R1 there may be zero or
   more LSRs which, for clarity, are not shown. The same is true for
   any other path between two DetNet entities shown in Figure 10.

   Relay node R4 has been configured to send one copy of the packet to
   Relay Node R2 (packet 1.2.1) and one copy to Edge Node EN2 (packet
   1.2.2).

   R2 receives packet copy 1.2.1 before packet copy 1.1 arrives, and,
   having been configured to perform packet elimination on this DetNet
   flow, forwards packet 1.2.1 to Relay Node R3. Packet copy 1.1 is of
   no further use and so is discarded by R2.

   Edge Node EN2 receives packet copy 1.2.2 from R4 before it receives
   packet copy 1.2.1 from R2 via relay Node R3. EN2 therefore strips
   any DetNet encapsulation from packet copy 1.2.2 and forwards the
   packet to CE2. When EN2 receives the later packet copy 1.2.1 this is
   discarded.

   The above is of course illustrative of many network scenarios that
   can be configured. Between a pair of relay nodes there may be one or
   more transit nodes that simply forward the DetNet traffic, but these
   are omitted for clarity.

5. DetNet MPLS Data Plane Considerations

   This section provides informative considerations related to providing
   DetNet service to flows which are identified based on their header
   information. At a high level, the following are provided on a per
   flow basis:

   Eliminating contention loss and jitter reduction:

      Use of allocated resources (queuing, policing, shaping) to ensure
      that the congestion-related loss and latency/jitter requirements
      of a DetNet flow are met.

   Explicit routes:

      Use of a specific path for a flow. This limits misordering and
      bounds latency.

   Service protection:

      Which in the case of this document primarily relates to
      replication and elimination. Changing the explicit path after a
      failure is detected in order to restore delivery of the required
      DetNet service characteristics is also possible. Path changes,
      even in the case of failure recovery, can lead to the out of order
      delivery of data.

   Load sharing:

      Generally, distributing packets of the same DetNet flow over
      multiple paths is not recommended. Such load sharing, e.g., via
      ECMP or UCMP, impacts ordering and possibly jitter.

   Troubleshooting:

      For example, to support identification of misbehaving flows.

   Recognize flow(s) for analytics:

      For example, increase counters.

   Correlate events with flows:

      For example, unexpected loss.

   The DetNet data plane also allows for the aggregation of DetNet
   flows, e.g., via MPLS hierarchical LSPs, to improved scaling. When
   DetNet flows are aggregated, transit nodes provide service to the
   aggregate and not on a per-DetNet flow basis. In this case, nodes
   performing aggregation will ensure that per-flow service requirements
   are achieved.

5.1. End-System Specific Considerations

   Data-flows requiring DetNet service are generated and terminated on
   end-systems. Encapsulation depends on application and its
   preferences. In a DetNet MPLS domain the DN functions use the d-CWs,
   S-Labels and F-Labels to provide DetNet services. However, an
   application may exchange further flow related parameters (e.g., time-
   stamp), which are not provided by DN functions.

   Specifics related to non-MPLS DetNet end station behavior are out
   side the scope of this document. For example, details on support for
   DetNet IP data flows can be found in [I-D.ietf-detnet-dp-sol-ip].
   This document is also useful for end stations that map IP flows to
   DetNet flows.

As a general rule, DetNet MPLS domains are capable of forwarding any
DetNet MPLS flows and the DetNet domain does not mandate the end-
system or edge system encapsulation format. Unless there is a proxy
of some form present, end-systems peer with similar end-systems using
the same application encapsulation format. For example, as shown in
Figure 11, IP applications peer with IP applications and Ethernet
L2VPN applications peer with Ethernet L2VPN applications.

```
          +-----+
          |  X  |                            +-----+
          +-----+                            |  X  |
          | Eth |             _____       +-----+
          +-----+      _____      /      \    | Eth |
               \    /      \__/         \___  +-----+
                \  /                       \  /
                 0======== tunnel-1 =======0_
                 |                           \
                  \                           |
                   0======== tunnel-2 ========0
                  / \                      __/ \
         +-----+   \__  DetNet MPLS domain /     \
         |  X  |      \         __        /      +-----+
         +-----+       _____/ _____/        |  X  |
         |  IP |                                 +-----+
         +-----+                                 |  IP |
                                                 +-----+
```
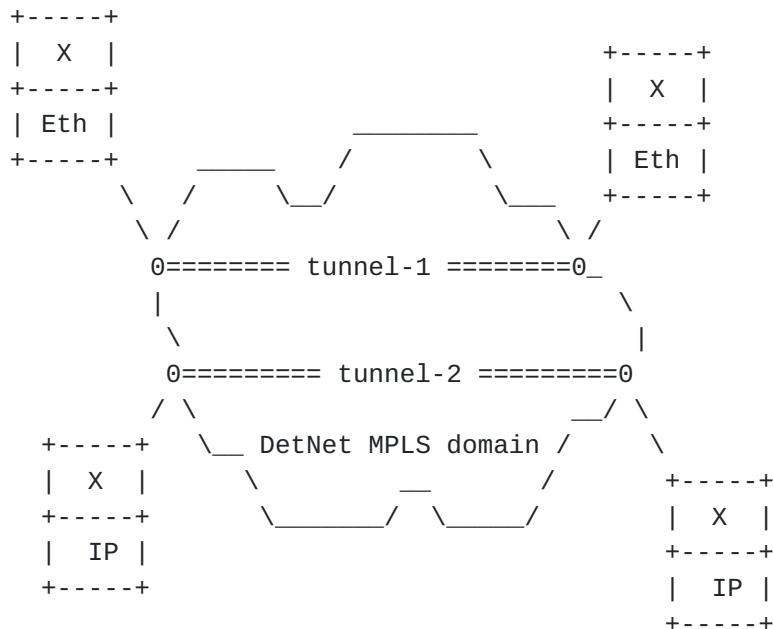
                Figure 11: End-Systems and The DetNet MPLS Domain

5.2. Sub-Network Considerations

As shown in Figure 2, MPLS nodes are interconnected by different sub-
network technologies, which may include point-to-point links. Each
of these need to provide appropriate service to DetNet flows. In
some cases, e.g., on dedicated point-to-point links or TDM
technologies, all that is required is for a DetNet node to
appropriately queue its output traffic. In other cases, DetNet nodes
will need to map DetNet flows to the flow semantics (i.e.,
identifiers) and mechanisms used by an underlying sub-network
technology. Figure 12 shows several examples of header formats that
can be used to carry DetNet MPLS flows over different sub-network
technologies. L2 represent a generic layer-2 encapsulation that
might be used on a point-to-point link. TSN represents the
encapsulation used on an IEEE 802.1 TSN network, as described in
Section 8. UDP/IP represents the encapsulation used on a DetNet IP
PSN, as described in Section 9 .

```
                           +------+  +------+  +------+
            App-Flow       |  X   |  |  X   |  |  X   |
                        +-----+======+--+======+--+======+-----+
            DetNet-MPLS    | d-CW |  | d-CW |  | d-CW |
                           +------+  +------+  +------+
                           |Labels|  |Labels|  |Labels|
                        +-----+======+--+======+--+======+-----+
            Sub-Network    |  L2  |  | TSN  |  | UDP  |
                           +------+  +------+  +------+
                                               |  IP  |
                                               +------+
                                               |  L2  |
                                               +------+
```
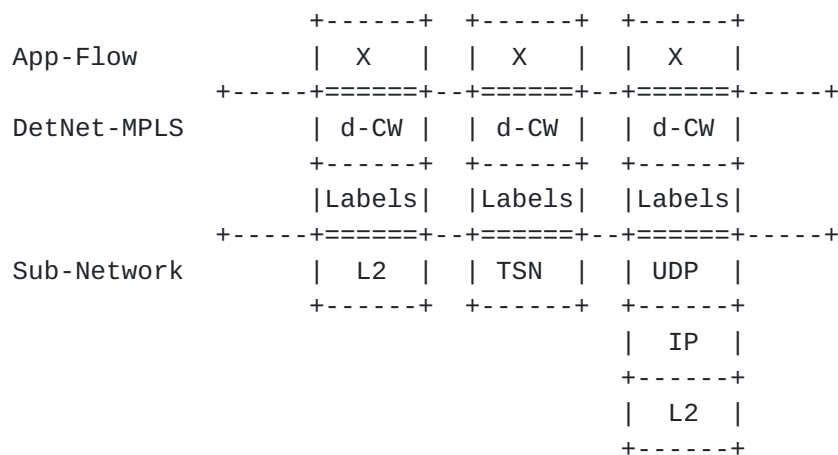
Figure 12: Example DetNet MPLS Sub-Network Formats

6. MPLS-Based DetNet Data Plane Solution

6.1. DetNet Over MPLS Encapsulation Components

To carry DetNet over MPLS the following is required:

1. A method of identifying the MPLS payload type.

2. A method of identifying the DetNet flow group to the processing
    element.

3. A method of distinguishing DetNet OAM packets from DetNet data
    packets.

4. A method of carrying the DetNet sequence number.

5. A suitable LSP to deliver the packet to the egress PE.

6. A method of carrying queuing and forwarding indication.

In this design an MPLS service label (the S-Label), similar to a
pseudowire (PW) label [RFC3985], is used to identify both the DetNet
flow identity and the payload MPLS payload type satisfying (1) and
(2) in the list above. OAM traffic discrimination happens through
the use of the Associated Channel method described in [RFC4385]. The
DetNet sequence number is carried in the DetNet Control word which
carries the Data/OAM discriminator. To simplify implementation and
to maximize interoperability two sequence number sizes are supported:
a 16 bit sequence number and a 28 bit sequence number. The 16 bit
sequence number is needed to support some types of legacy clients.
The 28 bit sequence number is used in situations where it is

   necessary ensure that in high speed networks the sequence number
   space does not wrap whilst packets are in flight.

