DetNet
Internet-Draft
Intended status: Informational
Expires: November 6, 2019

B. Varga, Ed.
J. Farkas
Ericsson
L. Berger
D. Fedyk
LabN Consulting, L.L.C.
A. Malis
S. Bryant
Huawei Technologies
J. Korhonen

May 5, 2019

DetNet Data Plane Framework
draft-ietf-detnet-data-plane-framework-00

Abstract

This document provides an overall framework for the Deterministic Networking data plane. It covers concepts and considerations that are generally common to any Deterministic Networking data plane specification.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Deterministic Networking (DetNet) provides a capability to carry specified unicast or multicast data flows for real-time applications with extremely low packet loss rates and assured maximum end-to-end delivery latency. A description of the general background and concepts of DetNet can be found in [I-D.ietf-detnet-architecture].

This document describes the concepts needed by any DetNet data plane specification and provides considerations for any corresponding implementation. It covers the building blocks that provide the DetNet service, the forwarding sub-layer functions, and the flow identification as described in the DetNet Architecture.

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

As part of the service sub-layer functions, this document describes typical DetNet node data plane operation. It describes the function and operation of the Packet Replication (PRF) Packet Elimination (PEF) and the Packet Ordering (POF) functions within the service sub-layer. It also describes the forwarding sub-layer that is used to eliminate (or reduce) contention loss and provide bounded latency for DetNet flows.

DetNet flows may be carried over network technologies that can provide the DetNet required level of service. For example, DetNet MPLS flows can be carried over IEEE 802.1 Time Sensitive Network (TSN) [IEEE802.1TSNTG] sub-networks. However, IEEE 802.1 TSN support is not required and some of the DetNet benefits can be gained by running over a data link layer that has not been specifically enhanced to support TSN.

Different traffic types, or application flows, can be mapped on top of DetNet. DetNet can optionally reuse header information provided by, or shared with, applications. An example of shared header fields can be found in [I-D.ietf-detnet-ip].

This document also covers concepts related to the controller plane and Operations, Administration, and Maintenance (OAM) functions.

2. Terminology

2.1. Terms Used in This Document

This document uses the terminology established in the DetNet architecture [I-D.ietf-detnet-architecture], and the reader is assumed to be familiar with that document and its terminology.

2.2. Abbreviations

The following abbreviations are used in this document:

CW	Control Word.
DetNet	Deterministic Networking.
L2	Layer 2.
L2VPN	Layer 2 Virtual Private Network.
LSR	Label Switching Router.
MPLS	Multiprotocol Label Switching.
MPLS-TE	Multiprotocol Label Switching - Traffic Engineering.
OAM	Operations, Administration, and Maintenance.
PEF	Packet Elimination Function.
PRF	Packet Replication Function.
PREOF	Packet Replication, Elimination and Ordering Functions.
POF	Packet Ordering Function.
PSN	Packet Switched Network.
PW	PseudoWire.
QoS	Quality of Service.
TSN	Time-Sensitive Network.

3. DetNet Data Plane Overview

This document describes how application flows, or app-flows, are carried over DetNet networks. The DetNet Architecture, [I-D.ietf-detnet-architecture], models the DetNet related data plane functions decomposed into two sub-layers: a service sub-layer and a forwarding sub-layer.

Figure 1 reproduced from the [I-D.ietf-detnet-architecture],shows a logical DetNet service with the two sub-layers.

```
             |  packets going  |         ^  packets coming   ^
             v down the stack  v         |   up the stack    |
          +----------------------+   +----------------------+
          |        Source        |   |      Destination     |
          +----------------------+   +----------------------+
          |   Service sub-layer:  |   |   Service sub-layer:  |
          |   Packet sequencing  |   | Duplicate elimination |
          |    Flow replication  |   |      Flow merging    |
          |    Packet encoding   |   |    Packet decoding   |
          +----------------------+   +----------------------+
          | Forwarding sub-layer: |   | Forwarding sub-layer: |
          |  Resource allocation |   |  Resource allocation |
          |    Explicit routes   |   |    Explicit routes   |
          +----------------------+   +----------------------+
          |     Lower layers     |   |     Lower layers     |
          +----------------------+   +----------------------+
                      v                           ^
                       \_________________________/
```

Figure 1: DetNet data plane protocol stack

The DetNet forwarding sub-layer may be directly provided by the DetNet service sub-layer, for example by IP or MPLS. Alternatively an overlay approach may be used in which the packet is natively carried between key nodes within the DetNet network (say between PREOF nodes) and a sub-layer is used to provide the information needed to reach the next hop in the overlay.

