DetNet
Internet-Draft
Intended status: Informational
Expires: September 14, 2017

J. Korhonen, Ed.
Broadcom
L. Andersson
Y. Jiang
Huawei
B. Varga
J. Farkas
Ericsson
CJ. Bernardos
UC3M
T. Mizrahi
Marvell

March 13, 2017

DetNet Data Plane solution
draft-dt-detnet-dp-sol-00

Abstract

This document specifies a PseudoWire-based Deterministic Networking data plane solution. The data plane solution 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 September 14, 2017.

Copyright Notice

Table of Contents

1. Introduction

This document specifies a Deterministic Networking (DetNet) data plane solution. The solution is based on PseudoWires (PW) [RFC3985] and makes use of the multi-segment (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 either an IP or MPLS [RFC4448][RFC6658] Packet Switched Network (PSN).

For the purpose of DetNet data plane, this document specifically specifies the PW encapsulation for DetNet flows, a DetNet Control Word (CW), a DetNet label, how MS-PW derived DetNet Relay and Edge nodes work, and as a specific new PW feature how the Packet Replication and Elimination function (PREF) is implemented using PWs. This document does not define the associated control plane functions, or operations and management (OAM).

2. Terminology

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 A DetNet aware PseudoWire Terminating Provider Edge (T-PE).

DA-S-PE A DetNet aware PseudoWire Switching Provider Edge (S-PE).

T-Label A hop-by-hop tunnel label layer between label switching routers (LSR).

L-Label A DetNet topology overlay label that is used between DA-*-PE devices.

flow-ID A DetNet flow identity that uniquely identifies a DetNet flow in a DetNet network. The flow-ID is part of the PseudoWire Encapsulation header.

local-ID A DA-T-PE and DA-S-PE node internal construct that uniquely identifies a DetNet flow. The local-ID can be equal to flow-ID or be derived using other means, e.g., programming required label to local-ID mappings directly into the label information base (LFIB).

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

PREF A Packet Replication and Elimination Function (PREF) does the replication and elimination processig of DetNet flow packets in DA-T-PE or DA-S-PE nodes. The replication function is essentially the existing 1+1

protection mechanism. The elimination function reuses and extends the existing [RFC3985] PseudoWire sequencing provided duplicate detection mechanism to operate over multiple (separate) PseudoWires that are sub-flows of a compound DetNet flow.

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

[Ed. to be written.. describe the scope here fot this document: this document only addresses the inter-connect case i.e., 802.1 over routed network (enlarge the layer-2 domain - EVPAN', and the native DetNet case.]

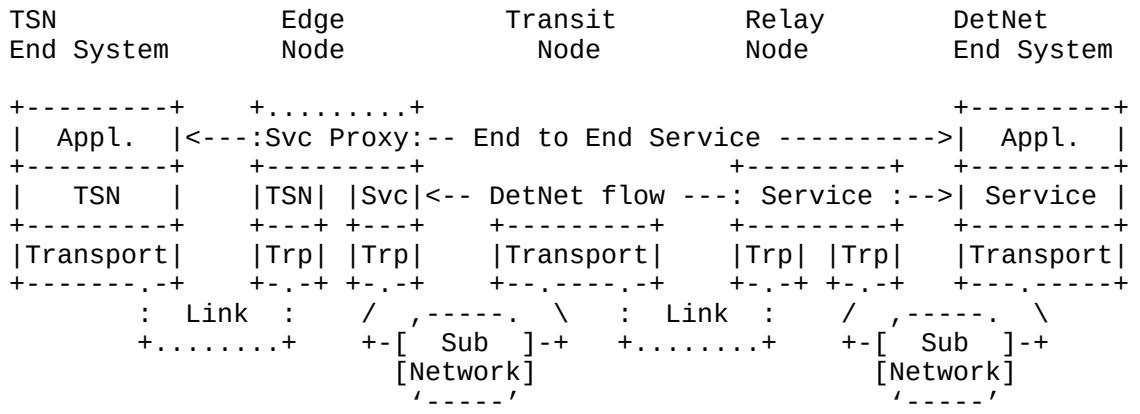

Figure 1: A simple DetNet enabled network architecture

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

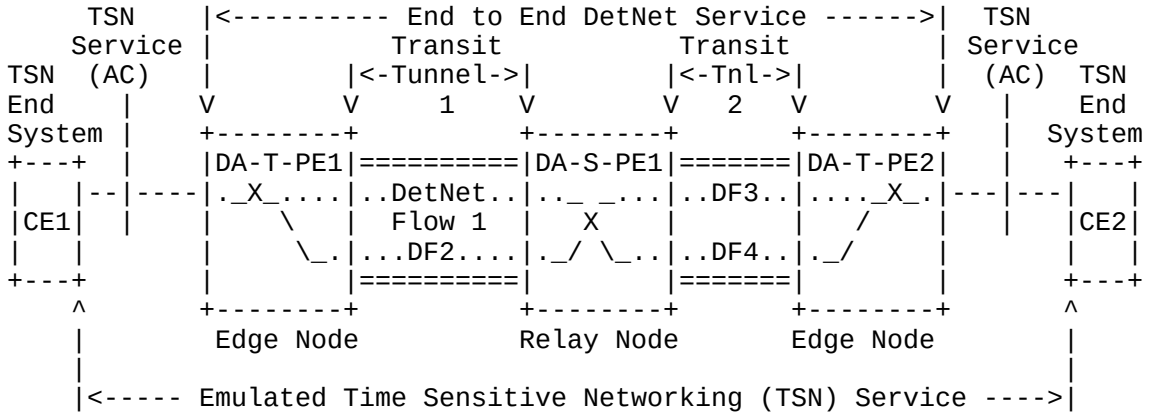

Figure 2: IEEE 802.1TSN over DetNet

Figure 3 illustrates how end to end DetNet service can be provided. In this case, the end systems are able to send and receive DetNet flows. For example, put application data in PseudoWire (PW) and encapsulated in IP. 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 replacing IP with MPLS or tunneling IP over MPLS, or simply interconnect network segments.

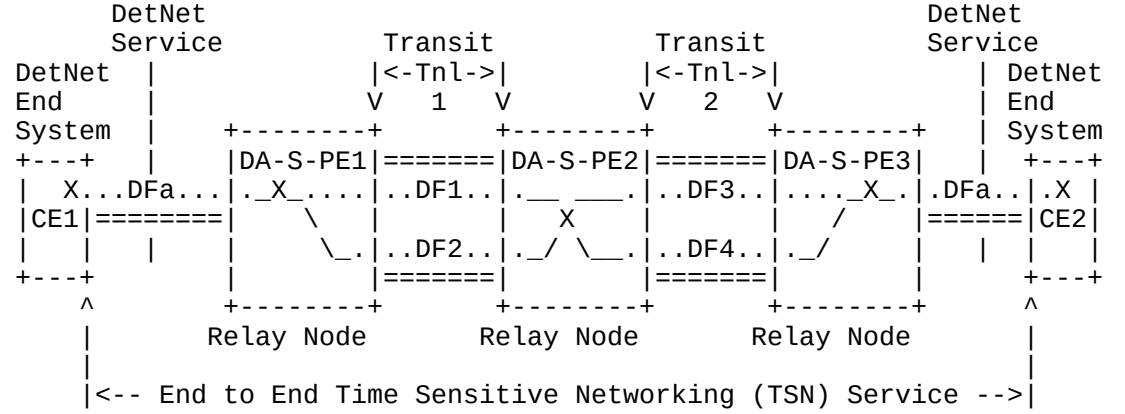

Figure 3: Native DetNet

4.1. DetNet data plane solution requirements

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

1. DetNet function related scenarios:

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

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

    * Service protection: recognize compound / member flows for replication an 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 (DA-T-PE, DA-S-PE and P) use a local-ID of the DetNet-(compound)-flow in order to accomplish its role during transport. Recognizing the flow-ID is more relaxed for DA-T-PE and DA-S-PE nodes, as they are fully aware of both the DetNet service and transport layers. The DetNet role of intermediate "P" nodes is limited to ensure congestion protection from the above listed DetNet functions. However, P nodes can usually recognize only "T-label" and cannot consider the whole label stack for flow recognition. Therefore, identifying each individual DetNet flow on a P node may not be achieved in some network scenarios.

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 DA-T-PE and DA-S-PE nodes. Local-ID MUST be unambiguously bound to the Flow-ID encoded in the DetNet packet.

5. DetNet data plane solution

5.1. DetNet Control Word

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 4.

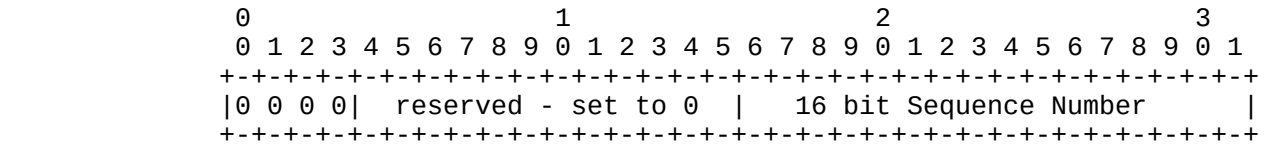

Figure 4: DetNet Control Word

[Editor's note: Shoudl we care about high speed links, here 16 bits of sequence number wraps fast? For example, in a case of 100Gb/s link, 16 bits of sequence number will wrap in ~6.6ms assuming 1250 octets of packets and ~3.3ms for 625 octets packets. Both numbers mean quite long fiber distances, though.]

5.2. DetNet flow identity word

The DetNet flow identity word (flow-ID) identifies a DetNet flow uniquely within a DetNet network. The flow-ID is also associated with the sequence number carried in the DetNet control word and used also for PREF purposes.

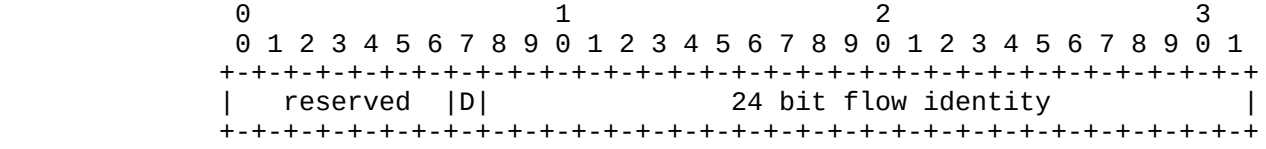

Figure 5: DetNet flow identity word

The management and assignment of the flow-IDs is outside of this specification. It is assumed that each DA-T-PE node have either a preconfigured flow-ID number space or some dynamic control plane protocol is able to coordinate the allocation of the flow-IDs across the DA-T-PE nodes, or a central entity, e.g., a Software Defined Networking controller.

The D bit defines the direction of the flow. The value 0 means 'east' and the value 1 means 'west'. The D bit can be used in ring topologies to allow DetNet flows with the same flow-ID cross with or without PREF processing taking place. Whether a DA-*-PE checks for

the D bit is based on a local policy setting. The support for D bit is optional. If a DA-*-PE does not support the D bit it MUST be treated as 0.

The reserved field in the flow-ID MUST be set to zero and ignored when received.

[Editor's note: we need some configuration knob defined for this feature.]

5.3. DetNet encapsulation

The DetNet data plane follows PW encapsulation. This document specifies a single encapsulation that can be used over both MPLS and IP packet switched Networks (PSN). The DetNet data plane encapsulation consists of a

- DetNet control word (d-CW): contains sequencing information for packet replication and duplicate elimination purposes. There is a separate sequence number space per each DetNet label.

- DetNet flow-ID (f-ID): uniquely identifies a DetNet flow within a DetNet network. Multiple DetNet PWs with different PW labels may have the same f-ID, which then implies the PWs are actually subflows of one compound flow.

