

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

J. Korhonen, Ed.

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

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

Abstract

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

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 1, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

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

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

1. Introduction

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

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

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

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

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

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

2. Terminology

2.1. Terms used in this document

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

The following terms are also used in this document:

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

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

2.2. Abbreviations

The following abbreviations used in this document:

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

3. Requirements language

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

4. DetNet data plane overview

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

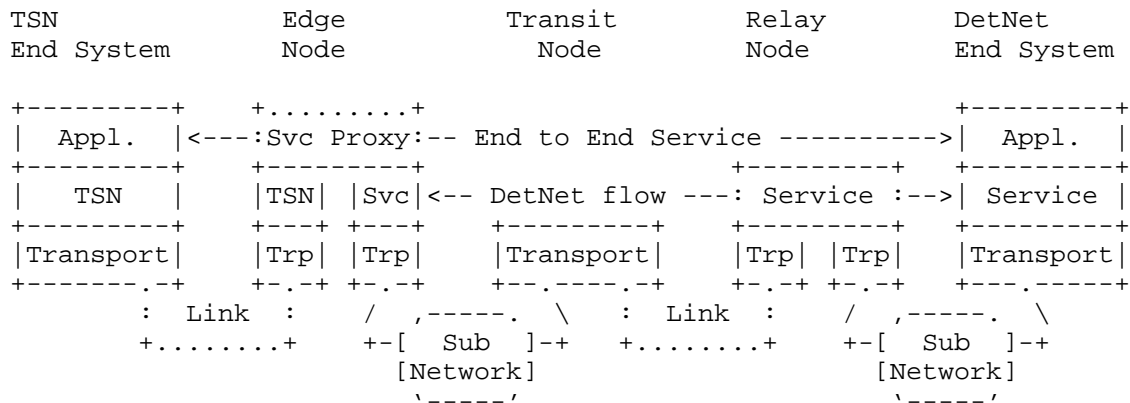


Figure 1: A simple DetNet enabled network architecture

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

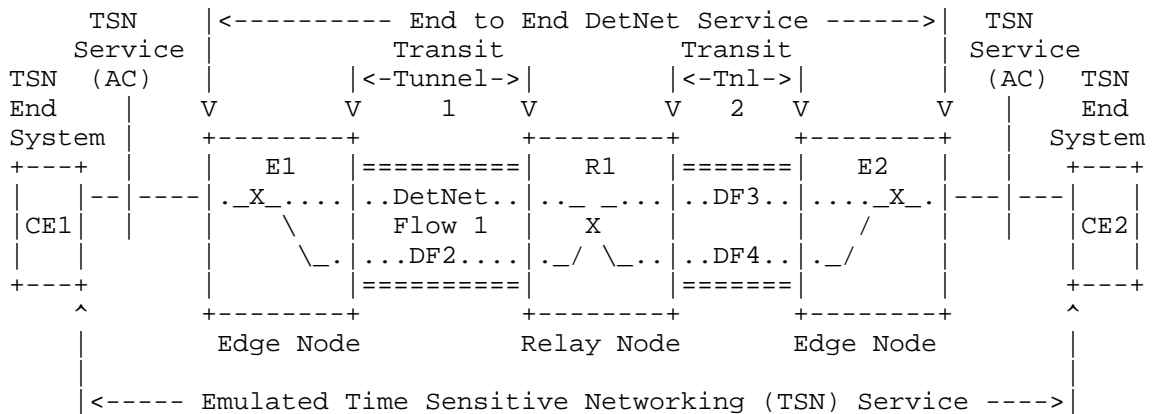


Figure 2: IEEE 802.1TSN over DetNet

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

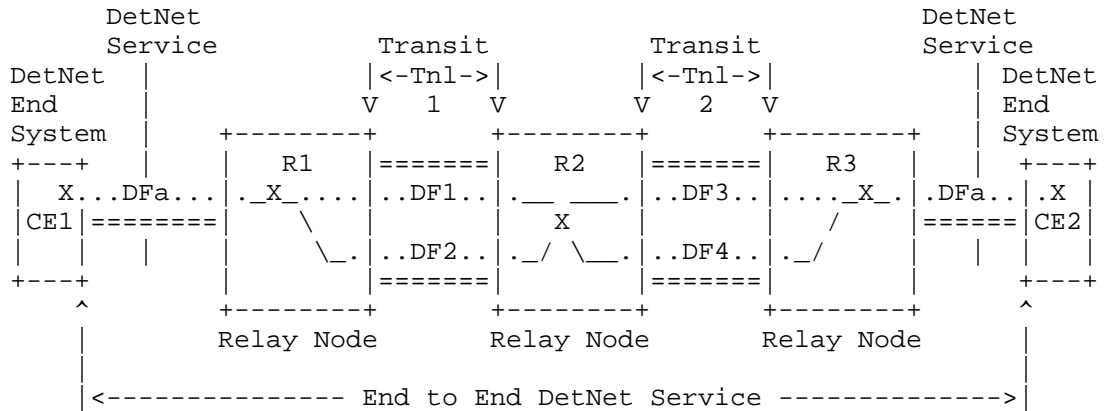


Figure 3: PW-Based Native DetNet

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

NOTE: This figures is TBD

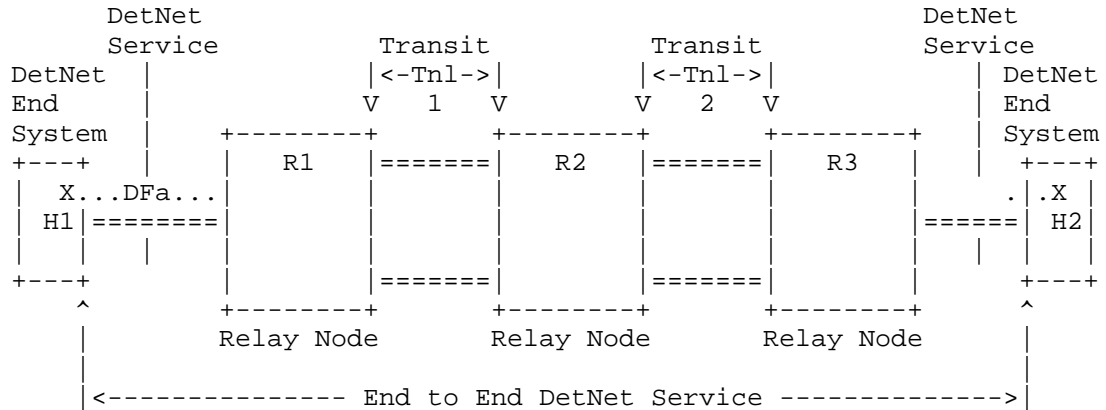


Figure 4: IP-Based Native DetNet

4.1. DetNet data plane encapsulation requirements

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

1. DetNet function related scenarios:

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

2. OAM function related scenarios:

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

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

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

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

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

5. DetNet data plane solution

5.1. DetNet specific packet fields

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

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

5.2. DetNet encapsulation

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

5.2.1. PseudoWire-based data plane encapsulation

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

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

The PW-based DetNet data plane encapsulation consists of:

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

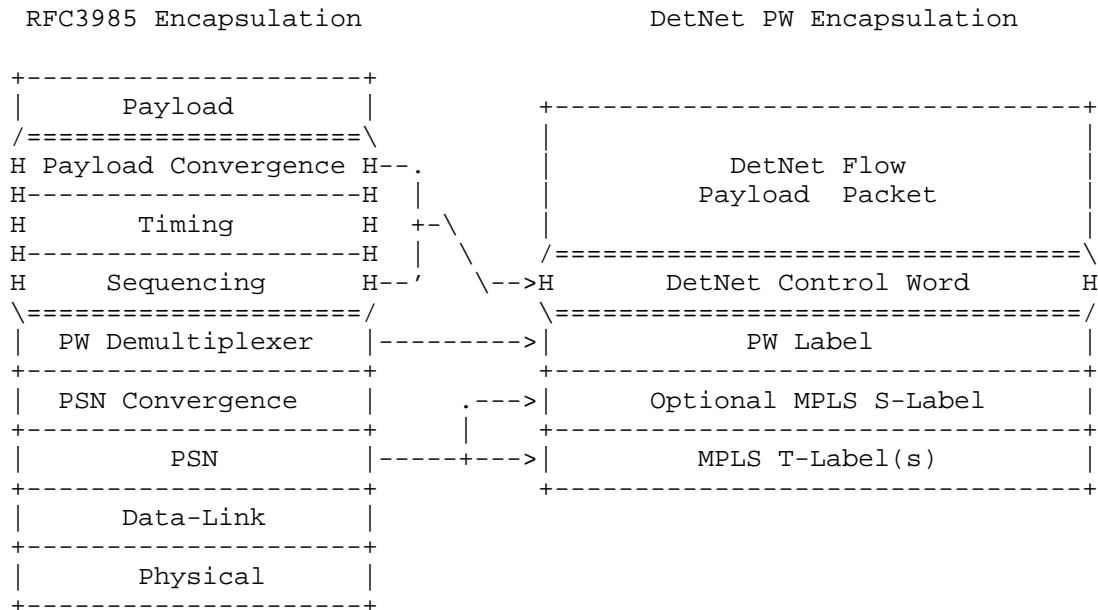


Figure 5: Encapsulation of a DetNet flow in a PW with MPLS(-TP) PSN

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

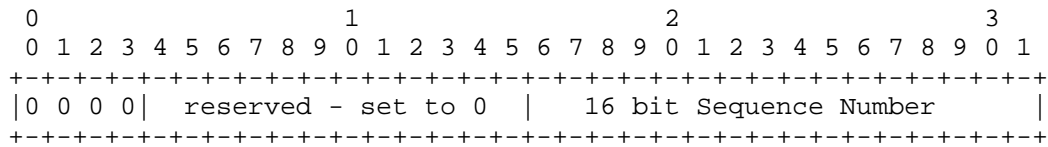


Figure 6: DetNet Control Word

5.2.2. Native IPv6-based data plane encapsulation

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

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

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

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

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

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

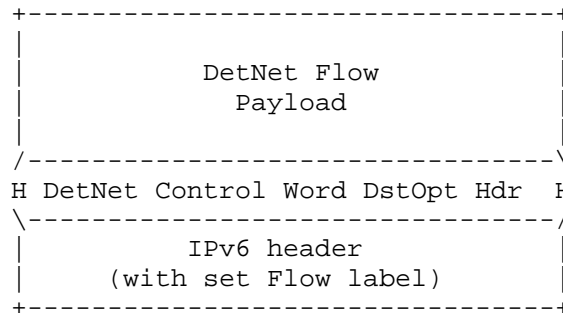


Figure 7: Encapsulation of a native IPv6 DetNet flow

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

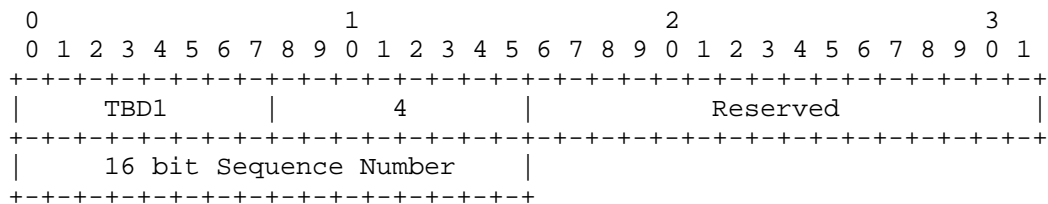


Figure 8: DetNet Destination Option

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

5.3. DetNet flow identification for duplicate detection

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

5.3.1. PseudoWire encapsulation

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

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

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

5.3.2. Native IPv6 encapsulation

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

6. PREF specific considerations

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

6.1. PseudoWire-based data plane

6.1.1. Forwarder clarifications

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

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

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

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

6.1.2. Edge node processing clarifications

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

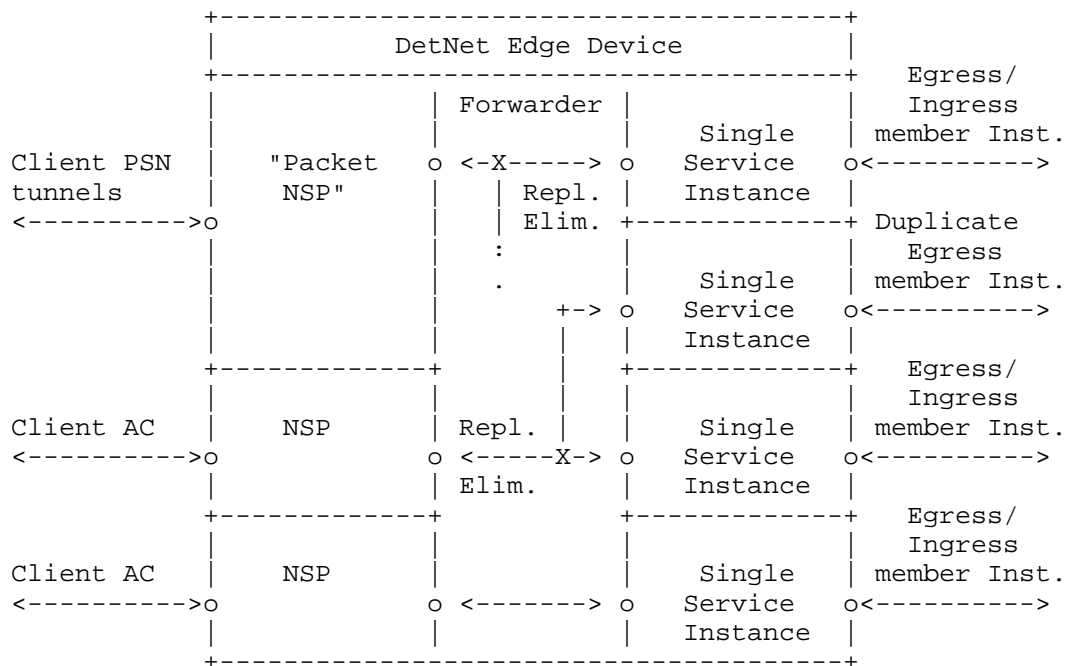


Figure 9: DetNet Edge Node processing

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

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

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

6.1.3. Relay node processing clarifications

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

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

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

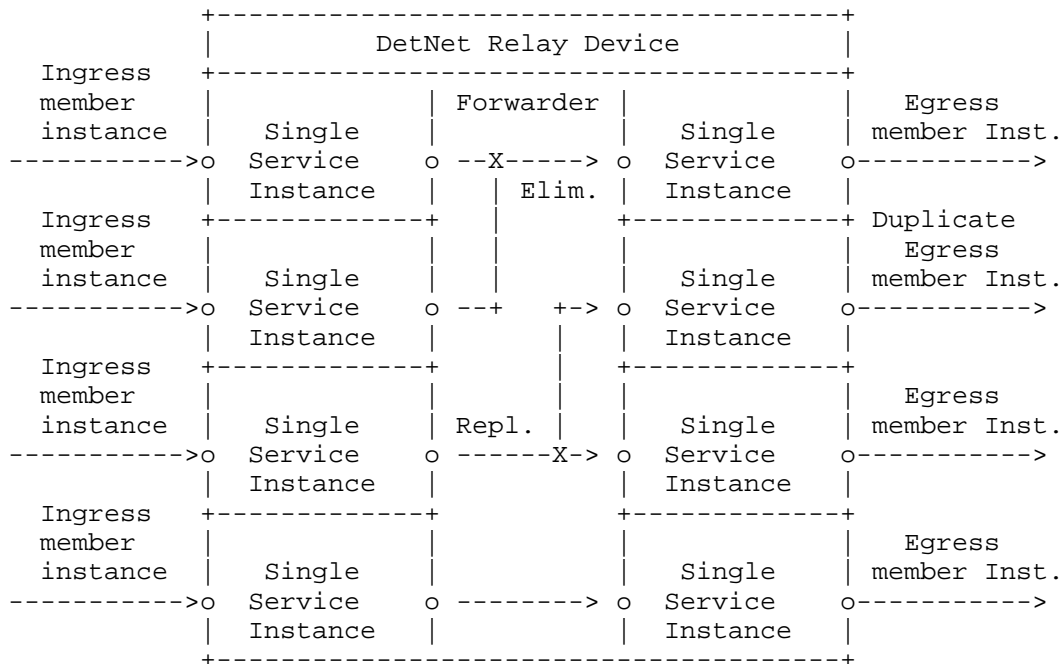


Figure 10: DetNet Relay Node processing

6.2. Native IPv6-based data plane

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

7. Other DetNet data plane considerations

7.1. Class of Service

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

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

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

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

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

7.2. Quality of Service

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

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

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

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

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

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

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

7.3. Cross-DetNet flow resource aggregation

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

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

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

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

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

7.4. Bidirectional traffic

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

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

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

7.5. Layer 2 addressing and QoS Considerations

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

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

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

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

7.6. Interworking between PW- and IPv6-based encapsulations

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

8. Time synchronization

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

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

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

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

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

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

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

9. Management and control considerations

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

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

9.1. PW Label and IPv6 Flow Label assignment and distribution

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

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

9.2. Packet replication and elimination

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

9.3. Explicit paths

[TBD: based on MPLS TE and SR.]

9.4. Congestion protection and latency control

[TBD]

9.5. Flow aggregation control

[TBD]

10. Security considerations

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

11. IANA considerations

TBD.

12. Acknowledgements

The author(s) ACK and NACK.

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

Jouni Korhonen

Janos Farkas

Norman Finn

Balazs Varga

Loa Andersson

Tal Mizrahi

David Mozes

Yuanlong Jiang

Carlos J. Bernardos

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

Lou Berger

Pat Thaler

13. References

13.1. Normative references

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC2211] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, DOI 10.17487/RFC2211, September 1997, <<http://www.rfc-editor.org/info/rfc2211>>.
- [RFC2212] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, DOI 10.17487/RFC2212, September 1997, <<http://www.rfc-editor.org/info/rfc2212>>.
- [RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <<http://www.rfc-editor.org/info/rfc2474>>.
- [RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <<http://www.rfc-editor.org/info/rfc3032>>.
- [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <<http://www.rfc-editor.org/info/rfc3168>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<http://www.rfc-editor.org/info/rfc3209>>.
- [RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, DOI 10.17487/RFC3270, May 2002, <<http://www.rfc-editor.org/info/rfc3270>>.

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

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

13.2. Informative references

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

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

Appendix A. Example of DetNet data plane operation

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

Appendix B. Example of pinned paths using IPv6

TBD.

Authors' Addresses

Jouni Korhonen (editor)

Email: jouni.nospam@gmail.com

Loa Andersson
Huawei

Email: loa@pi.nu

Yuanlong Jiang
Huawei

Email: jiangyuanlong@huawei.com

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

Email: norman.finn@mail01.huawei.com

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

Email: balazs.a.varga@ericsson.com

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

Email: janos.farkas@ericsson.com

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

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

Tal Mizrahi
Marvell
6 Hamada st.
Yokneam
Israel

Email: talmi@marvell.com

Lou Berger
LabN Consulting, L.L.C.

Email: lberger@labn.net

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

J. Farkas
B. Varga
Ericsson
R. Cummings
National Instruments
J. Yuanlong
Z. Yiyong
Huawei
June 30, 2017

DetNet Flow Information Model
draft-farkas-detnet-flow-information-model-01

Abstract

This document describes flow information model for Deterministic Networking (DetNet). The DetNet service is provided either for a Layer 3 or a Layer 2 flow. This document provides DetNet flow information model both for Layer 3 and Layer 2 flows in an integrated fashion.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 1, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Goals	3
1.2. Non Goals	4
2. Conventions Used in This Document	4
3. Terminology and Definitions	4
4. Naming Conventions	5
5. End System and DetNet domain	5
6. Flow	6
6.1. Identification and Specification of Flows	7
6.1.1. DetNet L3 Flow Identification and Specification at UNI	7
6.1.2. DetNet L2 Flow Identification and Specification at UNI	8
6.1.3. DetNetwork Flow Identification and Specification . .	8
6.2. Traffic Specification	9
6.3. Flow Rank	10
7. Source	10
8. Destination	11
9. Common Attributes of Source and Destination	12
9.1. End System Interfaces	12
9.2. Interface Capabilities	12
9.3. User to Network Requirements	13
10. DetNet Domain	14
10.1. DetNet Domain Capabilities	14
11. Status	15
11.1. Status Info	16
11.2. Interface Configuration	17
11.3. Failed Interfaces	17
12. Summary	17
13. IANA Considerations	17
14. Security Considerations	18
15. References	18
15.1. Normative References	18
15.2. Informative References	18
Authors' Addresses	19

1. Introduction

A Deterministic Networking (DetNet) service provides a capability to carry a unicast or a multicast data flow for an application with constrained requirements on network performance, e.g., low packet loss rate and/or latency. The DetNet service is provided either for a Layer 3 (L3) flow or a Layer 2 (L2) flow by an IP/MPLS network, see, e.g., [I-D.ietf-detnet-dp-alt]. Similarly, Time-Sensitive Networking (TSN) [IEEE8021TSN] can be used for L2 flows in a bridged network. DetNet and TSN have common architecture as expressed in [IETFDetNet] and [I-D.ietf-detnet-architecture]. DetNet service can be leveraged both by L3 and L2 flows, i.e., by DetNet L3 flows and DetNet L2 flows. Therefore, the DetNet flow information model provided by this document covers both DetNet L3 flows and DetNet L2 flows in an integrated fashion. Thus, the DetNet flow information model is based on [I-D.ietf-detnet-architecture] and on the data model specified by [IEEE8021Qcc]. Furthermore, the DetNet flow information model relies on the flow identification possibilities described in [IEEE8021CB], which is used by [IEEE8021Qcc] as well. In addition to TSN data model, [IEEE8021Qcc] also specifies configuration of TSN features (e.g., traffic scheduling specified by [IEEE8021Qbv]). Due to the common architecture and flow model, configuration features can be leveraged in certain deployment scenarios, e.g., when the network that provides the DetNet service includes both L3 and L2 network segments.