This forwarding sub-layer provides the quality underpin needed by the DetNet Service sub-layer. It may do this directly through the use of queuing techniques and traffic engineering methods, or it may do this through the assistance of its underlying connectivity. For example it may call upon Ethernet TSN capabilities defined in IEEE 802.1 TSN [IEEE802.1TSNTG].

The service sub-layer provides additional support beyond the connectivity function of the forwarding sub-layer. An example of

this is Packet Replication, Elimination, and Ordering (PREOF) function see Section 4.5.

The method of instantiating each of the layers is specific to the particular DetNet data plane method. There may be more than approach that is applicable to a given bearer network type.

3.1. Data Plane Characteristics

There are two major characteristics to the data plane:

1. How the data plane is constructed: The DetNet service sub-layer provides its functions for the DetNet application flows by using or applying existing standardized headers and/or encapsulations. The Detnet forwarding sub-layer may provide capabilities leveraging that same header or encapsulation technology e.g. Figure 2 or it may be achieved by other standardized technologies e.g. Figure 3.

2. Extensibility of that Data Plane: Whether or not the DetNet data plane includes the facility to carry additional information (metadata) that can be used to provide an enhanced service to the DetNet packet.

```
                DetNet        +-------+  +---------+
                Services      | DN IP |  | DN MPLS |
                              +-------+  +---------+
```

Figure 2: DetNet Services

```
                                                   +-----+
                                                   | TSN |
                              +-------+          +-+-----+-+
         DetNet               | DN IP |          | DN MPLS |
         Service           +--+--+----+----+   +-+---+-----+-+
         Examples          | TSN | DN MPLS |   | TSN | DN IP |
                           +-----+---------+   +-----+-------+
```

Figure 3: DetNet Service Encapsulations

3.2. Encapsulation

The encapsulation of the DetNet flows allows them to be sent over a data plane of a type other than their native type. Encapsulation is essential if, for example, it is required to send Ethernet TSN stream as a DetNet Application over a data plane such as MPLS. Figure 3 illustrates some encapsulation combinations.

The use of encapsulation is also required if additional information (meta-data) is needed by the DetNet data plane and their is either no ability to include it in the client data packet, or the specification of the client data plane does not permit the modification of the packet to include additional data. An example of such meta-data is the inclusion of a sequence number required by the PREOF function.

Encapsulation is also needed if the DetNet flow or aggregate flow is not easily recognised from its encapsulation.

3.3. Metadata

Metadata can be a useful way of identifying packets that need to be treated as a flow or flow aggregate. It is also useful as a way of including a sequence number the packet for use by the PREOF function or as a place to carry OAM indications or OAM information to instrument DetNet data plane operation.

Explicit inclusion of metadata is possible through the use of IP options or IP extension headers. New IP options are almost impossible to get standardized or to deploy in an operational network and will not be discussed further in this text. IPv6 extensions headers are finding popularity in current IPv6 development work, particularly in connection with Segment Routing of IPv6 (SRv6) and IP OAM. The design of a new IPv6 extension header or the modification of an existing one is a technique available in the tool box of the DetNet IP data plane designer.

Explicit inclusion of metadata in an IP packet is also possible through the inclusion of an MPLS label stack and the MPLS DetNet Control Word using one of the methods for carrying MPLS over IP [I-D.mpls-over-udp-ip]. This is described in more detail in Section 3.6.4.

Implicit metadata can be included through the use of the network programming paradigm [I-D.spring-srv6-network-programming] in which the suffix of an IPv6 address is used to encode additional information for use by the network of the receiving host. Examples of such information include the sequence number for use by the PREOF function, or even all the essential information being included into

the DetNet over MPLS label stack (the DetNet Control Word and the DetNet Service label).

3.4. DetNet IP Data Plane

An IP data plane may operate natively or through the use of an encapsulation. There are many IP encapsulations that may be interposed between the DetNet data plane IP header and the DetNet payload, and it is anticipated that more than one encapsulation may be deployed.