- PseudoWire Label (PW Label;): a standard PW label that identifies a PW Instance within a (DA-)T-PE or (DA-)S-PE device.

- DetNet topology overlay label (L-label): an optional label used between (DA-)T-PE or (DA-)S-PE nodes. The main use of L-labels is to tunnel PWs through a PE node and therefore effectively making a PE node to behave like a P node.

In a case of MPLS-based PSN, the tunnel labels between LSRs are referred as T-labels.

The DetNet CW and the Detnet flow-ID together constitute the DetNet PseudoWire encapsulation header.

> [Editor's note: The current design has the DetNet flow-ID as part of the every DetNet flow packet. The flow-ID identifies the flow uniquely within the DetNet network and together with the sequence number information from the DetNet control word is used for PREF purposes. The flow-ID makes is easy for the DA-*-PE node to associate different PWs into one compound flow and perform the elimination of duplicate packets. The flow-ID would point at the node internal construct that holds the received packet history for

each DetNet flow of interest. However, it could also be possible to associate multiple PWs into one DetNet flow just using the control plane provided information. In this case different PWs (using any PW label) would be mapped internally within a node to a local-ID (or similar construct), which again points at the internal per DetNet flow received packets history construct. The explicit in-band flow-ID is easy from the processing and control plane point of view. The local-ID approach does not need the in-band information (thus has less overhead) but requires more from the control plane and the mapping information has to be stored into the LFIB. Current design decision is the in-band flow-ID but may be changed to local-ID if there is a strong reason to do the change.]

Figure 6 illustrates a DetNet PseudoWire encapsulation using an MPLS PSN. Similarly, Figure 7 illustrates the DetNet PseudoWire encapsulation when IP PSN is used. The encapsulation is uniform above the PSN.

Depending on the network topology the "overlay label" (L-label) may be part of the label stack. The L-label tunnels guarantee PW labels remain unchanged between DA-*-PE nodes. Furthermore, L-labels tunnels allow selectively exposing the PW label to DA-*-PE nodes, which means some overlay topologies may just pass through specific DA-S-PEs without any DetNet specific processing.

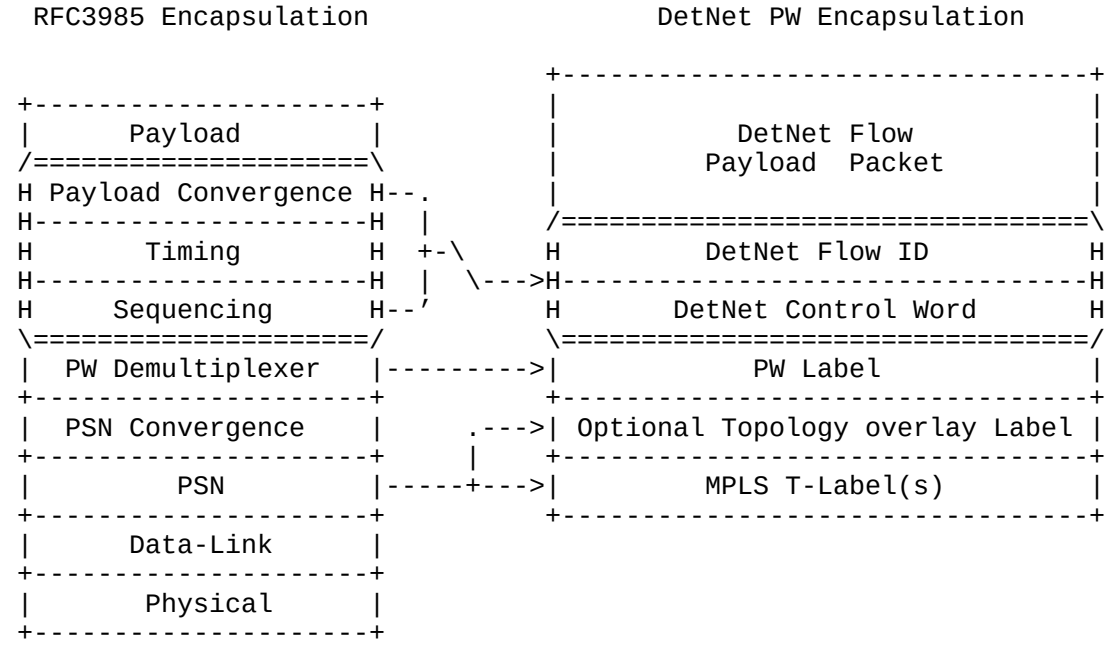

Figure 6: Encapsulation of a DetNet flow in a PW with MPLS(-TP) PSN

When IP PSN is used, the label stack it transports is only inspected when the IP packet destination address equals to the IP address of a DA-*-PE or a P node. Essentially there are one more IP tunnels between a number of DA-*-PE and/or P nodes. The LFIB and the forwarding information base (FIB) combination determines whether a PW gets terminated at the node or forwarded to another node within a new IP tunnel.

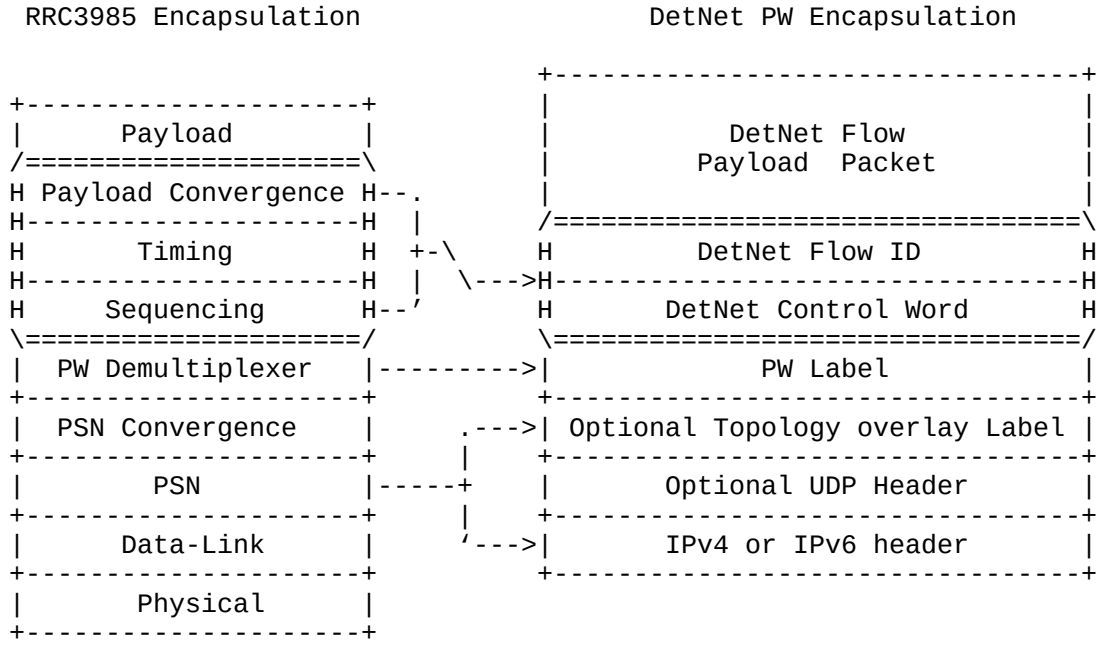

Figure 7: Encapsulation of a DetNet flow in a PW with IP PSN

6. PE reference model considerations

6.1. Forwarded clarifications

[Editor's note: The Detnet-aware "extended forwarder" does the heavy lifting on maintaining the sequence numbers associated with the DetNet labels. Extended forwarder is also responsible for packet replication and duplicate elimination. See the excerpt from RFC3985 Section 4.2.1. about forwarder's functions. We extend that to PREF:

> Some applications have to forward payload elements selectively from one or more ACs to one or more PWs. In such cases, there will also be a need to perform the inverse function on PWE3-PDUs received by a PE from the PSN. This is the function of the forwarder.

]

The DetNet specific new functionality in a DA-*-PE PW processing is the packet replication and duplication elimination function (PREF). This functional is a part of the "extended" forwarder. The PREF processing is triggered by the LFIB actions i.e., not all PWs receive

DetNet specific processing. Basically the LFIB has to be extended with a "PREF enabled" boolean configuration switch that is associated with the normal label actions (e.g., 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 packet (i.e., 1:M replication). The replicated packets MUST share the same DetNet PW control word sequence number and flow identity word flow-id.

The complex part of the DetNet PREF processing is tracking the history of received packets for multiple PWs. These PWs do not have the same PW label value while they still share the same PW sequence number counter and the history information. That is where the DetNet encapsulation header flow-ID plays an important role and binds the control word sequence number to the flow specific shared counter and history information within the PREF function.

The DetNet flow word contains a D flag bit (see Section 5.2), which makes the DA-*-PE node aware of the direction the flow-ID arrived from. If the node, based on the local policy, checks for the D bit setting that effectively means the sequence number history has to contain also the D bit information.

[Editor’s note: draw here an example of LFIB with the elimination action.]

6.2. DA-T-PE processing clarifications

The PW-based DetNet data plane solution overloads the T-PE with a DetNet Edge Node function. Such T-PE is referred as DA-T-PE and implies the T-PE is also aware of DetNet flows and may need to operate upon those. Figure 8 illustrates the overall DA-T-PE device functions. The figure shows both physical attachment circuit (AC) (e.g., Ethernet [RFC4448]) connecting to the PE, and a packet service connecting to the PE via an embedded label switching router (LSR) function [RFC6658]. Whether traffic flow from from a client AC and PSN LSP receives DetNet specific treatment is up to a local configuration and policy. A DA-T-PE can also serve as a normal T-PE.

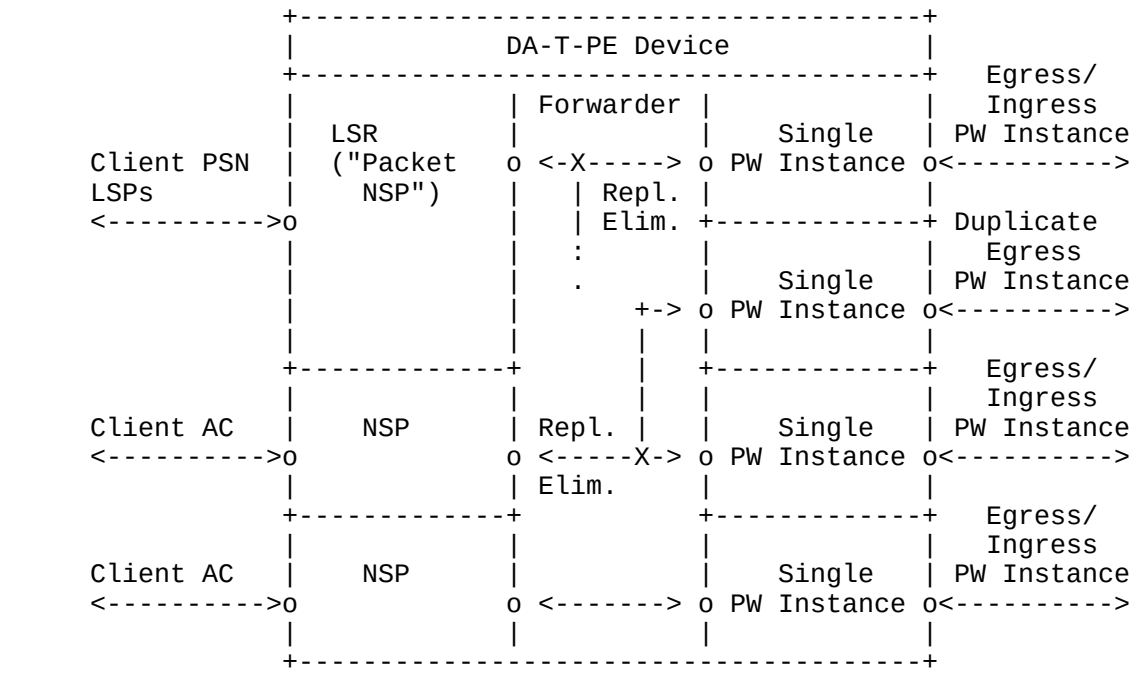

Figure 8: DetNet Edge Node as a DA-T-PE

A DA-T-PE 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 PW encapsulation and logic entirely, and thus is able to operate over unmodified PW implementation and deployment. The NSP approach works only between DA-T-PEs and cannot make use of DA-S-PEs (see Section 6.3).