Based on the DetNet architecture [I-D.ietf-detnet-architecture] (see Section 4), this document (this revision) only considers the Centralized Network / Distributed User Model out of the models specified by [IEEE8021Qcc]. That is, there is a User-Network Interface (UNI) between an end system and a network. Furthermore, there is a central entity for the control of the network. For instance, the central entity implements a Path Computation Element (PCE) for the calculation and establishment of paths needed for packet replication and elimination, if any.

[[NOTE (to be removed from a future revision): The Goals and Non goals subsections are only for initial revisions, they are to be removed from future revisions of this draft.]]

1.1. Goals

As it is expressed in the Charter [IETFDetNet], the DetNet WG collaborates with IEEE 802.1 TSN in order to define a common architecture for both Layer 2 and Layer 3, which is beneficial for various reasons, e.g., in order to simplify implementations. The flow information model should be also common along those lines. As the TSN flow information/data model specified by [IEEE8021Qcc] is

mature, the DetNet flow information model described in this document is based on [IEEE8021Qcc], which is an amendment to [IEEE8021Q].

The Centralized Network / Distributed User Model of [IEEE8021Qcc] is used in this revision as a start of the work. Further models can be also useful for DetNet, e.g., the Fully Centralized Model for the Industrial M2M use case [I-D.ietf-detnet-use-cases].

This document intends to specify flow information model only.

1.2. Non Goals

This document (this revision) does not intend to specify either flow data model or DetNet configuration. From these aspects, the goals of this document differ from the goals of [IEEE8021Qcc], which also specifies data model and configuration of certain TSN features.

2. Conventions Used in This Document

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].

The lowercase forms with an initial capital "Must", "Must Not", "Shall", "Shall Not", "Should", "Should Not", "May", and "Optional" in this document are to be interpreted in the sense defined in [RFC2119], but are used where the normative behavior is defined in documents published by SDOs other than the IETF.

3. Terminology and Definitions

This document uses the terminology established in Section 2 of the DetNet architecture document [I-D.ietf-detnet-architecture]. The DetNet <=> TSN dictionary of [I-D.ietf-detnet-architecture] is used to perform translation from [IEEE8021Qcc] to this document. Additional terms used in this document:

DetNet L3 Flow: Layer 3 (L3) flow leveraging DetNet service.

DetNet L2 Flow: Layer 2 (L2) flow leveraging DetNet service.

DetNetwork Flow: DetNet data plane specific encapsulated format of a DetNet L2 or L3 flow leveraging DetNet service.

4. Naming Conventions

The following naming conventions were used for naming information model components in this document. It is recommended that extensions of the model use the same conventions.

- o Names SHOULD be descriptive.
- o Names MUST start with uppercase letters.
- o Composed names MUST use capital letters for the first letter of each component. All other letters are lowercase, even for acronyms. Exceptions are made for acronyms containing a mixture of lowercase and capital letters, such as IPv6. Examples are SourceMacAddress and DestinationIPv6Address.

5. End System and DetNet domain

Deterministic service is required by time/loss sensitive application(s) running on an end system during communication with its peer(s). Such a data exchange has various requirements on delay and/or loss parameters.

The DetNet architecture [I-D.ietf-detnet-architecture] distinguishes two kinds of end systems: Source and Destination. The same distinction is applied for the DetNet flow information model. In addition to the end systems interested in a flow, the status information of the flow is also important. Therefore, the DetNet flow information model relies on four high level groups:

- o Source: an end system capable of sourcing a DetNet flow. The Source information group includes elements that specify the Source for a single flow. This information group is applied from the user to the network.
- o Destination: an end system that is a destination of a DetNet flow. The Destination information group includes elements that specify the Destination for a single flow. This information group is applied from the user to the network.
- o DetNet Domain: a network providing the DetNet service. The DetNet domain information group includes elements that specify the forwarding method for a single flow. This information group is applied within the network.
- o Status: the status of a DetNet flow. The status information group includes elements that specify the status of the flow in the network. This information group is applied from the network to

the user. This information group informs the user whether or not the flow is ready for use.

There are two operations for each flow with respect to a Source or a Destination:

- o Join: Source/Destination request to join the flow.
- o Leave: Source/Destination request to leave the flow.

[[NOTE (to be removed from a future revision): Adding Modify operation can be considered to address cases when a flow is slightly changed, e.g., only MaxPayloadSize (Section 6.2) has been changed. The advantage of having a Modify is that it allows to initiate a change of flow spec while leaving the current flow is operating until the change is accepted. If there is no linkage between the Join and the Leave, then in figuring out whether the new flow spec can be supported, the central entity has to assume that the resources committed to the current flow are in use. If it is a Modify the central entity knows that the resources supporting the current flow can be available for supporting the altered flow. Modify is considered to be an optional operation due to possible control-plane limitations.]]

As the DetNet UNI can provide service for both L3 and L2 flows, end systems may not need to implement the L3 <=> L2 Transfer Function specified by [IEEE8021CB] (see, e.g., subclause 6.3; see also subclause 46.1 in [IEEE8021Qcc]). An edge node may implement a function similar to the Transfer Function, see, e.g., the Svc Proxy in Figure 1 in [I-D.ietf-detnet-dp-alt].

6. Flow

The flows leveraging DetNet service can be unicast or multicast data flows for an application with constrained requirements on network performance, e.g., low packet loss rate and/or latency. Therefore, they can require different connectivity types: point-to-point (p2p) or point-to-multipoint (p2mp). The p2mp connectivity is created by a transport layer function (e.g., p2mp LSP) [I-D.ietf-detnet-dp-alt]. (Note that mp2mp connectivity is a superposition of p2mp connections.)

Many flows using DetNet service are periodic with fix packet size (i.e., Constant Bit Rate (CBR) flows), or periodic with variable packet size.

Delay and loss parameters are correlated because the effect of late delivery can result data loss for an application. However, not all

applications require hard limits on both parameters (delay and loss). For example, some real-time applications allow graceful degradation if loss happens (e.g., sample-based processing, media distribution). Some others may require high-bandwidth connections that make the usage of techniques like packet replication economically challenging or even impossible. Some applications may not tolerate loss, but are not delay sensitive (e.g., bufferless sensors). Time/loss sensitive applications may have somewhat special requirements especially for loss (e.g., no loss in two consecutive communication cycles; very low outage time, etc.).

Flows have the following attributes:

- a. DataFlowSpecification (Section 6.1)
- b. TrafficSpecification (Section 6.2)
- c. FlowRank (Section 6.3)

Flow attributes are described in the following sections.

6.1. Identification and Specification of Flows

Identification options for DetNet flows at the UNI and within the DetNet domain are specified as follows; see Section 6.1.1 for DetNet L3 flows (at UNI), Section 6.1.2 for DetNet L2 flows (at UNI) and Section 6.1.3 for DetNetwork flows (within the network).

6.1.1. DetNet L3 Flow Identification and Specification at UNI

DetNet L3 flows can be identified and specified by the following attributes:

- a. SourceIpAddress
- b. DestinationIpAddress
- c. IPv6FlowLabel
- d. Dscp
- e. Protocol
- f. SourcePort
- g. DestinationPort

6.1.2. DetNet L2 Flow Identification and Specification at UNI

DetNet L2 flows can be identified and specified by the following attributes:

- a. DestinationMacAddress
- b. SourceMacAddress
- c. Pcp
- d. VlanId
- e. EtherType

[[NOTE (to be removed from a future revision): The Multiple Stream Registration Protocol (MSRP) [IEEE8021Q] uses StreamID to match Talker registrations with their corresponding Listener registrations, i.e., to identify Streams (L2 TSN flows). The StreamID includes the following subcomponents:

- o A 48-bit MAC Address associated with the Talker sourcing the stream to the bridged network.
- o A 16-bit unsigned integer value, Unique ID, used to distinguish among multiple streams sourced by the same Talker.

]]

6.1.3. DetNetwork Flow Identification and Specification

DetNetwork flows can be identified and specified by the following attributes:

- a. SourceIpAddress
- b. DestinationIpAddress
- c. IPv6FlowLabel
- d. MplsLabel

[[NOTE (to be removed from a future revision): Attributes are based on latest data plane solution.]]

6.2. Traffic Specification

TrafficSpecification specifies how the Source transmits packets for the flow. This is effectively the promise/request of the Source to the network. The network uses this traffic specification to allocate resources and adjust queue parameters in network nodes.

TrafficSpecification has the following attributes:

- a. Interval: the period of time in which the traffic specification cannot be exceeded.
- b. MaxPacketsPerInterval: the maximum number of packets that the Source will transmit in one Interval.
- c. MaxPayloadSize: the maximum payload size that the Source will transmit.

[[NOTE (to be removed from a future revision): These attributes can be used to describe any type of traffic (e.g., CBR, VBR, etc.) and can be used during resource allocation to represent worst case scenarios. Further optional attributes can be considered to achieve more efficient resource allocation. Such optional attributes might be worth for flows with soft requirements (i.e., the flow is only loss sensitive or only delay sensitive, but not both delay-and-loss sensitive). Possible options how to extend TrafficSpecification attributes is for further discussion. Identified options are described in the following notes.]]

[[NOTE1: Based on the already defined attributes the most similar additional attributes for VBR type flows can be defined as follows:

- o AveragePacketsPerInterval: the average number of packets that the Source will transmit in one Interval.
- o AveragePayloadSize: the average payload size that the Source will transmit.

]]

[[NOTE2: another alternative to deal better with various traffic types can rely on [RFC6003], which describes the support of Metro Ethernet Forum (MEF) Ethernet traffic parameters for using for resource reservation purposes. Such a Bandwidth Profile can be also adapted to describe the set of traffic parameters for a Detnet flow. Committed Rate indicates the rate at which traffic commits to be sent by the source (described in terms of the CIR (Committed Information Rate) and CBS (Committed Burst Size) attributes.) Excess Rate

indicates the extent by which the traffic sent by the source exceeds the committed rate. The Excess Rate is described in terms of the EIR (Excess Information Rate) and EBS (Excess Burst Size) attributes.]]

[[NOTE3: a third alternative is to define application based traffic models such as [GPP22885] defines periodic and event-driven traffic model, and 5G PPP work defines traffic model for MTC (Machine Type Communication) use cases. Periodic traffic type is usually for status update between devices or devices transmit status report to a central unit in regular basis. TrafficPeriod, defines the period of the status update message. DataSize, defines the data size of the message which is constant. 3GPP also defines approximately-periodic transmission with variations on period and uncertainty in the time arrival of the packets. Event-triggered traffic type corresponds traffic being triggered by an MTC device event. MinIntervalBetweenEvent, defines the minimum interval between two events. Event-triggered transmission will not happen all the time, whenever an alert is sent, it waits until the issue being solved to be able to send another alert. MaxPacketPerEvent, defines the max number of packets within one message.]]

6.3. Flow Rank

FlowRank provides the rank of this flow relative to others flows in the network. This rank is used to determine success/failure of flow establishment. Rank (boolean) is used by the network to decide which flows can and cannot exist when network resources reach their limit. Rank is used to help to determine which flows can be dropped (i.e., removed from node configuration) if a port of a node becomes oversubscribed (e.g., due to network reconfiguration). The true value is more important than the false value (i.e., flows with false are dropped first).

[[NOTE (to be removed from a future revision): FlowRank specified by [IEEE8021Qcc] is according to L2 logic, where lower values are preferred.]]

7. Source

The Source object specifies:

- o The behavior of the Source for the flow (how/when the Source transmits).
- o The requirements of the Source from the network.
- o The capabilities of the interface(s) of the Source.

The Source object includes the following attributes:

- a. DataFlowSpecification (Section 6.1)
- b. TrafficSpecification (Section 6.2)
- c. FlowRank (Section 6.3)
- d. EndSystemInterfaces (Section 9.1)
- e. InterfaceCapabilities (Section 9.2)
- f. UserToNetworkRequirements (Section 9.3)

For the join operation, the DataFlowSpecification, FlowRank, EndSystemInterfaces, and TrafficSpecification SHALL be included within the Source. For the join operation, the UserToNetworkRequirements and InterfaceCapabilities groups MAY be included within the Source.

For the leave operation, the DataFlowSpecification and EndSystemInterfaces SHALL be included within the Source.

8. Destination

The Destination object includes the following attributes:

- a. DataFlowSpecification (Section 6.1)
- b. EndSystemInterfaces (Section 9.1)
- c. InterfaceCapabilities (Section 9.2)
- d. UserToNetworkRequirements (Section 9.3)

For the join operation, the DataFlowSpecification and EndSystemInterfaces SHALL be included within the Destination. For the join operation, the UserToNetworkRequirements and InterfaceCapabilities groups MAY be included within the Destination.

For the leave operation, the DataFlowSpecification and EndSystemInterfaces SHALL be included within the Destination.

[[NOTE (to be removed from a future revision): Should we add DestinationRank? It could distinguish the importance of Destinations if the flow cannot be provided for all Destinations.]]

9. Common Attributes of Source and Destination

Source and Destination end systems have the following common attributes in addition to DataFlowSpecification (Section 6.1).

9.1. End System Interfaces

EndSystemInterfaces is a list of identifiers, one for each physical interface (port) in the end system acting as a Source or Destination. An interface is identified by an IP or a MAC address.

EndSystemInterfaces can refer also to logical sub-Interfaces if supported by the end system, e.g., based on IfIndex parameter.

9.2. Interface Capabilities

InterfaceCapabilities specifies the network capabilities of all interfaces (ports) contained in the EndSystemInterfaces object (Section 9.1). These capabilities may be configured via the InterfaceConfiguration object (Section 11.2) of the Status object (Section 11).

Note that an end system may have multiple interfaces with different network capabilities. In this case, each interface should be specified in a distinct top-level Source or Destination object (i.e., one entry in EndSystemInterfaces (Section 9.1)). Use of multiple entries in EndSystemInterfaces is intended for network capabilities that span multiple interfaces (e.g., packet replication and elimination).";.

InterfaceCapabilities attributes:

- a. SubInterfaceCapable (sub-interface capable)
- b. PREF-Capable (packet replication and elimination capable)

[[NOTE (to be removed from a future revision): InterfaceCapabilities attributes are to be defined. For information, [IEEE8021Qcc] specifies the following attributes:

- o VlanTagCapable (Customer VLAN Tag capable)
- o CB-Capable (frame replication and elimination capable)
- o CB-StreamIdentTypeList (a list of the optional Stream Identification types supported by the interface as specified in [IEEE8021CB].)

- o CB-SequenceTypeList (a list of the optional Sequence Encode/Decode types supported by the interface as specified in [IEEE8021CB].)

]]

9.3. User to Network Requirements

UserToNetworkRequirements specifies user requirements for the flow, such as latency and reliability.

The UserToNetworkRequirements object includes the following attributes:

- a. NumReplicationTrees
- b. MaxLatency

NumReplicationTrees specifies the number of maximally disjoint trees that the network should configure to provide packet replication and elimination for the flow. NumReplicationTrees is provided by the Source only. Destinations SHALL set this element to one. Value zero and one indicate no packet replication and elimination for the flow. When NumReplicationTrees is greater than one, packet replication and elimination is to be used for the flow. If the Source sets this element to greater than one, and packet replication and elimination is not possible in the network (e.g., no disjoint paths, or the nodes do not support packet replication and elimination), then the FailureCode of the Status object is non-zero (Section 11.1).

MaxLatency is the maximum latency from Source to Destination(s) for a single packet of the flow. MaxLatency is specified as an integer number of nanoseconds. When this requirement is specified by the Source, it must be satisfied for all Destinations. When this requirement is specified by a Destination, it must be satisfied for that particular Destination only. If the UserToNetworkRequirements group is not provided within the Source or Destination object, then value zero SHALL be used for this element. Value zero represents a special use for the maximum latency requirement. Value zero locks-down the initial latency that the network provides in the AccumulatedLatency parameter of the Status object (Section 11) after the successful configuration of the flow, such that any subsequent increase in the latency beyond that initial value causes the flow to fail.

[[NOTE-1 (to be removed from a future revision): Should we add a parameter to specify the maximum packet loss rate that can be tolerated for the flow?]]

[[NOTE-2 (to be removed from a future revision): TrafficSpecification (Section 6.2) specifies the Peak Information Rate (PIR) of the flow, which is a kind of user requirement to the network. Should we add Committed Information Rate (CIR), i.e., the minimum rate the user requests to be guaranteed for the flow by the network?]]

10. DetNet Domain

The DetNet Domain may change the encapsulation of a DetNet L2 or L3 flow at the UNI. That impacts not only how a flow can be recognised inside the DetNet domain but also the resource reservation calculations.

The DetNet Domain object specifies:

- o The behavior of the flow (how/when it is transmitted).
- o The requirements of the flow from the network.
- o The capabilities of the DetNet domain.

The DetNet domain object includes the following attributes:

- a. DataFlowSpecification (Section 6.1)
- b. TrafficSpecification (Section 6.2)
- c. FlowRank (Section 6.3)
- d. DetnetDomainCapabilities (Section 10.1)
- e. UserToNetworkRequirements (Section 9.3)

10.1. DetNet Domain Capabilities

DetnetDomainCapabilities specifies the network capabilities, which can be used to provide DetNet service. DetNet Edge nodes may change the encapsulation of a flow according to the data plane used inside the DetNet domain.

DetnetDomainCapabilities object includes the following attributes:

- a. EncapsulationFormat (data plane specific encapsulation)
- b. PREF-Capable (packet replication and elimination capable)

11. Status

The Status object is provided by the network each Source and Destination of the flow. The Status object provides the status of the flow with respect to the establishment of the flow by the network. The Status object is delivered via the corresponding UNI to each Source and Destination end system of the flow. The Status is distinct for each Source or Destination because the AccumulatedLatency and InterfaceConfiguration objects are distinct, see below.

The Status object SHALL include the attributes a), b), c); and MAY include attributes d), e):

- a. DataFlowSpecification (Section 6.1)
- b. StatusInfo (Section 11.1)
- c. AccumulatedLatency (this section below)
- d. InterfaceConfiguration (Section 11.2)
- e. FailedInterfaces (Section 11.3)

DataFlowSpecification identifies the flow for which status is provided. DataFlowSpecification is described in (Section 6.1) If the Status object is provided without a Source or Destination object in a protocol message via a UNI, then the DataFlowSpecification object SHALL be included within the Status object for both join and leave operations. If the Status object immediately follows a Source or Destination object in the protocol message, then the DataFlowSpecification object is obtained from the Source/Destination object, and therefore DataFlowSpecification is not required within the Status object.

AccumulatedLatency provides the worst-case latency that a single packet of the flow can encounter along its current path(s) in the network. When provided to a Source, AccumulatedLatency is the worst-case latency for all Destinations (worst path). AccumulatedLatency is specified as an integer number of nanoseconds. Latency is measured using the time at which the data frame's message timestamp point passes the reference plane marking the boundary between the network media and PHY. The message timestamp point is specified by IEEE Std 802.1AS [IEEE8021AS] for various media. For a successful Status, the network returns a value less than or equal to the MaxLatency of the UserToNetworkRequirements (Section 9.3). If the NumReplicationTrees of the UserToNetworkRequirements (Section 9.3) is one, then the AccumulatedLatency SHALL provide the worst latency for

the current path from the Source to each Destination. If the path is changed (e.g., due to rerouting), then the AccumulatedLatency changes accordingly. If the NumReplicationTrees of the UserToNetworkRequirements (Section 9.3) is greater than one, AccumlatedLatency SHALL provide the worst latency for all paths in use from the Source to each Destination.

11.1. Status Info

StatusInfo provides information regarding the status of a flow's configuration in the network.

The StatusInfo object MAY include the following attributes:

- a. SourceStatus is an enumeration for the status of the flow's Source:
 - * None: no Source
 - * Ready: Source is ready
 - * Failed: Source failed
- b. DestinationStatus is an enumeration for the status of the flow's Destinations:
 - * None: no Destination
 - * Ready: all Destinations are ready
 - * PartialFailed: One or more Destinations ready, and one or more Listeners failed. The flow can be used tf the Source is Ready.
 - * Failed: All Destinations failed.
- c. FailureCode: A non-zero code that specifies the problem if the flow encounters a failure (e.g., packet replication and elimination is requested but not possible, or SourceStatus is Failed, or DestinationStatus is Failed, or DestinationStatus is PartialFailed).

[[NOTE (to be removed from a future revision): FailureCodes to be defined for DetNet. Table 46-1 of [IEEE8021Qcc] describes TSN failure codes.]]

11.2. Interface Configuration

InterfaceConfiguration provides configuration of interfaces in the Source/Destination. This configuration assists the network in meeting the requirements of the flow. The InterfaceConfiguration object is according to the capabilities of the interface. InterfaceConfiguration can be distinct for each Source or Destination of each flow. If the InterfaceConfiguration object is not provided within the Status object, then the network SHALL assume zero elements as the default (no interface configuration).

The InterfaceConfiguration object MAY include one or more the following attributes:

- a. MAC or IP Address to identify the interface
- b. DataFlowSpecification (Section 6.1)

11.3. Failed Interfaces

FailedInterfaces provides a list of one or more physical interfaces (ports) in the failed node when a failure occurs in network configuration.

The InterfaceConfiguration object includes the following attributes:

- a. MAC or IP Address to identify the interface
- b. InterfaceName

InterfaceName is the name of the interface (port) within the node. This interface name SHALL be persistent, and unique within the node.