One method of operating an IP DetNet data plane without encapsulation is to use "6-tuple" based flow identification, where "6-tuple" refers to information carried in IP and higher layer protocol headers. General background on the use of IP headers, and "5-tuples", to identify flows and support Quality of Service (QoS) can be found in [RFC3670]. [RFC7657] also provides useful background on the delivery differentiated services (DiffServ) and "tuple" based flow identification. DetNet flow aggregation may be enabled via the use of wildcards, masks, prefixes and ranges. The operation of this method is described in detail in [I-D.ietf-detnet-ip].

The DetNet forwarding plane may use explicit route capabilities and traffic engineering capabilities to provide a forwarding sub-layer that is responsible for providing resource allocation and explicit routes. It is possible to include metadata in a native IP packet explicitly, or implicitly.

3.5. DetNet MPLS Data Plane

MPLS provides the ability to forward traffic over implicit and explicit paths to the point in the network where the next DetNet service sub-layer action needs to take place. It does this through the use of a stack of one or more labels with various forwarding semantics.

MPLS also provides the ability to identify a service instance that is used to process the packet through the use of a label that maps the packet to a service instance.

In cases where metadata is needed to process an MPLS encapsulated packet at the service sub-layer, this has been been provided through the use of a shim layer also called a control word (CW) [RFC4385]. Although such CWs are frequently 32 bits long, there is no architectural constraint on its size of this structure, only the requirement that it is fully understood by all parties operating on it in the DetNet service sub-layer. The operation of this method is described in detail in [I-D.ietf-detnet-mpls].

3.6. Further DetNet Data Plane Considerations

This section needs further work.

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

Reservation and Allocation of resources:

Reservation of resources can allocate resources to specific DetNet flows. This can eliminate packet contention and loss for DetNet traffic. This also can reduce jitter for the DetNet traffic. DetNet flows are assumed to behave with respect to the reserved traffic profile. If other traffic shares the link resources, the use of (queuing, policing, shaping) policies can be used to ensure that the allocation of resources reserved for DetNet is met. Queuing and shaping of DetNet traffic could be required to ensure that DetNet traffic does not exceed its reserved profile but this would impact the DetNet service characteristics.

Explicit routes:

Use of a specific path for a flow. This allows control of the network delay by steering the packet with the ability to influence the physical path. Explicit routes complement reservation by ensuring that a consistent path can be associated with its resources for the duration of that path. Coupled with the traffic mechanism, this limits misordering and bounds latency. Explicit route computation can encompass a wide set of constraints and optimize the path for a certain characteristic e.g. highest bandwidth or lowest jitter. In these cases the "best" path for any set of characteristics may not be a shortest path. The selection of path can take into account multiple network metrics. Some of these metrics are measured and distributed by the routing system as traffic engineering metrics.

Service protection:

Use of multiple packet streams using multiple paths, for example 1+1 or 1:1 linear protection. For DetNet this primarily relates to packet replication and elimination capabilities. Changing the explicit path after a failure is detected to an already established path in order to restore delivery of the required DetNet service characteristics is another protection mechanism for example MPLS hitless protection. Path changes, even in the case of failure recovery, can lead to the out of order delivery of data

requiring packet ordering functions either within the DetNet service or at a high layer in the application traffic. Establishment of new paths after a failure is out of scope for DetNet services.

Network Coding:

Network Coding, not to be confused with network programming, comprises several techniques where multiple data flows are encoded. These resulting flows can then be sent on different paths. The encoding operation can combine flows and error recovery information. When the encoded flows are decoded and recombined the original flows can be recovered. Note that Network coding uses an alternative to packet by packet PREOF. Therefore, for certain network topologies and traffic loads, Network Coding can be used to improve a network's throughput, efficiency, latency, and scalability, as well as resilience to partition, attacks, and eavesdropping, as compared to traditional methods. DetNet could utilized Network coding as an alternative to other protection means. Network coding is often applied in wireless networks and is being explored for other network types.

Load sharing:

Use of packet by packet distribution of the same DetNet flow over multiple paths is not recommended except for the cases listed above where PREOF is utilized to improve protection of traffic and maintain order. Packet by packet load sharing, e.g., via ECMP or UCMP, impacts ordering and possibly jitter.

Troubleshooting:

Since Detnet leverages many different forwarding sub-layers, those technologies also support a number of tools to troubleshoot connectivity for example, to support identification of misbehaving flows. At the service layer again there are existing mechanisms to troubleshoot or monitor flows. Many of these mechanisms exist for IP and MPLS networks. A client of a DetNet service can introduce any monitoring applications which can detect and monitor delay and loss.