The DetNet-aware extended forwarder selects the egress segment PW based on the rules described in [RFC4448] and [RFC6658]. In both "normal AC" and "Packet AC" cases there is no DetNet encapsulation header available yet as it is the case with DA-S-PEs (see Section 6.3). It is the responsibility of the extended forwarder within the DA-T-PE to push the DetNet encapsulation header (i.e., the DetNet CW and the DetNet flow-ID) to the packet before forwarding it to the appropriate egress PW instance(s). The extended forwarder MAY copy the sequencing information from the native packet into the DetNet CW 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 flow-ID).

6.3. DA-S-PE processing clarifications

The PW-based DetNet data plane solution overloads a S-PE with a DetNet Relay function. Such S-PE device is referred as DA-S-PE and implies the S-PE is also aware of DetNet flows and may operate upon those. Figure 9 illustrates the overall DA-S-PE device functions.

A DA-S-PE participates to the packet replication and duplication elimination. This processing is done within an extended forwarder function. Whether an ingress PW receives DetNet specific processing depends on how the LFIB is programmed. For some PWs the DA-S-PE can act as a normal S-PE and for some apply the DetNet specific processing. It is also possible to treat the DA-S-PE as a P router using the L-label tunnels. Again, this is entirely up to how the LFIB has been programmed.

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

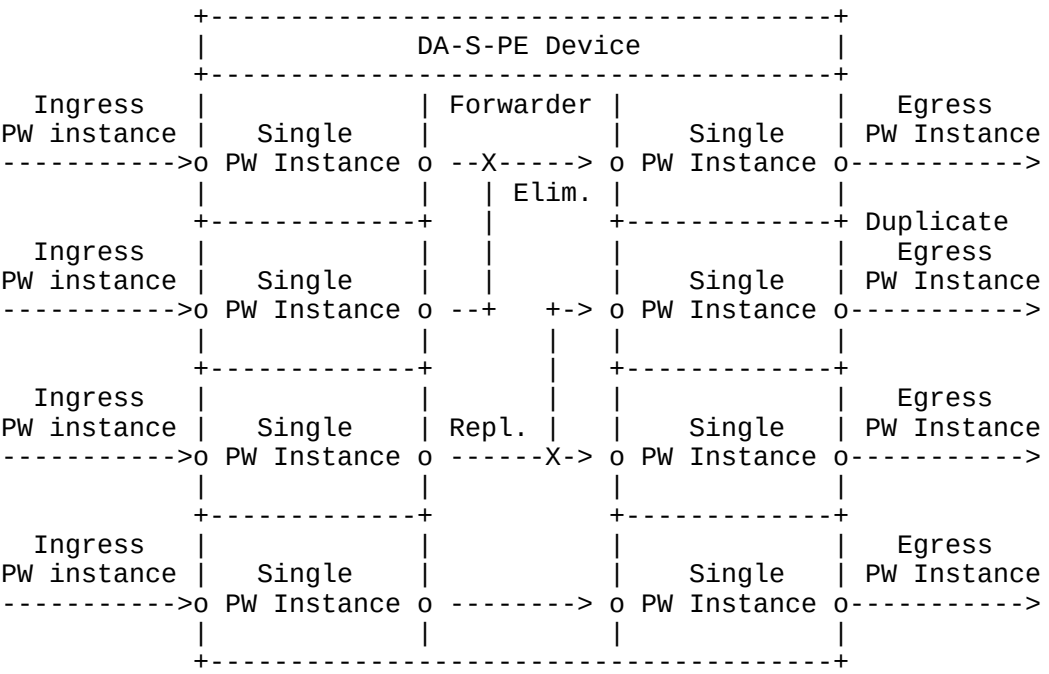

Figure 9: DetNet Relay Node as a DA-S-PE

7. Other DetNet considerations

7.1. Class of Service

[Editor’s note: Discuss the CoS.. and how that is archived when using MPLS or IP PSN.]

7.2. Quality of Service

[Editor’s note: Elaborate the QoS issues here..]

7.3. 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 a clock synchronization method is used, such as the Precision Time Protocol

(PTP) [IEEE1588]. In this case, there are a few possible approaches of how synchronization protocol packets are forwarded and handled by the network:

- PTP packets are sent as a normal DetNet flow: 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.

- PTP with on-path support: in this approach PTP packets are sent 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.

- 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 three 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.

7.4. Bidirectional traffic

Some DetNet applications generate bidirectional traffic and may require symmetric flows. There are already mechanisms that can be used to create bidirectional tunnels at the transport network level, such as MPLS-TP. The data plane solution SHOULD allow establishing bidirectional symmetric flows. Control plane mechanisms would need to also support this, though this is out of scope of this document. [Summary of existing mechanisms to create bidirectional tunnels that can be used.]

8. Control plane considerations

[Editor's note: discuss here what kind of enhancements are needed for DetNet and specifically for PREF.]

8.1. PW Label assignment and distribution

The PW label distribution follows the same mechanisms specified for MS-PW [RFC6073].

9. 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.

10. IANA Considerations

TBD.

11. 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

12. References

12.1. 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>.

[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>.

[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>.

[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>.

12.2. Informative 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.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>.

Appendix A. Example of DetNet data plane operation

[Editor’s note: 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..]

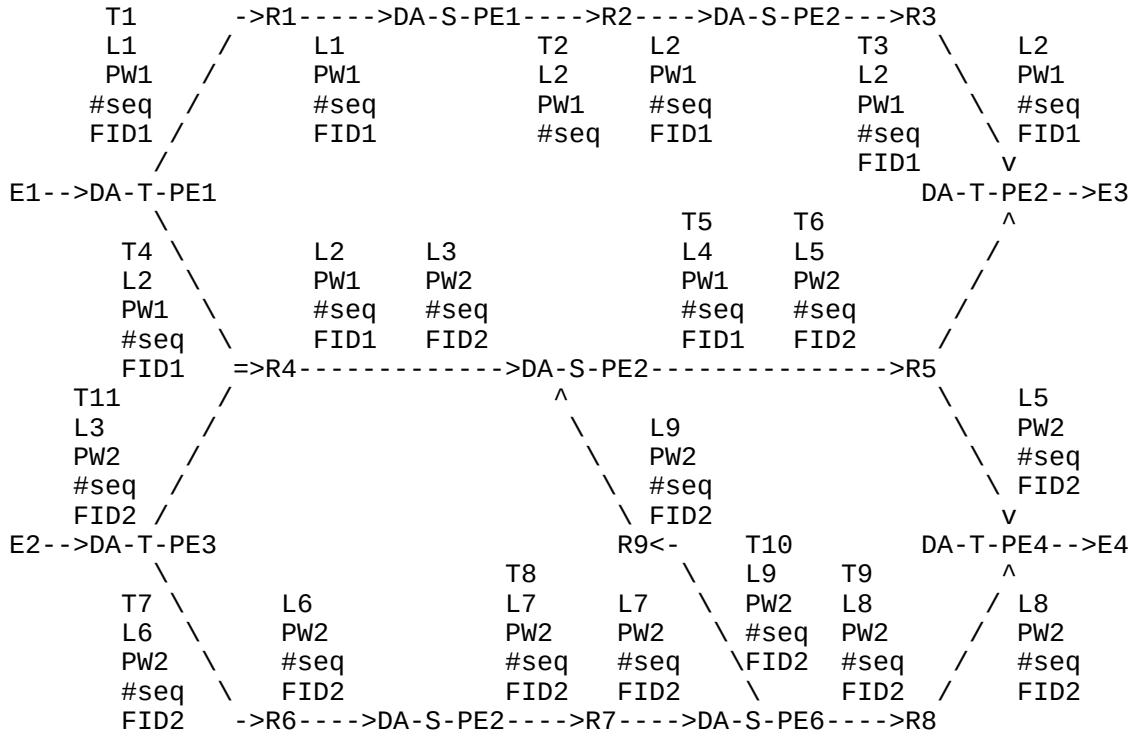

Figure 10: Replication and elimination example

[Editor’s note: the LFIB Figure 11 content to be updated.]

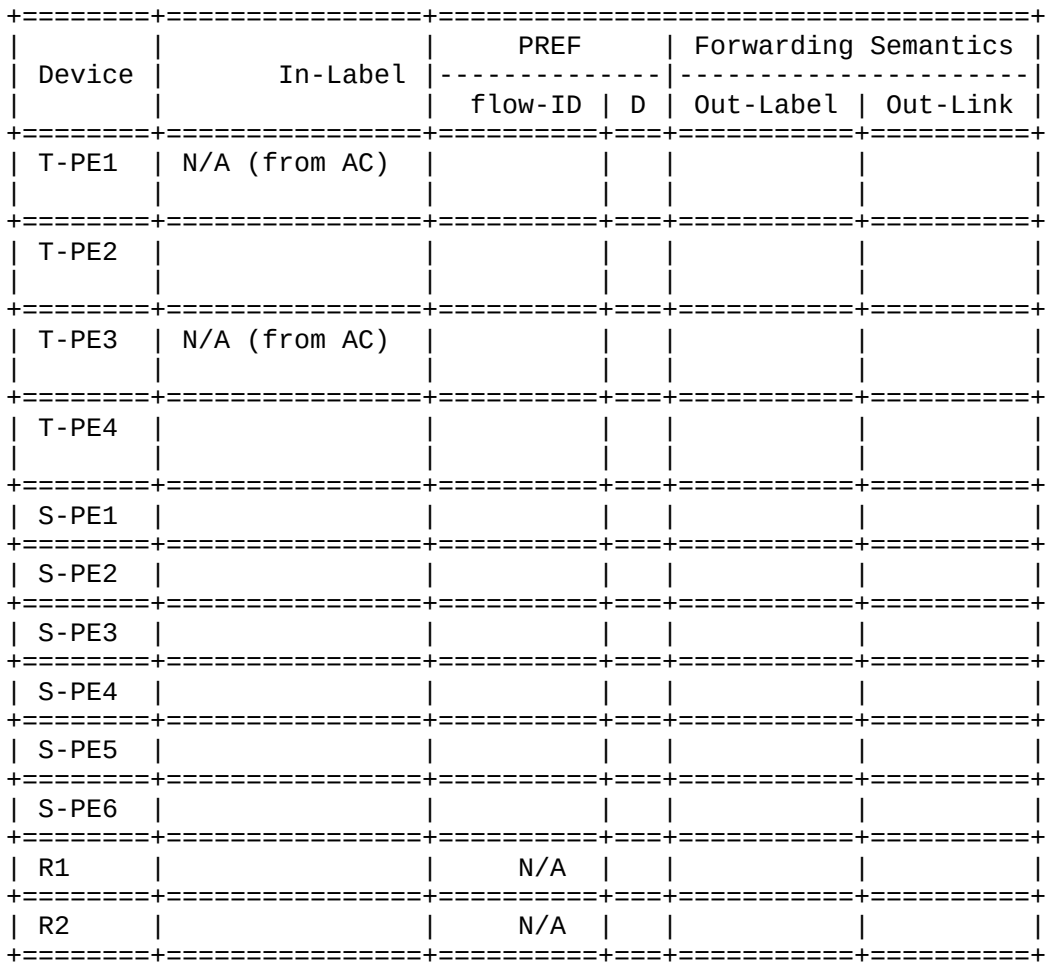

Figure 11: LFIB contents

Appendix B. Example of pinned paths using IP PSN

Authors' Addresses

Jouni Korhonen (editor)
Broadcom
3151 Zanker Road
San Jose, CA 95134
USA

Email: jouni.nospam@gmail.com

Loa Andersson
Huawei

Email: loa@pi.nu

Yuanlong Jiang
Huawei

Email: jiangyuanlong@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

DetNet
Internet-Draft
Intended status: Standards Track
Expires: September 13, 2017

J. Farkas
B. Varga
Ericsson
R. Cummings
National Instruments

March 12, 2017

DetNet Flow Information Model Based on TSN
draft-farkas-detnet-flow-information-model-00

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 September 13, 2017.