12. Summary

This document describes DetNet flow information model both for DetNet L3 flows and DetNet L2 flows based on the TSN data model specified by [IEEE8021Qcc]. This revision is extended with DetNet specific flow information model elements.

13. IANA Considerations

N/A.

14. Security Considerations

N/A.

15. References

15.1. Normative References

[I-D.ietf-detnet-architecture]

Finn, N. and P. Thubert, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-00 (work in progress), September 2016.

[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.ietf-detnet-use-cases]

Grossman, E., Gunther, C., Thubert, P., Wetterwald, P., Raymond, J., Korhonen, J., Kaneko, Y., Das, S., Zha, Y., Varga, B., Farkas, J., Goetz, F., and J. Schmitt, "Deterministic Networking Use Cases", draft-ietf-detnet-use-cases-01 (work in progress), February 2016.

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

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

15.2. Informative References

[GPP22885]

3GPP, "Study on LTE support for Vehicle-to-Everything (V2X) services", <<http://www.3gpp.org/DynaReport/22885.html>>.

[IEEE8021AS]

IEEE 802.1, "IEEE 802.1AS-2011: IEEE Standard for Local and metropolitan area networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks", 2011, <<http://standards.ieee.org/getieee802/download/802.1AS-2011.pdf>>.

[IEEE8021CB]

IEEE 802.1, "IEEE P802.1CB: IEEE Draft Standard for Local and metropolitan area networks - Frame Replication and Elimination for Reliability", 2017, <<http://www.ieee802.org/1/pages/802.1cb.html>>.

[IEEE8021Q]

IEEE 802.1, "IEEE 802.1Q-2014: IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks", 2014, <<http://standards.ieee.org/getieee802/download/802-1Q-2014.pdf>>.

[IEEE8021Qbv]

IEEE 802.1, "IEEE 802.1Qbv-2015: IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks -- Amendment 25: Enhancements for Scheduled Traffic", 2015, <<https://standards.ieee.org/findstds/standard/802.1Qbv-2015.html>>.

[IEEE8021Qcc]

IEEE 802.1, "IEEE P802.1Qcc-2015: IEEE Draft Standard for Local and metropolitan area networks - Bridges and Bridged Networks -- Amendment: Stream Reservation Protocol (SRP) Enhancements and Performance Improvements", 2017, <<http://www.ieee802.org/1/pages/802.1cc.html>>.

[IEEE8021TSN]

IEEE 802.1, "IEEE 802.1 Time-Sensitive Networking (TSN) Task Group", <<http://www.ieee802.org/1/pages/tsn.html>>.

[IETFDetNet]

IETF, "IETF Deterministic Networking (DetNet) Working Group", <<https://datatracker.ietf.org/wg/detnet/charter/>>.

Authors' Addresses

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

Email: janos.farkas@ericsson.com

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

Email: balazs.a.varga@ericsson.com

Rodney Cummings
National Instruments
11500 N. Mopac Expwy
Bldg. C
Austin, TX 78759-3504
USA

Email: rodney.cummings@ni.com

Jiang Yuanlong
Huawei

Email: jiangyuanlong@huawei.com

Zha Yiyong
Huawei

Email: zhayiyong@huawei.com

DetNet
Internet-Draft
Intended status: Standards Track
Expires: December 31, 2017

N. Finn
Huawei Technologies Co. Ltd
P. Thubert
Cisco
B. Varga
J. Farkas
Ericsson
June 29, 2017

Deterministic Networking Architecture
draft-ietf-detnet-architecture-02

Abstract

Deterministic Networking (DetNet) provides a capability to carry specified unicast or multicast data flows for real-time applications with extremely low data loss rates and bounded latency. Techniques used include: 1) reserving data plane resources for individual (or aggregated) DetNet flows in some or all of the intermediate nodes (e.g. bridges or routers) along the path of the flow; 2) providing explicit routes for DetNet flows that do not rapidly change with the network topology; and 3) distributing data from DetNet flow packets over time and/or space to ensure delivery of each packet's data in spite of the loss of a path. The capabilities can be managed by configuration, or by manual or automatic network management.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on December 31, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Terminology	4
2.1. Terms used in this document	4
2.2. IEEE 802 TSN to DetNet dictionary	6
3. Providing the DetNet Quality of Service	7
3.1. Primary goals defining the DetNet QoS	7
3.2. Mechanisms to achieve DetNet QoS	9
3.2.1. Congestion protection	9
3.2.2. Explicit routes	9
3.2.3. Jitter Reduction	10
3.2.4. Packet Replication and Elimination	11
3.3. Secondary goals for DetNet	12
3.3.1. Coexistence with normal traffic	12
3.3.2. Fault Mitigation	13
4. DetNet Architecture	14
4.1. DetNet stack model	14
4.1.1. Representative Protocol Stack Model	14
4.1.2. DetNet Data Plane Overview	16
4.1.3. Network reference model	18
4.2. DetNet systems	19
4.2.1. End system	19
4.2.2. DetNet edge, relay, and transit nodes	20
4.3. DetNet flows	21
4.3.1. DetNet flow types	21
4.3.2. Source guarantees	21
4.3.3. Incomplete Networks	23
4.4. Traffic Engineering for DetNet	23
4.4.1. The Application Plane	23
4.4.2. The Controller Plane	24
4.4.3. The Network Plane	24
4.5. Queuing, Shaping, Scheduling, and Preemption	25

4.6.	Service instance	26
4.7.	Flow identification at technology borders	27
4.7.1.	Exporting flow identification	27
4.7.2.	Flow attribute mapping between layers	29
4.7.3.	Flow-ID mapping examples	30
4.8.	Advertising resources, capabilities and adjacencies	32
4.9.	Provisioning model	32
4.9.1.	Centralized Path Computation and Installation	32
4.9.2.	Distributed Path Setup	32
4.10.	Scaling to larger networks	33
4.11.	Connected islands vs. networks	33
4.12.	Compatibility with Layer-2	33
5.	Open Questions	34
5.1.	Flat vs. hierarchical control	34
5.2.	Peer-to-peer reservation protocol	34
5.3.	Wireless media interactions	35
5.4.	Packet encoding for service protection	35
6.	Security Considerations	35
7.	Privacy Considerations	36
8.	IANA Considerations	36
9.	Acknowledgements	36
10.	Access to IEEE 802.1 documents	37
11.	Informative References	37
	Authors' Addresses	42

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. This is accomplished by dedicating network resources such as link bandwidth and buffer space to DetNet flows and/or classes of DetNet flows, and by replicating packets along multiple paths. Unused reserved resources are available to non-DetNet packets.

The Deterministic Networking Problem Statement

[I-D.ietf-detnet-problem-statement] introduces Deterministic Networking, and Deterministic Networking Use Cases

[I-D.ietf-detnet-use-cases] summarizes the need for it. See [I-D.dt-detnet-dp-alt] for a discussion of specific techniques that can be used to identify DetNet Flows and assign them to specific paths through a network.

A goal of DetNet is a converged network in all respects. That is, the presence of DetNet flows does not preclude non-DetNet flows, and the benefits offered DetNet flows should not, except in extreme cases, prevent existing QoS mechanisms from operating in a normal fashion, subject to the bandwidth required for the DetNet flows. A

single source-destination pair can trade both DetNet and non-DetNet flows. End systems and applications need not instantiate special interfaces for DetNet flows. Networks are not restricted to certain topologies; connectivity is not restricted. Any application that generates a data flow that can be usefully characterized as having a maximum bandwidth should be able to take advantage of DetNet, as long as the necessary resources can be reserved. Reservations can be made by the application itself, via network management, by an applications controller, or by other means.

Many applications of interest to Deterministic Networking require the ability to synchronize the clocks in end systems to a sub-microsecond accuracy. Some of the queue control techniques defined in Section 4.5 also require time synchronization among relay and transit nodes. The means used to achieve time synchronization are not addressed in this document. DetNet should accommodate various synchronization techniques and profiles that are defined elsewhere to solve exchange time in different market segments.

The present document is an individual contribution, but it is intended by the authors for adoption by the DetNet working group.

2. Terminology

2.1. Terms used in this document

The following special terms are used in this document in order to avoid the assumption that a given element in the architecture does or does not have Internet Protocol stack, functions as a router, bridge, firewall, or otherwise plays a particular role at Layer-2 or higher.

App-flow

The native format of a DetNet flow.

destination

An end system capable of receiving a DetNet flow.

DetNet domain

The portion of a network that is DetNet aware. It includes end systems and other DetNet nodes.

DetNet flow

A DetNet flow is a sequence of packets to which the DetNet service is to be provided.

DetNet compound flow and DetNet member flow

A DetNet compound flow is a DetNet flow that has been separated into multiple duplicate DetNet member flows, which

are eventually merged back into a single DetNet compound flow, at the DetNet transport layer. "Compound" and "member" are strictly relative to each other, not absolutes; a DetNet compound flow comprising multiple DetNet member flows can, in turn, be a member of a higher-order compound.

DetNet intermediate node

A DetNet relay node or transit node.

DetNet edge node

An instance of a DetNet relay node that includes either a DetNet service layer proxy function for DetNet service protection (e.g. the addition or removal of packet sequencing information) for one or more end systems, or starts or terminates congestion protection at the DetNet transport layer, analogous to a Label Edge Router (LER).

DetNet-UNI

User-to-Network Interface with DetNet specific functionalities. It is a packet-based reference point and may provide multiple functions like encapsulation, status, synchronization, etc.

end system

Commonly called a "host" or "node" in IETF documents, and an "end station" in IEEE 802 documents. End systems of interest to this document are either sources or destinations of DetNet flows. An end system may or may not be DetNet transport layer aware or DetNet service layer aware.

link

A connection between two DetNet nodes. It may be composed of a physical link or a sub-network technology that can provide appropriate traffic delivery for DetNet flows.

DetNet node

A DetNet aware end system, transit node, or relay node. "DetNet" may be omitted in some text.

DetNet relay node

A DetNet node including a service layer function that interconnects different DetNet transport layer paths to provide service protection. A DetNet relay node can be a bridge, a router, a firewall, or any other system that participates in the DetNet service layer. It typically incorporates DetNet transport layer functions as well, in which case it is collocated with a transit node.

reservation

A trail of configuration between source to destination(s) through transit nodes and subnets associated with a DetNet flow, to provide congestion protection.

DetNet service layer

The layer at which service protection is provided, either packet sequencing, replication, and elimination (Section 3.2.4) or network coding (Section 5.4).

source

An end system capable of sourcing a DetNet flow.

DetNet transit node

A node operating at the DetNet transport layer, that utilizes link layer and/or network layer switching across multiple links and/or sub-networks to provide paths for DetNet service layer functions. Optionally provides congestion protection over those paths. An MPLS LSR is an example of a DetNet transit node.

DetNet transport layer

The layer that optionally provides congestion protection for DetNet flows over paths provided by the underlying network.

TSN

Time-Sensitive Networking, TSN is a Task Group of the IEEE 802.1 Working Group.

2.2. IEEE 802 TSN to DetNet dictionary

This section also serves as a dictionary for translating from the terms used by the IEEE 802 Time-Sensitive Networking (TSN) Task Group to those of the DetNet WG.

Listener

The IEEE 802 term for a destination of a DetNet flow.

relay system

The IEEE 802 term for a DetNet intermediate node.

Stream

The IEEE 802 term for a DetNet flow.

Talker

The IEEE 802 term for the source of a DetNet flow.

3. Providing the DetNet Quality of Service

3.1. Primary goals defining the DetNet QoS

The DetNet Quality of Service can be expressed in terms of:

- o Minimum and maximum end-to-end latency from source to destination; timely delivery and jitter avoidance derive from these constraints
- o Probability of loss of a packet, under various assumptions as to the operational states of the nodes and links. A derived property is whether it is acceptable to deliver a duplicate packet, which is an inherent risk in highly reliable and/or broadcast transmissions

It is a distinction of DetNet that it is concerned solely with worst-case values for the end-to-end latency. Average, mean, or typical values are of no interest, because they do not affect the ability of a real-time system to perform its tasks. In general, a trivial priority-based queuing scheme will give better average latency to a data flow than DetNet, but of course, the worst-case latency can be essentially unbounded.

Three techniques are used by DetNet to provide these qualities of service:

- o Congestion protection (Section 3.2.1).
- o Explicit routes (Section 3.2.2).
- o Service protection (Section 3.2.4).

Congestion protection operates by reserving resources along the path of a DetNet Flow, e.g. buffer space or link bandwidth. Congestion protection greatly reduces, or even eliminates entirely, packet loss due to output packet congestion within the network, but it can only be supplied to a DetNet flow that is limited at the source to a maximum packet size and transmission rate.

Congestion protection addresses both of the DetNet QoS requirements (latency and packet loss). Given that DetNet nodes have a finite amount of buffer space, congestion protection necessarily results in a maximum end-to-end latency. It also addresses the largest contribution to packet loss, which is buffer congestion.

After congestion, the most important contributions to packet loss are typically from random media errors and equipment failures. Service protection is the name for the mechanisms used by DetNet to address

these losses. The mechanisms employed are constrained by the requirement to meet the users' latency requirements. Packet replication and elimination (Section 3.2.4) packet encoding Section 5.4 are described in this document to provide service protection; others may be found. Both mechanisms distribute the contents of DetNet flows over multiple paths in time and/or space, so that the loss of some of the paths does need not cause the loss of any packets. The paths are typically (but not necessarily) explicit routes, so that they cannot suffer temporary interruptions caused by the reconvergence of routing or bridging protocols.

These three techniques can be applied independently, giving eight possible combinations, including none (no DetNet), although some combinations are of wider utility than others. This separation keeps the protocol stack coherent and maximizes interoperability with existing and developing standards in this (IETF) and other Standards Development Organizations. Some examples of typical expected combinations:

- o Explicit routes plus service protection are exactly the techniques employed by [HSR-PRP]. Explicit routes are achieved by limiting the physical topology of the network, and the sequentialization, replication, and duplicate elimination are facilitated by packet tags added at the front or the end of Ethernet frames.
- o Congestion protection alone is offered by IEEE 802.1 Audio Video bridging [IEEE802.1BA-2011]. As long as the network suffers no failures, zero congestion loss can be achieved through the use of a reservation protocol (MSRP), shapers in every bridge, and a bit of network calculus.
- o Using all three together gives maximum protection.

There are, of course, simpler methods available (and employed, today) to achieve levels of latency and packet loss that are satisfactory for many applications. Prioritization and over-provisioning is one such technique. However, these methods generally work best in the absence of any significant amount of non-critical traffic in the network (if, indeed, such traffic is supported at all), or work only if the critical traffic constitutes only a small portion of the network's theoretical capacity, or work only if all systems are functioning properly, or in the absence of actions by end systems that disrupt the network's operations.

There are any number of methods in use, defined, or in progress for accomplishing each of the above techniques. It is expected that this DetNet Architecture will assist various vendors, users, and/or "vertical" Standards Development Organizations (dedicated to a single

industry) to make selections among the available means of implementing DetNet networks.

3.2. Mechanisms to achieve DetNet Qos

3.2.1. Congestion protection

The primary means by which DetNet achieves its QoS assurances is to reduce, or even completely eliminate, congestion at an output port as a cause of packet loss. Given that a DetNet flow cannot be throttled, this can be achieved only by the provision of sufficient buffer storage at each hop through the network to ensure that no packets are dropped due to a lack of buffer storage.

Ensuring adequate buffering requires, in turn, that the source, and every intermediate node along the path to the destination (or nearly every node -- see Section 4.3.3) be careful to regulate its output to not exceed the data rate for any DetNet flow, except for brief periods when making up for interfering traffic. Any packet sent ahead of its time potentially adds to the number of buffers required by the next hop, and may thus exceed the resources allocated for a particular DetNet flow.

The low-level mechanisms described in Section 4.5 provide the necessary regulation of transmissions by an end system or intermediate node to provide congestion protection. The reservation of the bandwidth and buffers for a DetNet flow requires the provisioning described in Section 4.9. A DetNet node may have other resources requiring allocation and/or scheduling, that might otherwise be over-subscribed and trigger the rejection of a reservation.

3.2.2. Explicit routes

In networks controlled by typical peer-to-peer protocols such as IEEE 802.1 ISIS bridged networks or IETF OSPF routed networks, a network topology event in one part of the network can impact, at least briefly, the delivery of data in parts of the network remote from the failure or recovery event. Thus, even redundant paths through a network, if controlled by the typical peer-to-peer protocols, do not eliminate the chances of brief losses of contact.

Many real-time networks rely on physical rings or chains of two-port devices, with a relatively simple ring control protocol. This supports redundant paths for service protection with a minimum of wiring. As an additional benefit, ring topologies can often utilize different topology management protocols than those used for a mesh network, with a consequent reduction in the response time to topology

changes. Of course, this comes at some cost in terms of increased hop count, and thus latency, for the typical path.

In order to get the advantages of low hop count and still ensure against even very brief losses of connectivity, DetNet employs explicit routes, where the path taken by a given DetNet flow does not change, at least immediately, and likely not at all, in response to network topology events. Service protection (Section 3.2.4 or Section 5.4) over explicit routes provides a high likelihood of continuous connectivity. Explicit routes are commonly used in MPLS TE LSPs.

3.2.3. Jitter Reduction

A core objective of DetNet is to enable the convergence of Non-IP networks onto a common network infrastructure. This requires the accurate emulation of currently deployed mission-specific networks, which typically rely on point-to-point analog (e.g. 4-20mA modulation) and serial-digital cables (or buses) for highly reliable, synchronized and jitter-free communications. While the latency of analog transmissions is basically the speed of light, legacy serial links are usually slow (in the order of Kbps) compared to, say, GigE, and some latency is usually acceptable. What is not acceptable is the introduction of excessive jitter, which may, for instance, affect the stability of control systems.

Applications that are designed to operate on serial links usually do not provide services to recover the jitter, because jitter simply does not exist there. Streams of information are expected to be delivered in-order and the precise time of reception influences the processes. In order to converge such existing applications, there is a desire to emulate all properties of the serial cable, such as clock transportation, perfect flow isolation and fixed latency. While minimal jitter (in the form of specifying minimum, as well as maximum, end-to-end latency) is supported by DetNet, there are practical limitations on packet-based networks in this regard. In general, users are encouraged to use, instead of, "do this when you get the packet," a combination of:

- o Sub-microsecond time synchronization among all source and destination end systems, and
- o Time-of-execution fields in the application packets.

Jitter reduction is provided by the mechanisms described in Section 4.5 that also provide congestion protection.

3.2.4. Packet Replication and Elimination

After congestion loss has been eliminated, the most important causes of packet loss are random media and/or memory faults, and equipment failures. Both causes of packet loss can be greatly reduced by spreading the data in a packet over multiple transmissions. One such method for service protection is described in this section, which sends the same packets over multiple paths. See also Section 5.4.

Packet replication and elimination, also known as seamless redundancy [HSR-PRP], or 1+1 hitless protection, is a function of the DetNet service layer. It involves three capabilities:

- o Providing sequencing information, once, at or near the source, to the packets of a DetNet compound flow. This may be done by adding a sequence number or time stamp as part of DetNet, or may be inherent in the packet, e.g. in a transport protocol, or associated to other physical properties such as the precise time (and radio channel) of reception of the packet. Section 3.2.2.
- o Replicating these packets into multiple DetNet member flows and, typically, sending them along at least two different paths to the destination(s), e.g. over the explicit routes of
- o Eliminating duplicated packets. This may be done at any step along the path to save network resources further down, in particular if multiple Replication points exist. But the most common case is to perform this operation at the very edge of the DetNet network, preferably in or near the receiver.

This function is a "hitless" version of, e.g., the 1+1 linear protection in [RFC6372]. That is, instead of switching from one flow to the other when a failure of a flow is detected, DetNet combines both flows, and performs a packet-by-packet selection of which to discard, based on sequence number.

In the simplest case, this amounts to replicating each packet in a source that has two interfaces, and conveying them through the network, along separate paths, to the similarly dual-homed destinations, that discard the extras. This ensures that one path (with zero congestion loss) remains, even if some intermediate node fails. The sequence numbers can also be used for loss detection and for re-ordering.

Detnet relay nodes in the network can provide replication and elimination facilities at various points in the network, so that multiple failures can be accommodated.

This is shown in the following figure, where the two relay nodes each replicate (R) the DetNet flow on input, sending the DetNet member flows to both the other relay node and to the end system, and eliminate duplicates (E) on the output interface to the right-hand end system. Any one link in the network can fail, and the Detnet compound flow can still get through. Furthermore, two links can fail, as long as they are in different segments of the network.

Packet replication and elimination

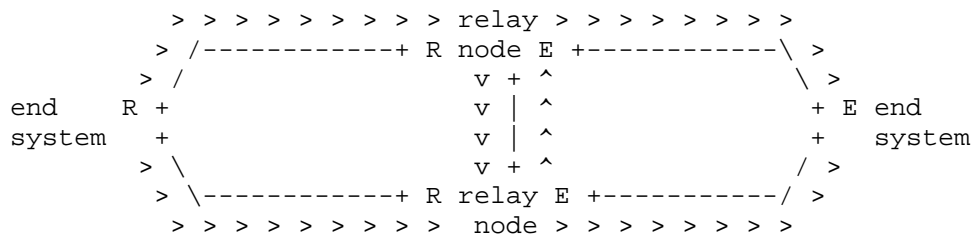


Figure 1

Note that packet replication and elimination does not react to and correct failures; it is entirely passive. Thus, intermittent failures, mistakenly created packet filters, or misrouted data is handled just the same as the equipment failures that are detected handled by typical routing and bridging protocols.

