Deterministic Networking (DetNet) Controller Plane Framework
draft-malis-detnet-controller-plane-framework-01

Abstract

This document provides a framework overview for the Deterministic Networking (DetNet) controller plane. It discusses concepts and requirements that will be basis for Detnet controller plane solution documents.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 6, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must
1. Introduction

Deterministic Networking (DetNet) provides the capability to carry specified unicast and/or multicast data flows for real-time applications with extremely low data loss rates and bounded latency within a network domain. As discussed in the Deterministic Networking Architecture [I-D.ietf-detnet-architecture], techniques used to provide this capability include reserving data plane resources for individual (or aggregated) DetNet flows in some or all of the intermediate nodes along the path of the flow, providing explicit routes for DetNet flows that do not immediately change with the network topology, and distributing data from DetNet flow packets over time and/or space to ensure delivery of each packet’s data in spite of the loss of a path.
The DetNet data plane is defined in a set of documents that are anchored by the DetNet Data Plane Framework [I-D.ietf-detnet-data-plane-framework] and the associated DetNet MPLS [I-D.ietf-detnet-mpls] and IP [I-D.ietf-detnet-ip] data plane specifications, with additional details and subnet mappings provided in [I-D.ietf-detnet-ip-over-mpls], [I-D.ietf-detnet-mpls-over-udp-ip], [I-D.ietf-detnet-mpls-over-tsn], [I-D.ietf-detnet-ip-over-tsn], and [I-D.ietf-detnet-tns-mpn-over-mpls].

While the Detnet Architecture and Data Plane Framework documents are primarily concerned with data plane operations, they do contain some references and requirements for functions that would be required in order to automate DetNet service provisioning and monitoring via a DetNet controller plane. The purpose of this document is to gather these references and requirements into a single document and discuss how various possible DetNet controller plane architectures could be used to satisfy these requirements, while not providing the actual protocol details for a DetNet controller plane solution. Such controller plane protocol solutions will be the subject of subsequent documents.

Note that in the DetNet overall architecture, the controller plane includes what are more traditionally considered separate control and management planes. Traditionally, the management plane is primarily involved with node and network provisioning, operational OAM for performance monitoring, and troubleshooting network behaviors and outages, while the control plane is primarily responsible for the instantiation and maintenance of flows, MPLS label allocation and distribution, and active in-band or out-of-band signaling to support these functions. In the DetNet architecture, all of this functionality is combined into a single Controller Plane. See Section 4.4.2 of [I-D.ietf-detnet-architecture] and the aggregation of Control and Management planes in [RFC7426] for further details.

1.1. Terminology

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

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.
2. DetNet Controller Plane Requirements

Other DetNet documents, including [I-D.ietf-detnet-architecture] and [I-D.ietf-detnet-data-plane-framework], contain requirements for the Controller Plane. For convenience, these requirements have been compiled here. The primary requirements of the DetNet Controller Plane are that it must be able to:

- Support the dynamic creation, modification, and deletion of DetNet flows. This may include some or all of explicit path determination, link bandwidth reservations, restricting flows to IEEE 802.1 Time-Sensitive Networking (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.
- Support DetNet flow aggregation and de-aggregation via the ability to dynamically create and delete flow aggregates (FAs), and be able to modify existing FAs by adding or deleting members.
- Operate in a converged network domain that contains both DetNet and non-DetNet flows.
- Allow flow instantiation requests to originate in an end application (via an Application Programming Interface (API), via static provisioning, or via a dynamic control plane, such as a centralized SDN controller or distributed signaling protocols. See Section 3 for further discussion of these options.
- In the case of the DetNet MPLS data plane, manage DetNet S-Label and F-Label allocation and distribution.
- Also in the case of the DetNet MPLS data plane, support packet replication, duplicate elimination, and packet ordering functions (PREOF), and to be able to place these functions at appropriate places in the network.
- Support applications that require the ability to synchronize the clocks in end systems to the extent supported by the DetNet data plane.
- Support queue control techniques defined in Section 4.5 of [I-D.ietf-detnet-architecture] and [I-D.finn-detnet-bounded-latency] that require time synchronization among network nodes.
- 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).

