

Deterministic Networking Working Group
Internet-Draft
Intended status: Informational
Expires: 23 November 2024

P. Liu
China Mobile
Y. Li
Huawei
T. Eckert
Futurewei Technologies USA
Q. Xiong
ZTE Corporation
J. Ryoo
ETRI
S. Zhu
New H3C Technologies
X. Geng
Huawei
22 May 2024

Requirements for Scaling Deterministic Networks
draft-ietf-detnet-scaling-requirements-06

Abstract

Aiming at scaling deterministic networks, this document describes the technical and operational requirements when the network has large variation in latency among hops, great number of flows and/or multiple domains without the same time source. Different deterministic levels of applications co-exist and are transported in such a network. This document also describes the corresponding Deterministic Networking (DetNet) data plane enhancement requirements.

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 23 November 2024.

Copyright Notice

Copyright (c) 2024 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 include Revised BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Revised BSD License.

Table of Contents

- [1.](#) Introduction [3](#)
- [2.](#) Conventions Used in This Document [4](#)
- [3.](#) Technical Requirements in Scaling Deterministic Networks . . [5](#)
 - [3.1.](#) Tolerate Time Asynchrony [5](#)
 - [3.1.1.](#) Support Asynchronous Clocks Across Domains [5](#)
 - [3.1.2.](#) Tolerate Clock Jitter & Wander within a Clock Synchronous Domain [6](#)
 - [3.1.3.](#) Provide Mechanisms not Requiring Strict Time Synchronization [6](#)
 - [3.1.4.](#) Provide Mechanisms not Requiring Synchronization . . [6](#)
 - [3.2.](#) Support Large Single-hop Propagation Latency [6](#)
 - [3.3.](#) Accommodate the Higher Link Speed [8](#)
 - [3.4.](#) Be Scalable to A Large Number of Flows and Tolerate High Utilization of Bandwidth [8](#)
 - [3.5.](#) Tolerate Failures of Links or Nodes and Topology Changes [9](#)
 - [3.6.](#) Prevent Flow Fluctuation [10](#)
 - [3.7.](#) Be Scalable to a Large Number of Hops with Complex Topology [11](#)
 - [3.8.](#) Support Multi-Mechanisms in Single Domain and Multi-Domains [12](#)
 - [3.8.1.](#) Support Provisioning of Multiple Mechanisms [13](#)
 - [3.8.2.](#) Support Mechanisms Switchover Crossing Multi-domains [14](#)
- [4.](#) Data Plane Enhancement Requirements [14](#)
 - [4.1.](#) Support Aggregated Flow Identification [15](#)
 - [4.2.](#) Support Information used by Functions Ensuring Deterministic Latency [15](#)
- [5.](#) Conclusion [16](#)
- [6.](#) Security Considerations [16](#)
- [7.](#) IANA Considerations [16](#)
- [8.](#) Acknowledgements [16](#)

9.	Contributors	16
10.	Informative References	17
Appendix A.	Examples of Scaling Deterministic Network Trials	20
	Authors' Addresses	21

1. Introduction

Packet networks are evolving from bandwidth-guaranteed Quality of Service (QoS) to latency-guaranteed QoS that guarantees bounded latency and definite latency. Bounded latency and definite latency can be further understood as in-time delivery, in which a packet arrives without exceeding a predetermined time, and on-time delivery, in which a packet arrives at a predetermined time, respectively. In addition, network survivability, which typically guarantees traffic recovery within 50 ms in the event of a network failure, is evolving to a level that guarantees lossless recovery. In order to realize the evolution of QoS and network survivability of these networks, Time-Sensitive Networking (TSN) technology and Deterministic Networking (DetNet) technology are considered to be essential.

TSN is a set of standards developed by the IEEE 802.1 TSN Task Group (TG) [[IEEE802.1TSN](#)] and specifies mechanisms and protocols necessary to realize highly available IEEE 802.1 networks with bounded latency to carry time-sensitive, real-time application traffic.

DetNet, of which architecture is defined in [RFC 8655](#) [[RFC8655](#)], provides a capability to carry specified unicast or multicast data flows for real-time applications with extremely low data loss rates and bounded latency under a single administrative control or within a closed group of administrative control. The overall framework for DetNet data plane is provided in [[RFC8938](#)], and various documents on different data plane technologies and their interworking technologies to extend the service range of data that TSN intends to deliver to the IP (Internet Protocol) and MPLS (Multi-Protocol Label Switching) networks have been standardized.