   The LSP used to forward the DetNet packet may be of any type (MPLS-
   LDP, MPLS-TE, MPLS-TP [RFC5921], or MPLS-SR
   [I-D.ietf-spring-segment-routing-mpls]). The LSP (F-Label) label
   and/or the S-Label may be used to indicate the queue processing as
   well as the forwarding parameters. Note that the possible use of
   Penultimate Hop Popping (PHP) means that the only label in a received
   label stack may be the S-Label.

6.2. MPLS Data Plane Encapsulation

   Figure 13 illustrates a DetNet data plane MPLS encapsulation. The
   MPLS-based encapsulation of the DetNet flows is a good fit for the
   scenarios described in Section 4.2.1 and Section 4.2.2. Furthermore,
   end to end DetNet service i.e., native DetNet deployment (see
   Section 4.2) is also possible if DetNet end systems are capable of
   initiating and termination MPLS encapsulated packets.

   The MPLS-based DetNet data plane encapsulation consists of:

   o DetNet control word (d-CW) containing sequencing information for
      packet replication and duplicate elimination purposes, and the OAM
      indicator.

   o DetNet service Label (S-Label) that identifies a DetNet flow at
      the receiving DetNet service sub-layer processing node.

   o Zero or more Detnet MPLS Forwarding label(s) (F-Label) used to
      direct the packet along the label switched path (LSP) to the next
      service sub-layer processing node along the path. When
      Penultimate Hop Popping is in use there may be no label F-Label in
      the protocol stack on the final hop.

   o The necessary data-link encapsulation is then applied prior to
      transmission over the physical media.

DetNet MPLS-based encapsulation

```
+--------------------------------+
|                                |
|          DetNet App-Flow       |
|          Payload  Packet       |
|                                |
+--------------------------------+ <--\
|        DetNet Control Word      |    |
+--------------------------------+    +--> DetNet data plane
|             S-Label             |    |     MPLS encapsulation
+--------------------------------+    |
|             F-Label(s)          |    |
+--------------------------------+ <--/
|             Data-Link           |
+--------------------------------+
|             Physical            |
+--------------------------------+
```
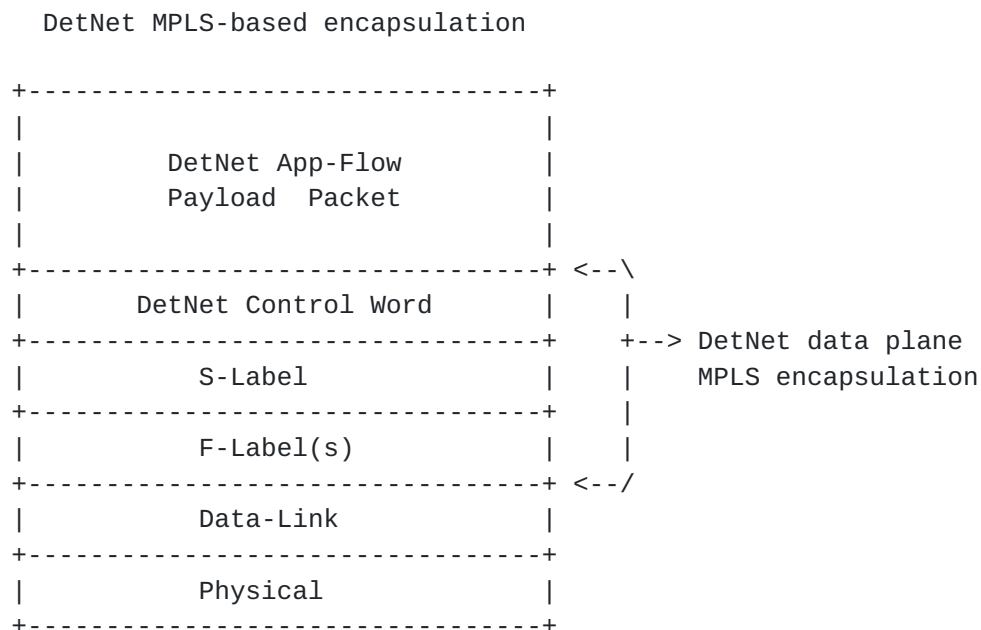
Figure 13: Encapsulation of a DetNet App-Flow in an MPLS(-TP) PSN

6.2.1. **DetNet Control Word and the DetNet Sequence Number**

A DetNet control word (d-CW) conforms to the Generic PW MPLS Control
Word (PWMCW) defined in [RFC4385]. The d-CW formatted as shown in
Figure 14 MUST be present in all DetNet packets containing app-flow
data.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 0|              Sequence Number                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
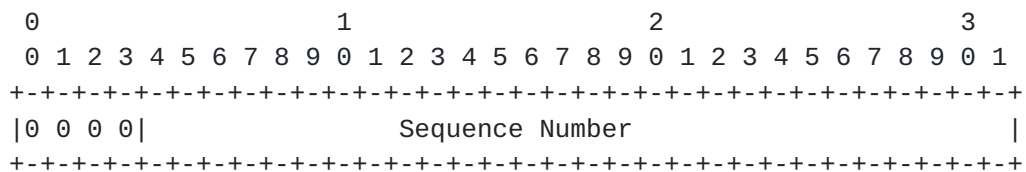
Figure 14: DetNet Control Word

(bits 0 to 3)

   Per [RFC4385], MUST be set to zero (0).

Sequence Number (bits 4 to 31)

   An unsigned value implementing the DetNet sequence number.

A separate sequence number space MUST be maintained by the the node
that adds the d-CW for each DetNet app-flow. The following sequence
number field lengths MUST be supported:

    0 bits

    16 bits

    28 bits

The sequence number length MUST be provisioned (configured) on a per
app-flow basis via configuration, e.g., the controller plane
described in Section 7.

A 0 bit sequence number field length indicates that there is no
DetNet sequence number used for the flow. When the length is zero,
the sequence number field MUST be set to zero (0) on all packets sent
for the flow.

When the sequence number field length is 16 or 28 bits for a flow,
the sequence number MUST be incremented by one for each new app-flow
packet sent. When the field length is 16 bits, d-CW bits 4 to 15
MUST be set to zero (0). This values carried in this field can wrap
and it is important to note that zero (0) is a valid field value.
For example, were the sequence number size is 16 bits, the sequence
will contain: 65535, 0, 1. In this case, zero (0) is an ordinary
sequence number. This differs from [RFC4448] where a sequence number
of zero (0) does not indicate that no sequence number field value is
in use.

The sequence number is optionally used during receive processing as
described below in Section 6.2.2.1 and Section 6.2.2.2.

6.2.2. S-Labels

App-flow identification at a DetNet service sub-layer is realized by
an S-Label. Each app-flow MUST be sent by the node that adds a d-CW
with a single specific S-Label value. MPLS-aware DetNet end systems
and edge nodes, which are by definition MPLS ingress and egress
nodes, MUST add and remove the d-CW and S-Label. Relay nodes MAY
swap S-Label values when processing an app-flow.

The S-Label value MUST be provisioned per app-flow via configuration,
e.g., via the controller plane described in Section 7. Note that
S-Labels provide app-flow identification at the downstream DetNet
service sub-layer receiver, not the sender. As such, S-Labels MUST
be allocated by the entity that controls the service sub-layer

receiving node's label space, and MAY be allocated from the platform label space [RFC3031].

The S-Label will normally be at the bottom of the label stack, immediately preceding the d-CW. To support service sub-layer level OAM, an OAM Associated Channel Header (ACH) [RFC4385] together with a Generic Associated Channel Label (GAL) [RFC5586] MAY be used in place of a d-CW.

