DetNet   J. Korhonen, Ed.
Internet-Draft   Nordic
Intended status: Standards Track   L. Andersson
Expires: May 3, 2018   Y. Jiang
                                                                  N. Finn
                                                                   Huawei
                                                                 B. Varga
                                                                J. Farkas
                                                                 Ericsson
                                                            CJ. Bernardos
                                                                     UC3M
                                                               T. Mizrahi
                                                                  Marvell
                                                                L. Berger
                                                                     LabN
                                                         October 30, 2017

                    DetNet Data Plane Encapsulation
                       draft-ietf-detnet-dp-sol-00

Abstract

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

   Comment #1: SB> An overarching comment is that the early part of the
     document is really fundamental architecture and perhaps belongs in
     the arch draft, leaving this draft to be more specific about
     solutions. Indeed if we cannot find a single solution that maps to
     both IP and MPLS underlays I wonder if we should publish two
     specialist RFCs?

   Discussion: One document at the beginning, split into two if/when
     needed. Would be post adoption in any case.

   Comment #2: SB> Whilst I think we should look for a common solution
     to IP and MPLS I do not think that this is where the DT ended up.

   Discussion: Agree.

Status of This Memo

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 extremely low
   packet loss rates and assured maximum end-to-end delivery latency.
   General background and concepts of DetNet can be found in
   [I-D.ietf-detnet-architecture].

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

   o   PW-based: One solution is based on PseudoWires (PW) [RFC3985] and
       [RFC5036] and makes use of multi-segment pseudowires (MS-PW)
       [RFC6073] to map DetNet Relay and Edge Nodes
       [I-D.ietf-detnet-architecture] [I-D.ietf-detnet-dp-alt] to PW
       architecture. The PW-based data plane can be run over an MPLS
       [RFC4448] [RFC6658] Packet Switched Network (PSN).

       Comment #3: SB> This is really an MPLS one. The world of IP PWs is
          a bit scruffy with some work in PWE3 and some in L2TPext which

   really went their own ways. There is for example no MS-PW IP
   design. The MS-PW approach needs to be examined more closely by
   the WG and thus at this stage be marked as provisional.

   Discussion: Agree. "can be" -> "is".

   Comment #3.1 LB> EVPN-based encapsulation is also a potential
   candidate. In general DetNet should look more closely to the
   delevopment around EVPN.

   Discussion Agree. EVPN could be a potential solution and the on-
   wire encapsuations are likely to be the same as PW-based
   encapsulation would be. EVPN even recommends using Control Word
   [RFC8214] (in the absence of entropy labels).

   o Native-IP: The other solution is based on IP header fields, namely
      on the IPv6 Flow Label and a new DetNet Control Word extension
      header option. It is targeted for native IPv6 networks.

   Comment #4: SB> The IP solution has not been properly examined by
   the WG and needs to be marked as provisional.

   Discussion: IP vs. MPLS is a deployment choice.

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

   Comment #5: SB> We clearly need to look at the implications of
   running this with a new IP header extension. Firstly we need input
   from 6man, but we also need to understand what happens in middle
   boxes, other components of the host stack etc.

   Discussion: A WG can develop their own extensions and then get
   approval from 6man. Sometimes that ends up redoing extensions in
   6man but not always.

   This document specifies the encapsulation for DetNet flows, including
   a DetNet Control Word (CW). Furthermore, it describes how DetNet
   flows are identified, how DetNet Relay and Edge nodes work, and how
   the Packet Replication and Elimination function (PREF) is implemented
   with these two data plane solutions. This document does not define
   the associated control plane functions, or Operations,
   Administration, and Maintenance (OAM). It also does not specify
   traffic handling capabilities required to deliver congestion

protection and latency control to DetNet flows as this is defined to
be provided by the underlying MPLS or IP network.

Comment #6: SB> OK, although I think that this may be a mistake.
  There may well be some coupling needed between the Detnet DP and
  the substrate/transport/underlay needed to make this work. If this
  is a genuine technical layering we should talk about it. If this
  is an artificial constraint imposed by the IESG we need to talk to
  them.

Discussion: The only interaction needed is that the flow
  identification is possible. That needs to be available for lower
  layers.

Comment #6.1: LA> Even though this document does not specify any OAM
  functions, we will need to verify that the GACh (Generalized
  Associate Channel) works correctly in a network that has
  replication and elimination.

Discussion: --

2. Terminology

2.1. Terms used in this document

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

The following terms are also used in this document:

DA-T-PE   MPLS based DetNet edge node: a DetNet-aware PseudoWire
               Terminating Provider Edge (T-PE).

DA-S-PE   MPLS based DetNet relay node: a DetNet-aware PseudoWire
               Switching Provider Edge (S-PE).

Comment #7 SB> We need to look at whether the S-PE concept is the
  best fit, or whether we should use introduce a Detnet relay to do
  this. An S-PE just swaps the PW label, but Detnet needs it to do
  more and a better model might be a new construct. However we
  could also discard the relay concept and run 1+n end to end, in
  which case the S-PEs would retain heir original function.

Discussion: Disagree of the dropping comment. However, the issues
  are most likely terminology related. The relay concept is part of
  the DetNet architecture A DA-S-PE was intended to be a DetNet
  relay, which may do more than just swapping labels (PREF

functionality). Current text is quite biased to MS-PW, which was
the starting point for the DetNet relay in a MPLS PW network.

T-Label   A label used to identify the LSP used to transport a
               DetNet flow across an MPLS PSN, e.g., a hop-by-hop
               label used between label switching routers (LSR).

S-Label   A DetNet node to DetNet node "service" label that is
               used between DA-*-PE devices.

PW Label   A PseudoWire label that is used to identify DetNet flow
               related PW Instances within a PE node.

Flow Label   IPv6 header field that is used to identify a DetNet
               flow (together with the source IP address field).

Comment #8 SB> If this is the IPv6 Flow label I think caution is
   needed as I don't think the handling of this is well defined or
   consistently implemented in networking equipment.

Discussion: DetNet specifies the use and discusses possible
   interaction with RFC6347 in this I-D.

local-ID   An edge and relay node internal construct that uniquely
               identifies a DetNet flow. It may be used to select
               proper forwarding and/or DetNet specific service
               function.

Comment #9 SB> Is this really an internal construct, or is it an on
   the wire construct? If it is needed end to end, I don't think it
   is correct to describe it as an internal construct. When you say
   "select" presumably you mean by potentially any DN aware node on
   the path?