Table of Contents

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 revision 00, 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.

Revision 00 is just a start; it is not complete. As this revision heavily relies on [IEEE8021Qcc], the need for further DetNet specific aspects is to be reviewed and missing pieces are to be added.

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.

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.

- Names SHOULD be descriptive.
- Names MUST start with uppercase letters.
- 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

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 three high level groups:

- 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.

- 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.

- 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:

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

- 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 MaxPacketSize (Section 6.2) has been changed.]]

As the DetNet UNI can provide both L3 and L2 services, 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 TSN flows are specified by [IEEE8021CB], which also includes IP flow identification, see, e.g., Table 6-1 in Clause 6. Therefore, the flow identification specified by [IEEE8021CB] is also applicable to DetNet flows.

[[NOTE (to be removed from a future revision): Extensions to the options specified by [IEEE8021CB] can be discussed.]]

DataFlowSpecification specifies DetNet flows as follows; see Section 6.1.1 for DetNet L3 flows and Section 6.1.2 for DetNet L2 flows.

6.1.1. DetNet L3 Flow Identification and Specification

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

a. SourceIpAddress

b. DestinationIpAddress

c. Dscp

d. Protocol

e. SourcePort

f. DestinationPort

g. MplsLabel

6.1.2. DetNet L2 Flow Identification and Specification

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

a. DestinationMacAddress

b. SourceMacAddress

c. Pcp

d. VlanId

[[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:

- 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.

]]

6.2. 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.

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 false value is more important than the true value (i.e., flows with true are dropped first).

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.

[[NOTE (to be removed from a future revision): Sub-Interfaces to be added, e.g., based on IfIndex.]]

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 10.2) of the Status object (Section 10).

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).";.

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

a. VlanTagCapable (Customer VLAN Tag capable)

b. CB-Capable (frame replication and elimination capable)

c. CB-StreamIdenTypeList (a list of the optional Stream Identification types supported by the interface as specified in [IEEE8021CB].)

d. 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 10.1).

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 10) 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. 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 10.1)

c. AccumulatedLatency (this section below)

d. InterfaceConfiguration (Section 10.2)

e. FailedInterfaces (Section 10.3)

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 AccumlatedLatency 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.

10.1. 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.]]

10.2. 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)

10.3. Failed Interfaces

FailedInterfaces provides a list of one or more physical interfaces (ports) in the failed node when a failure occurs in network configuration (i.e., non-zero FailureCode in StatusInfo object (Section 10.1)).

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.

11. 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 of the document is just to start the discussions; further work is needed.

12. IANA Considerations

N/A.

13. Security Considerations

N/A.

14. References

14.1. 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>.

14.2. Informative References

[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

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: September 14, 2017

T. Mizrahi
MARVELL
E. Grossman, Ed.
DOLBY
A. Hacker
MISTIQ
S. Das
Applied Communication Sciences
J. Dowdell
Airbus Defence and Space

March 13, 2017

Deterministic Networking (DetNet) Security Considerations
draft-sdt-detnet-security-00

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 4.

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/.

Copyright Notice

Table of Contents

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 othewise 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.

This initial version of this draft consists of a threat model and analysis, and in the future will be expanded to include mitigation strategies.

This draft also provides context for the DetNet security considerations by collecting into one place Section 4 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".

The DetNet technologies include ways to:

- 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
- Provide explicit routes for DetNet flows that do not rapidly change with the network topology
- 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

2. Abbreviations

IT Information technology (the application of computers to store, study, retrieve, transmit, and manipulate dataor 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, Discoverablility.

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 is briefly presented in this subsection.

The model classifies attackers based on two criteria:

- 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.

- 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.

3.2. Threat Analysis

3.2.1. Threats Related to 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. Threats Related to DetNet Flow Identification

3.2.2.1. DetNet Flow Modification or Spoofing

An attacker can modify some header fields of en 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. Threats Related to 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. Threats Related to 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:

- 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.

- 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.

- 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. Threats Related to 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. Threats Related to the 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. Threats Related to 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. Threats Related to Time Synchronization Mechanisms

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

3.3. Threat Summary

A summary of the threats that were discussed in this section is presented in Figure 1. For each threat, 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	Internal Inj.	External MITM	External 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. 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.

4.1. Architecture (draft 8)

4.1.1. 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.

4.1.2. 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:

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

4.2. Data Plane Alternatives (draft 4)

4.2.1. Security Considerations (sec 7)

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

4.3. Problem Statement (draft 5)

4.3.1. 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:

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

4.4. Use Cases (draft 11)

4.4.1. (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:

- Network or transport errors (e.g. malformed packets or excessive latency) can cause protocol failure.
- 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.
- Protocol commands may be available that are capable of restarting communications and otherwise interrupting processes.
- Protocol commands may be available that are capable of clearing, erasing, or resetting diagnostic information such as counters and diagnostic registers.
- Protocol commands may be available that are capable of requesting sensitive information about the controllers, their configurations, or other need-to-know information.
- 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.
- Protocol commands may be available that are capable of broadcasting messages to many devices at once (i.e. a potential DoS).

- Protocol commands may be available to query the device network to obtain defined points and their values (i.e. a configuration scan).

- Protocol commands may be available that will list all available function codes (i.e. a function scan).

- These inherent vulnerabilities, along with increasing connectivity between IT an OT networks, make network-based attacks very feasible.

- 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.

4.4.2. (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:

- Integration of distributed energy resources

- Proliferation of digital devices to enable management, automation, protection, and control

- Regulatory mandates to comply with standards for critical infrastructure protection

- Migration to new systems for outage management, distribution automation, condition-based maintenance, load forecasting, and smart metering

- 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:

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

- IP enables a rich set of features and capabilities to enhance the security posture
- 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.

4.4.3. (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.

4.4.4. (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.

4.4.5. (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.

4.4.6. (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.

5. IANA Considerations

This memo includes no requests from IANA.

6. Security Considerations

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

7. Informative References

[ARINC664P7]
ARINC, "ARINC 664 Aircraft Data Network, Part 7, Avionics Full-Duplex Switched Ethernet Network", 2009.

[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-11 (work in progress), October 2016.

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

[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@mistiqtech.com
URI: http://www.mistiqtech.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

DetNet
Internet Draft
Interned status: Standards Track
Expires: June 22, 2017

H. Wang
P. Wang
C. Zhang
Y. Yang
Chongqing University of
Posts and Telecommunications

December 19, 2016

Joint Scheduling Architecture for Deterministic Industrial Field/Backhaul Networks
draft-wang-detnet-backhaul-architecture-00

Abstract

Joint scheduling of industrial field network and backhaul network is significant for end-to-end deterministic delay requirements of data flows in factories. This document describes a joint scheduling architecture for deterministic industrial field and backhaul networks. Taking WIA-PA wireless field network and IPv6-based backhaul network as an example, this document shows how the joint scheduling architecture works.

Table of Contents

1. Introduction

Deterministic network is an essential element of the industrial network. Using deterministic network in the industrial field can enhance the network performance and greatly reduce the network packet loss. Thus, it is the future development direction of industrial network technology to use deterministic networks in the whole industrial network. Deterministic networks in industrial networks are mainly concentrated on the industrial field networks, such as ISA100.11a[IEC62734], WirelessHART[IEC62591] and WIA-PA[IEC62601], and there is little joint scheduling scheme that can be applied to industrial networks.

Nowadays, in the use case document[draft-bas-usecase-detnet] and architecture document[draft-finn-detnet-architecture] submitted by the IETF DetNet working group, a deterministic network based on Ethernet has already been researched. The document proposes a network architecture based on SDN technology, which can accurately control the transmission of data streams. However, the document does not consider the characteristics of the industrial backhaul networks and the actual situation of other industrial field deterministic networks. First of all, the data flow of industrial backhaul network is highly sensitive to the uncertainty of time. Therefore, it is very important that how to apply the deterministic networks based on Ethernet to industrial backhaul networks. Secondly, the existing deterministic networks in the industrial field have been widely deployed in the factory, and Deterministic network technology is already very mature, and direct replacement will consume a lot of manpower and material resources.

Based on existing work in the architecture document[draft-finn-detnet-architecture], this document proposes a joint scheduling architecture for deterministic industrial field networks. This framework will firstly replace the industrial backhaul networks and other non-deterministic networks of industrial networks into deterministic Ethernet-based network, and then on the basis of SDN technology, this document proposes a joint scheduler, which can be used for joint scheduling on other deterministic networks in deterministic Ethernet-based network and industrial field network. Through deploying the deterministic network throughout the industrial network based on the joint scheduling architecture, it can realize the end-to-end deterministic scheduling between different industrial field networks, and ensure data stream indicators as well as save manpower and material resources.

2. Joint Scheduling Architecture

For industrial networks, there are many network controllers in the network, which together constitute the control plane for the whole industrial network. The control plane is very important in the entire network, especially when it comes to cross domain transfer of time-sensitive data. So the control plane architecture will greatly affect the performance of the network, therefore it is becoming a research hotspot on how to give full play to the performance of their respective networks when the multiple controllers are in the joint cooperation. However, there is not a unified standard of joint architecture of multiple controllers in the industry at present. The main frameworks are the following two kinds: the distributed architecture and the centralized architecture. The WIA-PA network, which is the typical of WSNs standards which has become an

international standard for industrial field networks approved by IEC, is used as an example to illustrate these architectures.

2.1. Distributed Architecture

Distributed architecture is also known as East-West architecture. In the architecture, the status of all network controller is equal, these controllers are connected to each other to form an unstructured network, and achieve cross domain transfer task deployment through the mutual transmission of information, as shown in Figure 1.

In the distributed architecture, the controller can exchange different network topologies and the accessibility of information through the east-west interface, and each controller can build a global network topology. In the access to the global network topology, since each controller is equal, it can serve as a server role at the same time, as well as has the service capacity of starting deterministic cross-network transmission.