- Adapt to network topology changes such as links or nodes failures.
- Scale to handle the number of DetNet flows expected in a domain (which may require per-flow signaling or provisioning). This is similar to scalability requirements associated with network slicing [I-D.dong-spring-sr-for-enhanced-vpn].
- 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).
- Monitor the performance of DetNet flows to ensure that they are meeting required objectives.

3. DetNet Control Plane Architecture

As noted in the Introduction, the DetNet control plane is responsible for the instantiation and maintenance of flows, MPLS label allocation and distribution, and active in-band or out-of-band signaling to support these functions.

The following sections define three possible classes of DetNet control plane architectures: a fully distributed control plane utilizing dynamic signaling protocols, a fully centralized SDN-like control plane, and a hybrid control plane. They discuss the various information exchanges between entities in the network in each of these architectures and the advantages and disadvantages of each option.

In each of the following sections, examples are used to illustrate possible mechanisms that could be used in each of the architectures. These are not meant to be exhaustive or to preclude any other possible mechanism that could be used in place of those used in the examples.

3.1. Distributed Control Plane and Signaling Protocols

In a fully distributed configuration model, User-to-Network Interface (UNI) information is transmitted over a (to-be-defined) DetNet UNI protocol from the user side to the network side, and then UNI and network configuration information propagate in the network via distributed control plane signaling protocols. Using an RSVP-TE traffic-engineered MPLS network as an example:
1. An IGP collects topology information and DetNet capabilities of the network [draft-geng-detnet-info-distribution];

2. The control plane of the ingress edge node receives a flow establishment request from the UNI and calculates one or more valid path(s);

3. Using RSVP-TE [RFC3209], the ingress edge node sends a PATH message with an explicit route. After receiving the PATH message, the egress edge node sends a RESV message with the distributed label and resource reservation request.

Current reservation-oriented distributed control plane protocols, e.g., RSVP-TE and Stream Reservation Protocol (SRP) [IEEE.802.1Qcc-2018], can only reserve bandwidth along the path, while the configuration of a fine-grained schedule, e.g., Time Aware Shaping (TAS) [IEEE.802.1QBV_2015], is not supported. If RSVP-TE or SRP were to be used for a DetNet application, it would require extensions in order to support queue and scheduler reservations in addition to bandwidth reservation.

As discussed in Section 4.9 of [I-D.ietf-detnet-architecture], scalability is a primary concern for DetNet, given the large number of expected flows in a DetNet domain. This could potentially be much larger than, for example, the number of MPLS traffic tunnels in a network using MPLS traffic engineering, which would typically be \( N \times (N-1) \) tunnels, where \( N \) is the number of edge routers in the domain.

Even when flow aggregation is used, DetNet domains can be expected to support a very large number of flows that will need particular queuing disciplines and/or resource allocation, depending on the requirements for each flow. This could require a large amount of dynamic signaling, such as an RSVP-TE session to establish and maintain each flow. Other RSVP-TE scalability concerns are further discussed in [RFC5439].

All of the above tends to argue against a purely distributed control plane for DetNet domains.

3.2. SDN/Fully Centralized Control Plane

In the fully SDN/centralized configuration model, UNI information is transmitted from a Centralized User Configuration (CUC) or from applications via an API or northbound interface to a Centralized Controller, which is the sole source of routing and forwarding information for the domain. Configurations of nodes for DetNet flows are performed by the controller using a protocol such as NETCONF [RFC6241]/YANG [RFC6020] or PCE-CC [RFC8283]. For example:
1. The controller collects topology information and DetNet capabilities of the network via NETCONF/YANG;

2. The controller receives a flow establishment request from a UNI and calculates one or more valid path(s) through the network;

3. The controller chooses the optimal path and configures the devices along that path for flow transmission via PCE-CC.

### 3.3. Hybrid Control Plane

In the hybrid model, a controller and control plane protocols work together to provide DetNet services, and there are a number of possible combinations. For example:

1. A Centralized Controller collects topology information and DetNet capabilities of the network via an IGP and/or BGP-LS [RFC7752];

2. The controller receives a flow establishment request from a UNI and calculates one or more valid path(s) through the network;

3. Based on the calculation result, the CNC distributes flow path information to the ingress edge node and other information (e.g. replication/duplicate elimination) to the relevant nodes.