Discussion: It is an internal construct, so yes.

PREF   A Packet Replication and Elimination Function (PREF)
               does the replication and elimination processing of
               DetNet flow packets in edge or relay nodes. The
               replication function is essentially the existing 1+1
               protection mechanism. The elimination function reuses
               and extends the existing duplicate detection mechanism
               to operate over multiple (separate) DetNet member flows
               of a DetNet compound flow.

Comment #10 SB> I wonder if 1+1 is the right way to go, or whether
   1+n is better. A bunch of new techniques have emerged over the
   years and we really ought to look at creating paths with MRT.

   With 1+2 on main + the two MRT paths you have a two failure
   resiliency as far as it is possible to construct such paths in the
   underlying topology.

Discussion: As observed above, actually 1+n would be closer to what
   is needed. 1+1 was meant to be more an example showing there is
   existing work that can be leveraged.

2.2. Abbreviations

   The following abbreviations used in this document:

   AC   Attachment Circuit.

   CE   Customer Edge equipment.

   CoS   Class of Service.

   CW   Control Word.

   d-CW   DetNet Control Word.

   DetNet   Deterministic Networking.

   DF   DetNet Flow.

   L2VPN   Layer 2 Virtual Private Network.

   LSR   Label Switching Router.

   MPLS   Multiprotocol Label Switching.

   MPLS-TP   Multiprotocol Label Switching - Transport Profile.

   MS-PW   Multi-Segment PseudoWire (MS-PW).

   NSP   Native Service Processing.

   OAM   Operations, Administration, and Maintenance.

   PE   Provider Edge.

   PREF   Packet Replication and Elimination Function.

   PSN   Packet Switched Network.

   PW   PseudoWire.

   QoS   Quality of Service.

   TSN   Time-Sensitive Network.

3. Requirements language

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

4. DetNet data plane overview

   Comment #11 I am not sure whether this is a DP overview, or an
      architecture overview and hence whether this needs to be here or
      in the architecture draft.

   Discussion: Overview is more of an editorial matter and its final
      location can be discussed later on. Currently it is "no harm" to
      have it here but there are no binding reasons to keep the text in
      either.

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

```
  TSN              Edge            Transit         Relay          DetNet
  End System       Node             Node           Node           End System

  +---------+     +.........+                                    +---------+
  |  Appl.  |<---:Svc Proxy:-- End to End Service ---------->|  Appl.  |
  +---------+     +---------+                     +---------+   +---------+
  |   TSN   |     |TSN| |Svc|<-- DetNet flow ---: Service :-->| Service |
  +---------+     +---+ +---+     +---------+     +---------+   +---------+
  |Transport|     |Trp| |Trp|     |Transport|     |Trp| |Trp|   |Transport|
  +--------.-+     +-.-+ +-.-+     +--.----.-+     +-.-+ +-.-+   +---.-----+
        : Link   :    /  ,-----. \   : Link   :    /  ,-----. \
        +........+    +-[  Sub  ]-+   +........+    +-[  Sub  ]-+
                        [Network]                    [Network]
                         `-----'                      `-----'
```
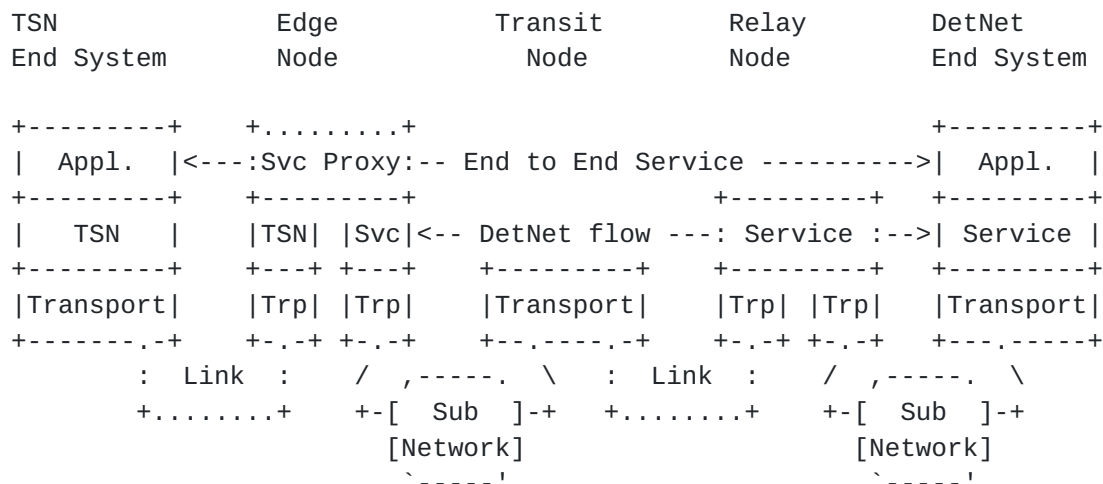
          Figure 1: A simple DetNet enabled network architecture

   Figure 2 illustrates how DetNet can provide services for IEEE
   802.1TSN end systems over a DetNet enabled network. The edge nodes

insert and remove required DetNet data plane encapsulation. The 'X'
in the edge and relay nodes represents a potential DetNet flow packet
replication and elimination point. This conceptually parallels L2VPN
services, and could leverage existing related solutions as discussed
below.

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

                  Figure 2: IEEE 802.1TSN over DetNet

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

```
           DetNet                                            DetNet
           Service              Transit          Transit     Service
     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    |
        |                                                         |
        |<--------------- End to End DetNet Service -------------->|
```