Similarly, an Entropy Label Indicator/Entropy Label (ELI/EL) [RFC6790] MAY be carried below the S-Label in the label stack in networks where DetNet flows would otherwise received ECMP treatment. When ELs are used, the same EL value SHOULD be used for all of the packets sent using a specific S-Label to force the flow to follow the same path. However, as previously stated in Section 5, the use of ECMP for DetNet flows is NOT RECOMMENDED. ECMP MAY be used for non-DetNet flows within a DetNet domain.

When receiving a DetNet MPLS flow, an implementation MUST identify the app-flow associated with the incoming packet based on the S-Label. When a node is using platform labels for S-Labels, no additional information is needed as the S-label uniquely identifies the app-flow. In the case where platform labels are not used, one or more F-Labels proceeding the S-Label MUST be used together with the S-Label to uniquely identify the incoming app-flows. When PHP is used, the incoming interface MAY also be used to together with any present F-Label(s) and the S-Label to uniquely identify an incoming app-flows. Note that choice to use platform label space for S-Label or S-Label plus one or more F-Labels to identify app flows is a local implementation choice, with one caveat. When one or more F-labels, or incoming interface, is needed together with an S-Label to uniquely identify, the controller plane MUST ensure that incoming DetNet MPLS packets arrive with the needed information (F-label(s) and/or incoming interface); the details of such are outside the scope of this document.

While NOT REQUIRED, the use of platform labels for S-Labels matches other pseudowire encapsulations. This implementation choice also impacts PEF and POF processing as described in the next section.

6.2.2.1. Packet Elimination Function Processing

Implementations MAY support the Packet Elimination Function (PEF) for received DetNet MPLS flows. When supported, use of the PEF for a particular app-flow MUST be provisioned via configuration, e.g., via the controller plane described in Section 7.

   After an app-flow is identified for a received DetNet MPLS packet, as
   described above, an implementation MUST check if PEF is configured
   for that app-flow. When configured the implementation MUST track the
   sequence number contained in received d-CWs and MUST ensure that
   duplicate (replicated) instances of a particular sequence number are
   discarded. The specific mechanisms used for an implementation to
   identify which received packets are duplicates and which are new is
   an implementation choice. Note that per Section 6.2.1 the sequence
   number field length may be 16 or 28 bits, and the field value can
   wrap.

   Note that an implementation MAY wish to constrain the maximum number
   sequence numbers that are tracked, on platform-wide or per flow
   basis. Some implementations MAY support the provisioning of the
   maximum number sequence numbers that are tracked number on either a
   platform-wide or per flow basis.

6.2.2.2. Packet Ordering Function Processing

   A function that is related to PEF is the Packet Ordering Function
   (POF). Implementations MAY support POF. When supported, use of the
   POF for a particular app-flow MUST be provisioned via configuration,
   e.g., via the controller plane described by Section 7.
   Implementations MAY required that PEF and POF be used in combination.
   There is no requirement related to the order of execution of the
   Packet Elimination and Ordering Functions in an implementation.

   After an app-flow is identified for a received DetNet MPLS packet, as
   described above, an implementation MUST check if POF is configured
   for that app-flow. When configured the implementation MUST track the
   sequence number contained in received d-CWs and MUST ensure that
   packets are processed in the order indicated in the received d-CW
   sequence number field, which may not be in the order the packets are
   received. As defined in Section 6.2.1 the sequence number field
   length may be 16 or 28 bits, is incremened by one (1) for each new
   app-flow packet sent, and the field value can wrap. The specific
   mechanisms used for an implementation to identify the order of
   received packets is an implementation choice.

   Note that an implementation MAY wish to constrain the maximum number
   of out of order packets that can be processed, on platform-wide or
   per flow basis. Some implementations MAY support the provisioning of
   this number on either a platform-wide or per flow basis. The number
   of out of order packets that can be processed also impacts the
   latency of a flow.

6.2.3. F-Labels

   F-Labels are support the DetNet forwarding sub-layer. F-Labels are
   used to provide LSP-based connectivity between DetNet service sub-
   layer processing nodes.

6.2.3.1. Service Sub-Layer and Packet Replication Function Processing

   DetNet MPLS end systems, edge nodes and relay nodes may operate at
   the DetNet service sub-layer with understand of app-flows and their
   requirements. As mentioned above, when operating at this layer such
   nodes can push, pop or swap (pop then push) S-Labels. In all cases,
   the F-Labels used for the app-flow are always replaced and this
   section applies.

   When sending a DetNet flow, Zero or more F-Labels MAY be added on top
   of an S-Label by the node pushing an S-Label. The F-Labels to be
   pushed when sending a particular app-flow MUST be provisioned per
   app-flow via configuration, e.g., via the controller plane discussed
   in Section 7. To allow for the omission of F-Labels, an
   implementation SHOULD also allow an outgoing interface to be
   provisioned.

   The Packet Replication Function (PRF) function MAY be supported by an
   implementation for outgoing DetNet flows. When supported, the same
   app-flow data will be sent over multiple outgoing forwarding sub-
   layer LSPs. To support PRF an implementation MUST support the
   setting of different sets of F-Labels. Hereto, to allow for the
   omission of F-Labels, an implementation SHOULD also allow multiple
   outgoing interfaces to be provisioned. PRF MUST NOT be used with
   app-flows configured with a d-CW sequence number field length of 0
   bits.

   When a single set of F-Labels is provisioned for a particular
   outgoing app-flow, that set of F-labels MUST be pushed after the
   S-Label is pushed. The outgoing packet is then forwarded as
   described below in Section 6.2.3.2. When a single outgoing interface
   is provisioned, the outgoing packet is then forwarded as described
   below in Section 6.2.3.2.

   When multiple sets of F-Labels or interfaces are provisioned for a
   particular outgoing app-flow, a copy of the outgoing packet,
   including the pushed S-Label, MUST be made per F-label set and
   outgoing interface. Each set of provisioned F-Labels are then pushed
   onto a copy of the packet. Each copy is then forwarded as described
   below in Section 6.2.3.2.

As described in the previous section, when receiving a DetNet MPLS
flow, an implementation identifies the app-flow associated with the
incoming packet based on the S-Label. When a node is using platform
labels for S-Labels, any F-Labels can be popped and the S-label
uniquely identifies the app-flow. In the case where platform labels
are not used, F-Label(s) MUST also be provisioned for incoming app-
flows. When PHP is used, incoming interface MUST be provisioned.
The provisioned information MUST then be used to identify incoming
app-flows based on the combination of S-Label and F-Label(s) or
incoming interface.

6.2.3.2. **Common F-Label Processing**

All DetNet aware MPLS nodes process F-Labels as needed to meet the
service requirements of the DetNet flow or flows carried in the LSPs
represented by the F-Labels. This includes normal push, pop and swap
operations. Such processing is essentially the same type of
processing enabled for TE LSPs, although the specific service
parameters, or traffic specification, can differ. When the DetNet
service parameters of the app-flow or flows carried in an LSP
represented by an F-Label can be met by an exiting TE mechanism, the
forwarding sub-layer processing node MAY be a DetNet unaware, i.e.,
standard, MPLS LSR. Such TE LSPs may provide LSP forwarding service
as defined in, but not limited to, [RFC3209], [RFC3270], [RFC3272],
[RFC3473], [RFC4875], [RFC5440], and [RFC6006].

More specifically, as mentioned above, the DetNet forwarding sub-
layer provides explicit routes and allocated resources, and F-Labels
are used to map to each. Explicit routes are supported based on the
topmost (outermost) F-Label that is pushed or swapped and the LSP
that corresponds to this label. This topmost (outgoing) label MUST
be associated with a provisioned outgoing interface and, for non-
point-to-point outgoing interfaces, a next hop LSR. Note that this
information MUST be provisioned via configuration or the controller
plane. In the previously mentioned special case where there is no
added F-labels and the outgoing interface is not a point-to-point
interface, the outgoing interface MUST also be associated with a next
hop LSR.

Resources may be allocated in a hierarchical fashion per LSP that is
represented by each F-Label. Each LSP MAY be provisioned with a
service parameters that dictates the specific traffic treatment to be
received by the traffic carried over that LSP. Implementations of
this document MUST ensure that traffic carried over each LSP
represented by an F-Label receives the traffic treatment provisioned
for that LSP. Typical mechanisms used to provide different treatment
to different flows includes the allocation of system resources (such
as queues and buffers) and provisioning or related parameters (such

as shaping, and policing). Support can also be provided via an
underlying network technology such IEEE802.1 TSN Section 8. Other
than in the TSN case, the specific mechanisms used by a DetNet node
to ensure DetNet service delivery requirements are met for supported
DetNet flows is outside the scope of this document.