If packet replication and elimination is used over paths providing congestion protection (Section 3.2.1), and member flows that take different-length paths through the network are combined, a merge point may require extra buffering to equalize the delays over the different paths. This equalization ensures that the resultant compound flow will not exceed its contracted bandwidth even after one or the other of the paths is restored after a failure.

3.3. Secondary goals for DetNet

Many applications require DetNet to provide additional services, including coexistence with other QoS mechanisms Section 3.3.1 and protection against misbehaving transmitters Section 3.3.2.

3.3.1. Coexistence with normal traffic

A DetNet network supports the dedication of a high proportion (e.g. 75%) of the network bandwidth to DetNet flows. But, no matter how much is dedicated for DetNet flows, it is a goal of DetNet to coexist with existing Class of Service schemes (e.g., DiffServ). It is also

important that non-DetNet traffic not disrupt the DetNet flow, of course (see Section 3.3.2 and Section 6). For these reasons:

- o Bandwidth (transmission opportunities) not utilized by a DetNet flow are available to non-DetNet packets (though not to other DetNet flows).
- o DetNet flows can be shaped or scheduled, in order to ensure that the highest-priority non-DetNet packet also is ensured a worst-case latency (at any given hop).
- o When transmission opportunities for DetNet flows are scheduled in detail, then the algorithm constructing the schedule should leave sufficient opportunities for non-DetNet packets to satisfy the needs of the users of the network. Detailed scheduling can also permit the time-shared use of buffer resources by different DetNet flows.

Ideally, the net effect of the presence of DetNet flows in a network on the non-DetNet packets is primarily a reduction in the available bandwidth.

3.3.2. Fault Mitigation

One key to building robust real-time systems is to reduce the infinite variety of possible failures to a number that can be analyzed with reasonable confidence. DetNet aids in the process by providing filters and policers to detect DetNet packets received on the wrong interface, or at the wrong time, or in too great a volume, and to then take actions such as discarding the offending packet, shutting down the offending DetNet flow, or shutting down the offending interface.

It is also essential that filters and service remarking be employed at the network edge to prevent non-DetNet packets from being mistaken for DetNet packets, and thus impinging on the resources allocated to DetNet packets.

There exist techniques, at present and/or in various stages of standardization, that can perform these fault mitigation tasks that deliver a high probability that misbehaving systems will have zero impact on well-behaved DetNet flows, except of course, for the receiving interface(s) immediately downstream of the misbehaving device. Examples of such techniques include traffic policing functions (e.g. [RFC2475]) and separating flows into per-flow rate-limited queues.

Packet sequencing

As part of DetNet service protection, supplies the sequence number for packet replication and elimination (Section 3.2.4). Peers with Duplicate elimination. This layer is not needed if a higher-layer transport protocol is expected to perform any packet sequencing and duplicate elimination required by the DetNet flow duplication.

Duplicate elimination

As part of the DetNet service layer, based on the sequenced number supplied by its peer, packet sequencing, Duplicate elimination discards any duplicate packets generated by DetNet flow duplication. It can operate on member flows, compound flows, or both. The duplication may also be inferred from other information such as the precise time of reception in a scheduled network. The duplicate elimination layer may also perform resequencing of packets to restore packet order in a flow that was disrupted by the loss of packets on one or another of the multiple paths taken.

Flow duplication

As part of DetNet service protection, replicates packets that belong to a DetNet compound flow into two or more DetNet member flows. Note that this function is separate from packet sequencing. Flow duplication can be an explicit duplication and remarking of packets, or can be performed by, for example, techniques similar to ordinary multicast replication. Peers with DetNet flow merging.

Network flow merging

As part of DetNet service protection, merges DetNet member flows together for packets coming up the stack belonging to a specific DetNet compound flow. Peers with DetNet flow duplication. DetNet flow merging, together with packet sequencing, duplicate elimination, and DetNet flow duplication, performs packet replication and elimination (Section 3.2.4).

Packet encoding

As part of DetNet service protection, as an alternative to packet sequencing and flow duplication, packet encoding combines the information in multiple DetNet packets, perhaps from different DetNet compound flows, and transmits that information in packets on different DetNet member Flows. Peers with Packet decoding.

Packet decoding

As part of DetNet service protection, as an alternative to flow merging and duplicate elimination, packet decoding takes packets from different DetNet member flows, and computes from those packets the original DetNet packets from the compound flows input to packet encoding. Peers with Packet encoding.

Congestion protection

The DetNet transport layer provides congestion protection. See Section 4.5. The actual queuing and shaping mechanisms are typically provided by underlying subnet layers, but since these can be closely associated with the means of providing paths for DetNet flows (e.g. MPLS LSPs or {VLAN, multicast destination MAC address} pairs), the path and the congestion protection are conflated in this figure.

Note that the packet sequencing and duplication elimination functions at the source and destination ends of a DetNet compound flow may be performed either in the end system or in a DetNet edge node. The reader must not confuse a DetNet edge function with other kinds of edge functions, e.g. an Label Edge Router, although the two functions may be performed together. The DetNet edge function is concerned with sequencing packets belonging to DetNet flows. The LER with encapsulating/decapsulating packets for transport, and is considered part of the network underlying the DetNet transport layer.

4.1.2. DetNet Data Plane Overview

A "Deterministic Network" will be composed of DetNet enabled nodes i.e., End Systems, Edge Nodes, Relay Nodes and collectively deliver DetNet services. DetNet enabled nodes are interconnected via Transit Nodes (i.e., routers) which support DetNet, but are not DetNet service aware. Transit nodes see DetNet nodes as end points. All DetNet enabled nodes are connect to sub-networks, where a point-to-point link is also considered as a simple sub-network. These sub-networks will provide DetNet compatible service for support of DetNet traffic. Examples of sub-networks include IEEE 802.1 TSN and OTN. Of course, multi-layer DetNet systems may also be possible, where one DetNet appears as a sub-network, and provides service to, a higher layer DetNet system. A simple DetNet concept network is shown in Figure 3.

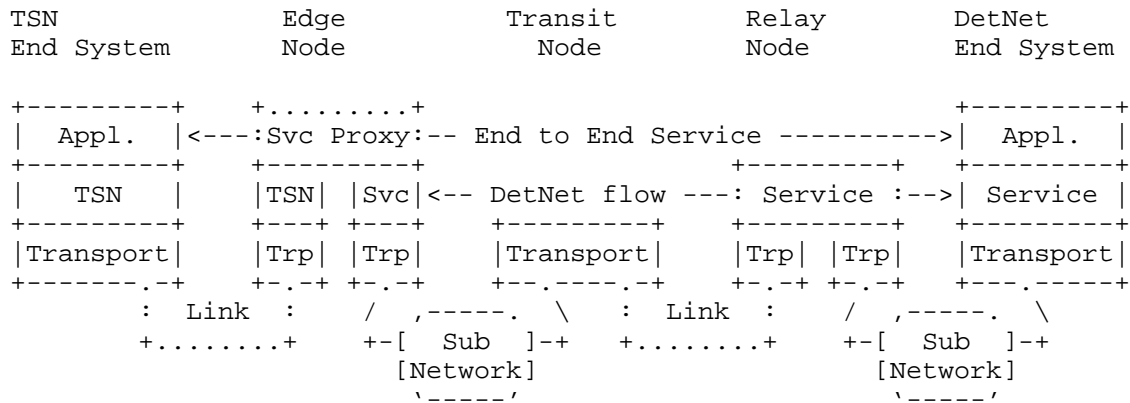


Figure 3: A Simple DetNet Enabled Network

Distinguishing the function of these two DetNet data plane layers, the DetNet service layer and the DetNet transport layer, helps to explore and evaluate various combinations of the data plane solutions available. This separation of DetNet layers, while helpful, should not be considered as formal requirement. For example, some technologies may violate these strict layers and still be able to deliver a DetNet service.

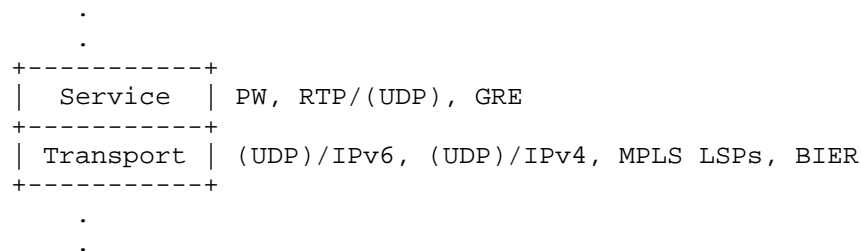


Figure 4: DetNet adaptation to data plane

In some networking scenarios, the end system initially provides a DetNet flow encapsulation, which contains all information needed by DetNet nodes (e.g., Real-time Transport Protocol (RTP) [RFC3550] based DetNet flow transported over a native UDP/IP network or PseudoWire). In other scenarios, the encapsulation formats might differ significantly. As an example, a CPRI "application's" I/Q data mapped directly to Ethernet frames may have to be transported over an MPLS-based packet switched network (PSN).

There are many valid options to create a data plane solution for DetNet traffic by selecting a technology approach for the DetNet

service layer and also selecting a technology approach for the DetNet transport layer. There are a high number of valid combinations.

One of the most fundamental differences between different potential data plane options is the basic addressing and headers used by DetNet end systems. For example, is the basic service a Layer 2 (e.g., Ethernet) or Layer 3 (i.e., IP) service. This decision impacts how DetNet end systems are addressed, and the basic forwarding logic for the DetNet service layer.

4.1.3. Network reference model

The figure below shows another view of the DetNet service related reference points and main components (Figure 5).

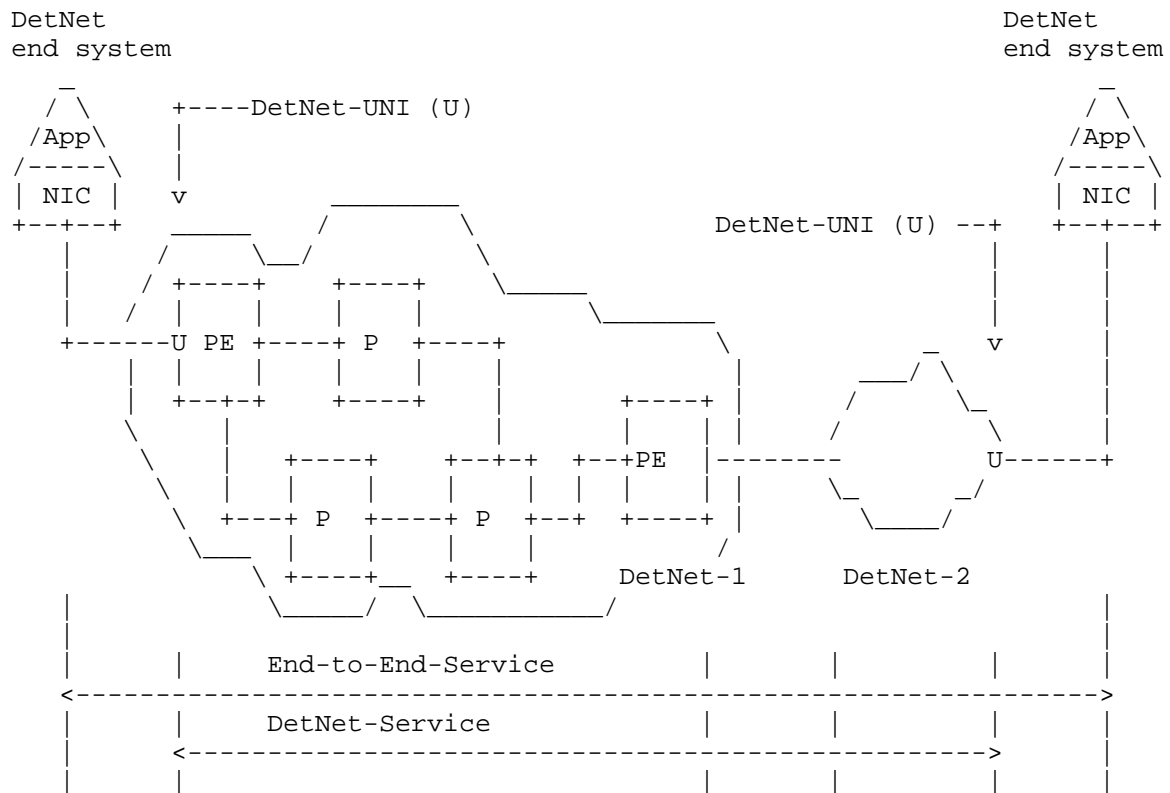


Figure 5: DetNet Service Reference Model (multi-domain)

DetNet-UNIs ("U" in Figure 5) are assumed in this document to be packet-based reference points and provide connectivity over the packet network. A DetNet-UNI may provide multiple functions, e.g.,

it may add networking technology specific encapsulation to the DetNet flows if necessary; it may provide status of the availability of the connection associated to a reservation; it may provide a synchronization service for the end system; it may carry enough signaling to place the reservation in a network without a controller, or if the controller only deals with the network but not the end points. Internal reference points of end systems (between the application and the NIC) are more challenging from control perspective and they may have extra requirements (e.g., in-order delivery is expected in end system internal reference points, whereas it is considered optional over the DetNet-UNI), therefore not covered in this document.

4.2. DetNet systems

4.2.1. End system

The native data flow between the source/destination end systems is referred to as application-flow (App-flow). The traffic characteristics of an App-flow can be CBR (constant bit rate) or VBR (variable bit rate) and can have L1 or L2 or L3 encapsulation (e.g., TDM (time-division multiplexing), Ethernet, IP). These characteristics are considered as input for resource reservation and might be simplified to ensure determinism during transport (e.g., making reservations for the peak rate of VBR traffic, etc.).

An end system may or may not be DetNet transport layer aware or DetNet service layer aware. That is, an end system may or may not contain DetNet specific functionality. End systems with DetNet functionalities may have the same or different transport layer as the connected DetNet domain. Grouping of end systems are shown in Figure 6.

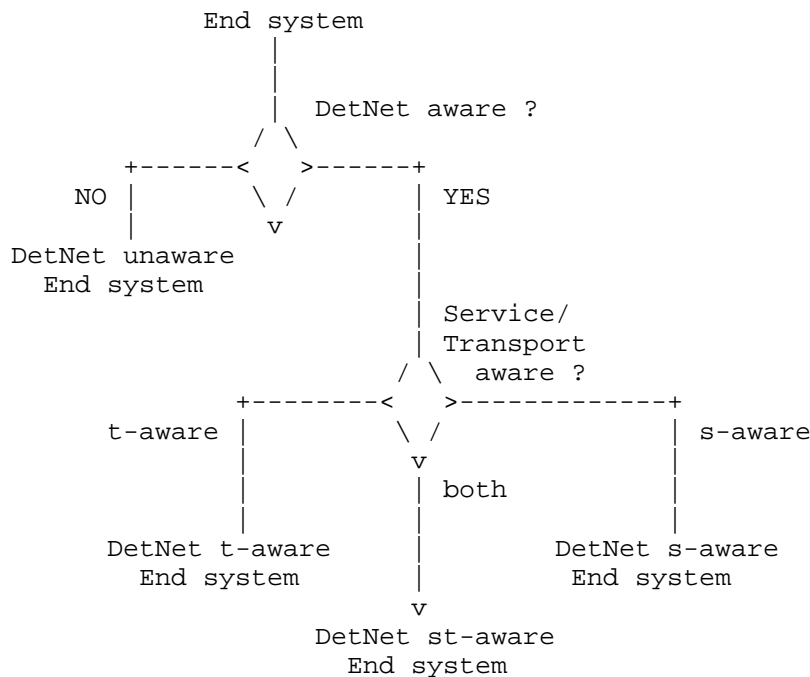


Figure 6: Grouping of end systems

Note some known use cases for end systems:

- o DetNet unaware: The classic case requiring network proxies.
- o DetNet t-aware: An extant TSN system. It knows about some TSN functions (e.g., reservation), but not about replication/elimination.
- o DetNet s-aware: An extant IEC 62439-3 system. It supplies sequence numbers, but doesn't know about zero congestion loss.
- o DetNet st-aware: A full functioning DetNet end station, it has DetNet functionalities and usually the same forwarding paradigm as the connected DetNet domain. It can be treated as an integral part of the DetNet domain.

4.2.2. DetNet edge, relay, and transit nodes

As shown in Figure 3, DetNet edge nodes providing proxy service and DetNet relay nodes providing the DetNet service layer are DetNet-aware, and DetNet transit nodes need only be aware of the DetNet transport layer.

In general, if a DetNet flow passes through one or more DetNet-unaware network node between two DetNet nodes providing the DetNet transport layer for that flow, there is a potential for disruption or failure of the DetNet QoS. A network administrator needs to ensure that the DetNet-unaware network nodes are configured to minimize the chances of packet loss and delay, and provision enough extra buffer space in the DetNet transit node following the DetNet-unaware network nodes to absorb the induced latency variations.

4.3. DetNet flows

4.3.1. DetNet flow types

A DetNet flow can have different formats during while it is transported between the peer end systems. Therefore, the following possible types / formats of a DetNet flow are distinguished in this document:

- o App-flow: native format of a DetNet flow. It does not contain any DetNet related attributes.
- o DetNet-t-flow: specific format of a DetNet flow. Only requires the congestion / latency features provided by the Detnet transport layer.
- o DetNet-s-flow: specific format of a DetNet flow. Only requires the replication/elimination feature ensured by the DetNet service layer.
- o DetNet-st-flow: specific format of a DetNet flow. It requires both DetNet service layer and DetNet transport layer functions during forwarding.

4.3.2. Source guarantees

For the purposes of congestion protection, DetNet flows can be synchronous or asynchronous. In synchronous DetNet flows, at least the intermediate nodes (and possibly the end systems) are closely time synchronized, typically to better than 1 microsecond. By transmitting packets from different DetNet flows or classes of DetNet flows at different times, using repeating schedules synchronized among the intermediate nodes, resources such as buffers and link bandwidth can be shared over the time domain among different DetNet flows. There is a tradeoff among techniques for synchronous DetNet flows between the burden of fine-grained scheduling and the benefit of reducing the required resources, especially buffer space.

In contrast, asynchronous DetNet flows are not coordinated with a fine-grained schedule, so relay and end systems must assume worst-case interference among DetNet flows contending for buffer resources. Asynchronous DetNet flows are characterized by:

- o A maximum packet size;
- o An observation interval; and
- o A maximum number of transmissions during that observation interval.

These parameters, together with knowledge of the protocol stack used (and thus the size of the various headers added to a packet), limit the number of bit times per observation interval that the DetNet flow can occupy the physical medium.

The source promises that these limits will not be exceeded. If the source transmits less data than this limit allows, the unused resources such as link bandwidth can be made available by the system to non-DetNet packets. However, making those resources available to DetNet packets in other DetNet flows would serve no purpose. Those other DetNet flows have their own dedicated resources, on the assumption that all DetNet flows can use all of their resources over a long period of time.

Note that there is no provision in DetNet for throttling DetNet flows (reducing the transmission rate via feedback); the assumption is that a DetNet flow, to be useful, must be delivered in its entirety. That is, while any useful application is written to expect a certain number of lost packets, the real-time applications of interest to DetNet demand that the loss of data due to the network is extraordinarily infrequent.

Although DetNet strives to minimize the changes required of an application to allow it to shift from a special-purpose digital network to an Internet Protocol network, one fundamental shift in the behavior of network applications is impossible to avoid: the reservation of resources before the application starts. In the first place, a network cannot deliver finite latency and practically zero packet loss to an arbitrarily high offered load. Secondly, achieving practically zero packet loss for unthrottled (though bandwidth limited) DetNet flows means that bridges and routers have to dedicate buffer resources to specific DetNet flows or to classes of DetNet flows. The requirements of each reservation have to be translated into the parameters that control each system's queuing, shaping, and scheduling functions and delivered to the hosts, bridges, and routers.

4.3.3. Incomplete Networks

The presence in the network of transit nodes or subnets that are not fully capable of offering DetNet services complicates the ability of the intermediate nodes and/or controller to allocate resources, as extra buffering, and thus extra latency, must be allocated at points downstream from the non-DetNet intermediate node for a DetNet flow.

4.4. Traffic Engineering for DetNet

Traffic Engineering Architecture and Signaling (TEAS) [TEAS] defines traffic-engineering architectures for generic applicability across packet and non-packet networks. From TEAS perspective, Traffic Engineering (TE) refers to techniques that enable operators to control how specific traffic flows are treated within their networks.

Because of its very nature of establishing explicit optimized paths, Deterministic Networking can be seen as a new, specialized branch of Traffic Engineering, and inherits its architecture with a separation into planes.

The Deterministic Networking architecture is thus composed of three planes, a (User) Application Plane, a Controller Plane, and a Network Plane, which echoes that of Figure 1 of Software-Defined Networking (SDN): Layers and Architecture Terminology [RFC7426].:

4.4.1. The Application Plane

Per [RFC7426], the Application Plane includes both applications and services. In particular, the Application Plane incorporates the User Agent, a specialized application that interacts with the end user / operator and performs requests for Deterministic Networking services via an abstract Flow Management Entity, (FME) which may or may not be collocated with (one of) the end systems.

At the Application Plane, a management interface enables the negotiation of flows between end systems. An abstraction of the flow called a Traffic Specification (TSpec) provides the representation. This abstraction is used to place a reservation over the (Northbound) Service Interface and within the Application plane. It is associated with an abstraction of location, such as IP addresses and DNS names, to identify the end systems and eventually specify intermediate nodes.