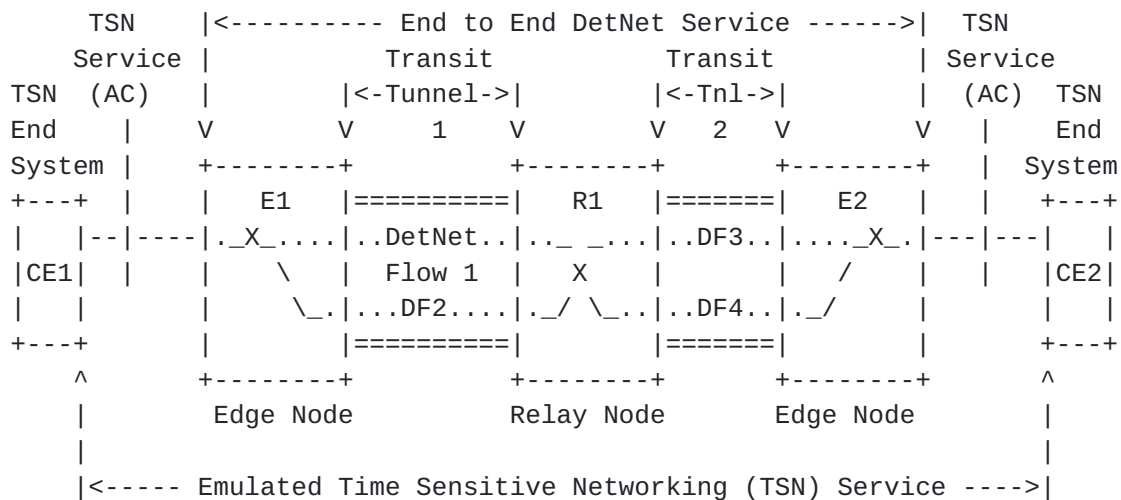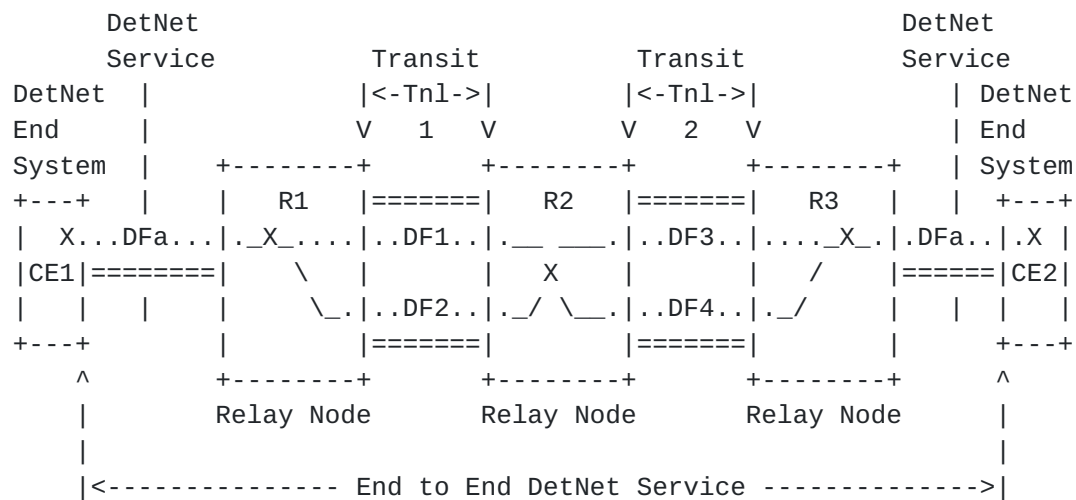   Figure 3: PW-Based Native DetNet

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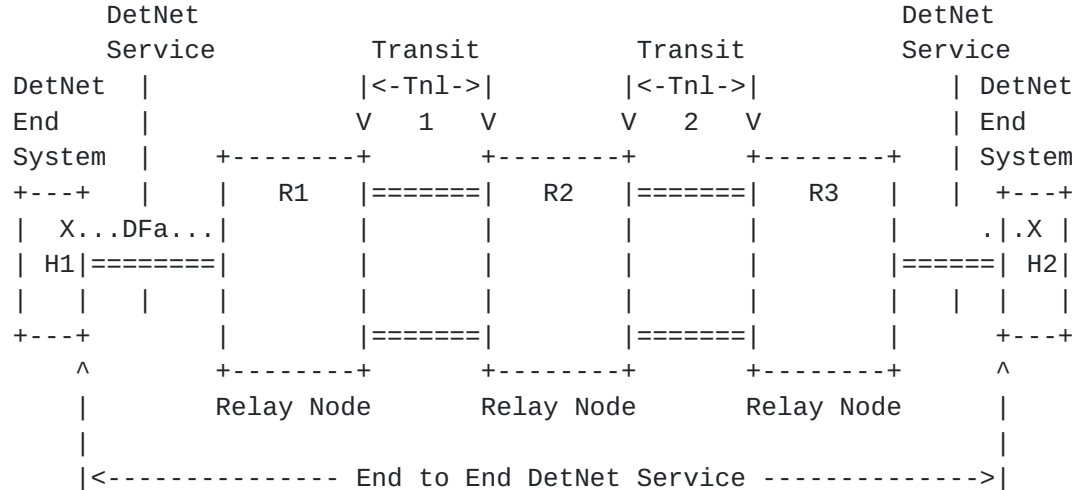
   NOTE: This figures is TBD

```
           DetNet                                            DetNet
           Service              Transit          Transit     Service
     DetNet   |          |<-Tnl->|          |<-Tnl->|          | DetNet
     End      |          V   1   V          V   2   V          | End
     System   |      +--------+      +--------+      +--------+ | System
     +---+    |      |   R1   |======|   R2   |======|   R3   | |   +---+
     | X...DFa...|      |        |        |        |        .|.X |
     | H1|=======|      |        |        |        |        |=====| H2|
     |   |   |   |      |        |        |        |        |   |   |   |
     +---+    |      |======|        |======|        |   |   +---+
        ^        +--------+      +--------+      +--------+    ^
        |          Relay Node      Relay Node      Relay Node    |
        |                                                         |
        |<--------------- End to End DetNet Service -------------->|
```

                    Figure 4: IP-Based Native DetNet

4.1. DetNet data plane encapsulation requirements

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

   1. DetNet function related scenarios:

      * Congestion protection and latency control: usage of allocated
         resources (queuing, policing, shaping).

      * Explicit routes: select/apply the flow specific path.

      * Service protection: recognize DetNet compound and member flows
         for replication an elimination.

      Comment #12 I am not sure whether the correct architectural
         construct is flow or flow group. Flow suggests that sharing/
         aggregation is not allowed but whether this is allowed or not
         is an application specific issue.

      Discussion: Agree that a flow group would be a better
         characterization.

      Comment #13 I think that there needs to be some clarification as
         to whether FG is is understood by the DN system exclusively or
         whether there is an expectation that it is understood by the
         underlay.

      Discussion: Agree that more detail is needed here. DetNet aware
         nodes need to understand flow groups. Underlay needs to be
         aware of flow groups at the resource allocation level.

   2. OAM function related scenarios:

      * troubleshooting (e.g., identify misbehaving flows, etc.)

      * recognize flow(s) for analytics (e.g., increase counters,
         etc.)