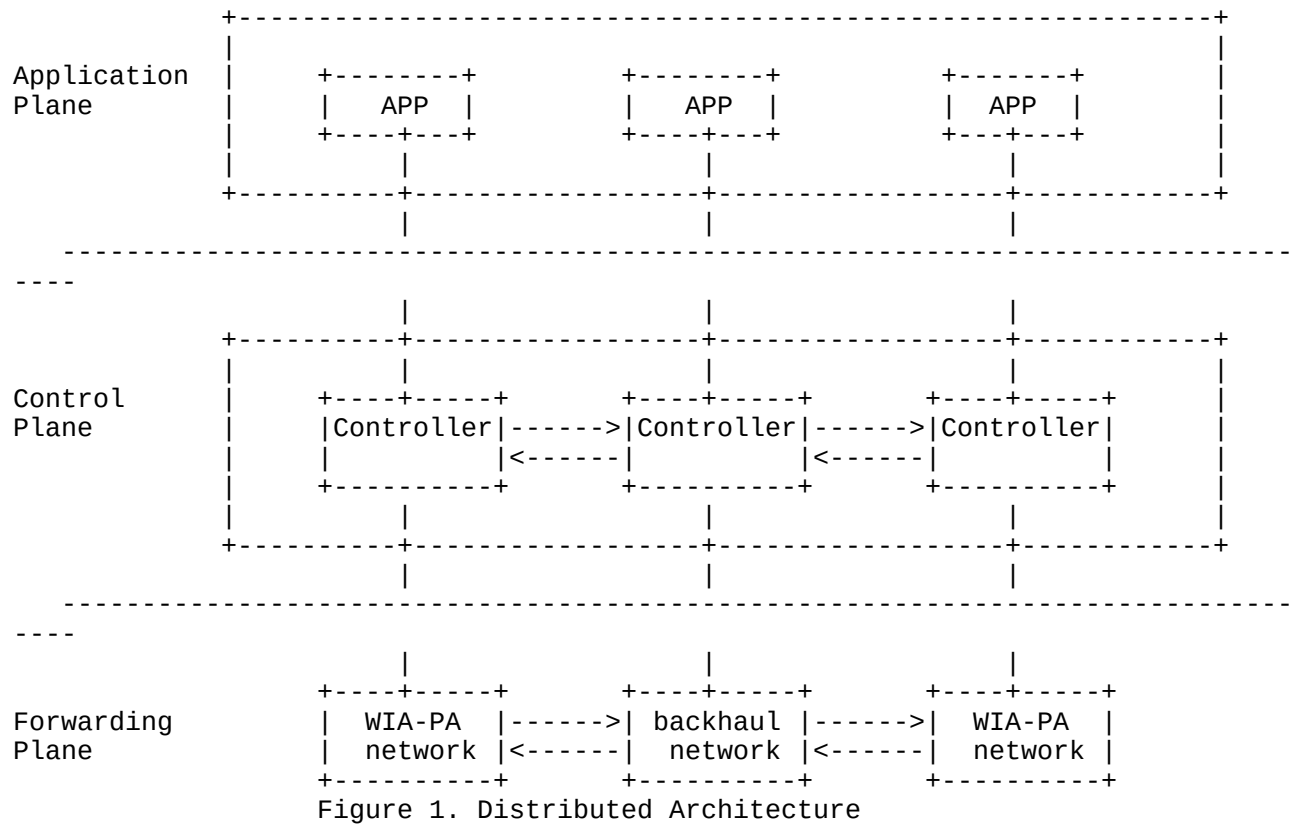

Figure 1. Distributed Architecture

2.2. Centralized Architecture

Centralized architecture is also known as vertical multi-level architecture. In this architecture, the control plane is divided into two parts, one is the basic control plane composed of a variety of network controllers; another part is a network controller composed of the main controller, which is responsible for controlling the basic control plane, as shown in Figure 2.

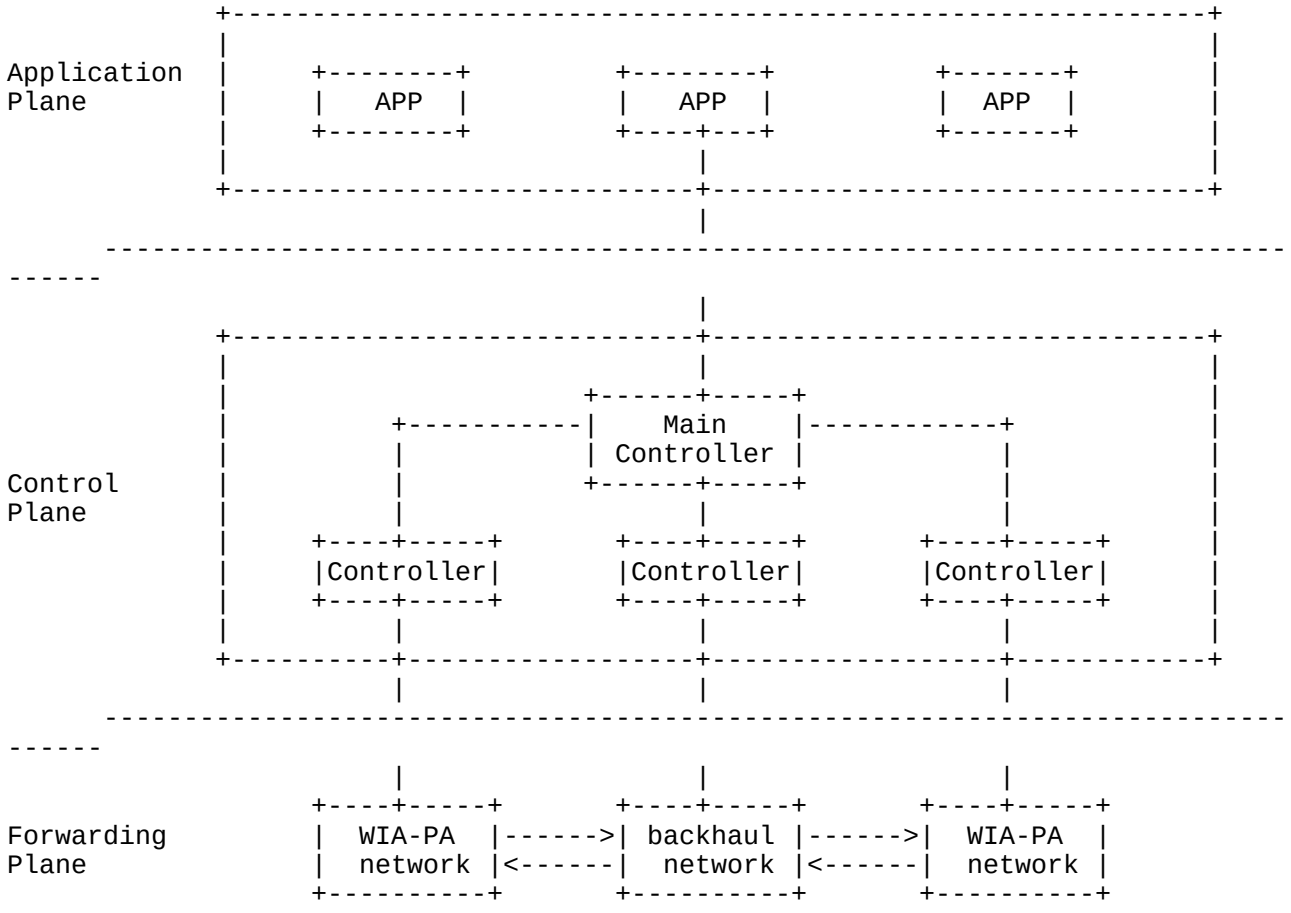

Figure 2. Centralized Architecture

The centralized architecture needn’t to expand the east-west interface. It only needs to establish a connection with the basic controllers through the southbound interface. After the connection

is established, the main controller obtains the every domain network topology through the API interface provided by the basic controllers, and storages global network topology on its own. It can also assign tasks to basic controllers through the API interface.

2.3. Joint Scheduling Architecture

In the practical application, distributed architecture not only needs to extend the east-west interface, but also maintains a global network topology in each controller. Only each controller maintains such a global network topology, it can ensure the deterministic control of the control plane for the whole network.

Though the centralized architecture does not have the above requirements, for the deterministic industrial network, the scale of the network is not very large, in the industrial backhaul network, a single SDN controller is sufficient to meet the control demands of industrial backhaul network. If centralized architecture is directly applied to an industrial network, it will not only be unable to give full play to the advantages of the architecture in multi controllers collaboration, but also cause meaningless information interaction between the controllers, which will waste network resource.

In view of the problems existing in these two architectures, this document takes the WIA-PA network as an example and proposes a joint scheduling architecture based on the architecture document[draft-finn-detnet-architecture]. The architecture is optimized according to the characteristics of deterministic industrial network, so that a single SDN controller can unite the WIA-PA network systems manager to manage the entire industrial network, and provide support for the deterministic scheduling of data streams across network transmission through industrial backhaul network located in different domains of WIA-PA network.

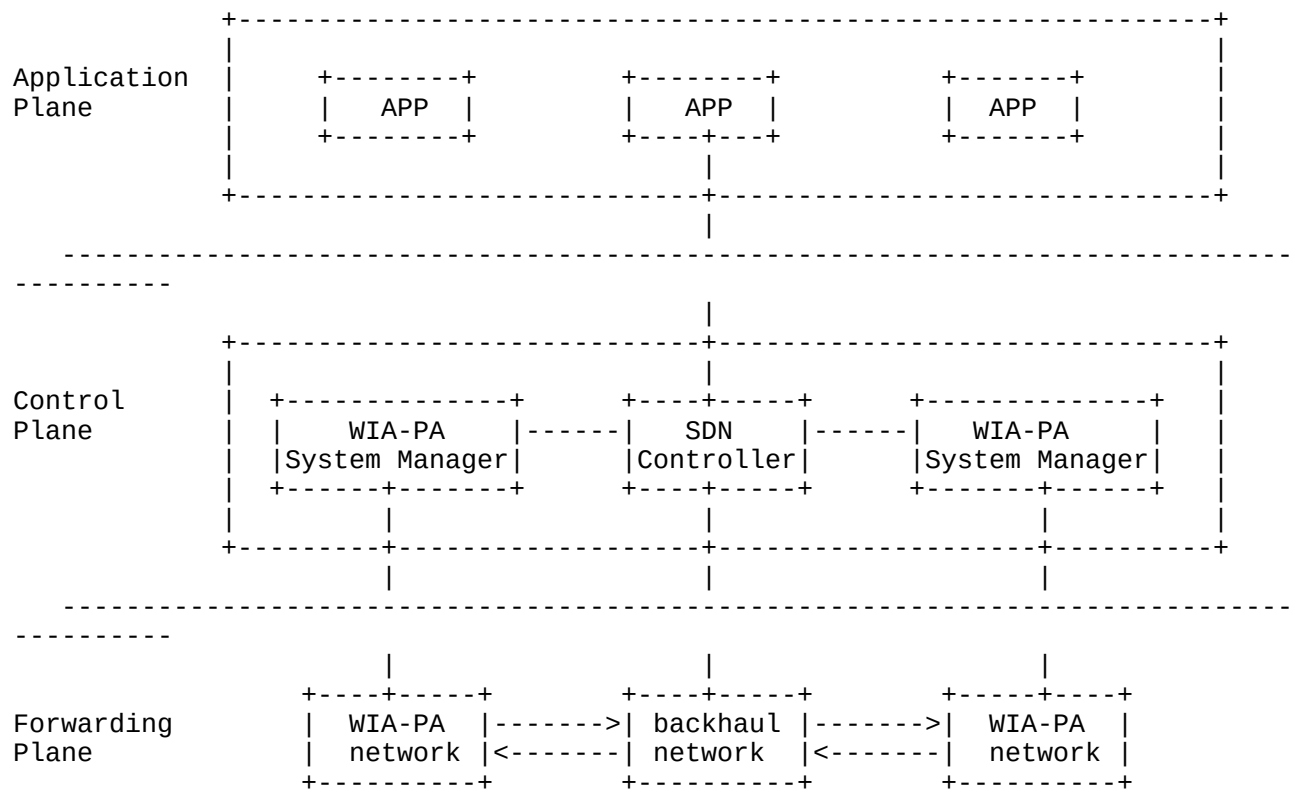

Figure 3. Joint scheduling architecture

As shown in Figure 3, joint scheduling architecture can be mainly classified into three planes:

- Forwarding plane: this plane contains various types of network equipment in different networks. It is the physical entities of the network transmission. In general, to achieve the desired network functions for the network manager, these devices are specific factors of management control operation, which makes their own resources abstract for their own control elements to manage and configure.