Recognize flow(s) for analytics:

To a large degree this follows the logic in the previous section. Analytics can be inherited from the two sub-layers. At the DetNet service edge packet and bit counters e.g. sent, received, dropped, and out of sequence are maintained.

Correlate events with flows:

The provider of a DetNet service may allow other capabilities to monitor flows such as more detail loss statistics and time stamping of events. The details of these capabilities are currently out of scope for this document.

Several of these capabilities are expanded upon in more detail below.

3.6.1. Service Protection

Service protection allow DetNet services to increase reliability and maintain a DetNet Service Assurance in the case of network congestion or some failures. Detnet relies on the underlying technology capabilities for various protection schemes. Protection schemes enable partial or complete coverage of the network paths and active protection with combinations of PRF, PRE, and POF.

3.6.1.1. Linear Service Protection

An example DetNet MPLS network fragment and packet flow is illustrated in Figure 4.

```
   1      1.1       1.1      1.2.1    1.2.1      1.2.2
CE1----EN1--------R1-------R2-------R3--------EN2-----CE2
        \          1.2.1 /                    /
         \1.2     /-----+                   /
          +------R4-----------------------+
                   1.2.2
```

Figure 4: Example Packet Flow in DetNet protected Network

In Figure 4 the numbers are used to identify the instance of a packet. Packet 1 is the original packet, and packets 1.1, and 1.2 are two first generation copies of packet 1. Packet 1.2.1 is a second generation copy of packet 1.2 etc. Note that these numbers never appear in the packet, and are not to be confused with sequence numbers, labels or any other identifier that appears in the packet. They simply indicate the generation number of the original packet so that its passage through the network fragment can be identified to the reader.

Customer Equipment CE1 sends a packet into the DetNet enabled network. This is packet (1). Edge Node EN1 encapsulates the packet as a DetNet Packet and sends it to Relay node R1 (packet 1.1). EN1

makes a copy of the packet (1.2), encapsulates it and sends this copy to Relay node R4.

Note that along the path from EN1 to R1 there may be zero or more nodes which, for clarity, are not shown. The same is true for any other path between two DetNet entities shown in Figure 4 .

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

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

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

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

This example also illustrates 1:1 protection scheme meaning there is traffic and path for each segment of the end to end path. Local DetNet relay nodes determine which packets are eliminated and which packets are forwarded. A 1+1 scheme where only one path is used for traffic at a time, could use the same topology. In this case there is no PRF function and traffic is switched upon detection of failure. An OAM scheme that monitors the paths detects the loss of path or traffic is required to initiate the switch. A POF may still be used in this case to prevent misordering of packets. In both cases the protection paths are established and maintained for the duration of the DetNet service.

3.6.1.2. Ring Service Protection

Ring protection may also be supported if the underlying technology supports it. Many of the same concepts apply however Rings are normally 1+1 protection for data efficiency reasons. [RFC8227] is an example of MPLS-TP data plane that supports Ring protection.

3.6.2. Aggregation Considerations

The DetNet data plane also allows for the aggregation of DetNet flows, to improved scaling by reducing the state per hop. How this is done is data plane or control plane dependent. When DetNet flows are aggregated, transit nodes provide service to the aggregate and not on a per-DetNet flow basis. When aggregating DetNet flows the flows should be compatible i.e. the same or very similar QoS and CoS characteristics. In this case, nodes performing aggregation will ensure that per-flow service requirements are achieved.

If bandwidth reservations are used, the sum of the reservations should be the sum of all the individual reservations, in other words, the reservations should not create an over subscription of bandwidth reservation. If maximum delay bounds are used the system should ensure that the aggregate does not exceed the delay bounds of the component flows.

DetNet encapsulation is a data plane mechanism that can be used to aggregate traffic. Encapsulation can either be in the same service type or in a different service type see Figure 3 for examples. When an encapsulation is used the choice of reserving a maximum resource level and then tracking the services in the aggregated service or adjusting the aggregated resources as the services are added is implementation and technology specific.

DetNet flows at edges must be able to handle rejection to an aggregation group due to lack of resources as well as conditions where general requirements are not satisfied.

3.6.2.1. IP Aggregation

IP aggregation has both data plane and controller plane aspects. For the data plane flows may be aggregated for treatment based on shared characteristics such as 5-tuple. Alternatively, an IP encapsulation may be used to tunnel an aggregate number of DetNet Flows between relay nodes.