4.4.2. The Controller Plane

The Controller Plane corresponds to the aggregation of the Control and Management Planes in [RFC7426], though Common Control and Measurement Plane (CCAMP) [CCAMP] makes an additional distinction between management and measurement. When the logical separation of the Control, Measurement and other Management entities is not relevant, the term Controller Plane is used for simplicity to represent them all, and the term controller refers to any device operating in that plane, whether is it a Path Computation entity or a Network Management entity (NME). The Path Computation Element (PCE) [PCE] is a core element of a controller, in charge of computing Deterministic paths to be applied in the Network Plane.

A (Northbound) Service Interface enables applications in the Application Plane to communicate with the entities in the Controller Plane.

One or more PCE(s) collaborate to implement the requests from the FME as Per-Flow Per-Hop Behaviors installed in the intermediate nodes for each individual flow. The PCEs place each flow along a deterministic sequence of intermediate nodes so as to respect per-flow constraints such as security and latency, and optimize the overall result for metrics such as an abstract aggregated cost. The deterministic sequence can typically be more complex than a direct sequence and include redundancy path, with one or more packet replication and elimination points.

4.4.3. The Network Plane

The Network Plane represents the network devices and protocols as a whole, regardless of the Layer at which the network devices operate. It includes Forwarding Plane (data plane), Application, and Operational Plane (control plane) aspects.

The network Plane comprises the Network Interface Cards (NIC) in the end systems, which are typically IP hosts, and intermediate nodes, which are typically IP routers and switches. Network-to-Network Interfaces such as used for Traffic Engineering path reservation in [RFC5921], as well as User-to-Network Interfaces (UNI) such as provided by the Local Management Interface (LMI) between network and end systems, are both part of the Network Plane, both in the control plane and the data plane.

A Southbound (Network) Interface enables the entities in the Controller Plane to communicate with devices in the Network Plane. This interface leverages and extends TEAS to describe the physical topology and resources in the Network Plane.

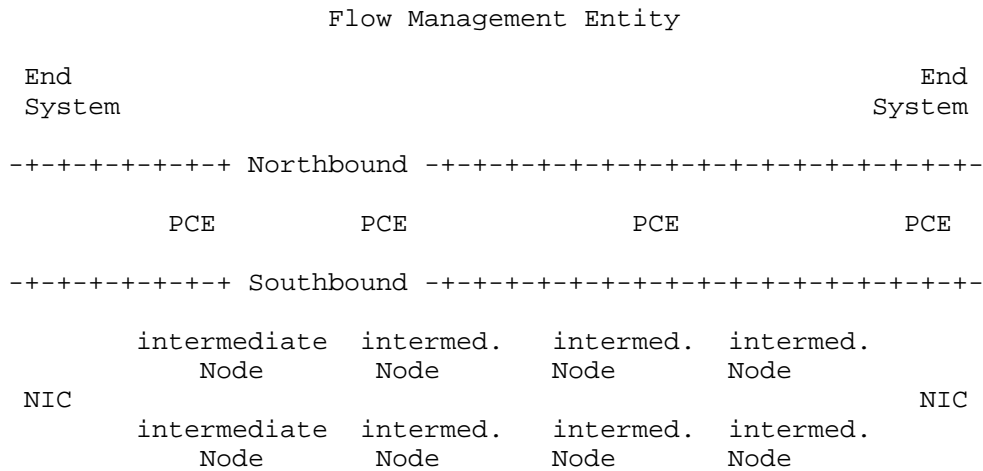


Figure 7

The intermediate nodes (and eventually the end systems NIC) expose their capabilities and physical resources to the controller (the PCE), and update the PCE with their dynamic perception of the topology, across the Southbound Interface. In return, the PCE(s) set the per-flow paths up, providing a Flow Characterization that is more tightly coupled to the intermediate node Operation than a TSpec.

At the Network plane, intermediate nodes may exchange information regarding the state of the paths, between adjacent systems and eventually with the end systems, and forward packets within constraints associated to each flow, or, when unable to do so, perform a last resort operation such as drop or declassify.

This specification focuses on the Southbound interface and the operation of the Network Plane.

4.5. Queuing, Shaping, Scheduling, and Preemption

DetNet achieves congestion protection and bounded delivery latency by reserving bandwidth and buffer resources at every hop along the path of the DetNet flow. The reservation itself is not sufficient, however. Implementors and users of a number of proprietary and standard real-time networks have found that standards for specific data plane techniques are required to enable these assurances to be made in a multi-vendor network. The fundamental reason is that latency variation in one system results in the need for extra buffer space in the next-hop system(s), which in turn, increases the worst-case per-hop latency.

Standard queuing and transmission selection algorithms allow a central controller to compute the latency contribution of each transit node to the end-to-end latency, to compute the amount of buffer space required in each transit node for each incremental DetNet flow, and most importantly, to translate from a flow specification to a set of values for the managed objects that control each relay or end system. The IEEE 802 has specified (and is specifying) a set of queuing, shaping, and scheduling algorithms that enable each transit node (bridge or router), and/or a central controller, to compute these values. These algorithms include:

- o A credit-based shaper [IEEE802.1Q-2014] Clause 34.
- o Time-gated queues governed by a rotating time schedule, synchronized among all transit nodes [IEEE802.1Qbv].
- o Synchronized double (or triple) buffers driven by synchronized time ticks. [IEEE802.1Qch].
- o Pre-emption of an Ethernet packet in transmission by a packet with a more stringent latency requirement, followed by the resumption of the preempted packet [IEEE802.1Qbu], [IEEE802.3br].

While these techniques are currently embedded in Ethernet and bridging standards, we can note that they are all, except perhaps for packet preemption, equally applicable to other media than Ethernet, and to routers as well as bridges.

4.6. Service instance

[Note: Service instance represents all the functions required on a node to allow the end-to-end service between the UNIs.]

The DetNet network reference model is shown in Figure 8 for a DetNet-Service scenario (i.e. between two DetNet-UNIs). In this figure, the end systems ("A" and "B") are connected directly to the edge nodes of the IP/MPLS network ("PE1" and "PE2"). End-systems participating DetNet communication may require connectivity before setting up an App-flow that requires the DetNet service. Such a connectivity related service instance and the one dedicated for DetNet service share the same access. Packets belonging to a DetNet flow are selected by a filter configured on the access ("F1" and "F2"). As a result, data flow specific access ("access-A + F1" and "access-B + F2") are terminated in the flow specific service instance ("SI-1" and "SI-2"). A tunnel is used to provide connectivity between the service instances.

The tunnel is used to transport exclusively the packets of the DetNet flow between "SI-1" and "SI-2". The service instances are configured to implement DetNet functions and a flow specific routing or bridging function depending on what connectivity the participating end systems require (L3 or L2). The service instance and the tunnel may or may not be shared by multiple DetNet flows. Sharing the service instance by multiple DetNet flows requires properly populated forwarding tables of the service instance.

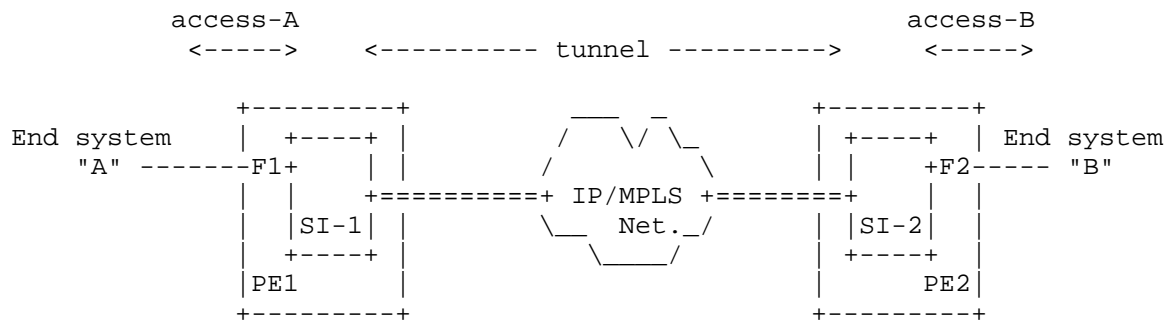


Figure 8: DetNet network reference model

[Note: The tunnel between the service instances may have some special characteristics. For example, in case of a "packet PW" based tunnel, there are differences in the usage of the packet PW for DetNet traffic compared to the network model described in [RFC6658]. In the DetNet scenario, the packet PW is used exclusively by the DetNet flow, whereas [RFC6658] states: "The packet PW appears as a single point-to-point link to the client layer. Network-layer adjacency formation and maintenance between the client equipments will follow the normal practice needed to support the required relationship in the client layer ... This packet pseudowire is used to transport all of the required layer 2 and layer 3 protocols between LSR1 and LSR2".]

[Note: Examples are provided in Annex 1 of [I-D.varga-detnet-service-model].]

4.7. Flow identification at technology borders

4.7.1. Exporting flow identification

An interesting feature of DetNet, and one that invites implementations that can be accused of "layering violations", is the need for lower layers to be aware of specific flows at higher layers,

in order to provide specific queuing and shaping services for specific flows. For example:

- o A non-IP, strictly L2 source end system X may be sending multiple flows to the same L2 destination end system Y. Those flows may include DetNet flows with different QoS requirements, and may include non-DetNet flows.
- o A router may be sending any number of flows to another router. Again, those flows may include DetNet flows with different QoS requirements, and may include non-DetNet flows.
- o Two routers may be separated by bridges. For these bridges to perform any required per-flow queuing and shaping, they must be able to identify the individual flows.
- o A Label Edge Router (LERs) may have a Label Switched Path (LSP) set up for handling traffic destined for a particular IP address carrying only non-DetNet flows. If a DetNet flow to that same address is requested, a separate LSP may be needed, in order that all of the Label Switch Routers (LSRs) along the path to the destination give that flow special queuing and shaping.

The need for a lower-level DetNet node to be aware of individual higher-layer flows is not unique to DetNet. But, given the endless complexity of layering and relayering over tunnels that is available to network designers, DetNet needs to provide a model for flow identification that is at least somewhat better than packet inspection. That is not to say that packet inspection to layer 4 or 5 addresses will not be used, or the capability standardized; but, there are alternatives.

A DetNet relay node can connect DetNet flows on different paths using different flow identification methods. For example:

- o A single unicast DetNet flow passing from router A through a bridged network to router B may be assigned a {VLAN, multicast destination MAC address} pair that is unique within that bridged network. The bridges can then identify the flow without accessing higher-layer headers. Of course, the receiving router must recognize and accept that multicast MAC address.
- o A DetNet flow passing from LSR A to LSR B may be assigned a different label than that used for other flows to the same IP destination.

In any of the above cases, it is possible that an existing DetNet flow can be used as a carrier for multiple DetNet sub-flows. (Not to

be confused with DetNet compound vs. member flows.) Of course, this requires that the aggregate DetNet flow be provisioned properly to carry the sub-flows.

Thus, rather than packet inspection, there is the option to export higher-layer information to the lower layer. The requirement to support one or the other method for flow identification (or both) is the essential complexity that DetNet brings to existing control plane models.

4.7.2. Flow attribute mapping between layers

Transport of DetNet flows over multiple technology domains may require that lower layers are aware of specific flows of higher layers. Such an "exporting of flow identification" is needed each time when the forwarding paradigm is changed on the transport path (e.g., two LSRs are interconnected by a L2 bridged domain, etc.). The three main forwarding methods considered for deterministic networking are:

- o IP routing
- o MPLS label switching
- o Ethernet bridging

Note: at the time of this publication, the exact format of flow identification is still WIP.

[Note: Seq-num attribute may require a similar functionality at technology border nodes.]

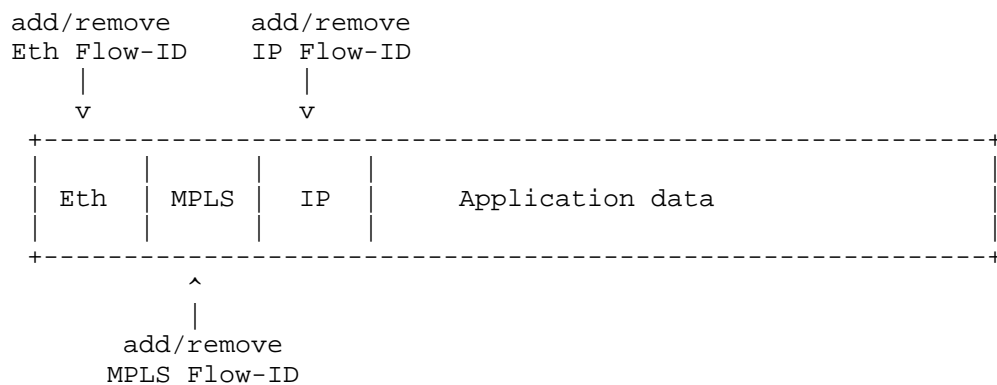


Figure 9: Packet with multiple Flow-IDs

The additional (domain specific) Flow-ID can be

- o created by a domain specific function or
- o derived from the Flow-ID added to the App-flow,

so that it must be unique inside the given domain. Note, that the Flow-ID added to the App-flow is still present in the packet, but transport nodes may lack the function to recognize it; that's why the additional Flow-ID is added (pushed).

4.7.3. Flow-ID mapping examples

IP nodes and MPLS nodes are assumed to be configured to push such an additional (domain specific) Flow-ID when sending traffic to an Ethernet switch (as shown in the examples below).

Figure 10 shows a scenario where an IP end system ("IP-A") is connected via two Ethernet switches ("ETH-n") to an IP router ("IP-1").

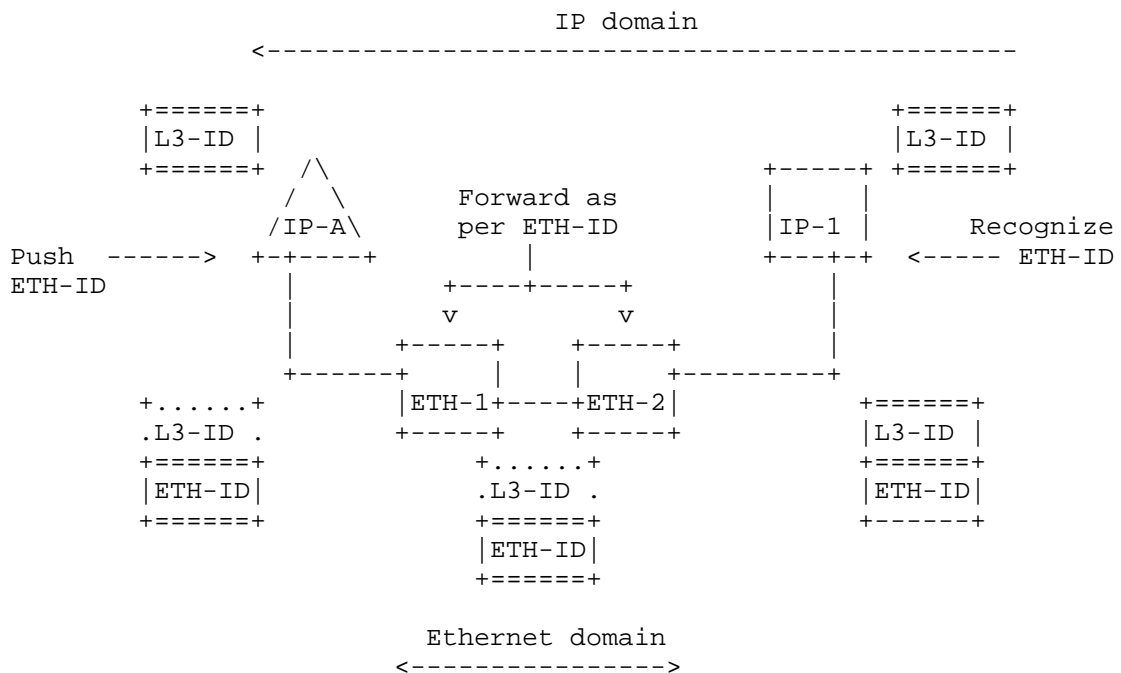


Figure 10: IP nodes interconnected by an Ethernet domain

End system "IP-A" uses the original App-flow specific ID ("L3-ID"), but as it is connected to an Ethernet domain it has to push an Ethernet-domain specific flow-ID ("VID + multicast MAC address", referred as "ETH-ID") before sending the packet to "ETH-1" node. Ethernet switch "ETH-1" can recognize the data flow based on the "ETH-ID" and it does forwarding toward "ETH-2". "ETH-2" switches the packet toward the IP router. "IP-1" must be configured to receive the Ethernet Flow-ID specific multicast stream, but (as it is an L3 node) it decodes the data flow ID based on the "L3-ID" fields of the received packet.

Figure 11 shows a scenario where MPLS domain nodes ("PE-n" and "P-m") are connected via two Ethernet switches ("ETH-n").

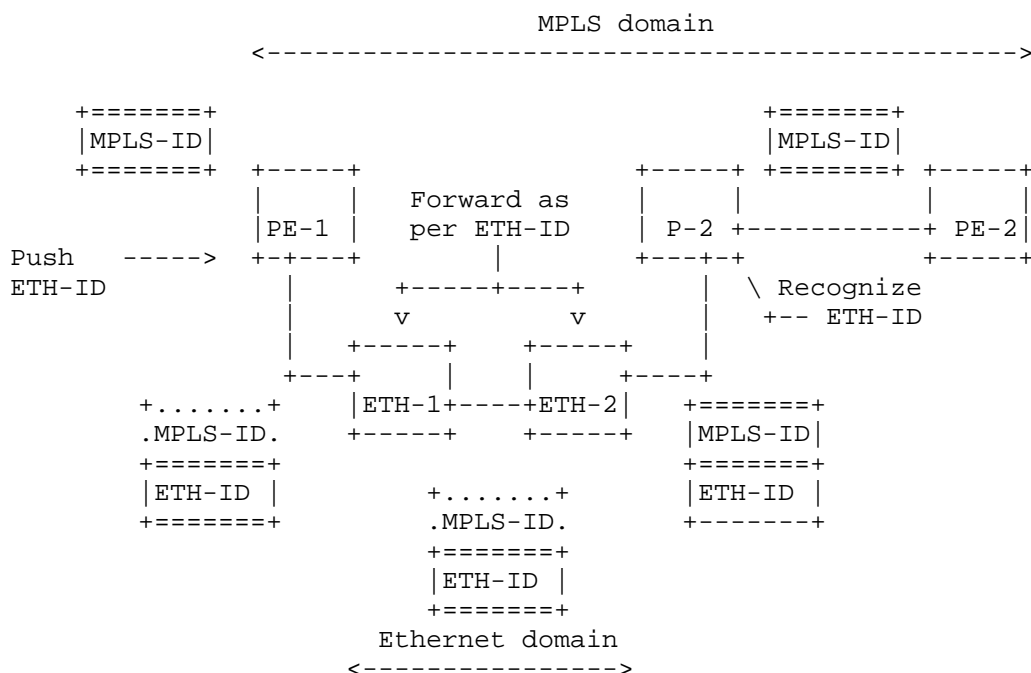


Figure 11: MPLS nodes interconnected by an Ethernet domain

"PE-1" uses the MPLS specific ID ("MPLS-ID"), but as it is connected to an Ethernet domain it has to push an Ethernet-domain specific flow-ID ("VID + multicast MAC address", referred as "ETH-ID") before sending the packet to "ETH-1". Ethernet switch "ETH-1" can recognize the data flow based on the "ETH-ID" and it does forwarding toward "ETH-2". "ETH-2" switches the packet toward the MPLS node ("P-2"). "P-2" must be configured to receive the Ethernet Flow-ID specific

multicast stream, but (as it is an MPLS node) it decodes the data flow ID based on the "MPLS-ID" fields of the received packet.

4.8. Advertising resources, capabilities and adjacencies

There are three classes of information that a central controller or decentralized control plane needs to know that can only be obtained from the end systems and/or transit nodes in the network. When using a peer-to-peer control plane, some of this information may be required by a system's neighbors in the network.

- o Details of the system's capabilities that are required in order to accurately allocate that system's resources, as well as other systems' resources. This includes, for example, which specific queuing and shaping algorithms are implemented (Section 4.5), the number of buffers dedicated for DetNet allocation, and the worst-case forwarding delay.
- o The dynamic state of an end or transit node's DetNet resources.
- o The identity of the system's neighbors, and the characteristics of the link(s) between the systems, including the length (in nanoseconds) of the link(s).

4.9. Provisioning model

4.9.1. Centralized Path Computation and Installation

A centralized routing model, such as provided with a PCE (RFC 4655 [RFC4655]), enables global and per-flow optimizations. (See Section 4.4.) The model is attractive but a number of issues are left to be solved. In particular:

- o Whether and how the path computation can be installed by 1) an end device or 2) a Network Management entity,
- o And how the path is set up, either by installing state at each hop with a direct interaction between the forwarding device and the PCE, or along a path by injecting a source-routed request at one end of the path.

4.9.2. Distributed Path Setup

Significant work on distributed path setup can be leveraged from MPLS Traffic Engineering, in both its GMPLS and non-GMPLS forms. The protocols within scope are Resource ReSerVation Protocol [RFC3209] [RFC3473](RSVP-TE), OSPF-TE [RFC4203] [RFC5392] and ISIS-TE [RFC5307]

[RFC5316]. These should be viewed as starting points as there are feature specific extensions defined that may be applicable to DetNet.

In a Layer-2 only environment, or as part of a layered approach to a mixed environment, IEEE 802.1 also has work, either completed or in progress. [IEEE802.1Q-2014] Clause 35 describes SRP, a peer-to-peer protocol for Layer-2 roughly analogous to RSVP [RFC2205]. [IEEE802.1Qca] defines how ISIS can provide multiple disjoint paths or distribution trees. Also in progress is [IEEE802.1Qcc], which expands the capabilities of SRP.