- Control plane: this plane is formed by the WIA-PA System Manager and the SDN controller. Joint scheduler is integrated into the SDN controller in the form of plugin, and other WIA-PA System Managers accept joint management scheduler by establishing a connection with the SDN controller. Meanwhile, inside the SDN controller, joint scheduler achieves the management of industrial backhaul network by directly calling the corresponding module of SDN controller.

o Application plane: this plane provides users with a unified interface about a variety of resources for the whole network. At the same time, it also provides users with an intuitive, user-friendly interface, which can shield the complex network information of the original.

Joint Scheduling Architecture defines an architecture that when industrial networks contain other deterministic networks, these deterministic networks and deterministic Ethernet-based networks are jointly scheduling. On the basis of this architecture, control and scheduling for the entire industrial network can be realized by joint scheduler, so as to provide a real-time protection for each data stream.

3. Joint Scheduling Scheme

Taking WIA-PA wireless field network and IPv6-based backhaul network as an example, this section shows how the joint scheduling architecture works. Existing WIA-PA scheduling scheme only applies to WIA-PA field network. Scheduling scheme will fail once the data is transferred to backhaul networks. Joint scheduling scheme is innovation and expansion of WIA-PA scheduling scheme.

Firstly, scheduling scheme based on SDN in industry backhaul network is added to the original scheduling scheme, so that data can flow in the industrial backhaul network, and the data can be identified and assigned existing backhaul network resource according to their requirements for the network resources.

Secondly, conducting an optimization for original WIA-PA scheduling scheme enables scheduling scheme based on WIA-PA networks plays together joint scheduler, and scheduling scheme can simultaneously apply to two non-adjacent domains so that it can be adapt to the cross-border joint operation based on SDN.

Thirdly, due to the specificity of cross-border transmission services, the joint scheduling scheme for WIA-PA network VCR_ID and Route ID is reclassified.

Finally, since the system manager allocates a short address to the field device on the basis of the network address information about its own domain in WIA-PA networks. Thus resulting in the entire network short address field device is uncertain. In order to identify the field device on different network domains and domain, the network identifier (PAN_ID) is applied to the joint scheduling scheme to identify WIA-PA network.

After the SDN controller initiates joint scheduling module, WIA-PA system manager will actively establish a connection with the united scheduler. After the scheduler receives a cross-border transmission request, joint scheduler will send a request for obtaining topology information and node information to WIA-PA System Manager. Then, the scheduler will assign paths and network resources according to this information by pre-defined scheduling algorithm.

After the routing and network resources have been calculated, joint scheduler will configure and deploy networks by the corresponding network controller.

3.1. WIA-PA Network Joint Scheduling

In the united scheduling process, path deployment and resource allocation for WIA-PA network are performed by calling the WIA-PA network system manager API interface. System manager will query the corresponding information of the field device in the network upon receiving the acquisition command of joint operation for the network information, and then return the received information to the united scheduler. The system manager will configure communication resources for the corresponding gateway device, routing equipment and field equipment if the system manager receives configuration commands from joint scheduler. After receiving a successful response, it will send a successful reply to the united scheduler.

3.2. Protocol Conversion

In the process of cross-border transmission, since industrial backhaul network is different from WIA-PA network, which is not an IP-based Ethernet. Protocol conversion of gateway for WIA-PA packet is needed when the data of WIA-PA network needs to transmit to another network through cross-border industrial backhaul. Meanwhile, according to the joint scheduling scheme, SDN controller is able to identify the WIA-PA Ethernet data stream, and allocate resources according to the data stream type and level of the data stream. Therefore, in the protocol conversion process of gateway, scheduling and control of WIA-PA data flow can be realized by SDN controller unless the VCR of WIA-PA data stream and the priority are filled in the IPv6 header.

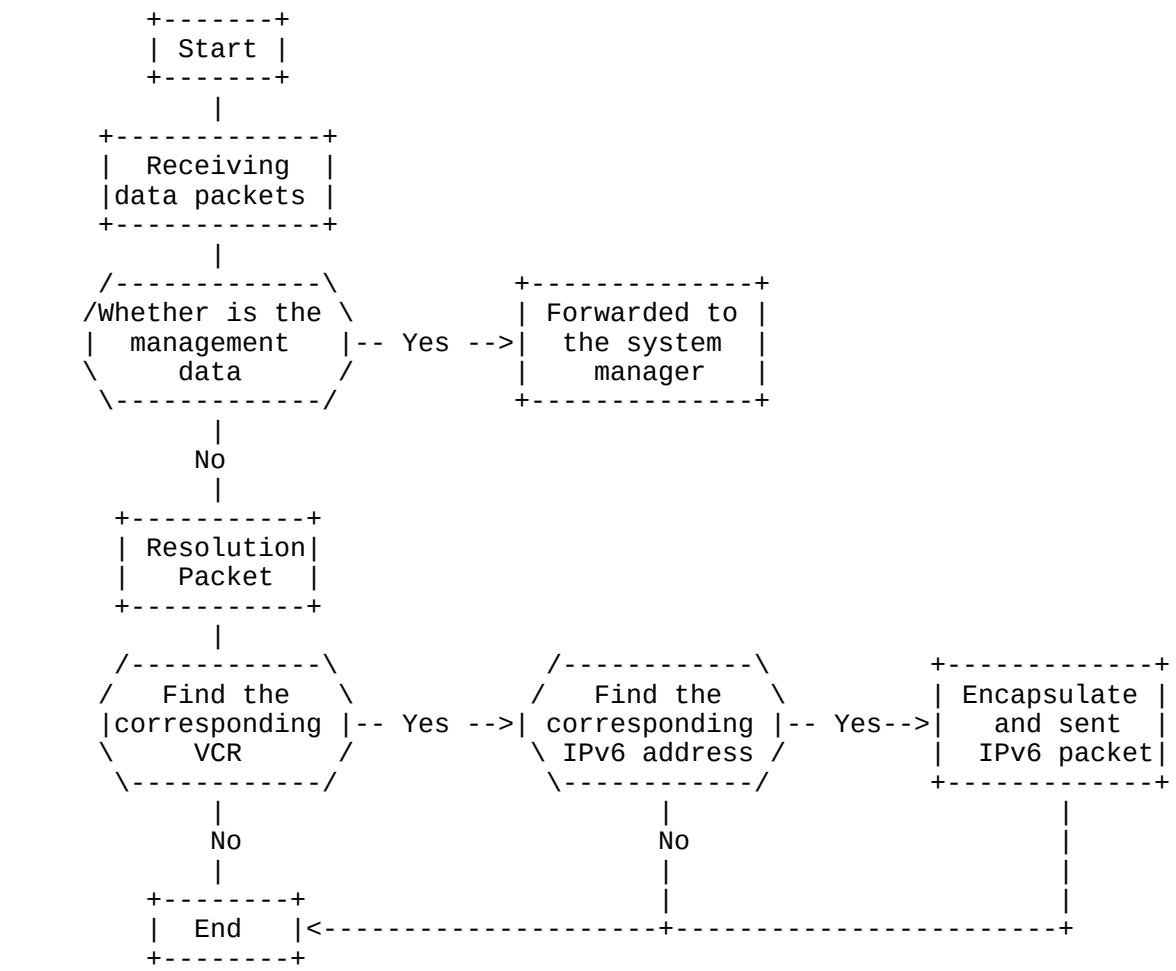

Figure 4. The conversion process of gateway protocol

As shown in Figure 4, according to the above section, the gateway will receive the address mapping of joint scheduler configure when configuration WIA-PA network. After that, VCR tables and IPv6 address-mapping tables will be formed according to this information. When the gateway receives WIA-PA packets, it will firstly parse out Route ID, Object ID and Instance ID, and find corresponding VCR from VCR tables. Meanwhile, the gateway finds the corresponding IPv6 address according to Route ID in IPv6 address mapping table. Then, the gateway begins to encapsulate WIA-PA packets based on IPv6 format, fill VCR_ID in IPv6 header flow label field, and fill the priority of WIA-PA packet in communication category of IPv6 header

fields, zero is used to fill up insufficient bytes. Then, the protocol conversion for WIA-PA data is completed.

When the gateway receives IPv6 packets from the industrial backhaul networks, the gateway will make out VCR_ID from IPv6 packet header, and find packets VCR in the domain WIA-PA network according to the VCR ID in its own maintenance VCR table, and replace it with the information of original packet. Then, the protocol conversion for IPv6 packet is completed.

3.3. Industrial Backhaul Network Scheduling

In deterministic network based on SDN, joint scheduler can recognize WIA-PA data stream through matching on IPv6 flow label field. According to priority of IPv6 and VCR_ID type, joint scheduling can allocate the necessary resources to communication, and ensure that the key data flow is not affected when adding new data flow in the existing network. It can also monitor the real-time data flow of the network. To protect critical data flows from affected, switching paths is also considered when necessary. The scheduling process of industrial backhaul network is shown in Figure 5.

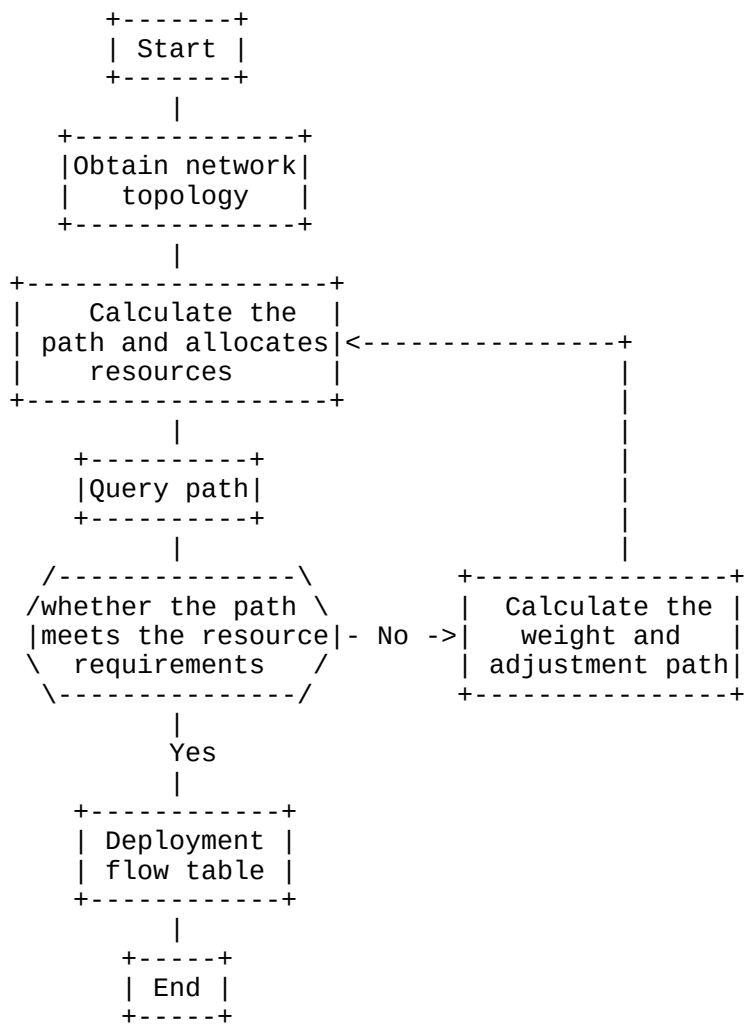

Figure 5. The scheduling process of Industrial backhaul network

After receiving the request for service, the joint scheduler will calculate the route information and network resource allocation. Once the path information and resource allocation are determined, joint dispatcher will confirm whether the resource path is capable of meeting business requirements through the inside module of SDN controller. If it meets business requirements, then the flow table is deployed by SDN controller. Otherwise, the path information and resource allocation are recalculated to choose the other paths to transmit data flow.

4. Security Considerations

5. IANA Considerations

This memo includes no request to IANA.

6. References

6.1. Normative References

6.2. Informative References

[IEC62734]
ISA/IEC, "ISA100.11a, Wireless Systems for Automation, also IEC 62734", 2011, <http://www.isa100wci.org/enUS/Documents/PDF/3405-ISA100-WirelessSystems-Future-brochWEB-ETSI.aspx>.