Packets that are marked in a way that does not correspond to
allocated resources, e.g., lack a provisioned F-Label, 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:

o MUST defend the DetNet QoS by discarding or remarking (to an
   allocated DetNet flow or non-competing non-DetNet flow) packets
   received that are not the subject of a completed resource
   allocation.

o MUST NOT use a DetNet allocated resource, e.g. a queue or shaper
   reserved for DetNet flows, for any packet that does match the
   corresponding DetNet flow.

o MUST ensure a QoS flow does not exceed its allocated resources or
   provisioned service level,

o MUST ensure a CoS flow or service class does not impact the
   service delivered to other flows. This requirement is similar to
   requirement for MPLS LSRs to that CoS LSPs do not impact the
   resources allocated to TE LSPs, e.g., via [RFC3473].

Subsequent sections provide additional considerations related to CoS
(Section 6.6.1), QoS (Section 6.6.2) and aggregation (Section 6.6.3).

6.3. OAM Indication

OAM follows the procedures set out in [RFC5085] with the restriction
that only Virtual Circuit Connectivity Verification (VCCV) type 1 is
supported.

As shown in Figure 3 of [RFC5085] when the first nibble of the d-CW
is 0x0 the payload following the d-CW is normal user data. However,
when the first nibble of the d-CW is 0X1, the payload that follows
the d-DW is an OAM payload with the OAM type indicated by the value
in the d-CW Channel Type field.

The reader is referred to [RFC5085] for a more detailed description
of the Associated Channel mechanism, and to the DetNet work on OAM
for more information DetNet OAM.

6.4. Flow Aggregation

[Author's note: need to revisit this section and ensure that we cover
(and fully specify) desired types of aggregation.]

1. Aggregate at the LSP (Forwarding)

2. Aggregating DetNet flows as a new DetNet flow

3. Simple Aggregation at the DetNet layer

The resource control and management aspects of aggregation (including
the queuing/shaping/ policing implications) will be covered in other
documents.

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. The DetNet data plane allows for the aggregation of DetNet
flows, to improved scaling. There are three methods of introducing
flow aggregation:

[Editor's note:]

The following review comments were received when this section was
committed to github.

General comment: We should points to the major issue of aggregation,
namely the Seq.Num related problem. The aggregated flows have their
own Seq.Num and those are independent. We should consider to group
the aggregation techniques as per their impact on what DetNet
functions they allow on a DetNet flow. (E.g., aggregation without
new Aggregate.Seq.Num would prohibit usage of FR, EF and in-order-
delivery function on the aggregate flow).

SR based aggregation can be treated as a form of H-LSP aggregation.
Should we differentiate them? What are the differences?

What are the issues when aggregating of different payload types?
Should we add an editor note on this?

Simple-aggregation-at-the-detnet-layer: is this not the same as
H-LSP? The A-label can be treated just as an additional F-Label.

[Editor's note: End of review comment.]

6.4.1. Aggregation at the LSP

   DetNet flows forwarded 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 falls 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.

   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.

6.4.2. Aggregating DetNet Flows as a new DetNet flow

   An aggregate can be built by layering DetNet flows as shown below:

```
    +---------------------------------+
    |                                 |
    |            DetNet Flow          |
    |          Payload  Packet        |
    |                                 |
    +---------------------------------+ <--\
    |         DetNet Control Word     |    |
    +=================================+    |
    |              S-Label            |    |
    +---------------------------------+    +----DetNet data plane
    |         DetNet Control Word     |    |    MPLS encapsulation
    +=================================+    |
    |              A-Label            |    |
    +---------------------------------+    |
    |            F-Label(s)           | <--/
    +---------------------------------+
    |             Data-Link           |
    +---------------------------------+
    |             Physical            |
    +---------------------------------+
```
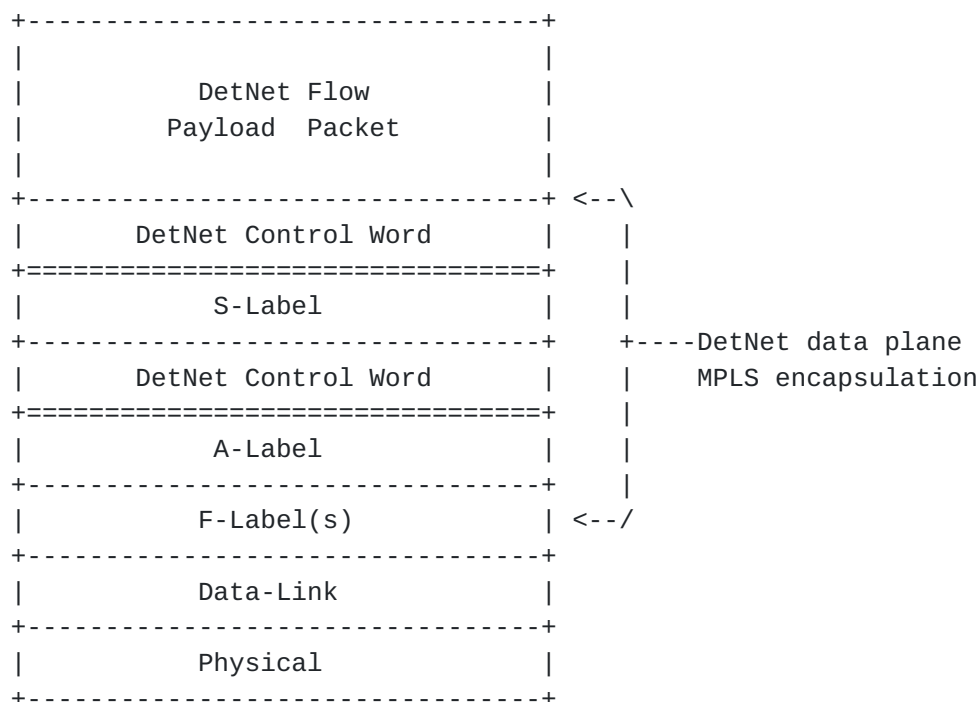
Both the Aggregation (A) label and the S-Label have their MPLS S bit
set indicating bottom of stack, and the d-CW allows the PREOF to
work.

It is a property of the A-label that what follows is d-CW followed by
an S-Label. A relay node processing the A-label would not know the
underlying payload type. This would only be known to a node that was
a peer of the node imposing the S-Label. However there is no real
need for it to know the payload type during aggregation processing.

6.4.3. Simple Aggregation at the DetNet Layer

Another approach would be not to include a d-CW for the aggregated
flow. This would be functionally similar to aggregation at the
forwarding sub-layer using H-LSPs, but would confine knowledge of the
aggregation to the DetNet layer. Such an approach shares the
disadvantage that PREOF operations would not be possible. OAM
operation in this mode is for further study.

```
 +--------------------------------+
 |                                |
 |            DetNet Flow         |
 |          Payload  Packet       |
 |                                |
 +--------------------------------+ <--\
 |         DetNet Control Word    |    |
 +================================+    |
 |              S-Label           |    +----DetNet data plane
 +--------------------------------+    |     MPLS encapsulation
 |              A-Label           |    |
 +--------------------------------+    |
 |             F-Label(s)         | <--/
 +--------------------------------+
 |              Data-Link         |
 +--------------------------------+
 |              Physical          |
 +--------------------------------+
```
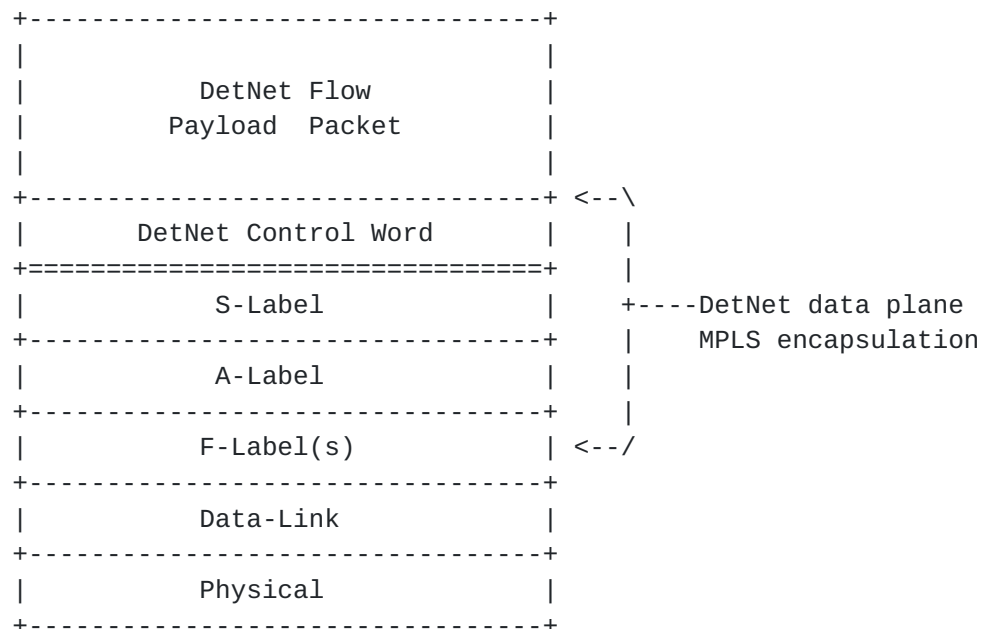
6.5. Service Sub-Layer Considerations

The edge and relay node internal procedures related to PREOF 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).