3.6.2.2. MPLS Aggregation

MPLS aggregation similarly has data plane and controller plane aspects. In the case of MPLS flows are often tunneled in a forwarding sub-layer and reservation is associated with that MPLS tunnel.

3.6.3. End-System Specific Considerations

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

As a general rule, DetNet domains are capable of forwarding any DetNet flows and the DetNet domain does not mandate the end-system or edge system encapsulation format. Unless there is a proxy of some form present, end-systems peer with similar end-systems using the same application encapsulation format. For example, as shown in Figure 5, IP applications peer with IP applications and Ethernet L2VPN applications peer with Ethernet L2VPN applications.

```
          +-----+
          |  X  |                               +-----+
          +-----+                               |  X  |
          | Eth |               ________        +-----+
          +-----+     _____    /        \       | Eth |
                 \   /     \__/          \___   +-----+
                  \ /                        \ /
                   0======== tunnel-1 ========0_
                   |                             \
                    \                             |
                    0========= tunnel-2 =========0
                   / \                        __/ \
            +-----+   \__ DetNet MPLS domain /     \
            |  X  |      \         __       /       +-----+
            +-----+       \_______/  \_____/        |  X  |
            |  IP |                                 +-----+
            +-----+                                 |  IP |
                                                    +-----+
```

Figure 5: End-Systems and The DetNet MPLS Domain

3.6.4. Sub-Network Considerations

Any of the DetNet service types may be transported by another DetNet service. MPLS nodes may interconnected by different sub-network technologies, which may include point-to-point links. Each of these sub-network technologies need to provide appropriate service to DetNet flows. In some cases, e.g., on dedicated point-to-point links or TDM technologies, all that is required is for a DetNet node to appropriately queue its output traffic. In other cases, DetNet nodes

will need to map DetNet flows to the flow semantics (i.e., identifiers) and mechanisms used by an underlying sub-network technology. Figure 6 shows several examples of header formats that can be used to carry DetNet MPLS flows over different sub-network technologies. L2 represent a generic layer-2 encapsulation that might be used on a point-to-point link. TSN represents the encapsulation used on an IEEE 802.1 TSN network, as described in [I-D.mpls-over-tsn]. UDP/IP represents the encapsulation used on a DetNet IP PSN, as described in [I-D.mpls-over-udp-ip].

```
                        +------+  +------+  +------+
       App-Flow         |  X   |  |  X   |  |  X   |
                  +-----+======+--+======+--+======+-----+
       DetNet-MPLS      | d-CW |  | d-CW |  | d-CW |
                        +------+  +------+  +------+
                        |Labels|  |Labels|  |Labels|
                  +-----+======+--+======+--+======+-----+
       Sub-Network      |  L2  |  | TSN  |  | UDP  |
                        +------+  +------+  +------+
                                            |  IP  |
                                            +------+
                                            |  L2  |
                                            +------+
```

Figure 6: Example DetNet MPLS Sub-Network Formats

4. Controller Plane (Management and Control) Considerations

4.1. DetNet Controller Plane Requirements

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

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

o Instantiate DetNet flows in a DetNet domain (which may include some or all of explicit path determination, link bandwidth reservations, restricting flows to IEEE 802.1 TSN links, node buffer and other resource reservations, specification of required queuing disciplines along the path, ability to manage bidirectional flows, etc.) as needed for a flow.

- In the case of MPLS Manage DetNet S-Label and F-Label allocation and distribution, where the DetNet MPLS encapsulation is in use see Section 4.4.

- The ability to support DetNet flow aggregation.

- Advertise static and dynamic node and link resources such as capabilities and adjacencies to other network nodes (for dynamic signaling approaches) or to network controllers (for centralized approaches).

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

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

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

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

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

4.2. Generic Controller Plane Considerations

This section covers control plane considerations that are independent of the data plane technology used for DetNet service delivery.

While management plane and control planes are traditionally considered separately, from the Data Plane perspective there is no practical difference based on the origin of flow provisioning information, and the DetNet architecture [I-D.ietf-detnet-architecture] refers to these collectively as the 'Controller Plane'. This document therefore does not distinguish between information provided by distributed control plane protocols, e.g., RSVP-TE [RFC3209] and [RFC3473], or by centralized network management mechanisms, e.g., RestConf [RFC8040], YANG [RFC7950], and the Path Computation Element Communication Protocol (PCEP) [I-D.ietf-pce-pcep-extension-for-pce-controller] or any combination thereof. Specific considerations and requirements for the DetNet Controller Plane are discussed in Section 4.1.