4. Using RSVP-TE, the ingress edge node sends a PATH message with an explicit route. After receiving the PATH message, the egress edge node sends a RESV message with the distributed label and resource reservation request.

or

1. The controller collects topology information and DetNet capability of the network via an IGP or BGP-LS;

2. The control plane of the ingress edge node receives a flow establishment request via a UNI;

3. The Ingress edge node sends the path establishment request to the controller through PCEP [RFC5440];

4. After path calculation, the CNC sends the path information of the flow to the ingress edge node via PCEP;

5. Using RSVP-TE, the ingress edge node sends a PATH message with an explicit route. After receiving the PATH message, the egress edge node sends a RESV message with the distributed label and resource reservation request.
There are many other variations that could be included in a hybrid control plane. This document cannot discuss all the possible control plane mechanisms that could be used in hybrid configuration models. Every solution has its own mechanisms and corresponding parameters that are required for it to work.

4. DetNet Control Plane Additional Details and Issues

This section discusses some additional DetNet control plane details and issues.

4.1. Explicit Paths

Explicit paths are required in DetNet to provide a stable transport service and guarantee that DetNet service is not affected when the network topology changes. The following features are necessary to have explicit paths in DetNet:

- Path computation: DetNet explicit paths need to meet the SLA (Service Level Agreement) requirements and/or resource guarantees from the application/client, which include bandwidth, maximum end-to-end delay, maximum end-to-end delay variation, maximum loss ratio, etc. In a distributed system with IGP-TE, CSPF (Constrained Shortest Path First) can be used to compute a set of feasible paths for a DetNet service. In a system with a network controller, a PCE (Path Computation Engine) can compute paths satisfying the requirements of DetNet with the network information collected from the DetNet domain.

- Path establishment: Once the path has been computed, the options discussed in Section 3 can be used to establish the path. Also see Section 4.4 and Section 4.6 for some additional considerations depending on the details of the network infrastructure.

- Strict or loose paths: An explicit path is strict when every intermediate hop is specified so that its route can’t change. An explicit path is loose when any IGP route is allowed along the path. Generally, end-to-end SLA guarantees require a strict explicit path in DetNet. However, when the IGP route is known to be able to meet the SLA requirements, loose explicit paths are also acceptable.

4.2. Resource Reservation

Network congestion could cause uncontrolled delay and/or packet loss. DetNet flows are supposed to be protected from congestion, so sufficient resource reservation for DetNet service is necessary. Resources in the network are complex and hard to quantize, and may
include such entities as packet processing resources, packet buffering, port and link bandwidth, and so on. The resources a particular flow requires are determined by the flow’s characteristics and SLA.

- Resource Allocation: Port bandwidth is one of the basic attributes of a network device which is easy to obtain or calculate. In current traffic engineering implementations, network resource allocation is synonymous with bandwidth allocation. A DetNet flow is characterized with a traffic specification as defined in [I-D.ietf-detnet-flow-information-model], including attributes such as Interval, Maximum Packets Per Interval, and Maximum Payload Size. The traffic specification describes the worst case, rather than the average case, for the traffic, to ensure that sufficient bandwidth and buffering resources are reserved to satisfy the traffic specification.

- Device configuration with or without flow discrimination: The resource allocation can be guaranteed by device configuration. For example, an output port bandwidth reservation can be configured as a parameter of queue management and the port scheduling algorithm. When DetNet flows are aggregated, a group of DetNet flows share the allocated resource in the network device. When the DetNet flows are treated independently, the device should maintains a mapping relationship between a DetNet flow and its corresponding resources.

4.3. PreOF Support

PreOF path redundancy is supported via packet replication and duplicate elimination (PreOF). A DetNet flow is replicated and goes through multiple networks paths to avoid packet loss caused by device or link failures. In general, current control plane mechanisms that can be used to establish an explicit path, whether distributed or centralized, support point-to-point (P2P) and point-to-multipoint (P2MP) path establishment. PreOF requires the ability to compute and establish a point-to-multipoint-to-point (P2MP2P) path. Protocol extensions will be required to support this new feature.

4.4. DetNet in a Traditional MPLS Domain

For the purposes of this document, "traditional MPLS" is defined as MPLS without the use of segment routing (see Section 4.6 for a discussion of MPLS with segment routing) or MPLS-TP [RFC5960].