It is important that the DetNet layer is configured such that a
DetNet node never receives its own replicated packets. If it were to
receive such packets the replication function would make the loop
more destructive of bandwidth than a conventional unicast loop.
Ultimately the TTL in the S-Label will cause the packet to die during
a transient, but given the sensitivity of applications to packet
latency the impact on the DetNet application would be severe.

6.5.1. Edge Node Processing

An edge node is resposable for matching ingress packets to the
service they require and encapsulating them accordingly.An edge node
may participate in the packet replication and duplication
elimination.

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.

The internal design of a relay node is out of scope of this document.
However the reader's attention is drawn to the need to make any PREOF
state available to the packet processor(s) dealing with packets to
which the PREOF functions must be applied, and to maintain that state
is such as way that it is available to the packet processor operation
on the next packet in the DetNet flow (which may be a duplicate, a
late packet, or the next packet in sequence.

[Editor's note: I think the rest of this section belongs in a new
"802.1 TSN (island Interconnect) over DetNet MPLS" section.]

This may be done in the DetNet layer, or where the native service
processing (NSP) [RFC3985] is IEEE 802.1CB [IEEE8021CB] capable, the
packet replication and duplicate elimination MAY entirely be done in
the NSP, bypassing the DetNet flow encapsulation and logic entirely.
This enables operating over unmodified implementations and
deployments. The NSP approach works only between edge nodes and
cannot make use of relay nodes.

The NSP approach is useful end to end tunnel and for for "island
interconnect" scenarios. However, when there is a need to do PREOF
in a middle of the network, such plain edge to edge operation is not
sufficient.

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).

6.5.2. Relay Node Processing

A DetNet Relay node operates in the DetNet forwarding sub-layer .
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. Some relay
nodes may be DetNet service aware, while others may be unmodified
LSRs that only understand how to swicth MPLS-TE LSPs.

It is also possible to treat the relay node as a transit node, see
Section 6.6.3. Again, this is entirely up to how the forwarding has
been programmed.

6.6. Forwarding Sub-Layer Considerations

6.6.1. Class of Service

Class and quality of service, i.e., CoS and QoS, are terms that are
often used interchangeably and confused with each other. 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].

6.6.2. Quality of Service

In addition to explicit routes, and packet replication and
elimination, described in Section 6 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. This includes Quality of Service (QoS) mechanisms at
the MPLS layer, that may be combined with the mechanisms defined by
the underlying network layer such as 802.1TSN.

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.

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 use to support the identification of flows requiring
DetNet QoS.

6.6.3. Cross-DetNet Flow Resource Aggregation

[Editor's NOTE: Isn't this section the same as "Aggregation at the
LSP". -- Address as part of aggregation section cleanup.]

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 forwarded 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 falls 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.

[NOTE: This needs to be revised:] 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.

6.6.4. Layer 2 Addressing and QoS Considerations

[Editor's NOTE: review and simplify this section. Doesn't this
belong in the TSN section? Alternatively, describe in generic/non
sub-network technology specific terms.]

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 IP encapsulations.

6.6.5. Time Synchronization

[Editor's Note: A detailed discussion of time synchronization is
outside the scope of this document, and the production of a
specialist text discussing this topic is encouraged. This section
will be updated/removed if such a document is available before
publication of this text.]

Time synchronization is important both from the perspective of
operating the DetNet network itself and from the perspective of
transferring time across the network between client applications.
Some clients may be able to use the DetNet as their provider of time
and frequency, others may require the DetNet to transfer time between
a client clock source and a client clock user.

For example, [RFC8169] describes a method of recording the packet
queuing time in an MPLS LSR on a packet by per packet basis and
forwarding this information to the egress edge system. This allows
compensation for any variable packet queuing delay to be applied at
the packet receiver. Other mechanisms for IP/MPLS networks are
defined based on IEEE Standard 1588 [IEEE1588], such as ITU-T
[G.8275.1] and [G.8275.2].

A more detailed discussion of time synchronization is outside the
scope of this document.

7. Controller Plane (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, and the DetNet architecture
[I-D.ietf-detnet-architecture] refers to these collectively as the
'Controller Plane'. This document therefore does not distinguish
between information provided by distributed control plane protocols,
e.g., RSVP-TE [RFC3209] and [RFC3473], or by centralized network

management mechanisms, e.g., RestConf [RFC8040], YANG [RFC7950], and
the Path Computation Element Communication Protocol (PCEP)
[I-D.ietf-pce-pcep-extension-for-pce-controller] or any combination
thereof. Specific considerations and requirements for the DetNet
Controller Plane are discussed in Section 7.6.

7.1. S-Label and F-Label Assignment and Distribution

[Editor's note - we may need additional text on resource allocation
in this section.]

DetNet S-Labels (see Section 6.2.2 for their definition) are similar
to other MPLS service labels that denote the contents of the MPLS
packet payload such as a layer 2 pseudowire, an IP packet that is
routed in a VPN context with a private address, or an Ethernet
virtual private network (EVPN) service.

S-Labels are expected to be allocated in the same manner as any other
service labels. S-Labels uniquely identify a particular DetNet flow,
and are local to the node on which the label is allocated. In the
DetNet service sub-layer the explicit route consists of the set of
Relay Nodes that the DetNet flow must traverse. They can be used to
identify the DetNet flow that a packet belongs to as it traverses a
particular node in a DetNet domain. Because labels are local to each
node rather than being a global identifier within a domain, they must
be advertised to their upstream DetNet service-aware peer nodes
(e.g., a DetNet MPLS End System as shown in Figure 3, or a DetNet
Relay or Edge Node as shown in Figure 7) and interpreted in the
context of their received F-Label.

As discussed in Section 4, the forwarding sub-layer uses one or more
F-Labels to forward DetNet packets between DetNet service-aware nodes
along explicitly defined routes at the DetNet forwarding sub-layer,
which in the context of this document is the MPLS layer. F-Labels
can also provide context for an S-Label. In the DetNet Forwarding
(MPLS) sub-layer the explicit route consists of the set of DetNet
nodes which are LSRs, links, and possibly link bundle members and
queues that the DetNet packets of a flow must traverse between nodes
in the DetNet service sub-layer (i.e. between a specific Edge Node
and the next hop Relay Node, between specific Relay Nodes, and
between a specific Relay node and the egress Edge Node. Resource
allocation corresponding to the set of Services supported over the
forwarding sub-layer, which may or may not include aggregation, is
required at this sub-layer. Explicit routes are used to ensure that
packets are routed through the resources that have been reserved for
them, and hence provide the DetNet application with the required
service. Multiple F-Labels may be pushed after an S-Label and there
is no requirement for all F-Labels to be controlled via the same

controller mechanisms. For example in EVPN, some labels are
distributed using BGP while others are distributed using LDP or RSVP.

Whether configuring, calculating and instantiating these routes is a
single-stage or multi-stage process, or in a centralized or
distributed manner, is out of scope of this document.

There are a number of approaches that could be used to provide
explicit routes and resource allocation in the MPLS layer:

o The path could be explicitly set up by a controller which
   calculates the path and explicitly configures each node along that
   path with the appropriate forwarding and resource allocation
   information.

o The path could be set up using RSVP-TE signaling.

o The path could be implemented using MPLS-based segment routing
   when extended to support resource allocation.

See Section 7.6 for further discussion of these alternatives.

Much like other MPLS labels, there are a number of alternatives
available for DetNet S-Label and F-Label advertisement to an upstream
peer node. These include distributed signaling protocols such as
RSVP-TE, centralized label distribution via a controller that manages
both the sender and the receiver using NETCONF/YANG, BGP, PCEP, etc.,
and hybrid combinations of the two. The details of the controller
plane solution required for the label distribution and the management
of the label number space are out of scope of this document, but as
mentioned above, there are particular DetNet considerations and
requirements that are discussed in Section 7.6.

7.2. Packet Replication, Elimination, and Ordering (PREOF)

The controller plane protocol solution required for managing the
PREOF processing is outside the scope of this document. That said,
it should be noted that the ability to determine, for a particular
flow, optimal packet replication and elimination points in the DetNet
domain requires explicit support. There are be capabilities that can
be used, or extended, for example GMPLS end-to-end recovery [RFC4872]
and GMPLS segment recovery [RFC4873].

7.3. Contention Loss and Jitter Reduction

As discussed in Section 1, this document does not specify the
mechanisms needed to eliminate contention loss or reduce jitter for
DetNet flows at the DetNet forwarding sub-layer. The ability to

manage node and link resources to be able to provide these functions
will be a necessary part of the DetNet controller plane. It will
also be necessary to be able to control the required queuing
mechanisms used to provide these functions along a flow's path
through the network. See Section 7.6 for further discussion of these
requirements.