      * correlate events with flows (e.g., volume above threshold,
         etc.)

      * etc.

   Each node (edge, relay and transit) use a local-ID of the DetNet-
   (compound)-flow in order to accomplish its role during transport.
   Recognizing the DetNet flow is more relaxed for edge and relay nodes,
   as they are fully aware of both the DetNet service and transport
   layers. The primary DetNet role of intermediate transport nodes is
   limited to ensuring congestion protection and latency control for the
   above listed DetNet functions.

   The DetNet data plane allows for the aggregation of DetNet flows,
   e.g., via MPLS hierarchical LSPs, to improved scaling. When DetNet
   flows are aggregated, transit nodes may have limited ability to

   provide service on per-flow DetNet identifiers. Therefore,
   identifying each individual DetNet flow on a transit node may not be
   achieved in some network scenarios, but DetNet service can still be
   assured in these scenarios through resource allocation and control.

   Comment #14 You could introduce the concept of a flow group
      identified into the packet. You may also include a flow id at a
      lower layer.

   Discussion: Agree on the identification properties. Adding a
      specific id into actual on-wire formats is not necessarily needed.

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

   Comment #15 I am confused as to what you mean by local ID. Is this
      an internal construct which packet parameters map to, in which
      case why is it being called up? IETF practise is to specify on
      the wire behaviour but not internal behaviour of equipments.

   Discussion: Local-id is an internal construct, which was intended to
      clarify the discussion in the I-D. Obviously it did not work as
      intended.

5. DetNet data plane solution

5.1. DetNet specific packet fields

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

   Comment #16 should, SHOULD, must or MUST?

   Discussion: SHOULD or MUST is ok. MUST is probably more
      appropriate.

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

5.2. DetNet encapsulation

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

5.2.1. **PseudoWire-based data plane encapsulation**

   Figure 5 illustrates a DetNet PW encapsulation over an MPLS PSN. The
   PW-based encapsulation of the DetNet flows fits perfectly for the
   Layer-2 interconnect deployment cases (see Figure 2). Furthermore,
   end to end DetNet service i.e., native DetNet deployment (see
   Figure 3) is also possible if DetNet-aware end systems are capable of
   initiating and termination MPLS encapsulated PWs. It is also
   possible use the same encapsulation format with a Packet PW over MPLS
   [RFC6658]. Transport of IP encapsulated DetNet flows, see
   Section 5.2.2, over DetNet PWs is also possible. Interworking
   between PW- and IPv6-based encapsulations is discussed further in
   Section 7.6.

   The PW-based DetNet data plane encapsulation consists of:

   o DetNet control word (d-CW) containing sequencing information for
      packet replication and duplicate elimination purposes. There is a
      separate sequence number space for each DetNet flow.

   o PseudoWire Label (PW Label) that is a standard PW label
      identifying a DetNet flow and a PW Instance within a (DA-)T-PE or
      (DA-)S-PE device.

   o An optional S-Label that represents DetNet Service LSP used
      between (DA-)T-PE or (DA-)S-PE nodes. One possible use of an
      S-Label is to identify the different DetNet member flows used to
      provide protection to a DetNet composite flow, perhaps even when
      both LSPs appear on the same link for some reason.

   Comment #17 This needs some discussion by the WG.

   Discussion: Agree, specifically if the I-D becomes WG document.

   o MPLS transport LSP label(s) (T-label) which may be a hop-by-hop
      label used between LSRs.

   Comment #18 Ordinarily this will of course be PHPed before arrival
      at an x-PE.

   Discussion: In most cases yes - depends on the network
      configuration. PHP is not mandatory and TP does not even have
      PHP.

       RFC3985 Encapsulation   DetNet PW Encapsulation

    +--------------------+
    | Payload | +---------------------------------+
    /====================\ | |
    H Payload Convergence H--. | DetNet Flow |
    H--------------------H | | Payload Packet |
    H   Timing   H +-\ | |
    H--------------------H | \   /===============================\
    H   Sequencing   H--'   \-->H   DetNet Control Word   H
    \====================/   \===============================/
    | PW Demultiplexer |--------->| PW Label |
    +--------------------+   +---------------------------------+
    | PSN Convergence | .--->| Optional MPLS S-Label |
    +--------------------+ | +---------------------------------+
    | PSN |-----+--->| MPLS T-Label(s) |
    +--------------------+   +---------------------------------+
    | Data-Link |
    +--------------------+
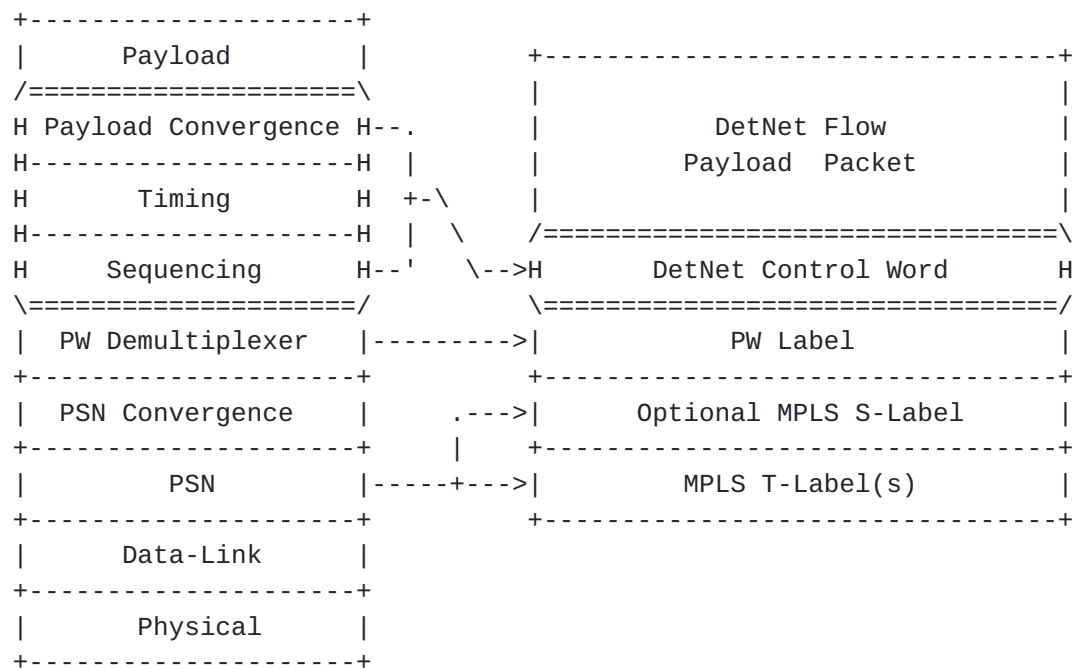 | Physical |
    +--------------------+