The integration/interaction of the DetNet control layer with an underlying IEEE 802.1 sub-network control layer will need to be defined.

4.10. Scaling to larger networks

Reservations for individual DetNet flows require considerable state information in each transit node, especially when adequate fault mitigation (Section 3.3.2) is required. The DetNet data plane, in order to support larger numbers of DetNet flows, must support the aggregation of DetNet flows into tunnels, which themselves can be viewed by the transit nodes' data planes largely as individual DetNet flows. Without such aggregation, the per-relay system may limit the scale of DetNet networks.

4.11. Connected islands vs. networks

Given that users have deployed examples of the IEEE 802.1 TSN TG standards, which provide capabilities similar to DetNet, it is obvious to ask whether the IETF DetNet effort can be limited to providing Layer-2 connections (VPNs) between islands of bridged TSN networks. While this capability is certainly useful to some applications, and must not be precluded by DetNet, tunneling alone is not a sufficient goal for the DetNet WG. As shown in the Deterministic Networking Use Cases draft [I-D.ietf-detnet-use-cases], there are already deployments of Layer-2 TSN networks that are encountering the well-known problems of over-large broadcast domains. Routed solutions, and combinations routed/bridged solutions, are both required.

4.12. Compatibility with Layer-2

Standards providing similar capabilities for bridged networks (only) have been and are being generated in the IEEE 802 LAN/MAN Standards Committee. The present architecture describes an abstract model that can be applicable both at Layer-2 and Layer-3, and over links not defined by IEEE 802. It is the intention of the authors (and

hopefully, as this draft progresses, of the DetNet Working Group) that IETF and IEEE 802 will coordinate their work, via the participation of common individuals, liaisons, and other means, to maximize the compatibility of their outputs.

DetNet enabled end systems and intermediate nodes can be interconnected by sub-networks, i.e., Layer-2 technologies. These sub-networks will provide DetNet compatible service for support of DetNet traffic. Examples of sub-networks include 802.1TSN and a point-to-point OTN link. Of course, multi-layer DetNet systems may be possible too, where one DetNet appears as a sub-network, and provides service to, a higher layer DetNet system.

5. Open Questions

There are a number of architectural questions that will have to be resolved before this document can be submitted for publication. Aside from the obvious fact that this present draft is subject to change, there are specific questions to which the authors wish to direct the readers' attention.

5.1. Flat vs. hierarchical control

Boxes that are solely routers or solely bridges are rare in today's market. In a multi-tenant data center, multiple users' virtual Layer-2/Layer-3 topologies exist simultaneously, implemented on a network whose physical topology bears only accidental resemblance to the virtual topologies.

While the forwarding topology (the bridges and routers) are an important consideration for a DetNet Flow Management Entity (Section 4.4.1), so is the purely physical topology. Ultimately, the model used by the management entities is based on boxes, queues, and links. The authors hope that the work of the TEAS WG will help to clarify exactly what model parameters need to be traded between the intermediate nodes and the controller(s).

5.2. Peer-to-peer reservation protocol

As described in Section 4.9.2, the DetNet WG needs to decide whether to support a peer-to-peer protocol for a source and a destination to reserve resources for a DetNet stream. Assuming that enabling the involvement of the source and/or destination is desirable (see Deterministic Networking Use Cases [I-D.ietf-detnet-use-cases]), it remains to decide whether the DetNet WG will make it possible to deploy at least some DetNet capabilities in a network using only a peer-to-peer protocol, without a central controller.

(Note that a UNI (see Section 4.4.3) between an end system and a DetNet edge node, for sources and/or listeners to request DetNet services, can be either the first hop of a per-to-peer reservation protocol, or can be deflected by the DetNet edge node to a central controller for resolution. Similarly, a decision by a central controller can be effected by the controller instructing the end system or DetNet edge node to initiate a per-to-peer protocol activity.)

5.3. Wireless media interactions

Deterministic Networking Use Cases [I-D.ietf-detnet-use-cases] illustrates cases where wireless media are needed in a DetNet network. Some wireless media in general use, such as IEEE 802.11 [IEEE802.1Q-2014], have significantly higher packet loss rates than typical wired media, such as Ethernet [IEEE802.3-2012]. IEEE 802.11 includes support for such features as MAC-layer acknowledgements and retransmissions.

The techniques described in Section 3 are likely to improve the ability of a mixed wired/wireless network to offer the DetNet QoS features. The interaction of these techniques with the features of specific wireless media, although they may be significant, cannot be addressed in this document. It remains to be decided to what extent the DetNet WG will address them, and to what extent other WGs, e.g. 6TiSCH, will do so.

5.4. Packet encoding for service protection

There are methods for using multiple paths to provide service protection that involve encoding the information in a packet belonging to a DetNet flow into multiple transmission units, typically combining information from multiple packets into any given transmission unit. Such techniques may be applicable for use as a DetNet service protection technique, assuming that the DetNet users' needs for timeliness of delivery and freedom from interference with misbehaving DetNet flows can be met.

No specific mechanisms are defined here, at this time. This section will either be enhanced or removed. Contributions are invited.

6. Security Considerations

Security in the context of Deterministic Networking has an added dimension; the time of delivery of a packet can be just as important as the contents of the packet, itself. A man-in-the-middle attack, for example, can impose, and then systematically adjust, additional delays into a link, and thus disrupt or subvert a real-time

application without having to crack any encryption methods employed. See [RFC7384] for an exploration of this issue in a related context.

Furthermore, in a control system where millions of dollars of equipment, or even human lives, can be lost if the DetNet QoS is not delivered, one must consider not only simple equipment failures, where the box or wire instantly becomes perfectly silent, but bizarre errors such as can be caused by software failures. Because there is essential no limit to the kinds of failures that can occur, protecting against realistic equipment failures is indistinguishable, in most cases, from protecting against malicious behavior, whether accidental or intentional. See also Section 3.3.2.

Security must cover:

- o the protection of the signaling protocol
- o the authentication and authorization of the controlling systems
- o the identification and shaping of the DetNet flows

7. Privacy Considerations

DetNet is provides a Quality of Service (QoS), and as such, does not directly raise any new privacy considerations.

However, the requirement for every (or almost every) node along the path of a DetNet flow to identify DetNet flows may present an additional attack surface for privacy, should the DetNet paradigm be found useful in broader environments.

8. IANA Considerations

This document does not require an action from IANA.

9. Acknowledgements

The authors wish to thank Jouni Korhonen, Erik Nordmark, George Swallow, Rudy Klecka, Anca Zamfir, David Black, Thomas Watteyne, Shitanshu Shah, Craig Gunther, Rodney Cummings, Balazs Varga, Wilfried Steiner, Marcel Kiessling, Karl Weber, Janos Farkas, Ethan Grossman, Pat Thaler, Lou Berger, and especially Michael Johas Teener, for their various contribution with this work.

10. Access to IEEE 802.1 documents

To access password protected IEEE 802.1 drafts, see the IETF IEEE 802.1 information page at <https://www.ietf.org/proceedings/52/slides/bridge-0/tsld003.htm>.

11. Informative References

- [AVnu] <http://www.avnu.org/>, "The AVnu Alliance tests and certifies devices for interoperability, providing a simple and reliable networking solution for AV network implementation based on the Audio Video Bridging (AVB) standards.".
- [CCAMP] IETF, "Common Control and Measurement Plane",
<<https://datatracker.ietf.org/doc/charter-ietf-ccamp/>>.
- [HART] www.hartcomm.org, "Highway Addressable Remote Transducer, a group of specifications for industrial process and control devices administered by the HART Foundation".
- [HSR-PRP] IEC, "High availability seamless redundancy (HSR) is a further development of the PRP approach, although HSR functions primarily as a protocol for creating media redundancy while PRP, as described in the previous section, creates network redundancy. PRP and HSR are both described in the IEC 62439 3 standard.",
<<http://webstore.iec.ch/webstore/webstore.nsf/artnum/046615!opendocument>>.
- [I-D.dt-detnet-dp-alt] Korhonen, J., Farkas, J., Mirsky, G., Thubert, P., Zhuangyan, Z., and L. Berger, "DetNet Data Plane Protocol and Solution Alternatives", draft-dt-detnet-dp-alt-04 (work in progress), September 2016.
- [I-D.ietf-6tisch-architecture] Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-11 (work in progress), January 2017.
- [I-D.ietf-6tisch-tsch] Watteyne, T., Palattella, M., and L. Grieco, "Using IEEE802.15.4e TSCH in an IoT context: Overview, Problem Statement and Goals", draft-ietf-6tisch-tsch-06 (work in progress), March 2015.

- [I-D.ietf-detnet-problem-statement]
Finn, N. and P. Thubert, "Deterministic Networking Problem Statement", draft-ietf-detnet-problem-statement-01 (work in progress), September 2016.
- [I-D.ietf-detnet-use-cases]
Grossman, E., Gunther, C., Thubert, P., Wetterwald, P., Raymond, J., Korhonen, J., Kaneko, Y., Das, S., Zha, Y., Varga, B., Farkas, J., Goetz, F., Schmitt, J., Vilajosana, X., Mahmoodi, T., Spirou, S., and P. Vizarreta, "Deterministic Networking Use Cases", draft-ietf-detnet-use-cases-12 (work in progress), April 2017.
- [I-D.ietf-roll-rpl-industrial-applicability]
Phinney, T., Thubert, P., and R. Assimiti, "RPL applicability in industrial networks", draft-ietf-roll-rpl-industrial-applicability-02 (work in progress), October 2013.
- [I-D.svshah-tsvwg-deterministic-forwarding]
Shah, S. and P. Thubert, "Deterministic Forwarding PHB", draft-svshah-tsvwg-deterministic-forwarding-04 (work in progress), August 2015.
- [I-D.varga-detnet-service-model]
Varga, B. and J. Farkas, "DetNet Service Model", draft-varga-detnet-service-model-02 (work in progress), May 2017.
- [IEEE802.11-2012]
IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", 2012, <<http://standards.ieee.org/getieee802/download/802.11-2012.pdf>>.
- [IEEE802.1AS-2011]
IEEE, "Timing and Synchronizations (IEEE 802.1AS-2011)", 2011, <<http://standards.ieee.org/getIEEE802/download/802.1AS-2011.pdf>>.
- [IEEE802.1BA-2011]
IEEE, "AVB (Audio Video Bridging) Systems (IEEE 802.1BA-2011)", 2011, <<http://standards.ieee.org/getIEEE802/download/802.1BA-2011.pdf>>.

- [IEEE802.1CB]
IEEE, "Frame Replication and Elimination for Reliability (IEEE Draft P802.1CB)", 2016,
<<http://www.ieee802.org/1/files/private/cb-drafts/>>.
- [IEEE802.1Q-2014]
IEEE, "MAC Bridges and VLANs (IEEE 802.1Q-2014", 2014,
<<http://standards.ieee.org/getieee802/download/802-1Q-2014.pdf>>.
- [IEEE802.1Qbu]
IEEE, "Frame Preemption", 2016,
<<http://www.ieee802.org/1/files/private/bu-drafts/>>.
- [IEEE802.1Qbv]
IEEE, "Enhancements for Scheduled Traffic", 2016,
<<http://www.ieee802.org/1/files/private/bv-drafts/>>.
- [IEEE802.1Qca]
IEEE 802.1, "IEEE 802.1Qca Bridges and Bridged Networks - Amendment 24: Path Control and Reservation", IEEE P802.1Qca/D2.1 P802.1Qca, June 2015,
<<https://standards.ieee.org/findstds/standard/802.1Qca-2015.html>>.
- [IEEE802.1Qcc]
IEEE, "Stream Reservation Protocol (SRP) Enhancements and Performance Improvements", 2016,
<<http://www.ieee802.org/1/files/private/cc-drafts/>>.
- [IEEE802.1Qch]
IEEE, "Cyclic Queuing and Forwarding", 2016,
<<http://www.ieee802.org/1/files/private/ch-drafts/>>.
- [IEEE802.1TSNTG]
IEEE Standards Association, "IEEE 802.1 Time-Sensitive Networks Task Group", 2013,
<<http://www.IEEE802.org/1/pages/avbridges.html>>.
- [IEEE802.3-2012]
IEEE, "IEEE Standard for Ethernet", 2012,
<<http://standards.ieee.org/getieee802/download/802.3-2012.pdf>>.
- [IEEE802.3br]
IEEE, "Interspersed Express Traffic", 2016,
<<http://www.ieee802.org/3/br/>>.

- [IEEE802154]
IEEE Standard for Information Technology, "IEEE 802.15.4, Part. 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks", June 2011.
- [IEEE802154e]
IEEE Standard for Information Technology, "IEEE 802.15.4e, Part. 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 1: MAC sublayer", April 2012.
- [ISA100.11a]
ISA/IEC, "ISA100.11a, Wireless Systems for Automation, also IEC 62734", 2011, < <http://www.isa100wci.org/en-US/Documents/PDF/3405-ISA100-WirelessSystems-Future-broch-WEB-ETSI.aspx>>.
- [ISA95]
ANSI/ISA, "Enterprise-Control System Integration Part 1: Models and Terminology", 2000, <<https://www.isa.org/isa95/>>.
- [ODVA]
<http://www.odva.org/>, "The organization that supports network technologies built on the Common Industrial Protocol (CIP) including EtherNet/IP."
- [PCE]
IETF, "Path Computation Element", <<https://datatracker.ietf.org/doc/charter-ietf-pce/>>.
- [Profinet]
<http://us.profinet.com/technology/profinet/>, "PROFINET is a standard for industrial networking in automation.", <<http://us.profinet.com/technology/profinet/>>.
- [RFC2205]
Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, DOI 10.17487/RFC2205, September 1997, <<http://www.rfc-editor.org/info/rfc2205>>.
- [RFC2475]
Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, DOI 10.17487/RFC2475, December 1998, <<http://www.rfc-editor.org/info/rfc2475>>.
- [RFC3209]
Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<http://www.rfc-editor.org/info/rfc3209>>.

- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <<http://www.rfc-editor.org/info/rfc3473>>.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, DOI 10.17487/RFC3550, July 2003, <<http://www.rfc-editor.org/info/rfc3550>>.
- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<http://www.rfc-editor.org/info/rfc4203>>.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <<http://www.rfc-editor.org/info/rfc4655>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <<http://www.rfc-editor.org/info/rfc5307>>.
- [RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, DOI 10.17487/RFC5316, December 2008, <<http://www.rfc-editor.org/info/rfc5316>>.
- [RFC5392] Chen, M., Zhang, R., and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5392, DOI 10.17487/RFC5392, January 2009, <<http://www.rfc-editor.org/info/rfc5392>>.
- [RFC5673] Pister, K., Ed., Thubert, P., Ed., Dwars, S., and T. Phinney, "Industrial Routing Requirements in Low-Power and Lossy Networks", RFC 5673, DOI 10.17487/RFC5673, October 2009, <<http://www.rfc-editor.org/info/rfc5673>>.
- [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, <<http://www.rfc-editor.org/info/rfc5921>>.

- [RFC6372] Sprecher, N., Ed. and A. Farrel, Ed., "MPLS Transport Profile (MPLS-TP) Survivability Framework", RFC 6372, DOI 10.17487/RFC6372, September 2011, <<http://www.rfc-editor.org/info/rfc6372>>.
- [RFC6658] Bryant, S., Ed., Martini, L., Swallow, G., and A. Malis, "Packet Pseudowire Encapsulation over an MPLS PSN", RFC 6658, DOI 10.17487/RFC6658, July 2012, <<http://www.rfc-editor.org/info/rfc6658>>.
- [RFC7384] Mizrahi, T., "Security Requirements of Time Protocols in Packet Switched Networks", RFC 7384, DOI 10.17487/RFC7384, October 2014, <<http://www.rfc-editor.org/info/rfc7384>>.
- [RFC7426] Haleplidis, E., Ed., Pentikousis, K., Ed., Denazis, S., Hadi Salim, J., Meyer, D., and O. Koufopavlou, "Software-Defined Networking (SDN): Layers and Architecture Terminology", RFC 7426, DOI 10.17487/RFC7426, January 2015, <<http://www.rfc-editor.org/info/rfc7426>>.
- [TEAS] IETF, "Traffic Engineering Architecture and Signaling", <<https://datatracker.ietf.org/doc/charter-ietf-teas/>>.
- [WirelessHART]
www.hartcomm.org, "Industrial Communication Networks - Wireless Communication Network and Communication Profiles - WirelessHART - IEC 62591", 2010.

Authors' Addresses

Norman Finn
Huawei Technologies Co. Ltd
3755 Avocado Blvd.
PMB 436
La Mesa, California 91941
US

Phone: +1 925 980 6430
Email: norman.finn@mail01.huawei.com

Pascal Thubert
Cisco Systems
Village d'Entreprises Green Side
400, Avenue de Roumanille
Batiment T3
Biot - Sophia Antipolis 06410
FRANCE

Phone: +33 4 97 23 26 34
Email: pthubert@cisco.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

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 3, 2018

T. Mizrahi
MARVELL
E. Grossman, Ed.
DOLBY
A. Hacker
MISTIQ
S. Das
Applied Communication Sciences
J. Dowdell
Airbus Defence and Space
H. Austad
Cisco Systems
K. Stanton
INTEL
N. Finn
HUAWEI
July 2, 2017

Deterministic Networking (DetNet) Security Considerations
draft-sdt-detnet-security-01

Abstract

A deterministic network is one that can carry data flows for real-time applications with extremely low data loss rates and bounded latency. Deterministic networks have been successfully deployed in real-time operational technology (OT) applications for some years (for example [ARINC664P7]). However, such networks are typically isolated from external access, and thus the security threat from external attackers is low. IETF Deterministic Networking (DetNet) specifies a set of technologies that enable creation of deterministic networks on IP-based networks of potentially wide area (on the scale of a corporate network) potentially bringing the OT network into contact with Information Technology (IT) traffic and security threats that lie outside of a tightly controlled and bounded area (such as the internals of an aircraft). These DetNet technologies have not previously been deployed together on a wide area IP-based network, and thus can present security considerations that may be new to IP-based wide area network designers. This draft, intended for use by DetNet network designers, provides insight into these security considerations. In addition, this draft collects all security-related statements from the various DetNet drafts (Architecture, Use Cases, etc) into a single location Section 7.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 3, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	4
2. Abbreviations	5
3. Security Threats	6
3.1. Threat Model	6
3.2. Threat Analysis	7
3.2.1. Delay	7
3.2.1.1. Delay Attack	7
3.2.2. DetNet Flow Identification	7
3.2.2.1. DetNet Flow Modification or Spoofing	7
3.2.3. Resource Segmentation or Slicing	7
3.2.3.1. Inter-segment Attack	7
3.2.4. Packet Replication and Elimination	7
3.2.4.1. Replication: Increased Attack Surface	8
3.2.4.2. Replication-related Header Manipulation	8

3.2.5.	Path Choice	8
3.2.5.1.	Path Manipulation	8
3.2.5.2.	Path Choice: Increased Attack Surface	8
3.2.6.	Control Plane	9
3.2.6.1.	Control or Signaling Packet Modification	9
3.2.6.2.	Control or Signaling Packet Injection	9
3.2.7.	Scheduling or Shaping	9
3.2.7.1.	Reconnaissance	9
3.2.8.	Time Synchronization Mechanisms	9
3.3.	Threat Summary	9
4.	Security Threat Impacts	10
4.1.	Delay-Attacks	10
4.1.1.	Data Plane Delay Attacks	11
4.1.2.	Control Plane Delay Attacks	11
4.2.	Flow Identification and Spoofing	11
4.2.1.	Flow identification	11
4.2.2.	Spoofing	12
4.2.2.1.	Dataplane Spoofing	12
4.2.2.2.	Control Plane Spoofing	12
4.3.	Segmentation attacks (injection)	12
4.3.1.	Data Plane Segmentation	12
4.3.2.	Control Plane segmentation	13
4.4.	Replication and Elimination	13
4.4.1.	Increased Attack Surface	13
4.4.2.	Header Manipulation at Elimination Bridges	13
4.5.	Impact of Attacks to Path Choice	13
4.6.	Impact of Attacks by Use Case Industry	13
5.	Security Threat Mitigation	15
5.1.	Path Redundancy	16
5.2.	Integrity Protection	16
5.3.	DetNet Node Authentication	16
5.4.	Encryption	17
5.5.	Control and Signaling Message Protection	17
5.6.	Dynamic Performance Analytics	17
5.7.	Mitigation Summary	18
6.	Association of Attacks to Use Cases	19
6.1.	Use Cases by Common Themes	19
6.1.1.	Network Layer - AVB/TSN Ethernet	19
6.1.2.	Central Administration	19
6.1.3.	Hot Swap	20
6.1.4.	Data Flow Information Models	20
6.1.5.	L2 and L3 Integration	20
6.1.6.	End-to-End Delivery	20
6.1.7.	Proprietary Deterministic Ethernet Networks	20
6.1.8.	Replacement for Proprietary Fieldbuses	20
6.1.9.	Deterministic vs Best-Effort Traffic	21
6.1.10.	Deterministic Flows	21
6.1.11.	Unused Reserved Bandwidth	21

6.1.12. Interoperability	21
6.1.13. Cost Reductions	21
6.1.14. Insufficiently Secure Devices	22
6.1.15. DetNet Network Size	22
6.1.16. Multiple Hops	22
6.1.17. Level of Service	22
6.1.18. Bounded Latency	23
6.1.19. Low Latency	23
6.1.20. Symmetrical Path Delays	23
6.1.21. Reliability and Availability	23
6.1.22. Redundant Paths	24
6.1.23. Security Measures	24
6.2. Attack Types by Use Case Common Theme	24
7. Appendix A: DetNet Draft Security-Related Statements	26
7.1. Architecture (draft 8)	27
7.1.1. Fault Mitigation (sec 4.5)	27
7.1.2. Security Considerations (sec 7)	27
7.2. Data Plane Alternatives (draft 4)	28
7.2.1. Security Considerations (sec 7)	28
7.3. Problem Statement (draft 5)	28
7.3.1. Security Considerations (sec 5)	28
7.4. Use Cases (draft 11)	29
7.4.1. (Utility Networks) Security Current Practices and Limitations (sec 3.2.1)	29
7.4.2. (Utility Networks) Security Trends in Utility Networks (sec 3.3.3)	30
7.4.3. (BAS) Security Considerations (sec 4.2.4)	32
7.4.4. (6TiSCH) Security Considerations (sec 5.3.3)	32
7.4.5. (Cellular radio) Security Considerations (sec 6.1.5)	32
7.4.6. (Industrial M2M) Communication Today (sec 7.2)	33
8. IANA Considerations	33
9. Security Considerations	33
10. Informative References	33
Authors' Addresses	34

1. Introduction

Security is of particularly high importance in DetNet networks because many of the use cases which are enabled by DetNet [I-D.ietf-detnet-use-cases] include control of physical devices (power grid components, industrial controls, building controls) which can have high operational costs for failure, and present potentially attractive targets for cyber-attackers.