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 forwarding
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 reservations may
differ in each direction. The concepts of associated bidirectional
flows and co-routed bidirectional flows can be applied to DetNet
flows.

While the MPLS data plane must support bidirectional DetNet flows,
there are no special bidirectional features with respect to the data
plane other than the need for the two directions of a co-routed
bidirectional flow to take the same path. 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 MPLS Labels
in each direction.

DetNet's use of PREOF may increase the complexity of using co-routing
bidirectional flows, since if PREOF is used, then the replication
points in one direction would have to match the elimination points in
the other direction, and vice versa, and the optimal points for these
functions in one direction may not match the optimal points in the
other.

Control and management mechanisms will need to support bidirectional
flows, but the specification of such mechanisms are out of scope of
this document. Related control plan mechanisms have been defined in
[RFC3473], [RFC6387] and [RFC7551].

This is further discussed in Section 7.6.

7.5. Flow Aggregation Control

Section 6.4 discusses the use of flow aggregation in DetNet. It
includes flow aggregation accomplished through the use of
hierarchical LSPs, aggregating multiple DetNet flows into a single
new DetNet flow, and simple aggregation at the DetNet layer. It will
be the responsibility of the DetNet controller plane to be able to
properly provision the use of these mechanisms. These requirements
are included in the next section.

7.6. DetNet Controller (Control and Management) Plane Requirements

While the definition of controller plane for DetNet is out of the
scope of this document, there are particular considerations and
requirements for such that result from the unique characteristics of
the DetNet architecture [I-D.ietf-detnet-architecture] and data plane
as defined herein.

The primary requirements of the DetNet controller plane are that it
must be able to:

o Instantiate DetNet flows in a DetNet domain (which may include
   some or all of explicit path and PREOF replication and elimination
   node determination, link bandwidth reservations, 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.

o Manage DetNet S-Label and F-Label allocation and distribution,
   when the DetNet MPLS encapsulation is in use

o The ability to support DetNet flow aggregation

o 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)

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

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

These requirements, as stated earlier, could be satisfied using
distributed control protocol signaling (such as RSVP-TE), centralized
network management provisioning mechanisms (such as BGP, PCEP, YANG
[I-D.ietf-detnet-flow-information-model], etc.) or hybrid
combinations of the two, and could also make use of MPLS-based
segment routing.

In the abstract, the results of either distributed signaling or
centralized provisioning are equivalent from a DetNet data plane
perspective - flows are instantiated, explicit routes are determined,
resources are reserved, and packets are forwarded through the domain
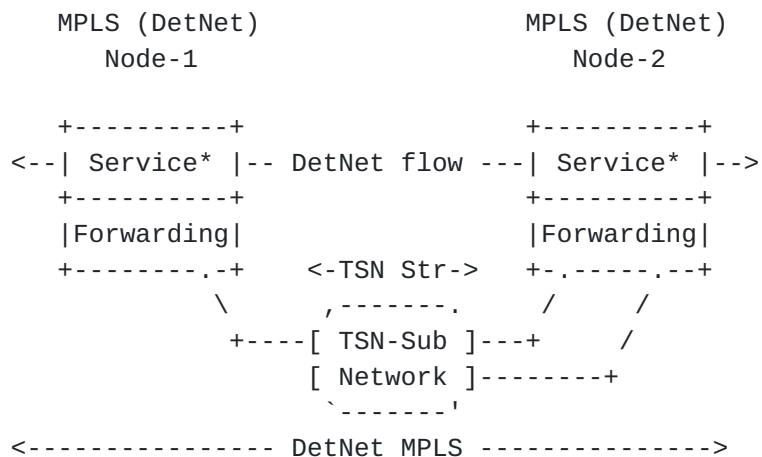using the MPLS data plane.

However, from a practical and implementation standpoint, they are not
equivalent at all. Some approaches are more scalable than others in
terms of signaling load on the network. Some can take advantage of
global tracking of resources in the DetNet domain for better overall
network resource optimization. Some are more resilient than others
if link, node, or management equipment failures occur. While a
detailed analysis of the control plane alternatives is out of the
scope of this document, the requirements from this document can be
used as the basis of a later analysis of the alternatives.

8. DetNet MPLS Operation Over IEEE 802.1 TSN Sub-Networks

[Editor's note: this is a place holder section. A standalone section
on MPLS over IEEE 802.1 TSN. Includes RFC2119 Language.]

This section covers how DetNet MPLS flows operate over an IEEE 802.1
TSN sub-network. Figure 15 illustrates such a scenario, where two
MPLS (DetNet) nodes are interconnected by a TSN sub-network. Node-1
is single homed and Node-2 is dual-homed. MPLS nodes can be (1)
DetNet MPLS End System, (2) DetNet MPLS Edge or Relay node or (3)
MPLS Transit node.

Note: in case of MPLS Transit node there is no DetNet Service sub-
layer processing.

```
      MPLS (DetNet)                   MPLS (DetNet)
         Node-1                          Node-2


      +----------+                  +----------+
   <--| Service* |-- DetNet flow ---| Service* |-->
      +----------+                  +----------+
      |Forwarding|                  |Forwarding|
      +--------.-+    <-TSN Str->   +-.------.--+
               \      ,-------.     /     /
              +----[ TSN-Sub ]---+     /
                   [ Network ]--------+
                    `-------'
   <--------------- DetNet MPLS --------------->
```

Note: * no service sub-layer required for transit nodes

        Figure 15: DetNet Enabled MPLS Network Over a TSN Sub-Network

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. Furthermore IEEE 802.1CB
[IEEE8021CB] defines frame replication and elimination functions for
reliability that should prove both compatible with and useful to,
DetNet networks. All these functions have to identify flows those
require TSN treatment.

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 may need a CoS marking, such as the priority field of an IEEE
Std 802.1Q VLAN tag, to give the packet proper service.

The challange for MPLS DeNet flows is that the protocol interworking
function defined in IEEE 802.1CB [IEEE8021CB] works only for IP
flows. The aim of the protocol interworking function is to convert
an ingress flow to use a specific multicast destination MAC address
and VLAN, for example to direct the packets through a specific path
inside the bridged network. A similar interworking pair at the other
end of the TSN sub-network would restore the packet to its original
destination MAC address and VLAN.

As protocol interworking function defined in [IEEE8021CB] does not
work for MPLS labeled flows, the DetNet MPLS nodes MUST ensure proper
TSN sub-network specific Ethernet encapsulation of the DetNet MPLS
packets. For a given TSN Stream (i.e., DetNet flow) an MPLS (DetNet)
node MUST behave as a TSN-aware Talker or a Listener inside the TSN
sub-network.

8.1. Mapping of TSN Stream ID and Sequence Number

TSN capable MPLS (DetNet) nodes are TSN-aware Talker/Listener as
shown in Figure 16. MPLS (DetNet) node MUST provide the TSN sub-
network specific Ethernet encapsulation over the link(s) towards the
sub-network. An TSN-aware MPLS (DetNet) node MUST support the
following TSN components:

1. For recognizing flows:

    * Stream Identification (MPLS-flow-aware)

2. For FRER used inside the TSN domain, additonaly:

    * Sequencing function (MPLS-flow-aware)

    * Sequence encode/decode function

3. For FRER when the node is a TSN replication or elimination point,
    additionally:

    * Stream splitting function

    * Individual recovery function

[Editor's note: Should we added here requirements regarding IEEE
802.1Q C-VLAN component?]

The Stream Identification and The Sequencing functions are slightly
modified for frames passed down the protocol stack from the upper
layers.

Stream Identification MUST pair MPLS flows and TSN Streams and encode
that in data plane formats as well. The packet's stream_handle
subparameter (see IEEE 802.1CB [IEEE8021CB]) inside the Talker/
Listener is defined based on the Flow-ID used in the upper DetNet
MPLS layer. Stream Identification function MUST encode Ethernet
header fields namely (1) the destination MAC-address, (2) the VLAN-ID
and (3) priority parameters with TSN sub-network specific values.
Encoding is provided for the frame passed down the stack from the
upper layers.