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

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

      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| reserved - set to 0 | 16 bit Sequence Number |
      +-+
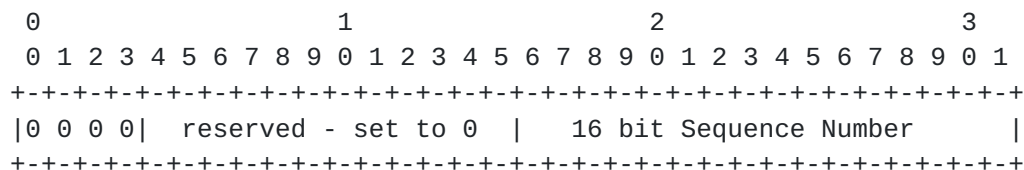

                     Figure 6: DetNet Control Word

   Comment #19 We need to think about whether "identical is the correct
      term. The Ethernet S/N skips zero (it uses zero to mean no S/N in
      use). is that the behaviour that we want?

   Discussion: Good point. Identical is a wrong statement. The format
      is the same the behaviour of SN is slightly different as 0 is
      assumed to be a valid SN.

5.2.2. Native IPv6-based data plane encapsulation

   Comment #20 SB> This part of the design needs to be marked as
      provisional until it has a more thorough WG review.

   Discussion: Ok, however, this is still an individual I-D.

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

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

   o   IPv6 header as the transport protocol.

   o   IPv6 header Flow Label that is used to help to identify a DetNet
       flow (i.e., roughly an equivalent to the PW Label). A Flow Label
       together with the IPv6 source address uniquely identifies a DetNet
       flow.

   Comment #21 SB> Have we validated that it is unconditionally safe to
      make this assumption about the use of the FL?

   Discussion: RFC6437 does not restrict such use and DetNet DP
      solution can always define their own use of flow label. It should
      be noted that a DetNet aware node will always contain new code and
      is not a load balancer.

   o   DetNet Control Word IPv6 Destination Option containing sequencing
       information for packet replication and duplicate elimination
       function (PREF) purposes. The DetNet Destination Option is
       equivalent to the DetNet Control Word.

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

Comment #22 SB> We will probably need to agree an option ordering,
    and need to note that the 6man IPv6 solution already operates on
    the edge of the ability of h/w to see that far into the packet.

Discussion: RFC8200 describes extension header ordering - there is
    not much to agree there. Agree on the hardware lookup challenges.
    However, the issues of SR header are not this I-D to fix.

Comment #23 SB> I am not sure the above is needed since it is by
    definition correct.

Discussion: (next header) agree.

A DetNet-aware end station (a host) or an intermediate node receiving
an IPv6 packet destined to it and containing a DetNet Destination
Option does the appropriate processing of the packet. This may
involve packet duplication and elimination (PREF processing),
terminating a tunnel or delivering the packet to the upper layers/
Applications.

```
          +--------------------------------+
          |                                |
          |          DetNet Flow           |
          |            Payload             |
          |                                |
          /--------------------------------\
          H DetNet Control Word DstOpt Hdr  H
          \--------------------------------/
          |            IPv6 header          |
          |       (with set Flow label)     |
          +--------------------------------+
```
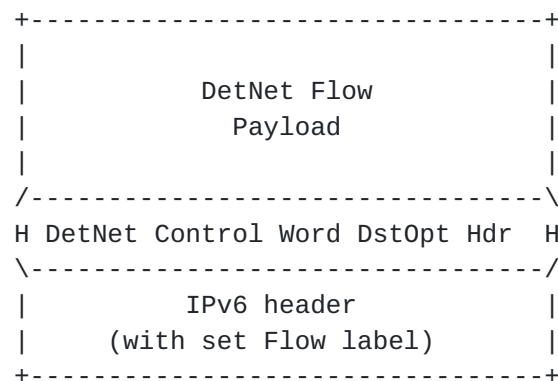
         Figure 7: Encapsulation of a native IPv6 DetNet flow

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

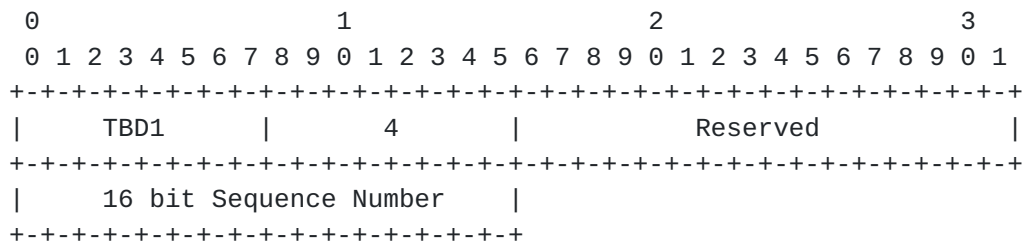Comment #24 SB> Can an SR node look at a DO?

Discussion: Yes.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     TBD1      |      4        |              Reserved         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     16 bit Sequence Number   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: DetNet Destination Option

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

5.3. DetNet flow identification for duplicate detection

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

Comment #25 SB> So I wonder if the right place to put the FI is in
   the IPv6 FI, or in the IPv6 address itself?

Discussion: Each flow having different address is challenging if we
   want to terminate multiple flows into the same node with one
   address or originate multiple flows from a node with one address
   (note, we are aware of the one /64 per node discussion but cannot
   assume it here, at least not yet).

5.3.1. PseudoWire encapsulation

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

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

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

Comment #26 SB> I am not sure of the utility of this reference. In
   FRR you should not receive packets concurrently on two paths. It
   seems fine to state the the requirement that a single label is
   used for both paths.

Discussion: OK with the same label comment. OK to remove the FRR
   reference here.

5.3.2. Native IPv6 encapsulation

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

Comment #27 SB> See earlier. If there are enough IPv6 addresses to
   address video fragments, why not DN flows? Then this problem goes
   away.

Discussion: See the earlier comment #25 discussion. If nodes get
   their addressies via DHCPv6 basically ruins this mechanism. Also
   the assumption for this to work is that the node has a full /64 to
   use, which is not always the case. Otherwise the idea is just
   fine.