This situation is even more acute given that one of the goals of DetNet is to provide a "converged network", i.e. one that includes both IT traffic and OT traffic, thus exposing potentially sensitive OT devices to attack in ways that were not previously common (usually

because they were under a separate control system or otherwise isolated from the IT network). Security considerations for OT networks is not a new area, and there are many OT networks today that are connected to wide area networks or the Internet; this draft focuses on the issues that are specific to the DetNet technologies and use cases.

The DetNet technologies include ways to:

- o Reserve data plane resources for DetNet flows in some or all of the intermediate nodes (e.g. bridges or routers) along the path of the flow
- o Provide explicit routes for DetNet flows that do not rapidly change with the network topology
- o Distribute data from DetNet flow packets over time and/or space to ensure delivery of each packet's data in spite of the loss of a path

This draft includes sections on threat modeling and analysis, threat impact and mitigation, and the association of various attacks with various use cases both by industry and based on the Use Case Common Themes section of the DetNet Use Cases draft [I-D.ietf-detnet-use-cases].

This draft also provides context for the DetNet security considerations by collecting into one place Section 7 the various remarks about security from the various DetNet drafts (Use Cases, Architecture, etc). This text is duplicated here primarily because the DetNet working group has elected not to produce a Requirements draft and thus collectively these statements are as close as we have to "DetNet Security Requirements".

2. Abbreviations

IT Information technology (the application of computers to store, study, retrieve, transmit, and manipulate data or information, often in the context of a business or other enterprise - Wikipedia).

OT Operational Technology (the hardware and software dedicated to detecting or causing changes in physical processes through direct monitoring and/or control of physical devices such as valves, pumps, etc. - Wikipedia)

MITM Man in the Middle

SN Sequence Number

STRIDE Addresses risk and severity associated with threat categories: Spoofing identity, Tampering with data, Repudiation, Information disclosure, Denial of service, Elevation of privilege.

DREAD Compares and prioritizes risk represented by these threat categories: Damage potential, Reproducibility, Exploitability, how many Affected users, Discoverability.

PTP Precision Time Protocol [IEEE1588]

3. Security Threats

This section presents a threat model, and analyzes the possible threats in a DetNet-enabled network.

We distinguish control plane threats from data plane threats. The attack surface may be the same, but the types of attacks are different. For example, a delay attack is more relevant to data plane than to control plane. There is also a difference in terms of security solutions: the way you secure the data plane is often different than the way you secure the control plane.

3.1. Threat Model

The threat model used in this memo is based on the threat model of Section 3.1 of [RFC7384]. This model classifies attackers based on two criteria:

- o Internal vs. external: internal attackers either have access to a trusted segment of the network or possess the encryption or authentication keys. External attackers, on the other hand, do not have the keys and have access only to the encrypted or authenticated traffic.
- o Man in the Middle (MITM) vs. packet injector: MITM attackers are located in a position that allows interception and modification of in-flight protocol packets, whereas a traffic injector can only attack by generating protocol packets.

DetNet-Service, one of the service scenarios described in [I-D.varga-detnet-service-model], is the case where a service connects DetNet networking islands, i.e. two or more otherwise independent DetNet network domains are connected via a link that is not intrinsically part of either network. This implies that there could be DetNet traffic flowing over a non-DetNet link, which may provide an attacker with an advantageous opportunity to tamper with DetNet traffic. The security properties of non-DetNet links are outside of the scope of DetNet Security, but it should be noted that

use of non-DetNet services to interconnect DetNet networks merits security analysis to ensure the integrity of the DetNet networks involved.

3.2. Threat Analysis

3.2.1. Delay

3.2.1.1. Delay Attack

An attacker can maliciously delay DetNet data flow traffic. By delaying the traffic, the attacker can compromise the service of applications that are sensitive to high delays or to high delay variation.

3.2.2. DetNet Flow Identification

3.2.2.1. DetNet Flow Modification or Spoofing

An attacker can modify some header fields often route packets in a way that causes the DetNet flow identification mechanisms to misclassify the flow. Alternatively, the attacker can inject traffic that is tailored to appear as if it belongs to a legitimate DetNet flow. The potential consequence is that the DetNet flow resource allocation cannot guarantee the performance that is expected when the flow identification works correctly.

Note that in some cases there may be an explicit DetNet header, but in some cases the flow identification may be based on fields from the L3/L4 headers. If L3/L4 headers are involved, for purposes of this draft we assume they are encrypted and/or integrity-protected from external attackers.

3.2.3. Resource Segmentation or Slicing

3.2.3.1. Inter-segment Attack

An attacker can inject traffic, consuming network device resources, thereby affecting DetNet flows. This can be performed using non-DetNet traffic that affects DetNet traffic, or by using DetNet traffic from one DetNet flow that affects traffic from different DetNet flows.

3.2.4. Packet Replication and Elimination

3.2.4.1. Replication: Increased Attack Surface

Redundancy is intended to increase the robustness and survivability of DetNet flows, and replication over multiple paths can potentially mitigate an attack that is limited to a single path. However, the fact that packets are replicated over multiple paths increases the attack surface of the network, i.e., there are more points in the network that may be subject to attacks.

3.2.4.2. Replication-related Header Manipulation

An attacker can manipulate the replication-related header fields (R-TAG). This capability opens the door for various types of attacks. For example:

- o Forward both replicas - malicious change of a packet SN (Sequence Number) can cause both replicas of the packet to be forwarded. Note that this attack has a similar outcome to a replay attack.
- o Eliminate both replicas - SN manipulation can be used to cause both replicas to be eliminated. In this case an attacker that has access to a single path can cause packets from other paths to be dropped, thus compromising some of the advantage of path redundancy.
- o Flow hijacking - an attacker can hijack a DetNet flow with access to a single path by systematically replacing the SNs on the given path with higher SN values. For example, an attacker can replace every SN value S with a higher value $S+C$, where C is a constant integer. Thus, the attacker creates a false illusion that the attacked path has the lowest delay, causing all packets from other paths to be eliminated. Once the flow is hijacked the attacker can either replace en route packets with malicious packets, or simply injecting errors, causing the packets to be dropped at their destination.

3.2.5. Path Choice

3.2.5.1. Path Manipulation

An attacker can maliciously change, add, or remove a path, thereby affecting the corresponding DetNet flows that use the path.

3.2.5.2. Path Choice: Increased Attack Surface

One of the possible consequences of a path manipulation attack is an increased attack surface. Thus, when the attack described in the

previous subsection is implemented, it may increase the potential of other attacks to be performed.

3.2.6. Control Plane

3.2.6.1. Control or Signaling Packet Modification

An attacker can maliciously modify en route control packets in order to disrupt or manipulate the DetNet path/resource allocation.

3.2.6.2. Control or Signaling Packet Injection

An attacker can maliciously inject control packets in order to disrupt or manipulate the DetNet path/resource allocation.

3.2.7. Scheduling or Shaping

3.2.7.1. Reconnaissance

A passive eavesdropper can gather information about en route DetNet flows, e.g., the number of DetNet flows, their bandwidths, and their schedules. The gathered information can later be used to invoke other attacks on some or all of the flows.

3.2.8. Time Synchronization Mechanisms

An attacker can use any of the attacks described in [RFC7384] to attack the synchronization protocol, thus affecting the DetNet service.

3.3. Threat Summary

A summary of the attacks that were discussed in this section is presented in Figure 1. For each attack, the table specifies the type of attackers that may invoke the attack. In the context of this summary, the distinction between internal and external attacks is under the assumption that a corresponding security mechanism is being used, and that the corresponding network equipment takes part in this mechanism.

Attack	Attacker Type			
	Internal MITM	Inj.	External MITM	Inj.
Delay attack	+		+	
DetNet Flow Modification or Spoofing	+	+		
Inter-segment Attack	+	+		
Replication: Increased Attack Surface	+	+	+	+
Replication-related Header Manipulation	+			
Path Manipulation	+	+		
Path Choice: Increased Attack Surface	+	+	+	+
Control or Signaling Packet Modification	+			
Control or Signaling Packet Injection		+		
Reconnaissance	+		+	
Attacks on Time Sync Mechanisms	+	+	+	+

Figure 1: Threat Analysis Summary

4. Security Threat Impacts

This section describes the impact of the attacks described in Section 3. Mitigations are discussed further in Section 5.

In computer security, the impact (or consequence) of an incident can be measured in loss of confidentiality, integrity or availability of information. In other words, this section describes the effect of a successful attack. The scope is limited to the effect of a successful attack on DetNet itself, not the applications that _use_ Detnet as this is highly application specific.

4.1. Delay-Attacks

4.1.1. Data Plane Delay Attacks

Dropped messages can result in stream instability. If only a single path is used, the entire stream can be disrupted. In a multipath scenario, large delays on one stream can lead to increased buffer and CPU resources on the elimination bridge.

If the attack is carried out on a sole link (i.e. no multipath), the DetNet stream can be interrupted and result in outages.

4.1.2. Control Plane Delay Attacks

In and of itself, this is not directly a threat, the effects of delaying control messages can have quite adverse effects later.

Delayed messages for tear-down can lead to resource leakage if a stream is not torn down at the correct time. This can in turn result in failure to allocate new streams giving rise to a denial of service attack.

In the case where an End-point should be added to a multicast, failure to deliver said signalling message will prevent the new EP from receiving expected frames.

Likewise, when an EP should be removed from a multicast group, delaying such messages can lead to loss of privacy as the EP will continue to receive messages even after it is removed.

4.2. Flow Identification and Spoofing

4.2.1. Flow identification

Of all the attacks, this is one of the most difficult to detect and counter. Often, an attacker will start out by observing the traffic going through the network and use the knowledge gathered in this phase to mount future attacks.

The attacker can, at their leisure, observe over time all aspects of the messaging and signalling, learning the intent and purpose of all traffic flows. At some later date, possibly at an important time in an operational context, the attacker can launch a multi-faceted attack, possibly in conjunction with some demand for ransom.

The flow-id in the header of the data plane-messages gives an attacker a very reliable identifier for DetNet traffic, and this traffic has a high probability of going to lucrative targets.

4.2.2. Spoofing

4.2.2.1. Dataplane Spoofing

Spoofing dataplane messages can result in increased resource consumptions on the bridges throughout the network as it will increase buffer usage and CPU utilization. This can lead to resource exhaustion and/or increased delay.

If the attacker manages to create valid headers, the false messages can be forwarded through the network, using part of the allocated bandwidth. This in turn can cause legitimate messages to be dropped when the budget has been exhausted.

Finally, the endpoint will have to deal with invalid messages being delivered to the endpoint instead of (or in addition to) a valid message.

4.2.2.2. Control Plane Spoofing

A successful control plane spoofing-attack has a very large potential. It can do anything from modifying existing streams by changing the available bandwidth, add or remove endpoints or drop the stream altogether. It would also be possible to falsely create new streams, which could give an attacker the ability to exhaust the systems resources, or just enable a high quality DetNet stream outside the Network engineer's control.

4.3. Segmentation attacks (injection)

4.3.1. Data Plane Segmentation

Injection of false messages in a DetNet stream could lead to exhaustion of the available bandwidth for a stream if the bridges accounts false messages to the stream's budget.

In a multipath scenario, injected messages will cause an increased CPU utilization on elimination bridges and if enough paths are subject to malicious injection, the legitimate messages could be dropped. Likewise it can cause an increase in buffer usage. In total, this will consume more resources on the bridges than normal, giving rise to a potential resource exhaustion attack on the bridges.

If a stream is interrupted, the end application will be affected by what is now a non-deterministic stream.

4.3.2. Control Plane segmentation

A successful Control Plane segmentation attack will cause control messages to be interpreted by nodes in the network. This has the potential to create new streams (exhausting resources), drop existing (denial of service), add/remove end-stations to a multicast group (loss of privacy) or modify the stream attributes (reducing available bandwidth, or increasing it so that new streams cannot reserve a path).

In short, this means that you cannot trust the stream reservation properties or the network itself.

As with spoofing, if an attacker is able to inject control-plane messages and the receiving end does not detect it, the receiving station must be able to.

4.4. Replication and Elimination

The Replication and Elimination is relevant only to Data Plane messages as Signalling is not subject to multipath routing.

4.4.1. Increased Attack Surface

Covered briefly in Section 4.3

4.4.2. Header Manipulation at Elimination Bridges

Covered briefly in Section 4.3

4.5. Impact of Attacks to Path Choice

This is covered in part in Section 4.3, and as with Replication and Elimination (Section 4.4, this is relevant for DataPlane messages.

4.6. Impact of Attacks by Use Case Industry

This section rates the severity of various components of the impact of a successful vulnerability exploit to use cases by industry as described in [I-D.ietf-detnet-use-cases], including Pro Audio, Electrical Utilities, Building Automation, Wireless for Industrial, Cellular Radio, and Industrial M2M (split into two areas, M2M Data Gathering and M2M Control Loop).

Components of Impact (left column) include Criticality of Failure, Effects of Failure, Recovery, and DetNet Functional Dependence. Criticality of failure summarizes the seriousness of the impact. The impact of a resulting failure can affect many different metrics that

vary greatly in scope and severity. In order to reduce the number of variables, the following were included: Financial, Health and Safety, People well being, Affect on a single organization, and affect on multiple organizations. Recovery outlines how long it would take for an affected use case to get back to its pre-failure state (Recovery time objective, RTO), and how much of the original service would be lost in between the time of service failure and recovery to original state (Recovery Point Objective, RPO). DetNET dependence maps how much the following DetNet service objectives contribute to impact of failure: Time dependency, data integrity, source node integrity, availability, latency/jitter.

The scale of the Impact mappings is low, medium, and high. In some use cases there may be a multitude of specific applications in which DetNET is used. For simplicity this section attempts to average the varied impacts of different applications. This section does not address the overall risk of a certain impact which would require the likelihood of a failure happening.

In practice any such ratings will vary from case to case; the ratings shown here are given as examples.

	Pro A	Util	Bldg	Wire-less	Cell	M2M Data	M2M Ctrl
Criticality	Med	Hi	Low	Med	Med	Med	Med
Effects							
Financial	Med	Hi	Med	Med	Low	Med	Med
Health/Safety	Med	Hi	Hi	Med	Med	Med	Med
People WB	Med	Hi	Hi	Low	Hi	Low	Low
Effect 1 org	Hi	Hi	Med	Hi	Med	Med	Med
Effect >1 org	Med	Hi	Low	Med	Med	Med	Med
Recovery							
Recov Time Obj	Med	Hi	Med	Hi	Hi	Hi	Hi
Recov Point Obj	Med	Hi	Low	Med	Low	Hi	Hi
DetNet Dependence							
Time Dependency	Hi	Hi	Low	Hi	Med	Low	Hi
Latency/Jitter	Hi	Hi	Med	Med	Low	Low	Hi
Data Integrity	Hi	Hi	Med	Hi	Low	Hi	Low
Src Node Integ	Hi	Hi	Med	Hi	Med	Hi	Hi
Availability	Hi	Hi	Med	Hi	Low	Hi	Hi

Figure 2: Impact of Attacks by Use Case Industry

5. Security Threat Mitigation

This section describes a set of measures that can be taken to mitigate the attacks described in Section 3. These mitigations should be viewed as a toolset that includes several different and diverse tools. Each application or system will typically use a subset of these tools, based on a system-specific threat analysis.

5.1. Path Redundancy

Description

A DetNet flow that can be forwarded simultaneously over multiple paths. Path replication and elimination [I-D.ietf-detnet-architecture] provides resiliency to dropped or delayed packets. This redundancy improves the robustness to failures and to man-in-the-middle attacks.

Related attacks

Path redundancy can be used to mitigate various man-in-the-middle attacks, including attacks described in Section 3.2.1, Section 3.2.2, Section 3.2.3, and Section 3.2.8.

5.2. Integrity Protection

Description

An integrity protection mechanism, such as a Hash-based Message Authentication Code (HMAC) can be used to mitigate modification attacks. Integrity protection can be used on the data plane header, to prevent its modification and tampering. Integrity protection in the control plane is discussed in Section 5.5.

Related attacks

Integrity protection mitigates attacks related to modification and tampering, including the attacks described in Section 3.2.2 and Section 3.2.4.

5.3. DetNet Node Authentication

Description

Source authentication verifies the authenticity of DetNet sources, allowing to mitigate spoofing attacks. Note that while integrity protection (Section 5.2) prevents intermediate nodes from modifying information, authentication verifies the source of the information.

Related attacks

DetNet node authentication is used to mitigate attacks related to spoofing, including the attacks of Section 3.2.2, and Section 3.2.4.

5.4. Encryption

Description

DetNet flows can be forwarded in encrypted form.

Related attacks

While confidentiality is not considered an important goal with respect to DetNet, encryption can be used to mitigate recon attacks (Section 3.2.7).

5.5. Control and Signaling Message Protection

Description

Control and signaling messages can be protected using authentication and integrity protection mechanisms.

Related attacks

These mechanisms can be used to mitigate various attacks on the control plane, as described in Section 3.2.6, Section 3.2.8 and Section 3.2.5.

5.6. Dynamic Performance Analytics

Description

Information about the network performance can be gathered in real-time in order to detect anomalies and unusual behavior that may be the symptom of a security attack. The gathered information can be based, for example, on per-flow counters, bandwidth measurement, and monitoring of packet arrival times. Unusual behavior or potentially malicious nodes can be reported to a management system, or can be used as a trigger for taking corrective actions. The information can be tracked by DetNet end systems and transit nodes, and exported to a management system, for example using NETCONF.

Related attacks

Performance analytics can be used to mitigate various attacks, including the ones described in Section 3.2.1, Section 3.2.3, and Section 3.2.8.

5.7. Mitigation Summary

The following table maps the attacks of Section 3 to the impacts of Section 4, and to the mitigations of the current section. Each row specifies an attack, the impact of this attack if it is successfully implemented, and possible mitigation methods.

Attack	Impact	Mitigations
Delay Attack	-Non-deterministic delay -Data disruption -Increased resource consumption	-Path redundancy -Performance analytics
DetNet Flow Modification or Spoofing	-Increased resource consumption -Data disruption	-Path redundancy -Integrity protection -DetNet Node authentication
Inter-Segment Attack	-Increased resource consumption -Data disruption	-Path redundancy -Performance analytics
Replication: Increased attack surface	-All impacts of other attacks	-Integrity protection -DetNet Node authentication
Replication-related Header Manipulation	-Non-deterministic delay -Data disruption	-Integrity protection -DetNet Node authentication
Path Manipulation	-Enabler for other attacks	-Control message protection
Path Choice: Increased Attack Surface	-All impacts of other attacks	-Control message protection
Control or Signaling Packet Modification	-Increased resource consumption -Non-deterministic delay -Data disruption	-Control message protection
Control or Signaling Packet Injection	-Increased resource consumption	-Control message protection

	-Non-deterministic delay -Data disruption	
Reconnaissance	-Enabler for other attacks	-Encryption
Attacks on Time Sync Mechanisms	-Non-deterministic delay -Increased resource consumption -Data disruption	-Path redundancy -Control message protection -Performance analytics

Figure 3: Mapping Attacks to Impact and Mitigations

6. Association of Attacks to Use Cases

6.1. Use Cases by Common Themes

Different attacks can have different impact and/or mitigation depending on the use case, so we would like to make this association in our analysis. However since there is a potentially unbounded list of use cases, we categorize the attacks with respect to the common themes of the use cases as identified in the Use Case Common Themes section of the DetNet Use Cases draft [I-D.ietf-detnet-use-cases]. We describe each theme and its associated attacks, impacts and mitigations.

6.1.1. Network Layer - AVB/TSN Ethernet

Presumably it will be possible to run DetNet over other underlying network layers besides Ethernet, but Ethernet is explicitly supported. Is the attack specific to the Ethernet AVB/TSN protocols? Does the threat affect only Ethernet, or any underlying network layer?

6.1.2. Central Administration

A DetNet network is expected to be controlled by a centralized network configuration and control system. Such a system may be in a single central location, or it may be distributed across multiple control entities that function together as a unified control system for the network. Is the attack directed at threat the central control system of the network? Does it interfere with OAM?

6.1.3. Hot Swap

A DetNet network is not expected to be "plug and play" - it is expected that there is some centralized network configuration and control system. However, the ability to "hot swap" components (e.g. due to malfunction) is similar enough to "plug and play" that this kind of behavior may be expected in DetNet networks, depending on the implementation. Does the attack target "hot swap" ("plug and play") operation in the network?