[IEC62591]
IEC, "Industrial Communication Networks - Wireless Communication Network and Communication Profiles - WirelessHART - IEC 62591", 2010, <https://webstore.iec.ch/p-preview/info_iec62591%7Bed1.0%7Den.pdf>

[IEC62601]
IEC, "Industrial networks - Wireless communication network and communication profiles - WIA-PA - IEC 62601", 2015, <https://webstore.iec.ch/preview/info_iec62601%7Bed2.0%7Db.pdf>

[I-D.finn-detnet-problem-statement]
Finn, N. and P. Thubert, "Deterministic Networking Problem Statement", draft-finn-detnet-problem-statement-04 (work in progress), October 2015.

[I-D.finn-detnet-architecture]
Finn, N., Thubert, P., and M. Teener, "Deterministic Networking Architecture", draft-finn-detnetarchitecture-03 (work in progress), March 2016.

[I-D.bas-usecase-detnet]
Kaneko, Y., Toshiba and Das, S, "Building Automation Use Cases and Requirements for Deterministic Networking", draft-bas-usecase-detnet-00 (work in progress), April 2016.

Authors' Addresses

Heng Wang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6248-7845
Email: wangheng@cqupt.edu.cn

Ping Wang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6246-1061
Email: wangping@cqupt.edu.cn

Chang Zhang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6246-1061
Email: zc910522@126.com

Yi Yang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6246-1061
Email: 15023705316@163.com

DetNet
Internet Draft
Interned status: Standards Track
Expires: May 19, 2017

H. Wang
P. Wang
H. Yang
Chongqing University of
Posts and Telecommunications
November 15, 2016

Joint Real-Time Scheduling Methods for Deterministic Industrial Field/Backhaul Networks
draft-wang-detnet-joint-scheduling-00

Abstract

In industrial field/backhaul networks, the joint real-time scheduling method is important to keep end-to-end data streams meeting the deadline. This document proposes four joint scheduling methods, the four methods consider time slotting the industrial backhaul network, regarding industrial backhaul network as a black box system, ignoring delay of industrial backhaul and establishing latency model of an industrial backhaul network.

Table of Contents

1. Introduction

Industrial field network is a network that can be deployed in industrial process and monitor industrial field equipment and systems to achieve the target of control and management. It can improve production efficiency, reduce human intervention to industrial production process and decrease the cost of production. It has significant importance for industrial modernization.

Industrial field bus and industrial ethernet are two kinds of common solutions to industrial automation with the development of industrial field network, however they are both wired network. If they can combine the technology of wireless sensor network, a new network, industrial wireless network, can free from being bonded to wires and cables, and is more easy and flexible to deployment. Industrial wireless network is a communication network which is oriented toward building automation, and process automation, and industrial automation. There are three major international standards (ISA100[ISA100.11a], WirelessHART[WirelessHART],WIA-PA[WIA-PA]) in the area of industrial wireless network currently.

Industrial backhaul network is a transition network, which combines industrial field network with higher level network to achieve the goal of interconnection. It mainly solves the problem of access of industrial field network data to higher level network. Industrial field network is generally limited to a specific region, such as a plant. By this network, transaction data of industrial field network can be transferred to internet or other industrial field networks. Industrial backhaul network is a medium-sized network, which can cover from a few kilometers to tens of kilometers. The major technology of industrial wireless backhaul network consists of Wi-Fi, WiMAX and LET.

In order to adapt the presentation and development of industry 4.0, which is aimed to elevate the level of manufacturing, industrial field network should not be confined to a plant network only. Therefore, it is necessary to introduce the technology of industrial backhaul network to break the restrictions of interconnection between different networks, and to form a mixed network of industrial field network and backhaul network. Figure 1 indicates a typical network architecture of the mixed network. It is a type of deterministic network, and had been illustrated about use cases and architecture in the drafts proposed by DetNet Workgroup of IETF of draft-bas-usecase-detnet-02 and draft-finn-detnet-architecture-04.

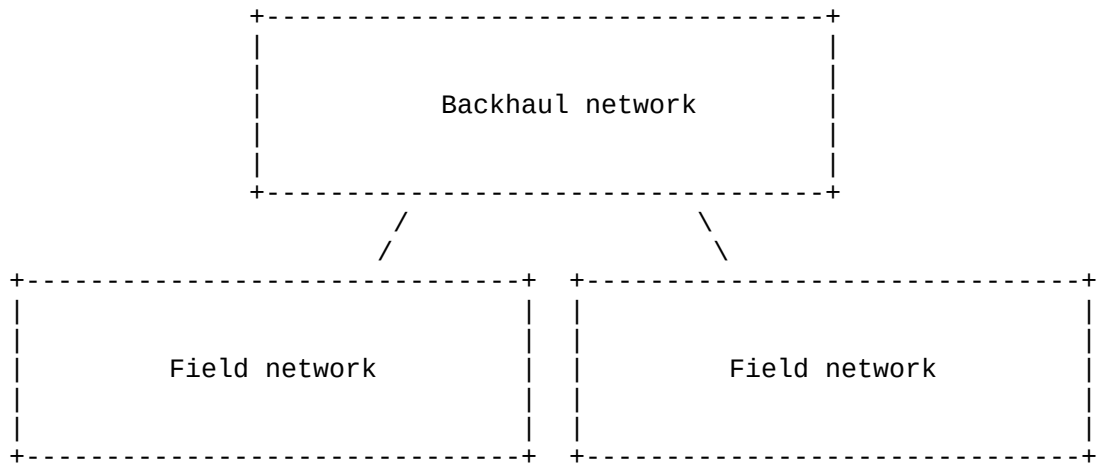

Figure 1. Typical industrial field-backhaul network

In this mixed network architecture of industrial field network and backhaul network, field network is made up of ISA100, which is industrial wireless sensor network protocol. In the former network, a node deployed in a plant can communicate with a node deployed in another plant through a backhaul network.

2. Deterministic industrial field-backhaul network requirement

The draft of draft-finn-detnet-problem-statement put forward by DetNet Workgroup of IETF had described the requirement of deterministic network and deterministic scheduling partially. Because industrial field network directly faces the monitoring of industrial process, it is a difference between industrial field network data and general network data. Industrial field network has high demands about the deterministic delay bounds. It will affect the productivity, and even generate industrial accidents, when there are high packet loss and latency in a field network. For instance, real-time monitoring of level measurement and control are required to avoid overfilling of oil tanks that may lead to serious economic loss and environmental threats.

So, it is needed that a deterministic joint scheduling method can guarantee the determination of network data in such a new network architecture.

3. Deterministic Industrial field-backhaul network Joint Scheduling Key Technology

3.1. End-to-end Network Data Stream

In an industrial field network, end-to-end network data stream indicates a complete transmission path that a source device node transfers to a destination device node (common node or gateway). While in an industrial field-backhaul network, it indicates a complete transmission path that a field network source device node transfers through an industrial backhaul to another field network destination device node.

Industrial field-backhaul network data stream have following features:

- o Period. Every data stream in network generates period data.
- o Deterministic. Every data stream in network has a deadline, network scheduling should ensure every data stream arrive at destination node before its deadline.
- o Sequential. A path of an end-to-end network data stream are made up of every two sequential node transmission link. In the process of scheduling, it must be scheduled by the order of sequence of links in the path.

3.2. Network Communication Resource

In the deterministic industrial field networks with backhaul network architecture, schedulable network communication resources are time slot, channel and link. If the backhaul network is SDN architecture, then the SDN controller could schedule the bandwidth and cache of switch. Therefore, bandwidth and cache resources can be included in schedulable network communication resources.

- o Time slot. Time slot is the basic unit in the TDMA based network communications. The length of time slots is settled and is the same in the entire network. Only one packet or ACK can be transmitted in one time slot.
- o Channel. In order to increase network throughput, industrial field network standards provide a number of channels of different frequencies. If the links do not interfere with each other, then we can use different channels to transmit simultaneously.

o Link. Link refers to a direct communication link between one node and another and no intermediate switching nodes. The network data stream is composed of a lot of links. The devices in the industrial field network devices are half-duplex, so the links in the industrial field network are unidirectional.

3.3. Network Time Slot Scheduling

In TDMA-based industrial field network, time is divided into time slots of the same length. One transmission can be conducted in each time slot and the links using different channels to transmit if they do not interfere each other.

In the time slot scheduling process, it will cause link collision when a node arranged to transmit and receive simultaneously, and it will cause channel collision when the same channel is used within a certain range. AS shown in figure 2, the network time slot scheduling process should avoid such collisions.

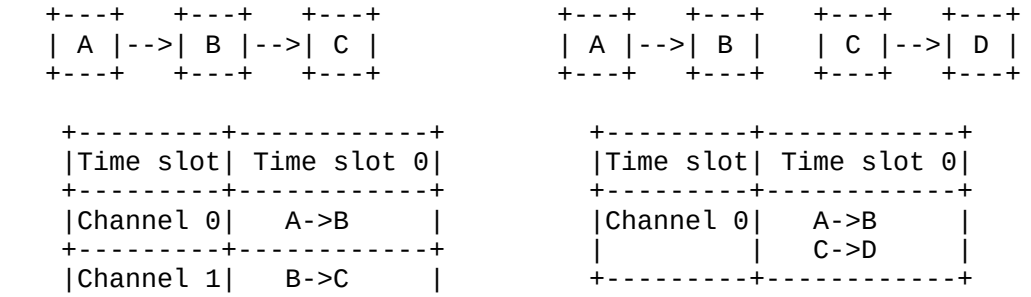

Figure 2. Link Collision & Channel Collision

4. Joint real-Time scheduling methods for deterministic industrial field-backhaul network

Joint real-time scheduling methods for deterministic industrial field/backhaul networks, which cross networks, are intend to solve the deterministic problem of industrial field /backhaul networks. Since the current network infrastructure imports backhaul network, the deterministic scheduling algorithm need to collaborate with backhaul network to conduct joint scheduling to ensure data certainty. The proposal put forward the following solutions.

4.1. Time-Slotted Industrial Backhaul Networks

In order to ensure determinism, industrial field networks utilize TDMA to make the network time-slotted. If the industrial backhaul

network can also be time-slotted, then the deterministic scheduling algorithm can jointly schedule with minor alterations. Industrial backhaul network can be built with a variety of network standards such as Wi-Fi, WiMAX, LTE and so on. But in consideration of the high cost and poor feasibility of time-slotted WiMAX and LTE, we assume that the IEEE802.11 can be time-slotted. Wi-Fi network has various networking modes, such as peer to peer networking mode, point to multi-point networking mode and the relay network mode. Here we consider the hierarchical network constructed in point to multi-point networking mode, as shown in Figure 3.