The sequence generation function resides in the Sequencing function.
It generates a sequence_number subparameter for each packet of a
Stream passed down to the lower layers. Sequencing function MUST
copy sequence information from the MPLS d-CW of the packet to the
sequence_number subparameter for the frame passed down the stack from
the upper layers.

```
        MPLS (DetNet)
           Node-1
        <---------->


        +----------+
     <--| Service  |-- DetNet flow -----------------
        +----------+
        |Forwarding|
        +----------+    +--------------+
        | L2 with  |<---| L2 Relay with |---- TSN ----
        |   TSN    |    | TSN function  |    Stream
        +-----.----+    +--.---------.--+
               _____/           _____


         TSN-aware
          Talker /           TSN-Bridge
          Listener             Relay


             <--------- TSN sub-network ------------
```
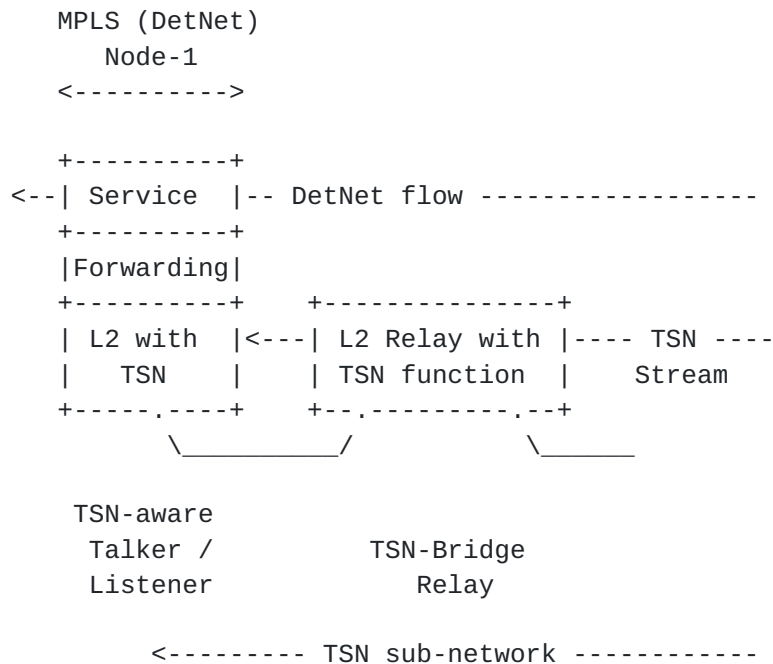
              Figure 16: MPLS (DetNet) Node with TSN Functions

   The Sequence encode/decode function MUST support the Redundancy tag
   (R-TAG) format as per Clause 7.8 of IEEE 802.1CB [IEEE8021CB].

8.2. TSN Usage of FRER

   TSN Streams supporting DetNet flows may use Frame Replication and
   Elimination for Redundancy (FRER) [802.1CB] based on the loss service
   requirements of the TSN Stream, which is derived from the DetNet
   service requirements of the DetNet mapped flow. The specific
   operation of FRER is not modified by the use of DetNet and follows
   IEEE 802.1CB [IEEE8021CB].

   FRER function and the provided service recovery is available only
   within the TSN sub-network however as the Stream-ID and the TSN
   sequence number are paired with the MPLS flow parameters they can be
   combined with PREOF functions.

8.3. Management and Control Implications

   [Editor's note: This section is TBD Covers Creation, mapping, removal
   of TSN Stream IDs, related parameters and,when needed, configuration
   of FRER. Supported by management/control plane.]

9. **DetNet MPLS Operation over DetNet IP PSNs**

   This section specifies the DetNet encapsulation over an IP network.
   The approach is modeled on the operation of MPLS and PseudoWires (PW)
   over an IP Packet Switched Network (PSN) [RFC3985][RFC4385][RFC7510].
   It maps the MPLS data plane encapsulation described in Section 6.2 to
   the DetNet IP data plane define in [I-D.ietf-detnet-dp-sol-ip].

   To carry DetNet with full functionality at the DetNet layer over an
   IP network, the following components are required (these are a subset
   of the requirements for MPLS encapsulation listed in Section 6.1):

   1. A method of identifying the DetNet flow group to the processing
       element.

   2. A method of carrying the DetNet sequence number.

   3. A method of distinguishing DetNet OAM packets from DetNet data
       packets.

   4. A method of carrying queuing and forwarding indication.

   These requirements are satisfied by the DetNet over MPLS
   Encapsulation described in Section 6.2.

   This document builds on the the specification of MPLS over UDP and IP
   defined in [RFC7510]. It replaces the the F-Label(s) used in
   Section 6.2 with UDP and IP headers. The UDP and IP header
   information is used to identify DetNet flows, including member flows,
   per [I-D.ietf-detnet-dp-sol-ip]. The resulting encapsulation is
   shown in Figure 17.

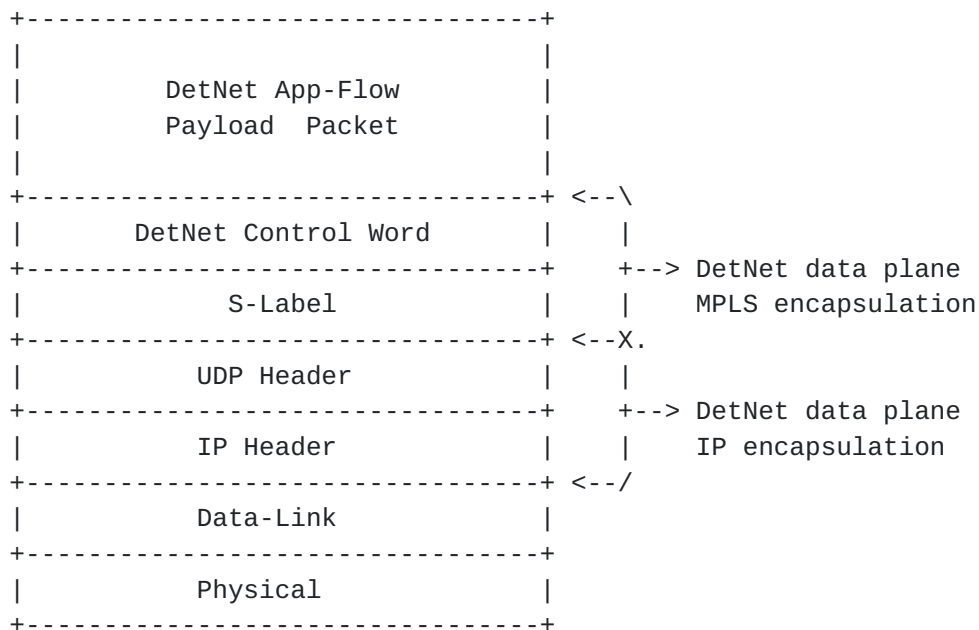   Note that this encapsulation works equally well with IPv4 and IPv6.

```
       +-------------------------------+
       |                               |
       |        DetNet App-Flow        |
       |        Payload  Packet        |
       |                               |
       +-------------------------------+ <--\
       |      DetNet Control Word      |    |
       +-------------------------------+    +--> DetNet data plane
       |            S-Label            |    |     MPLS encapsulation
       +-------------------------------+ <--X.
       |          UDP Header           |    |
       +-------------------------------+    +--> DetNet data plane
       |          IP Header            |    |     IP encapsulation
       +-------------------------------+ <--/
       |          Data-Link            |
       +-------------------------------+
       |          Physical             |
       +-------------------------------+
```

             Figure 17: IP Encapsulation of DetNet MPLS

   d-CW and and S-Labels are used as defined in Section 6.2 and are not
   modified by this section.

   To support outgoing DetNet MPLS over IP, an implementation MUST
   support the provisioning of IP/UDP header information in place of
   sets of F-Labels. Note that multiple sets of F-Labels can be
   provisioned to support PRF on transmitted DetNet flows and therefore,
   when PRF is supported, multiple IP/UDP headers MAY be provisioned.
   When multiple IP/UDP headers are provisioned for a particular
   outgoing app-flow, a copy of the outgoing packet, including the
   pushed S-Label, MUST be made for each. The headers for each outgoing
   packet MUST be based on the configuration information and as defined
   in [RFC7510], with one exception. The one exceptions is that the UDP
   Source Port value MUST be set to uniquely identify the DetNet
   (forwarding sub-layer) flow. The packet MUST then be handed as a
   DetNet IP packet, per [I-D.ietf-detnet-dp-sol-ip].

   To support receive processing an implementation MUST also support the
   provisioning of received IP/UDP header information. When S-Labels
   are taken from platform label space, all that is required is to
   provision that receiving IP/UDP encapsulated DetNet MPLS packets is
   permitted. Once the IP/UDP header is stripped, the S-label uniquely
   identifies the app-flow. When S-Labels are not taken from platform
   label space, IP/UDP header information MUST be provisioned. The
   provisioned information MUST then be used to identify incoming app-

flows based on the combination of S-Label and incoming IP/UDP header.
Normal receive processing, including PEOF can then take place.

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

This document makes no IANA requests.

12. Contributors

RFC7322 limits the number of authors listed on the front page of a
draft to a maximum of 5, far fewer than the many individuals below
who made important contributions to this draft. The editor wishes to
thank and acknowledge each of the following authors for contributing
text to this draft. See also Section 13.