4.2.1. Flow Aggregation Control

Flow aggregation includes aggregation accomplished through the use of hierarchical LSPs in MPLS and tunnels, in the case of IP, MPLS and TSN, both of which aggregate multiple DetNet flows into a single new DetNet flow. It can also be grouping of IP flows that share 5-tuple of 6-tuple attributes or flow identifiers at the DetNet sub-layer.

Control of aggregation involves a set of procedures not necessarily in a strict order:

o Traffic engineering resource collection and distribution:

Available resources are tracked through control plane or management plane databases and distributed amongst controllers or nodes that can manage resources.

o Path computation and resource allocation:

When DetNet services are provisioned or requested one or more paths meeting the requirements are selected and the resources verified and recorded.

o Resource assignment and data plane co-ordination:

The assignment of resources along the path depends on the technology and it includes assignment of specific links and coordination of the queuing and other traffic management capabilities.

o Assigned Resource recording and updating:

Depending on the specific technology the assigned resources are updated and distributed in the databases preventing over subscription.

4.2.2. Explicit Routes

Explicit routes are used to ensure that packets are routed through the resources that have been reserved for them, and hence provide the DetNet application with the required service. A requirement for the DetNet Controller Plane will be the ability to assign a particular identified DetNet IP flow to a path through the DetNet domain that has been assigned the required nodal resources. This provides the appropriate traffic treatment for the flow and also includes particular links as a part of the path that are able to support the DetNet flow. For example, by using IEEE 802.1 TSN links (as discussed in [I-D.mpls-over-tsn]) DetNet parameters can be maintained. Further considerations and requirements for the DetNet Controller Plane are discussed in Section 4.1.

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

There are several of approaches that could be used to provide explicit routes and resource allocation in the DetNet forwarding sub-layer. For example:

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

- o The path could use a distributed control plane such as RSVP [RFC2205] or RSVP-TE [RFC3473] extended to support DetNet IP flows.

- o The path could be implemented using IPv6-based segment routing when extended to support resource allocation.

See Section 4.1 for further discussion of these alternatives. In addition, [RFC2386] contains useful background information on QoS-based routing, and [RFC5575] discusses a specific mechanism used by BGP for traffic flow specification and policy-based routing.

4.2.3. Contention Loss and Jitter Reduction

As discussed in Section 1, this document does not specify the mechanisms needed to eliminate packet contention, packet loss or reduce jitter for DetNet flows at the DetNet forwarding sub-layer. The ability to manage node and link resources to be able to provide these functions is a necessary part of the DetNet controller plane. It is also necessary to be able to control the required queuing mechanisms used to provide these functions along a flow's path through the network. See [I-D.ietf-detnet-ip] --> and Section 4.1 for further discussion of these requirements.

4.2.4. Bidirectional Traffic

DetNet applications typically generate bidirectional traffic. IP and MPLS typically treat each direction separately and do not force interdependence of each direction. MPLS has considered bidirectional traffic requirements and the MPLS definitions from [RFC5654] are useful to illustrate terms such as associated bidirectional flows and co-routed bidirectional flows. MPLS defines a point-to-point associated bidirectional LSP as consisting of two unidirectional point-to-point LSPs, one from A to B and the other from B to A, which are regarded as providing a single logical bidirectional forwarding path. This is analogous to standard IP routing. MPLS defines a point-to-point co-routed bidirectional LSP as an associated bidirectional LSP which satisfies the additional constraint that its two unidirectional component LSPs follow the same path (in terms of both nodes and links) in both directions. An important property of co-routed bidirectional LSPs is that their unidirectional component LSPs share fate. In both types of bidirectional LSPs, resource reservations may differ in each direction. The concepts of associated bidirectional flows and co-routed bidirectional flows can also be applied to DetNet IP flows.

While the DetNet IP data plane must support bidirectional DetNet flows, there are no special bidirectional features with respect to the data plane other than the need for the two directions of a co-routed bidirectional flow to take the same path. That is to say that bidirectional DetNet flows are solely represented at the management and control plane levels, without specific support or knowledge within the DetNet data plane. Fate sharing and associated or co-routed bidirectional flows, can be managed at the control level.