6. PREF specific considerations

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

6.1. PseudoWire-based data plane

6.1.1. Forwarder clarifications

The DetNet specific new functionality in an edge or relay node
processing is the packet replication and duplication elimination
function (PREF). This function is a part of the DetNet-aware
"extended" forwarder. The PREF processing is triggered by the
received packet of a DetNet flow.

Comment #28 SB> I am not sure what you mean by triggered here.
   Hopefully we are not thinking of dataplane triggered
   configuration?

Discussion: "Initiated" is probably more appropriate wording.

Basically the forwarding entry has to be extended with a "PREF
enabled" boolean configuration switch that is associated with the
normal forwarding actions (e.g., in case of MPLS a swap, push, pop,
..). The output of the PREF elimination function is always a single

packet. The output of the PREF replication function is always one or
more packets (i.e., 1:M replication). The replicated packets MUST
share the same DetNet control word sequence number.

The complex part of the DetNet PREF processing is tracking the
history of received packets for multiple DetNet member flows. These
ingress DetNet member flows (to a node) MUST have the same local-ID
if they belong to the same DetNet-(compound)-flow and share the same
sequence number counter and the history information.

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

Comment #29: SB> There needs to be some text about preventing a node
   ever receiving its own replicated packets. Indeed that would
   suggest that the flow id should be changed and replication should
   only take place on configured flow IDs. I have a feeling that
   this would all be a lot safer if replication only happened at
   ingress and we managed the diversity of the paths.

Discussion: Agree on hardening the loopback text considerations.

6.1.2. Edge node processing clarifications

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

Comment #30: SB> Shouldn't the behaviour simply be a property of a
   given PW?
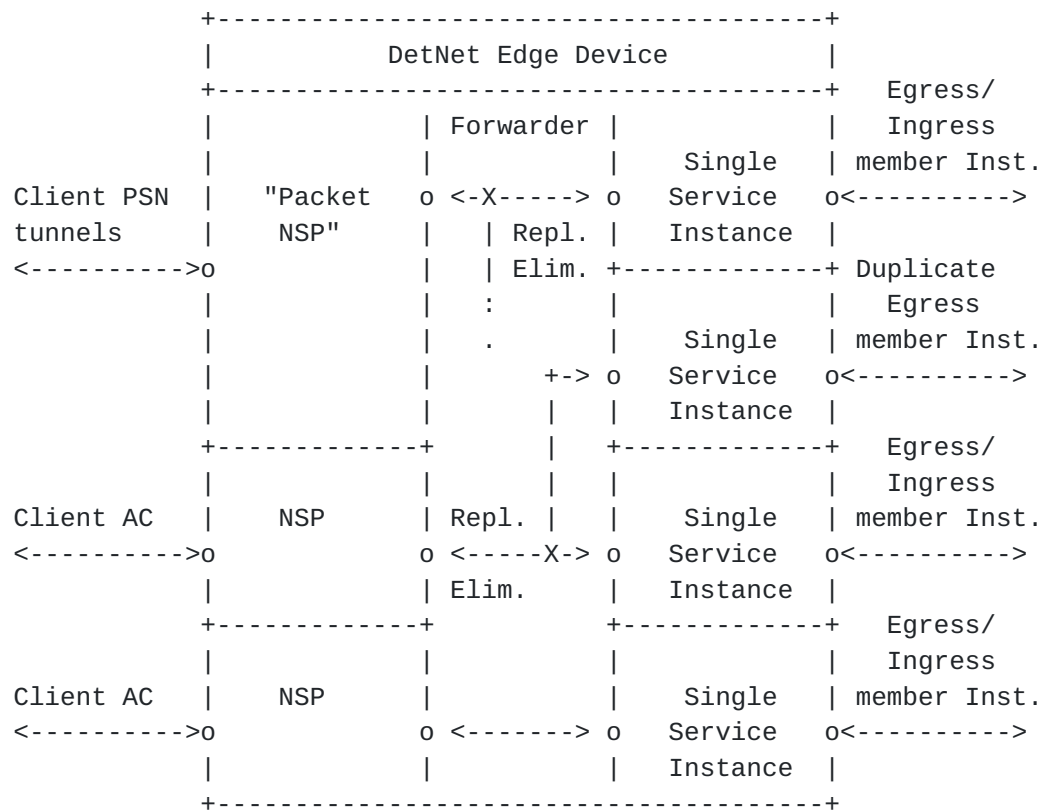
Discussion: Agree in principle.

```
                +----------------------------------------+
                |             DetNet Edge Device         |
                +----------------------------------------+  Egress/
                |                 | Forwarder |         |    Ingress
                |                 |           |  Single | member Inst.
  Client PSN    |    "Packet   o <-X-----> o   Service  o<---------->
  tunnels       |     NSP"     |  | Repl. |   Instance  |
  <---------->o               |  | Elim. +-------------+ Duplicate
                |              |  | :         |         |  Egress
                |              |  | .         |  Single | member Inst.
                |              |     +-> o   Service  o<---------->
                |              |      |  | Instance  |
                +-------------+       |  +-------------+  Egress/
                |             |       |  |         |    Ingress
  Client AC     |    NSP      | Repl. |  |  Single | member Inst.
  <---------->o              o <-----X-> o   Service  o<---------->
                |             | Elim.    | Instance  |
                +-------------+          +-------------+  Egress/
                |             |          |         |    Ingress
  Client AC     |    NSP      |          |  Single | member Inst.
  <---------->o              o <-------> o   Service  o<---------->
                |             |          | Instance  |
                +----------------------------------------+
```

                Figure 9: DetNet Edge Node processing

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

   Comment #31 SB> This would be a fine way to operate the PW system -
      edge to edge.

   Discussion: When it comes to use of NSPs, agree. Also for "island
      interconnect" this is a fine. However, when there is a need to do
      PREF in a middle, plain edge to edge is not enough.

   The DetNet-aware extended forwarder selects the egress DetNet member
   flow based on the DetNet forwarding rules. In both "normal AC" and
   "Packet AC" cases there may be no DetNet encapsulation header
   available yet as it is the case with relay nodes (see Section 6.1.3).

It is the responsibility of the extended forwarder within the edge node to push the DetNet specific encapsulation (if not already present) to the packet before forwarding it to the appropriate egress DetNet member flow instance(s).

Comment #32 SB> I am not convinced of the wisdom of having a mid-point node convert a flow into a DN flow, which is what you are implying here. This seems like an ingress function.