    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

    Janos Farkas
    Ericsson
    Magyar Tudosok krt. 11.
    Budapest 1117
    Hungary
    Email: janos.farkas@ericsson.com

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

      Spain
      Email: cjbc@it.uc3m.es

      Tal Mizrahi
      Marvell
      6 Hamada st.
      Yokneam
      Israel
      Email: talmi@marvell.com

      Lou Berger
      LabN Consulting, L.L.C.
      Email: lberger@labn.net

      Stewart Bryant
      Huawei Technologies
      Email: stewart.bryant@gmail.com

      Mach Chen
      Huawei Technologies
      Email: mach.chen@huawei.com

      Andrew G. Malis
      Huawei Technologies
      Email: agmalis@gmail.com

[13]. Acknowledgements

Andrew Malis

Carlos J. Bernardos

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

Lou Berger

Pat Thaler

Thanks for Stewart Bryant for his extensive review of the previous
versions of the document.

14. References

14.1. Normative References

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

[RFC2211] Wroclawski, J., "Specification of the Controlled-Load
           Network Element Service", RFC 2211, DOI 10.17487/RFC2211,
           September 1997, <https://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,
           <https://www.rfc-editor.org/info/rfc2212>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol
           Label Switching Architecture", RFC 3031,
           DOI 10.17487/RFC3031, January 2001,
           <https://www.rfc-editor.org/info/rfc3031>.

[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,
           <https://www.rfc-editor.org/info/rfc3032>.

[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>.

   [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,
              <https://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,
              <https://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,
              <https://www.rfc-editor.org/info/rfc3473>.

   [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,
              <https://www.rfc-editor.org/info/rfc4206>.

   [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>.

   [RFC5085] Nadeau, T., Ed. and C. Pignataro, Ed., "Pseudowire Virtual
              Circuit Connectivity Verification (VCCV): A Control
              Channel for Pseudowires", RFC 5085, DOI 10.17487/RFC5085,
              December 2007, <https://www.rfc-editor.org/info/rfc5085>.

   [RFC5129] Davie, B., Briscoe, B., and J. Tay, "Explicit Congestion
              Marking in MPLS", RFC 5129, DOI 10.17487/RFC5129, January
              2008, <https://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, <https://www.rfc-editor.org/info/rfc5462>.

   [RFC7510] Xu, X., Sheth, N., Yong, L., Callon, R., and D. Black,
              "Encapsulating MPLS in UDP", RFC 7510,
              DOI 10.17487/RFC7510, April 2015,
              <https://www.rfc-editor.org/info/rfc7510>.

   [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>.

14.2. Informative References

   [G.8275.1]
              International Telecommunication Union, "Precision time
              protocol telecom profile for phase/time synchronization
              with full timing support from the network", ITU-T
              G.8275.1/Y.1369.1 G.8275.1, June 2016,
              <https://www.itu.int/rec/T-REC-G.8275.1/en>.

   [G.8275.2]
              International Telecommunication Union, "Precision time
              protocol telecom profile for phase/time synchronization
              with partial timing support from the network", ITU-T
              G.8275.2/Y.1369.2 G.8275.2, June 2016,
              <https://www.itu.int/rec/T-REC-G.8275.2/en>.

   [I-D.ietf-detnet-architecture]
              Finn, N., Thubert, P., Varga, B., and J. Farkas,
              "Deterministic Networking Architecture", draft-ietf-
              detnet-architecture-11 (work in progress), February 2019.

   [I-D.ietf-detnet-dp-sol-ip]
              Korhonen, J., Varga, B., "DetNet IP Data Plane
              Encapsulation", 2018.

   [I-D.ietf-detnet-flow-information-model]
              Farkas, J., Varga, B., Cummings, R., and Y. Jiang, "DetNet
              Flow Information Model", draft-ietf-detnet-flow-
              information-model-03 (work in progress), March 2019.

   [I-D.ietf-pce-pcep-extension-for-pce-controller]
              Zhao, Q., Li, Z., Negi, M., and C. Zhou, "PCEP Procedures
              and Protocol Extensions for Using PCE as a Central
              Controller (PCECC) of LSPs", draft-ietf-pce-pcep-
              extension-for-pce-controller-01 (work in progress),
              February 2019.

   [I-D.ietf-spring-segment-routing-mpls]
              Bashandy, A., Filsfils, C., Previdi, S., Decraene, B.,
              Litkowski, S., and R. Shakir, "Segment Routing with MPLS
              data plane", draft-ietf-spring-segment-routing-mpls-18
              (work in progress), December 2018.

   [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, <https://www.rfc-editor.org/info/rfc2205>.

   [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,
             <https://www.rfc-editor.org/info/rfc2474>.

   [RFC3272] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., and X.
             Xiao, "Overview and Principles of Internet Traffic
             Engineering", RFC 3272, DOI 10.17487/RFC3272, May 2002,
             <https://www.rfc-editor.org/info/rfc3272>.

   [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>.

   [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,
             <https://www.rfc-editor.org/info/rfc4448>.

   [RFC4872] Lang, J., Ed., Rekhter, Y., Ed., and D. Papadimitriou,
              Ed., "RSVP-TE Extensions in Support of End-to-End
              Generalized Multi-Protocol Label Switching (GMPLS)
              Recovery", RFC 4872, DOI 10.17487/RFC4872, May 2007,
              <https://www.rfc-editor.org/info/rfc4872>.

   [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel,
              "GMPLS Segment Recovery", RFC 4873, DOI 10.17487/RFC4873,
              May 2007, <https://www.rfc-editor.org/info/rfc4873>.

   [RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S.
              Yasukawa, Ed., "Extensions to Resource Reservation
              Protocol - Traffic Engineering (RSVP-TE) for Point-to-
              Multipoint TE Label Switched Paths (LSPs)", RFC 4875,
              DOI 10.17487/RFC4875, May 2007,
              <https://www.rfc-editor.org/info/rfc4875>.

   [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation
              Element (PCE) Communication Protocol (PCEP)", RFC 5440,
              DOI 10.17487/RFC5440, March 2009,
              <https://www.rfc-editor.org/info/rfc5440>.

   [RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed.,
              "MPLS Generic Associated Channel", RFC 5586,
              DOI 10.17487/RFC5586, June 2009,
              <https://www.rfc-editor.org/info/rfc5586>.

   [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, <https://www.rfc-editor.org/info/rfc5654>.

   [RFC5921] Bocci, M., Ed., Bryant, S., Ed., Frost, D., Ed., Levrau,
              L., and L. Berger, "A Framework for MPLS in Transport
              Networks", RFC 5921, DOI 10.17487/RFC5921, July 2010,
              <https://www.rfc-editor.org/info/rfc5921>.

   [RFC6003] Papadimitriou, D., "Ethernet Traffic Parameters",
              RFC 6003, DOI 10.17487/RFC6003, October 2010,
              <https://www.rfc-editor.org/info/rfc6003>.

   [RFC6006] Zhao, Q., Ed., King, D., Ed., Verhaeghe, F., Takeda, T.,
              Ali, Z., and J. Meuric, "Extensions to the Path
              Computation Element Communication Protocol (PCEP) for
              Point-to-Multipoint Traffic Engineering Label Switched
              Paths", RFC 6006, DOI 10.17487/RFC6006, September 2010,
              <https://www.rfc-editor.org/info/rfc6006>.

   [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M.
              Aissaoui, "Segmented Pseudowire", RFC 6073,
              DOI 10.17487/RFC6073, January 2011,
              <https://www.rfc-editor.org/info/rfc6073>.

   [RFC6387] Takacs, A., Berger, L., Caviglia, D., Fedyk, D., and J.
              Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label
              Switched Paths (LSPs)", RFC 6387, DOI 10.17487/RFC6387,
              September 2011, <https://www.rfc-editor.org/info/rfc6387>.

   [RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and
              L. Yong, "The Use of Entropy Labels in MPLS Forwarding",
              RFC 6790, DOI 10.17487/RFC6790, November 2012,
              <https://www.rfc-editor.org/info/rfc6790>.

   [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,
              <https://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,
              <https://www.rfc-editor.org/info/rfc7950>.

   [RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF
              Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017,
              <https://www.rfc-editor.org/info/rfc8040>.

   [RFC8169] Mirsky, G., Ruffini, S., Gray, E., Drake, J., Bryant, S.,
              and A. Vainshtein, "Residence Time Measurement in MPLS
              Networks", RFC 8169, DOI 10.17487/RFC8169, May 2017,
              <https://www.rfc-editor.org/info/rfc8169>.

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 PREOF.
   The figure is subject to change depending on the further DT decisions
   on the label handling..]

Authors' Addresses

   Jouni Korhonen (editor)

   Email: jouni.nospam@gmail.com

Balazs Varga (editor)
Ericsson
Magyar Tudosok krt. 11.
Budapest 1117
Hungary

Email: balazs.a.varga@ericsson.com