In traditional MPLS domains, a dynamic control plane using distributed signaling protocols is typically used for the distribution of MPLS labels used for forwarding MPLS packets. The
dynamic signaling protocols most commonly used for label distribution are LDP [RFC5036], RSVP-TE, and BGP [RFC8277] (which enables BGP/MPLS-based Layer 3 VPNs [RFC4384] and Layer 2 VPNs [RFC7432]).

Any of these protocols could be used to distribute DetNet Service Labels (S-Labels) and Aggregation Labels (A-Labels) [I-D.ietf-detnet-mpls]. As discussed in [I-D.ietf-detnet-data-plane-framework], S-Labels are similar to other MPLS service labels, such as pseudowire, L3 VPN, and L2 VPN labels, and could be distributed in a similar manner, such as through the use of targeted LDP or BGP. If these were to be used for DetNet, they would require extensions to support DetNet-specific features such as PREOF, aggregation (A-Labels), node resource allocation, and queue placement.

However, as discussed in Section 3.1, distributed signaling protocols may have difficulty meeting DetNet’s scalability requirements. MPLS also allows SDN-like centralized label management and distribution as an alternative to distributed signaling protocols, using protocols such as PCEP and OpenFlow [OPENFLOW].

PCEP, particularly when used as a part of PCE-CC, is a possible candidate protocol to use for centralized management of traditional MPLS-based DetNet domains. However, PCE path calculation algorithms would need to be extended to include the location determination for PREOF nodes in a path, and the means to signal the necessary resource reservation and PREOF function placement information to network nodes. See ((?I-D.ietf-pce-pcep-extension-for-pce-controller)) for further discussion of PCE-CC and PCEP for centralized control of an MPLS domain.

4.5. IP

In a later revision of this document, this section will discuss necessary protocol extensions to existing IP routing protocols such as IS-IS and BGP. It should be noted that a DetNet IP domain is simpler than a DetNet MPLS domain, and doesn’t support PREOF, so only one path per flow or flow aggregate is required, with no path merging.

4.6. DetNet with Segment Routing (SR)

Segment Routing [RFC8402] is a scalable approach to building network domains that utilizes a combination of source routing in packet headers and centralized network control to compute paths through the network and distribute those paths with associated policy to network edge nodes for use in packet headers. It greatly reduces the amount of network signaling associated with distributed signaling protocols.
such as RSVP-TE, and also greatly reduces the amount of state in core nodes compared with that required for traditional MPLS and IP routing, as the state is now in the packets rather than in the routers. This is especially useful for DetNet, where a very large number of flows through a network domain are expected, which would otherwise require the instantiation of state for each flow traversing each node in the network.

The DetNet MPLS and IP data planes were specifically constructed to allow the use of DetNet with both types of segment routing, SR-MPLS \[I-D.ietf-spring-segment-routing-mpls\] and SRv6 \[I-D.ietf-6man-segment-routing-header\].

In the DetNet context, DetNet in an SR-MPLS or SRv6 data plane could be used in conjunction with centralized flow management and complete label stack distribution to Detnet domain entry nodes via a centralized controller. Extensions to PCEP to allow the use of PCE-CC with SR-MPLS.

One possible architecture is PCE-CC combined with SR-MPLS or SRv6. Extensions to PCEP to allow the use of PCE-CC with SR-MPLS are described in \[I-D.zhao-pce-pcep-extension-pce-controller-sr\], with SRv6 in \[I-D.dhody-pce-pcep-extension-pce-controller-srv6\].

This approach would allow the details of packet or flow treatment to be encoded directly in the SIDs on each packet in a flow to reduce the amount of state in network nodes. This approach also allows the integration of DetNet domains with general SR-based backbone networks in a converged domain. In this approach, a new set of functions for DetNet queuing treatments available in the DetNet domain would need to be defined for inclusion in the SR stack.

This is not the only possible approach. There is ongoing work on a number of alternative signaling mechanisms for MPLS-SR and SRv6, including extensions to IGP and BGP to support distributed signaling. In addition, BGP-LS and BGP route reflectors could be added for a hybrid solution.

A possible mostly centralized hybrid approach could be to use a PCE-CC to push paths represented by SID lists while using BGP-LS to collect network topology and link state information. An IGP is used for the usual link state flooding in order to establish adjacencies, but not for DetNet flow path calculations, only for best effort traffic as usual.