Discussion: OK. The text here has issues and seems to mix relay and edge.

The extended forwarder MAY copy the sequencing information from the native DetNet packet into the DetNet sequence number field and vice versa. If there is no existing sequencing information available in the native packet or the forwarder chose not to copy it from the native packet, then the extended forwarder MUST maintain a sequence number counter for each DetNet flow (indexed by the DetNet flow identification).

6.1.3. Relay node processing clarifications

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

Comment #33 SB> I don't think that a relay node in not a normal construct so I am not sure "overload" is the right term here.

Discussion: Agree. There is a terminology issue here.

A DetNet Relay node participates to the packet replication and duplication elimination. This processing is done within an extended forwarder function. Whether an ingress DetNet member flow receives DetNet specific processing depends on how the forwarding is programmed. For some DetNet member flows the relay node can act as a normal relay node and for some apply the DetNet specific processing (i.e., PREF).

Comment #34 SB> Again relay node is not a normal term, so am not sure what it does in the absence of a PREF function.

Discussion: Relay node was a DetNet aware S-PE originally, which is not explicitly stated here anymore, thus slightly confusing text here. The text here needs to clarify the roles of PREF and switching functions. A DetNet relay is described in the

architecture document. However, there is definitely room for
termonilogy and text improvements.

It is also possible to treat the relay node as a transit node, see
Section 7.3. Again, this is entirely up to how the forwarding has
been programmed.

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

```
                +---------------------------------------+
                |            DetNet Relay Device        |
    Ingress     +---------------------------------------+
    member      |              | Forwarder |            |   Egress
    instance    |   Single     |           |   Single   | member Inst.
 ----------->o  Service     o --X-----> o  Service     o----------->
                |   Instance   |  | Elim. |   Instance  |
    Ingress     +-------------+   |       +-------------+ Duplicate
    member      |              |  |       |             |   Egress
    instance    |   Single     |  |       |   Single   | member Inst.
 ----------->o  Service     o --+   +-> o  Service     o----------->
                |   Instance   |      |  |   Instance  |
    Ingress     +-------------+      |   +-------------+
    member      |              |  |  |   |             |   Egress
    instance    |   Single     | Repl. |  |   Single   | member Inst.
 ----------->o  Service     o ------X-> o  Service     o----------->
                |   Instance   |         |   Instance  |
    Ingress     +-------------+          +-------------+
    member      |              |         |             |   Egress
    instance    |   Single     |         |   Single   | member Inst.
 ----------->o  Service     o --------> o  Service     o----------->
                |   Instance   |         |   Instance  |
                +---------------------------------------+
```

Figure 10: DetNet Relay Node processing

Comment #35 SB> Somewhere in the dp document there needs to be a
note of the requirement for interfaces to do fast exchange of
counter state, and a note to those planning the network and

designing the control plane that they need to provide support for
this.

Discussion: We kinf of agree but also think the above exchange or
synchronization of counter states is not in our scope to solve.

6.2. Native IPv6-based data plane

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

7. Other DetNet data plane considerations

7.1. Class of Service

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

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

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

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

7.2. **Quality of Service**

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

   In addition to path pinning and packet replication and elimination,
   described in Section 5 above, DetNet provides zero congestion loss
   and bounded latency and jitter.

   Comment #36 SB> I just searched from the beginning of the document
      and this was the the first reference I found to pinning.

   Discussion: Terminology isssue. Should use, for example, explicit
      paths which is used in the architecture I-D.

   As described in [I-D.ietf-detnet-architecture], there are different
   mechanisms that maybe used separately or in combination to deliver a
   zero congestion loss service. These mechanisms are provided by the
   either the MPLS or IP layers, and may be combined with the mechanisms
   defined by the underlying network layer such as 802.1TSN.

   A baseline set of QoS capabilities for DetNet flows carried in PWs
   and MPLS can provided by MPLS with Traffic Engineering (MPLS-TE)
   [RFC3209] and [RFC3473]. TE LSPs can also support explicit routes
   (path pinning). Current service definitions for packet TE LSPs can
   be found in "Specification of the Controlled Load Quality of
   Service", [RFC2211], "Specification of Guaranteed Quality of
   Service", [RFC2212], and "Ethernet Traffic Parameters", [RFC6003].
   Additional service definitions are expected in future documents to
   support the full range of DetNet services. In all cases, the
   existing label-based marking mechanisms defined for TE-LSPs and even
   E-LSPs are use to support the identification of flows requiring
   DetNet QoS.

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

combination of the Flow Label together with the IP source address
uniquely identifies a DetNet flow.

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

Comment #37 SB> Why not MUST?

Discussion: OK with MUST.

o Defend the DetNet QoS by discarding or remarking (to a non-DetNet
   CoS) packets received that are not the subject of a completed
   reservation.

o Not use a DetNet reserved resource, e.g. a queue or shaper
   reserved for DetNet flows, for any packet that does not carry a
   DetNet Class of Service marker.

7.3. Cross-DetNet flow resource aggregation

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

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

DetNet flows transported via IP have more limited aggregation
options, due to the available traffic flow identification fields of

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

Comment #38 SB> I am sure we can do better than this with SR, or the
   use of routing techniques that map certain addresses to certain
   paths.

Discussion: --

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

7.4. Bidirectional traffic

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

While the IPv6 and MPLS data planes must support bidirectional DetNet
flows, there are no special bidirectional features with respect to
the data plane other than need for the two directions take the same
paths. Fate sharing and associated vs co-routed bidirectional flows
can be managed at the control level. Note, that there is no stated
requirement for bidirectional DetNet flows to be supported using the
same IPv6 Flow Labels or MPLS Labels in each direction. Control

   mechanisms will need to support such bidirectional flows for both
   IPv6 and MPLS, but such mechanisms are out of scope of this document.
   An example control plane solution for MPLS can be found in [RFC7551].

7.5. Layer 2 addressing and QoS Considerations

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

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

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

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

7.6. Interworking between PW- and IPv6-based encapsulations

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

8. Time synchronization

   Comment #39 SB> This section should point the reader to RFC8169
      (residence time in MPLS n/w. We need to consider if we need to
      introduce the same concept in IP.

   Discussion: agree.

[Editor's note: describe a bit of issues and deployment
considerations related to time-synchronization within DetNet. Refer
to DT discussion and the slides that summarize different approaches
and rough synchronization performance numbers. Finally, scope time-
synchronization solution outside data plane.]