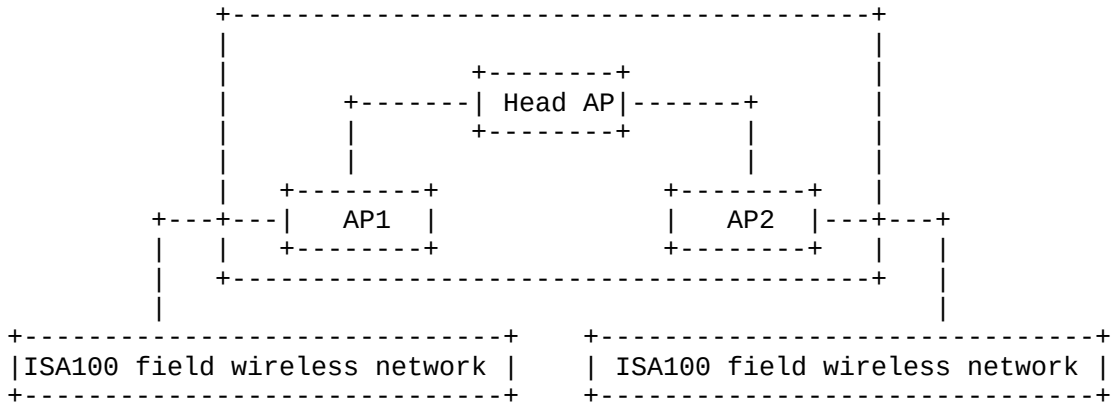

Figure 3. Industrial Backhaul Network consisting of WIFI

Although IEEE802.11 also supports 13 channels, but the AP was not free to switch channels, which means that the AP cannot use a channel in the current time slot and use another channel the next time slot. However, we assume that the network architecture, the following points AP under head AP, which are AP1 and AP2 in FIG 1, can transmit packets simultaneously as long as their transmission task do not contain the same AP, i.e. head AP. For example, when a data stream of field network is transmitting packets to AP1 in a time slot, AP2 is able to receive packets from head AP, or send packets to field network in the same time slot. Therefore, the backhaul network constructed with wireless APs can be considered as a single-channel linear network, which is shown in Figure 4.

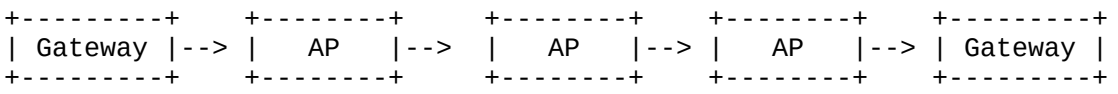

Figure 4. A single-channel linear network

Therefore, the data stream in industrial field/ backhaul network can be deemed to be equivalent to the data stream in field network, only that every piece of data streams need to go through the linear network consisting of wireless APs. So the scheduling process is proposed as follows:

1. Abstract end to end data stream in the entire network, and initialize a different priority for each stream.

2. Establish the delay model of network data stream. If collisions happened between different priority data stream, the low-priority data stream will be delayed by high-priority data stream, so a model can be built under the worst circumstances that the low-priority data streams impacted by higher priority data streams.

3. Estimate the network schedulability. A data stream is schedulable if the minimum time for the data stream to complete transmission, plus the worst delay time caused by higher priority data streams, is less than or equal to deadline, In the current priority allocation scheme, if each data stream is schedulable, the network can be considered as schedulable. If the data stream cannot be scheduled, then change the priority allocation scheme and estimate again until a corresponding scheme is found.

4. Allocate time slot and channel for every data stream. Traverse data streams according to their priority, and each data stream should allocate the next link that is about to be released in each time slot to the greatest extent. According to the rule that low-priority data streams should give way to high-priority data streams, the spare channels can be utilized if there is no collision. However, if collisions happened between data streams of different priority, then the lower-priority data stream should be placed in the next time slot until there are no unallocated higher priority data streams. Follow these rules until the whole network scheduling is completed.

The scheduling process is shown in Figure 5:

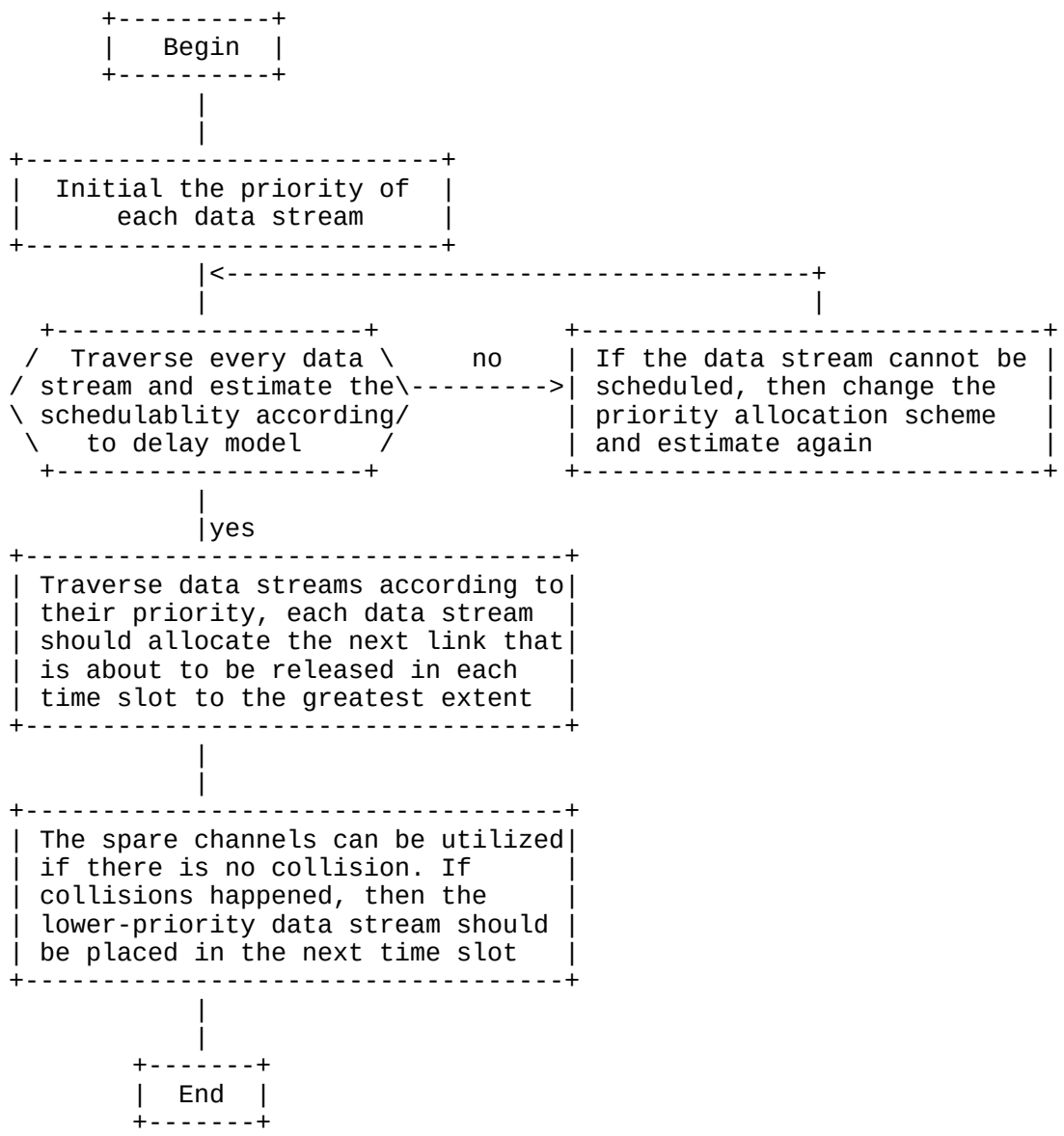

Figure 5. Scheduling of times-slotted industrial backhaul network

4.2. Consider Industrial Backhaul Network as a Black Box

In order to solve the deterministic problem of industrial backhaul network, industrial backhaul can be deemed as white box to conduct fine controls through inner mechanism. While it can also be regarded as a black box so that we can only consider its delay impacts and ignore its internal details.

When the packet goes through the industrial backhaul network, we can give it a timestamp at the application layer and read it after the transmission completed. Then delay caused by the backhaul network can be figured out and a fitting curve of delay can be worked out by collecting large amount of data. It has been verified experimentally that the delay is concentrated in a numerical range despite its randomness. Therefore, we can get the approximate delay of packets caused by the industrial backhaul network.

After that, a few of scheduling paths of different priority can be implemented in the industrial field network. A main scheduling path can be configured according to the average delay of the backhaul network. And some redundant paths should be pre-configured in case the delay of the main path is too high.

The scheduling process of industrial field/backhaul network can be divided into three periods, as shown in Figure 6:

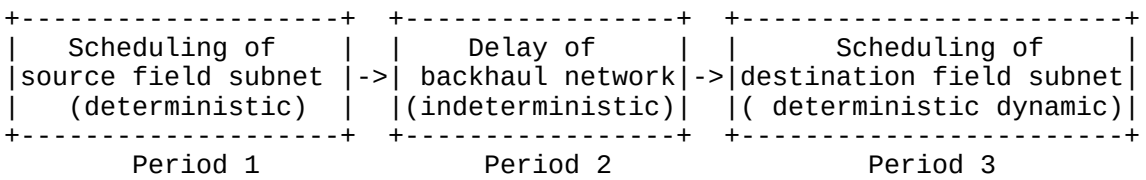

Figure 6. Three periods of scheduling

In source field subnet we can apply the deterministic scheduling algorithm of field network to conduct deterministic polymerization and get the time spent by each data stream to go through the source subnet. Then the data stream goes through the backhaul network, which is a black box and it will cause indeterministic delay which is in a numerical range. When the data stream comes out the backhaul network, the timestamp should be parsed. If the deadline is missed, it indicates that the packet has gone through poor network and need to be retransmitted. If there is time left, scheduling path can be dynamically selected at downward gateway to get the schedulability of the end to end data stream.

4.3. Ignore the Delay of Industrial Backhaul Network

Since the field network is slow-speed (256 KB/s), while industrial backhaul network is a high-speed, if the industrial backhaul networks adopt IEEE802.11, gigabit wireless routers supporting IEEE802.11 ac can make the delay of industrial backhaul network quite small. As a result, the joint deterministic scheduling of the entire network only needs to cover the field network that is located at the ends of the backhaul network.

4.4. Build Delay Model of Industrial Backhaul Network

If industrial backhaul network is built with IEEE802.11, the network access delay test model under IEEE802.11 DCF mode can be established by using Markov chain or queuing theory. At the same time, the model under IEEE802.11 PCF mode can be established based on queuing theory.

Therefore, the field network only need to build the delay model of backhaul network that follows one delay model, then the total transmission scheduling delay will follow certain regularity. The total transmission delay will meet delay requirements with specified probability by scheduling, in other words, the unsuccessfulness of scheduling is acceptable, but the scheduling success rate should be in a range of 90% ~ 95%.

5. Security Considerations

This memo includes no request to IANA.

6. IANA Considerations

This memo includes no request to IANA.

7. References

7.1. Normative References

7.2. Informative References

[ISA100.11a]
ISA/IEC, "ISA100.11a, Wireless Systems for Automation, also IEC 62734", 2011, < http://www.isa100wci.org/enUS/Documents/PDF/3405-ISA100-WirelessSystems-Future-brochWEB-ETSI.aspx>.

[WirelessHART]
www.hartcomm.org, "Industrial Communication Networks - Wireless Communication Network and Communication Profiles - WirelessHART - IEC 62591", 2010.

[WIA-PA]
CN-GB. GB/T 26790.1-2011. Industrial wireless networks WIA specification.Part 1: WIA System architecture and communication specification for process automation (WIA-PA)[S]. China: CN-GB, 2011.

[I-D.finn-detnet-problem-statement]
Finn, N. and P. Thubert, "Deterministic Networking Problem Statement", draft-finn-detnet-problem-statement-04 (work in progress), October 2015.

[I-D.finn-detnet-architecture]
Finn, N., Thubert, P., and M. Teener, "Deterministic Networking Architecture", draft-finn-detnetarchitecture-03 (work in progress), March 2016.

[I-D.bas-usecase-detnet]
Kaneko, Y., Toshiba and Das, S, "Building Automation Use Cases and Requirements for Deterministic Networking", draft-bas-usecase-detnet-00 (work in progress), April 2016.

Authors' Addresses

Heng Wang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6248-7845
Email: wangheng@cqupt.edu.cn

Ping Wang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6246-1061
Email: wangping@cqupt.edu.cn

Hang Yang
Chongqing University of Posts and Telecommunications
2 Chongwen Road
Chongqing, 400065
China

Phone: (86)-23-6246-1061
Email: 18716322620@163.com