A similar approach for network slicing that could be leveraged for DetNet is described in \[I-D.dong-spring-sr-for-enhanced-vpn\].
Also, note that SR cannot currently support DetNet PREOF functionality without extensions. One possible approach could be to combine SR with BIER-TE, as discussed in [I-D.ietf-bier-te-arch]. Another possible approach specific to SRv6 is discussed in [I-D.geng-detnet-dp-sol-srv6].

5. Management Plane Overview

The Management Plane includes the ability to statically provision network nodes and to use OAM to monitor DetNet performance and detect outages or other issues at the DetNet layer.

5.1. Provisioning

Static provisioning in a Detnet network will be performed via the use of appropriate YANG models, including [I-D.ietf-detnet-yang] and [I-D.ietf-detnet-topology-yang].

5.2. DetNet Operations, Administration and Maintenance (OAM)

The overall framework and requirements for DetNet OAM are discussed in [I-D.mirsky-detnet-oam]. This document currently includes additional OAM details that may eventually be merged into that document.

5.2.1. OAM for Performance Monitoring (PM)

5.2.1.1. Active PM

Active PM is performed by injecting OAM packets into the network to estimate the performance of the network by measuring the performance of the OAM packets. Adding extra traffic can affect the delay and throughput performance of the network, and for this reason active PM is not recommended for use in operational DetNet domains. However, it is a useful test tool when commissioning a new network.

5.2.1.2. Passive PM

Passive PM monitors the actual service traffic in a network domain in order to measure its performance without having a detrimental affect on the network. As compared to Active PM, Passive PM is much preferred for use in DetNet domains.

A proposal for DetNet passive performance measurement is contained in [I-D.chen-detnet-loss-delay].
5.2.2. OAM for Fault/Defect Management (FM)

[I-D.mirsky-detnet-oam] contains requirements for fault/defect detection and management in a DetNet domain.

6. IANA Considerations

This document has no actions for IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

7. Security Considerations

The overall security considerations of DetNet are discussed in
[I-D.ietf-detnet-architecture] and [I-D.ietf-detnet-security]. For DetNet networks that make use of Segment Routing (whether SR-MPLS or SRv6), the security considerations in [RFC8402] also apply.

DetNet networks that make use of a centralized controller plane may be threatened by the loss of connectivity (whether accidental or malicious) between the central controller and the network nodes, and/or the spoofing of control messages from the controller to the network nodes. This is important since such networks depend on centralized controllers to calculate flow paths and instantiate flow state in the network nodes. For networks that use both DetNet and Segment Routing with a centralized controller, this would also include the calculation of SID lists and their installation in edge/border routers.

In both cases, such threats may be mitigated through redundant controllers, the use of authentication between the controller(s) and the network nodes, and other mechanisms for protection against DOS attacks. A mechanism for supporting one or more alternative central controllers and the ability to fail over to such an alternative controller will be required.

8. Acknowledgments

Thanks to Jim Guichard, Donald Eastlake, and Stewart Bryant for their review comments.

9. References
9.1. Normative References

[I-D.ietf-detnet-architecture]

[I-D.ietf-detnet-data-plane-framework]

[I-D.ietf-detnet-flow-information-model]

[I-D.ietf-detnet-ip]

[I-D.ietf-detnet-mpls]

[I-D.ietf-detnet-security]


9.2. Informative References

[I-D.chen-detnet-loss-delay]

[I-D.dhody-pce-pcep-extension-pce-controller-srv6]

[I-D.dong-spring-sr-for-enhanced-vpn]

[I-D.finn-detnet-bounded-latency]

[I-D.geng-detnet-dp-sol-srv6]

[I-D.ietf-6man-segment-routing-header]

[I-D.ietf-bier-te-arch]


Mirsky, G. and M. Chen, "Operations, Administration and Maintenance (OAM) for Deterministic Networks (DetNet)", draft-mirsky-detnet-oam-03 (work in progress), May 2019.


Authors’ Addresses

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

Xuesong Geng
Huawei
Email: gengxuesong@huawei.com

Mach (Guoyi) Chen
Huawei
Email: mach.chen@huawei.com

Fengwei Qin
China Mobile
Email: qinfengwei@chinamobile.com