6.1.4. Data Flow Information Models

Data Flow Information Models specific to DetNet networks are to be specified by DetNet. Thus they are "new" and thus potentially present a new attack surface. Does the threat take advantage of any aspect of our new Data Flow Info Models?

6.1.5. L2 and L3 Integration

A DetNet network is intended to integrate between Layer 2 (bridged) network(s) (e.g. AVB/TSN LAN) and Layer 3 (routed) network(s) (e.g. using IP-based protocols). Does the attack target L2? L3? Both? The interaction between the two?

6.1.6. End-to-End Delivery

Packets sent over DetNet are guaranteed not to be dropped by the network due to congestion. (Packets may however be dropped for intended reasons, e.g. per security measures). Does the attack result in packets (which should be delivered) not being delivered? Does it result in packets that should not be delivered being delivered?

6.1.7. Proprietary Deterministic Ethernet Networks

There are many proprietary non-interoperable deterministic Ethernet-based networks currently available; DetNet is intended to provide an open-standards-based alternative to such networks. Does the threat relate to a specific such network that is being "emulated" or "replaced" by DetNet, for example by exploiting specific commands specific to that network protocol?

6.1.8. Replacement for Proprietary Fieldbuses

There are many proprietary "field buses" used in today's industrial and other industries; DetNet is intended to provide an open-standards-based alternative to such buses. Does the threat relate to a specific fieldbus that is being "emulated" or "replaced" by DetNet,

for example by exploiting specific commands specific to that network protocol?

6.1.9. Deterministic vs Best-Effort Traffic

DetNet is intended to support coexistence of time-sensitive operational (OT, deterministic) traffic and information (IT, "best effort") traffic on the same ("unified") network. Does the attack affect only IT or only OT or both types of traffic? Does the threat affect any interaction between IT and OT traffic, e.g. by changing relative priority or handling of IT vs. OT packets?

6.1.10. Deterministic Flows

Reserved bandwidth data flows (deterministic flows) must be isolated from each other and from best-effort traffic, so that even if the network is saturated with best-effort and/or reserved bandwidth traffic the configured flows are not adversely affected. Does the attack affect the isolation of one (reserved) flow from another?

6.1.11. Unused Reserved Bandwidth

If bandwidth reservations are made for a stream but the associated bandwidth is not used at any point in time, that bandwidth is made available on the network for best-effort traffic. If the owner of the reserved stream then starts transmitting again, the bandwidth is no longer available for best-effort traffic, on a moment-to-moment basis. (Such "temporarily available" bandwidth is not available for time-sensitive traffic, which must have its own reservation). Does the attack affect the system's ability to allocate unused reserved BW to best-effort traffic?

6.1.12. Interoperability

The DetNet network specifications are intended to enable an ecosystem in which multiple vendors can create interoperable products, thus promoting device diversity and potentially higher numbers of each device manufactured. Does the threat take advantage of differences in implementation of "interoperable" products made by different vendors?

6.1.13. Cost Reductions

The DetNet network specifications are intended to enable an ecosystem in which multiple vendors can create interoperable products, thus promoting higher numbers of each device manufactured, promoting cost reduction and cost competition among vendors. Does the threat take

advantage of "low cost" HW or SW components or other "cost-related shortcuts" that might be present in devices?

6.1.14. Insufficiently Secure Devices

The DetNet network specifications are intended to enable an ecosystem in which multiple vendors can create interoperable products, thus promoting device diversity and potentially higher numbers of each device manufactured. Does the threat attack "naivete" of SW, for example SW that was not designed to be sufficiently secure (or secure at all) but is deployed on a DetNet network that is intended to be highly secure? (For example IoT exploits like the Mirai video-camera botnet ([MIRAI])).

6.1.15. DetNet Network Size

DetNet networks range in size from very small, e.g. inside a single industrial machine, to very large, for example a Utility Grid network spanning a whole country, and involving many "hops" over various kinds of links for example radio repeaters, microwave links, fiber optic links, etc.. Does the attack affect DetNet networks of only certain sizes, e.g. very large networks, or very small? This might be related to how the attack is introduced into the network, for example if the entire network is local, there is a threat that power can be cut to the entire network. If the network is large, perhaps only a part of the network is attacked. Does the threat take advantage of attack vectors that are specific to network size?

6.1.16. Multiple Hops

DetNet networks range in size from very small, e.g. inside a single industrial machine, to very large, for example a Utility Grid network spanning a whole country, and involving many "hops" over various kinds of links for example radio repeaters, microwave links, fiber optic links, etc.. Does the attack exploit the presence of more than one "hop"? Does the threat exploit the presence of more than one type of "hop", e.g. between radio and microwave links? Does the threat exploit a specific type of "hop", e.g. something specific to a fiber optic link, or other type of link?

6.1.17. Level of Service

A DetNet is expected to provide means to configure the network that include querying network path latency, requesting bounded latency for a given stream, requesting worst case maximum and/or minimum latency for a given path or stream, and so on. It is an expected case that the network cannot provide a given requested service level. In such cases the network control system should reply that the requested

service level is not available (as opposed to accepting the parameter but then not delivering the desired behavior). Does the attack affect any querying or replying to such service-level-related traffic? Can the attack cause incorrect responses from the system regarding timing-related configuration? For example replying that a requested level of service is available when it isn't, or that the requested level of service is not available when it actually is available?

6.1.18. Bounded Latency

Does the threat affect the network's ability to deliver packets within the agreed-upon latency boundaries?

6.1.19. Low Latency

Applications may require "extremely low latency" however depending on the application these may mean very different latency values; for example "low latency" across a Utility grid network is on a different time scale than "low latency" in a motor control loop in a small machine. The intent is that the mechanisms for specifying desired latency include wide ranges, and that architecturally there is nothing to prevent arbitrarily low latencies from being implemented in a given network. Does the threat affect the network's ability to deliver packets within the agreed-upon low latency?

6.1.20. Symmetrical Path Delays

Some applications would like to specify that the transit delay time values be equal for both the transmit and return paths. Does the attack affect the network's ability to provide matched transmit and return path delays (latencies)?

6.1.21. Reliability and Availability

DetNet based systems are expected to be implemented with essentially arbitrarily high availability (for example 99.9999% up time, or even 12 nines). The intent is that the DetNet designs should not make any assumptions about the level of reliability and availability that may be required of a given system, and should define parameters for communicating these kinds of metrics within the network. Does the attack affect the reliability of the DetNet network? Is it a large or small change, e.g. the difference between completely taking down the network for some period of time, vs reducing its reliability by just one "nine"? Does the threat affect the availability of the DetNet network?

6.1.22. Redundant Paths

DetNet based systems are expected to be implemented with essentially arbitrarily high reliability/availability. A strategy used by DetNet for providing such extraordinarily high levels of reliability is to provide redundant paths that can be seamlessly switched between, all the while maintaining the required performance of that system. Does the attack affect the configuration or operation of redundant paths?

6.1.23. Security Measures

A DetNet network must be made secure against devices failures, attackers, misbehaving devices, and so on. Does the threat affect such security measures themselves, e.g. by attacking SW designed to protect against device failure?

6.2. Attack Types by Use Case Common Theme

The following table lists the attacks of Section 3, assigning a number to each type of attack. That number is then used as a short form identifier for the attack in Figure 5.

Attack	Section
1 Delay Attack	Section 3.2.1
2 DetNet Flow Modification or Spoofing	Section 3.2.2
3 Inter-Segment Attack	Section 3.2.3
4 Replication: Increased attack surface	Section 3.2.4.1
5 Replication-related Header Manipulation	Section 3.2.4.2
6 Path Manipulation	Section 3.2.5.1
7 Path Choice: Increased Attack Surface	Section 3.2.5.2
8 Control or Signaling Packet Modification	Section 3.2.6.1
9 Control or Signaling Packet Injection	Section 3.2.6.2
10 Reconnaissance	Section 3.2.7
11 Attacks on Time Sync Mechanisms	Section 3.2.8

Figure 4: List of Attacks

The following table maps the use case themes presented in this memo to the attacks of Figure 4. Each row specifies a theme, and the attacks relevant to this theme are marked with a '+'.

Theme	Attack										
	1	2	3	4	5	6	7	8	9	10	11
Network Layer - AVB/TSN Eth.	+	+	+	+	+	+	+	+	+	+	+
Central Administration						+	+	+	+	+	+
Hot Swap		+	+								+
Data Flow Information Models											
L2 and L3 Integration					+	+					

End-to-end Delivery				+	+								
Proprietary Deterministic Ethernet Networks			+			+	+	+	+				
Replacement for Proprietary Fieldbuses			+			+	+	+	+				
Deterministic vs. Best-Effort Traffic			+										
Deterministic Flows			+										
Unused Reserved Bandwidth			+										
Interoperability													
Cost Reductions													
Insufficiently Secure Devices													
DetNet Network Size	+					+	+						+
Multiple Hops	+	+				+	+						+
Level of Service								+	+	+			
Bounded Latency	+												+
Low Latency	+												+
Symmetric Path Delays	+												+
Reliability and Availability	+	+	+	+	+	+	+	+	+	+	+	+	+
Redundant Paths				+	+			+	+				
Security Measures													

Figure 5: Mapping Between Themes and Attacks

7. Appendix A: DetNet Draft Security-Related Statements

This section collects the various statements in the currently existing DetNet Working Group drafts. For each draft, the section name and number of the quoted section is shown. The text shown here

is the work of the original draft authors, quoted verbatim from the drafts. The intention is to explicitly quote all relevant text, not to summarize it.

7.1. Architecture (draft 8)

7.1.1. Fault Mitigation (sec 4.5)

One key to building robust real-time systems is to reduce the infinite variety of possible failures to a number that can be analyzed with reasonable confidence. DetNet aids in the process by providing filters and policers to detect DetNet packets received on the wrong interface, or at the wrong time, or in too great a volume, and to then take actions such as discarding the offending packet, shutting down the offending DetNet flow, or shutting down the offending interface.

It is also essential that filters and service remarking be employed at the network edge to prevent non-DetNet packets from being mistaken for DetNet packets, and thus impinging on the resources allocated to DetNet packets.

There exist techniques, at present and/or in various stages of standardization, that can perform these fault mitigation tasks that deliver a high probability that misbehaving systems will have zero impact on well-behaved DetNet flows, except of course, for the receiving interface(s) immediately downstream of the misbehaving device. Examples of such techniques include traffic policing functions (e.g. [RFC2475]) and separating flows into per-flow rate-limited queues.

7.1.2. Security Considerations (sec 7)

Security in the context of Deterministic Networking has an added dimension; the time of delivery of a packet can be just as important as the contents of the packet, itself. A man-in-the-middle attack, for example, can impose, and then systematically adjust, additional delays into a link, and thus disrupt or subvert a real-time application without having to crack any encryption methods employed. See [RFC7384] for an exploration of this issue in a related context.

Furthermore, in a control system where millions of dollars of equipment, or even human lives, can be lost if the DetNet QoS is not delivered, one must consider not only simple equipment failures, where the box or wire instantly becomes perfectly silent, but bizarre errors such as can be caused by software failures. Because there is essential no limit to the kinds of failures that can occur, protecting against realistic equipment failures is indistinguishable,

in most cases, from protecting against malicious behavior, whether accidental or intentional.

Security must cover:

- o Protection of the signaling protocol
- o Authentication and authorization of the controlling nodes
- o Identification and shaping of the flows

7.2. Data Plane Alternatives (draft 4)

7.2.1. Security Considerations (sec 7)

This document does not add any new security considerations beyond what the referenced technologies already have.

7.3. Problem Statement (draft 5)

7.3.1. Security Considerations (sec 5)

Security in the context of Deterministic Networking has an added dimension; the time of delivery of a packet can be just as important as the contents of the packet, itself. A man-in-the-middle attack, for example, can impose, and then systematically adjust, additional delays into a link, and thus disrupt or subvert a real-time application without having to crack any encryption methods employed. See [RFC7384] for an exploration of this issue in a related context.

Typical control networks today rely on complete physical isolation to prevent rogue access to network resources. DetNet enables the virtualization of those networks over a converged IT/OT infrastructure. Doing so, DetNet introduces an additional risk that flows interact and interfere with one another as they share physical resources such as Ethernet trunks and radio spectrum. The requirement is that there is no possible data leak from and into a deterministic flow, and in a more general fashion there is no possible influence whatsoever from the outside on a deterministic flow. The expectation is that physical resources are effectively associated with a given flow at a given point of time. In that model, Time Sharing of physical resources becomes transparent to the individual flows which have no clue whether the resources are used by other flows at other times.

Security must cover:

- o Protection of the signaling protocol

- o Authentication and authorization of the controlling nodes
- o Identification and shaping of the flows
- o Isolation of flows from leakage and other influences from any activity sharing physical resources

7.4. Use Cases (draft 11)

7.4.1. (Utility Networks) Security Current Practices and Limitations (sec 3.2.1)

Grid monitoring and control devices are already targets for cyber attacks, and legacy telecommunications protocols have many intrinsic network-related vulnerabilities. For example, DNP3, Modbus, PROFIBUS/PROFINET, and other protocols are designed around a common paradigm of request and respond. Each protocol is designed for a master device such as an HMI (Human Machine Interface) system to send commands to subordinate slave devices to retrieve data (reading inputs) or control (writing to outputs). Because many of these protocols lack authentication, encryption, or other basic security measures, they are prone to network-based attacks, allowing a malicious actor or attacker to utilize the request-and-respond system as a mechanism for command-and-control like functionality. Specific security concerns common to most industrial control, including utility telecommunication protocols include the following:

- o Network or transport errors (e.g. malformed packets or excessive latency) can cause protocol failure.
- o Protocol commands may be available that are capable of forcing slave devices into inoperable states, including powering-off devices, forcing them into a listen-only state, disabling alarming.
- o Protocol commands may be available that are capable of restarting communications and otherwise interrupting processes.
- o Protocol commands may be available that are capable of clearing, erasing, or resetting diagnostic information such as counters and diagnostic registers.
- o Protocol commands may be available that are capable of requesting sensitive information about the controllers, their configurations, or other need-to-know information.

- o Most protocols are application layer protocols transported over TCP; therefore it is easy to transport commands over non-standard ports or inject commands into authorized traffic flows.
- o Protocol commands may be available that are capable of broadcasting messages to many devices at once (i.e. a potential DoS).
- o Protocol commands may be available to query the device network to obtain defined points and their values (i.e. a configuration scan).
- o Protocol commands may be available that will list all available function codes (i.e. a function scan).
- o These inherent vulnerabilities, along with increasing connectivity between IT and OT networks, make network-based attacks very feasible.
- o Simple injection of malicious protocol commands provides control over the target process. Altering legitimate protocol traffic can also alter information about a process and disrupt the legitimate controls that are in place over that process. A man-in-the-middle attack could provide both control over a process and misrepresentation of data back to operator consoles.

7.4.2. (Utility Networks) Security Trends in Utility Networks (sec 3.3.3)

Although advanced telecommunications networks can assist in transforming the energy industry by playing a critical role in maintaining high levels of reliability, performance, and manageability, they also introduce the need for an integrated security infrastructure. Many of the technologies being deployed to support smart grid projects such as smart meters and sensors can increase the vulnerability of the grid to attack. Top security concerns for utilities migrating to an intelligent smart grid telecommunications platform center on the following trends:

- o Integration of distributed energy resources
- o Proliferation of digital devices to enable management, automation, protection, and control
- o Regulatory mandates to comply with standards for critical infrastructure protection

- o Migration to new systems for outage management, distribution automation, condition-based maintenance, load forecasting, and smart metering
- o Demand for new levels of customer service and energy management

This development of a diverse set of networks to support the integration of microgrids, open-access energy competition, and the use of network-controlled devices is driving the need for a converged security infrastructure for all participants in the smart grid, including utilities, energy service providers, large commercial and industrial, as well as residential customers. Securing the assets of electric power delivery systems (from the control center to the substation, to the feeders and down to customer meters) requires an end-to-end security infrastructure that protects the myriad of telecommunications assets used to operate, monitor, and control power flow and measurement.

"Cyber security" refers to all the security issues in automation and telecommunications that affect any functions related to the operation of the electric power systems. Specifically, it involves the concepts of:

- o Integrity : data cannot be altered undetectably
- o Authenticity : the telecommunications parties involved must be validated as genuine
- o Authorization : only requests and commands from the authorized users can be accepted by the system
- o Confidentiality : data must not be accessible to any unauthenticated users

When designing and deploying new smart grid devices and telecommunications systems, it is imperative to understand the various impacts of these new components under a variety of attack situations on the power grid. Consequences of a cyber attack on the grid telecommunications network can be catastrophic. This is why security for smart grid is not just an ad hoc feature or product, it's a complete framework integrating both physical and Cyber security requirements and covering the entire smart grid networks from generation to distribution. Security has therefore become one of the main foundations of the utility telecom network architecture and must be considered at every layer with a defense-in-depth approach. Migrating to IP based protocols is key to address these challenges for two reasons:

- o IP enables a rich set of features and capabilities to enhance the security posture
- o IP is based on open standards, which allows interoperability between different vendors and products, driving down the costs associated with implementing security solutions in OT networks.

Securing OT (Operation technology) telecommunications over packet-switched IP networks follow the same principles that are foundational for securing the IT infrastructure, i.e., consideration must be given to enforcing electronic access control for both person-to-machine and machine-to-machine communications, and providing the appropriate levels of data privacy, device and platform integrity, and threat detection and mitigation.

7.4.3. (BAS) Security Considerations (sec 4.2.4)

When BAS field networks were developed it was assumed that the field networks would always be physically isolated from external networks and therefore security was not a concern. In today's world many BASs are managed remotely and are thus connected to shared IP networks and so security is definitely a concern, yet security features are not available in the majority of BAS field network deployments .

The management network, being an IP-based network, has the protocols available to enable network security, but in practice many BAS systems do not implement even the available security features such as device authentication or encryption for data in transit.

7.4.4. (6TiSCH) Security Considerations (sec 5.3.3)

On top of the classical requirements for protection of control signaling, it must be noted that 6TiSCH networks operate on limited resources that can be depleted rapidly in a DoS attack on the system, for instance by placing a rogue device in the network, or by obtaining management control and setting up unexpected additional paths.

7.4.5. (Cellular radio) Security Considerations (sec 6.1.5)

Establishing time-sensitive streams in the network entails reserving networking resources for long periods of time. It is important that these reservation requests be authenticated to prevent malicious reservation attempts from hostile nodes (or accidental misconfiguration). This is particularly important in the case where the reservation requests span administrative domains. Furthermore, the reservation information itself should be digitally signed to

reduce the risk of a legitimate node pushing a stale or hostile configuration into another networking node.

Note: This is considered important for the security policy of the network, but does not affect the core DetNet architecture and design.

7.4.6. (Industrial M2M) Communication Today (sec 7.2)

Industrial network scenarios require advanced security solutions. Many of the current industrial production networks are physically separated. Preventing critical flows from be leaked outside a domain is handled today by filtering policies that are typically enforced in firewalls.

8. IANA Considerations

This memo includes no requests from IANA.

9. Security Considerations

The security considerations of DetNet networks are presented throughout this document.

10. Informative References

[ARINC664P7]

ARINC, "ARINC 664 Aircraft Data Network, Part 7, Avionics Full-Duplex Switched Ethernet Network", 2009.

[I-D.ietf-detnet-architecture]

Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-01 (work in progress), March 2017.

[I-D.ietf-detnet-use-cases]

Grossman, E., Gunther, C., Thubert, P., Wetterwald, P., Raymond, J., Korhonen, J., Kaneko, Y., Das, S., Zha, Y., Varga, B., Farkas, J., Goetz, F., Schmitt, J., Vilajosana, X., Mahmoodi, T., Spirou, S., and P. Vizarreta, "Deterministic Networking Use Cases", draft-ietf-detnet-use-cases-12 (work in progress), April 2017.

[I-D.varga-detnet-service-model]

Varga, B. and J. Farkas, "DetNet Service Model", draft-varga-detnet-service-model-02 (work in progress), May 2017.

- [IEEE1588] IEEE, "IEEE 1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems Version 2", 2008.
- [MIRAI] krebsonsecurity.com, "<https://krebsonsecurity.com/2016/10/hacked-cameras-dvrs-powered-todays-massive-internet-outage/>", 2016.
- [RFC7384] Mizrahi, T., "Security Requirements of Time Protocols in Packet Switched Networks", RFC 7384, DOI 10.17487/RFC7384, October 2014, <<http://www.rfc-editor.org/info/rfc7384>>.

Authors' Addresses

Tal Mizrahi
Marvell

Email: talmi@marvell.com

Ethan Grossman (editor)
Dolby Laboratories, Inc.
1275 Market Street
San Francisco, CA 94103
USA

Phone: +1 415 645 4726
Email: ethan.grossman@dolby.com
URI: <http://www.dolby.com>

Andrew J. Hacker
MistiQ Technologies, Inc
Harrisburg, PA
USA

Email: ajhacker@mistiqttech.com
URI: <http://www.mistiqttech.com>

Subir Das
Applied Communication Sciences
150 Mount Airy Road, Basking Ridge
New Jersey, 07920
USA

Email: sdas@appcomsci.com

John Dowdell
Airbus Defence and Space
Celtic Springs
Newport NP10 8FZ
United Kingdom

Email: john.dowdell.ietf@gmail.com

Henrik Austad
Cisco Systems
Philip Pedersens vei 1
Lysaker 1366
Norway

Email: henrik@austad.us

Kevin Stanton
Intel

Email: kevin.b.stanton@intel.com

Norman Finn
Huawei

Email: norman.finn@mail01.huawei.com