DetNet's use of PREOF may increase the complexity of using co-routing bidirectional flows, since if PREOF is used, then the replication points in one direction would have to match the elimination points in the other direction, and vice versa, and the optimal points for these

functions in one direction may not match the optimal points in the other subsequent to the network and traffic constraints.

Control and management mechanisms need to support bidirectional flows, but the specification of such mechanisms are out of scope of this document. An example control plane solution for MPLS can be found in . Related control plan mechanisms have been defined in [RFC3473] , [RFC6387] and [RFC7551].

This is further discussed in Section 4.1.

4.3. IP-Specific Controller Plane Considerations

This section covers IP data plane specific control plane considerations.

4.3.1. Flow Identification and Aggregation

Section 3 discussed the use of the IP "6-tuple" for flow identification, and goes on to discuss how identified flows use specific QoS mechanisms for flow-specific traffic treatment, including path control and resource allocation. [I-D.ietf-detnet-ip] contains detailed DetNet IP flow identification procedures. Flow identification will play an important role for the DetNet controller plane.

Section 3.6.2 and Section 3.6.2.1 discuss the use of flow aggregation in DetNet. Flow aggregation can be accomplished using any of the 6-tuple fields defined in [I-D.ietf-detnet-ip] , using a DSCP identified traffic class or other field. It will be the responsibility of the DetNet controller plane to be able to properly provision the use of these aggregation mechanisms. These requirements are included in Section 4.1.

4.4. MPLS-Specific Controller Plane Considerations

This section covers MPLS data plane specific control plane considerations. This section needs generalizing.

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

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

DetNet S-Labels [I-D.ietf-detnet-mpls] for their definition) are similar to other MPLS service labels that denote the contents of the MPLS packet payload such as a layer 2 pseudowire, an IP packet that

is routed in a VPN context with a private address, or an Ethernet virtual private network (EVPN) service.

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

As discussed in Section 3, the forwarding sub-layer uses one or more F-Labels to forward DetNet packets between DetNet service-aware nodes along explicitly defined routes at the DetNet forwarding sub-layer, which in the context of this document is MPLS. F-Labels can also provide context for an S-Label. In the DetNet Forwarding (MPLS) sub-layer the explicit route consists of the set of DetNet nodes which are LSRs, links, and possibly link bundle members and queues that the DetNet packets of a flow must traverse between nodes in the DetNet service sub-layer (i.e. between a specific Edge Node and the next hop Relay Node, between specific Relay Nodes, and between a specific Relay node and the egress Edge Node. Resource allocation corresponding to the set of Services supported over the forwarding sub-layer, which may or may not include aggregation, is required at this sub-layer. Explicit routes are used to ensure that packets are routed through the resources that have been reserved for them, and hence provide the DetNet application with the required service. Multiple F-Labels may be pushed after an S-Label and there is no requirement for all F-Labels to be controlled via the same controller mechanisms. For example in EVPN, some labels are distributed using BGP while others are distributed using LDP or RSVP.

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

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

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

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

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

Much like other MPLS labels, there are a number of alternatives available for DetNet S-Label and F-Label advertisement to an upstream peer node. These include distributed signaling protocols such as RSVP-TE, centralized label distribution via a controller that manages both the sender and the receiver using NETCONF/YANG, BGP, PCEP, etc., and hybrid combinations of the two. The details of the controller plane solution required for the label distribution and the management of the label number space are out of scope of this document, but as mentioned above, there are particular DetNet considerations and requirements that are discussed in Section 4.1.

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

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

4.6. Contention Loss and Jitter Reduction

As discussed in Section 1, this document does not specify the mechanisms needed to eliminate contention loss or reduce jitter for DetNet flows at the DetNet forwarding sub-layer. The ability to manage node and link resources to be able to provide these functions will be a necessary part of the DetNet controller plane. It will also be necessary to be able to control the required queuing mechanisms used to provide these functions along a flow's path through the network. See Section 4.1 for further discussion of these requirements.

5. 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.

6. IANA Considerations

This document makes no IANA requests.

7. Contributors

RFC7322 limits the number of authors listed on the front page of a draft to a maximum of 5, far fewer than the many individuals below who made important contributions to this draft. The editor wishes to thank and acknowledge each of the following authors for contributing text to this draft. See also Section 8.