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

o PTP packets can be sent either as DetNet flows or as high-priority
   best effort packets. Using DetNet for PTP packets requires
   careful consideration to prevent unwanted interactions between
   clock-synchronized network nodes and the packets that synchronize
   the clocks.

o PTP packets are sent as a normal DetNet flow through network nodes
   that are not time-synchronized: in this approach PTP traffic is
   forwarded as a DetNet flow, and as such it is forwarded in a way
   that allows a low delay variation. However, since intermediate
   nodes do not take part in the synchronization protocol, this
   approach provides a relatively low degree of accuracy.

o PTP with on-path support: in this approach PTP packets are sent as
   ordinary or as DetNet flows, and intermediate nodes take part in
   the protocol as Transparent Clocks or Boundary Clocks [IEEE1588].
   The on-path PTP support by intermediate nodes provides a higher
   degree of accuracy than the previous approach. The actual
   accuracy depends on whether all intermediate nodes are PTP-
   capable, or only a subset of them.

o Time-as-a-service: in this approach accurate time is provided as-
   a-service to the DetNet source and destination, as well as the
   intermediate nodes. Since traffic between the source and
   destination is sent over a provider network, if the provider
   supports time-as-a-service, then accurate time can be provided to
   both the source and the destination of DetNet traffic. This
   approach can potentially provide the highest degree of accuracy.

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

It should be noted that in all four approaches it is not recommended to use replication and elimination for synchronization packets; the replication/elimination approach may in some cases reduce the synchronization accuracy, since the observed path delay will be bivalent.

Comment #40 SB> I am not sure why we sould not use PREP. We should explain to the reader.

Discussion: Agree that a this can be opened a bit more in detail. The issue is explained briefly in the last sentence but it could be more clear.

9. Management and control considerations

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

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

9.1. PW Label and IPv6 Flow Label assignment and distribution

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

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

[9.2](#). **Packet replication and elimination**

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

[9.3](#). **Explicit paths**

[TBD: based on MPLS TE and SR.]

[9.4](#). **Congestion protection and latency control**

[TBD]

[9.5](#). **Flow aggregation control**

[TBD]

[10](#). **Security considerations**

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

[11](#). **IANA considerations**

TBD.

[12](#). **Acknowledgements**

The author(s) ACK and NACK.

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

Jouni Korhonen

Janos Farkas

Norman Finn

Balazs Varga

Loa Andersson

Tal Mizrahi

David Mozes

Yuanlong Jiang

Carlos J. Bernardos

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

Lou Berger

Pat Thaler

13. References

13.1. Normative references

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
            Requirement Levels", BCP 14, RFC 2119,
            DOI 10.17487/RFC2119, March 1997,
            <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>.

[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>.

[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>.

[RFC3168]   Ramakrishnan, K., Floyd, S., and D. Black, "The Addition
            of Explicit Congestion Notification (ECN) to IP",
            RFC 3168, DOI 10.17487/RFC3168, September 2001,
            <https://www.rfc-editor.org/info/rfc3168>.

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

   [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>.

   [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>.

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

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

   [RFC6658] Bryant, S., Ed., Martini, L., Swallow, G., and A. Malis,
              "Packet Pseudowire Encapsulation over an MPLS PSN",
              RFC 6658, DOI 10.17487/RFC6658, July 2012,
              <https://www.rfc-editor.org/info/rfc6658>.

13.2. Informative references

   [I-D.ietf-6man-segment-routing-header]
              Previdi, S., Filsfils, C., Raza, K., Leddy, J., Field, B.,
              daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d.,
              Matsushima, S., Leung, I., Linkova, J., Aries, E., Kosugi,
              T., Vyncke, E., Lebrun, D., Steinberg, D., and R. Raszuk,
              "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-
              segment-routing-header-07 (work in progress), July 2017.

   [I-D.ietf-detnet-architecture]
              Finn, N., Thubert, P., Varga, B., and J. Farkas,
              "Deterministic Networking Architecture", draft-ietf-
              detnet-architecture-03 (work in progress), August 2017.

   [I-D.ietf-detnet-dp-alt]
              Korhonen, J., Farkas, J., Mirsky, G., Thubert, P.,
              Zhuangyan, Z., and L. Berger, "DetNet Data Plane Protocol
              and Solution Alternatives", draft-ietf-detnet-dp-alt-00
              (work in progress), October 2016.

   [I-D.sdt-detnet-security]
              Mizrahi, T., Grossman, E., Hacker, A., Das, S.,
              "Deterministic Networking (DetNet) Security
              Considerations, draft-sdt-detnet-security, work in
              progress", 2017.

   [IEEE1588]
              IEEE, "IEEE 1588 Standard for a Precision Clock
              Synchronization Protocol for Networked Measurement and
              Control Systems Version 2", 2008.

   [IEEE8021CB]
              Finn, N., "Draft Standard for Local and metropolitan area
              networks - Seamless Redundancy", IEEE P802.1CB
              /D2.1 P802.1CB, December 2015,
              <http://www.ieee802.org/1/files/private/cb-drafts/
              d2/802-1CB-d2-1.pdf>.

[IEEE8021Q]
          IEEE 802.1, "Standard for Local and metropolitan area
          networks--Bridges and Bridged Networks (IEEE Std 802.1Q-
          2014)", 2014, <http://standards.ieee.org/about/get/>.

[RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S.
          Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1
          Functional Specification", RFC 2205, DOI 10.17487/RFC2205,
          September 1997, <https://www.rfc-editor.org/info/rfc2205>.

[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>.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast
          Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090,
          DOI 10.17487/RFC4090, May 2005,
          <https://www.rfc-editor.org/info/rfc4090>.

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

[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>.

[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>.

Appendix A. Example of DetNet data plane operation

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

[Appendix B](#). Example of pinned paths using IPv6

TBD.

Authors' Addresses

   Jouni Korhonen (editor)
   Nordic Semiconductor

   Email: jouni.nospam@gmail.com

   Loa Andersson
   Huawei

   Email: loa@pi.nu

   Yuanlong Jiang
   Huawei

   Email: jiangyuanlong@huawei.com

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

   Email: norman.finn@mail01.huawei.com

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

   Email: balazs.a.varga@ericsson.com

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

   Email: janos.farkas@ericsson.com

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

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

   Tal Mizrahi
   Marvell
   6 Hamada st.
   Yokneam
   Israel

   Email: talmi@marvell.com

   Lou Berger
   LabN Consulting, L.L.C.

   Email: lberger@labn.net