Loa Andersson
Huawei
Email: loa@pi.nu

Yuanlong Jiang
Huawei
Email: jiangyuanlong@huawei.com

Norman Finn
Huawei
3101 Rio Way
Spring Valley, CA 91977
USA
Email: norman.finn@mail01.huawei.com

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

Carlos J. Bernardos
Universidad Carlos III de Madrid
Av. Universidad, 30
Leganes, Madrid 28911
Spain
Email: cjbc@it.uc3m.es

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

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

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

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

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

Don Fedyk
LabN Consulting, L.L.C.
Email: dfedyk@labn.net

8. 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

Andrew Malis

Carlos J. Bernardos

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

Lou Berger

Pat Thaler

9. References

9.1. Normative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <https://www.rfc-editor.org/info/rfc3473>.

[RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, DOI 10.17487/RFC4385, February 2006, <https://www.rfc-editor.org/info/rfc4385>.

9.2. Informative References

[I-D.ietf-detnet-architecture]
Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-12 (work in progress), March 2019.

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

[I-D.ietf-detnet-ip]
Korhonen, J., Varga, B., "DetNet Data Plane: IP", 2019.

[I-D.ietf-detnet-mpls]
Korhonen, J., Varga, B., "DetNet Data Plane: MPLS", 2019.

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

[I-D.mpls-over-tsn]
Korhonen, J., Varga, B., "DetNet Data Plane: MPLS over IEEE 802.1 Time Sensitive Networking (TSN)", 2019.

[I-D.mpls-over-udp-ip]
Korhonen, J., Varga, B., "DetNet Data Plane: MPLS over IP", 2019.

[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.

[I-D.spring-srv6-network-programming]
Filsfils, C., Camarillo, P., "SRv6 Network Programming, draft-filsfils-spring-srv6-network-programming, work in progress", 2019.

[IEEE802.1TSNTG]
IEEE Standards Association, "IEEE 802.1 Time-Sensitive Networking Task Group", <http://www.ieee802.org/1/tsn>.

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

[RFC2386] Crawley, E., Nair, R., Rajagopalan, B., and H. Sandick, "A Framework for QoS-based Routing in the Internet", RFC 2386, DOI 10.17487/RFC2386, August 1998, <https://www.rfc-editor.org/info/rfc2386>.

[RFC3670] Moore, B., Durham, D., Strassner, J., Westerinen, A., and W. Weiss, "Information Model for Describing Network Device QoS Datapath Mechanisms", RFC 3670, DOI 10.17487/RFC3670, January 2004, <https://www.rfc-editor.org/info/rfc3670>.

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

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

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <https://www.rfc-editor.org/info/rfc5575>.

[RFC5654] Niven-Jenkins, B., Ed., Brungard, D., Ed., Betts, M., Ed., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, DOI 10.17487/RFC5654, September 2009, <https://www.rfc-editor.org/info/rfc5654>.

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

[RFC7551] Zhang, F., Ed., Jing, R., and R. Gandhi, Ed., "RSVP-TE Extensions for Associated Bidirectional Label Switched Paths (LSPs)", RFC 7551, DOI 10.17487/RFC7551, May 2015, <https://www.rfc-editor.org/info/rfc7551>.

[RFC7657] Black, D., Ed. and P. Jones, "Differentiated Services (Diffserv) and Real-Time Communication", RFC 7657, DOI 10.17487/RFC7657, November 2015, <https://www.rfc-editor.org/info/rfc7657>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/info/rfc7950>.

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

[RFC8227] Cheng, W., Wang, L., Li, H., van Helvoort, H., and J. Dong, "MPLS-TP Shared-Ring Protection (MSRP) Mechanism for Ring Topology", RFC 8227, DOI 10.17487/RFC8227, August 2017, <https://www.rfc-editor.org/info/rfc8227>.

Authors' Addresses

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

Email: balazs.a.varga@ericsson.com

Janos Farkas
Ericsson
Magyar Tudosok krt. 11.
Budapest 1117
Hungary

Email: janos.farkas@ericsson.com

Lou Berger
LabN Consulting, L.L.C.

Email: lberger@labn.net

Don Fedyk
LabN Consulting, L.L.C.

Email: dfedyk@labn.net

Andrew G. Malis
Huawei Technologies

Email: agmalis@gmail.com

Stewart Bryant
Huawei Technologies

Email: stewart.bryant@gmail.com

Jouni Korhonen

Email: jouni.nospam@gmail.com