When deterministic networks become large and/or multiple domains should be stitched, DetNet solutions need to take into consideration one or more of the following points:

- * There is relaxed clock synchronization or no clock synchronization in different domains. ([Section 3.1](#))
- * The end to end path is a combination of low and high latency hops. ([Section 3.2](#))
- * There are various transmission rates supported at different ports and on different network nodes. ([Section 3.3](#))

- * There are a large number of flows which may be difficult to identify per-flow state and cause the high utilization of bandwidth. ([Section 3.4](#))
- * The topology changes and failures of links or nodes might be more common. ([Section 3.5](#))
- * The flow fluctuation caused by a large number of flows may happen frequently. ([Section 3.6](#))
- * The number of hops might be large with complex topology. ([Section 3.7](#))
- * The mechanisms used to ensure bounded latency (e.g. queuing mechanism) may be multiple or have different configuration/parameters in multi-domains. ([Section 3.8](#))

Such domains are normally within a single administrative control network or multiple cooperating administrative networks within a closed group of administrative control [[RFC8655](#)]. However they may belong to different Autonomous System (AS) domains and be controlled by multiple peering or cascaded controllers, and at the same time they may not have the same time clock source. Maintaining per flow status becomes impractical in the large scale network. Aggregation and disaggregation of flows take place at the boundaries of DetNet domains as well as within a DetNet domain with various rules. The flow identification and packet treatment should take care of one or combined changes introduced by scaling deterministic networks.

Based on the consideration above, this document describes requirements for scaling deterministic networks which demands the enhancement based on the existing bounded latency mechanisms and the corresponding data plane to ensure the DetNet service for a single administrative network or multiple (cooperating) administrative networks that defined in the DetNet charter. The deterministic network for open Internet is not within current scope.

2. Conventions Used in This Document

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.

While [RFC2119] and [RFC8174] describe interpretations of these key words in terms of protocol specifications and implementations, they are used in this document to describe technical and operational requirements to realize scaling deterministic networks.

3. Technical Requirements in Scaling Deterministic Networks

Due to the attributes described in Introduction Section, the corresponding technical requirements should be considered.

3.1. Tolerate Time Asynchrony

A deterministic network may span over multiple networks with one or more cooperating administrative domains. There are many types of network nodes in the multiple domains which may introduce disparate local time variation, which require the tolerance of time asynchrony.

3.1.1. Support Asynchronous Clocks Across Domains

One of DetNet's objectives is to stitch TSN islands together. All devices inside a TSN domain are time-synchronized, and most of TSN technologies rely on precise time synchronization [IEEE802.1Qbv][IEEE802.1Qch][IEEE802.1Qav]. However, different TSN islands may have different clocks which are not synchronized as shown in Figure 2, where the time difference of two TSN domains is D. DetNet needs to connect these two TSN domains together and provide end-to-end deterministic latency service. The mechanism adopted by a deterministic network MUST be prepared to cross multiple non-synced TSN domains. This can be done, for example, by putting extra buffer space at the ingress of a new domain, increasing the dead time as a guard band [IEEE802.1Qdv], or using some timing compensation mechanism. This document does not intend to list all the potential ways.

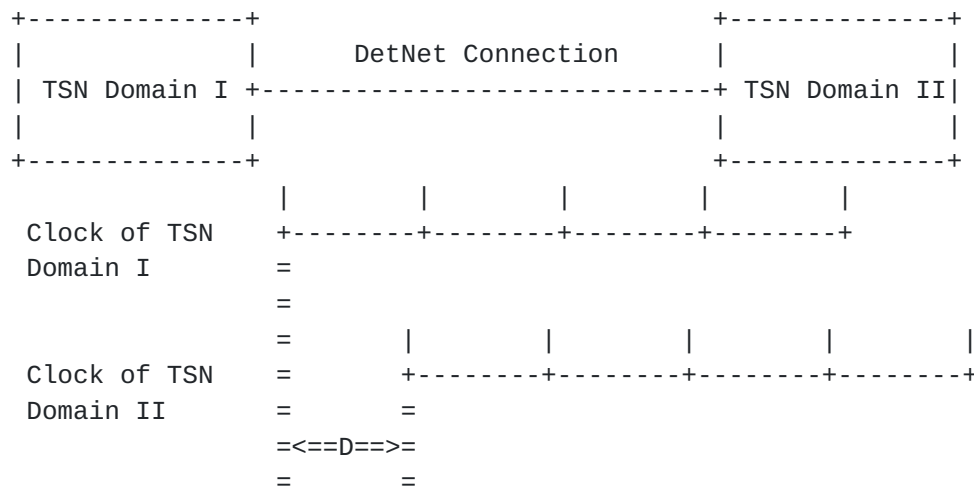


Figure 1: Clock asynchrony between two TSN islands

3.1.2. Tolerate Clock Jitter & Wander within a Clock Synchronous Domain

Within a single time synchronization domain, different clock accuracy is expected, for example the crystal oscillator in Ethernet is specified at 100 ppm [[Fast-Ethernet-MII-clock](#)], Synchronous Ethernet (SyncE) can achieve 50 ppb [[G.8262](#)], and more precise time synchronization [[G.8273](#)] is expected in 5G mobile backhaul. The clocks experience different jitter and wander. It may cause different level of asymmetry of the path. The deterministic networks SHOULD be able to recover or absorb such time variance within a domain.

3.1.3. Provide Mechanisms not Requiring Strict Time Synchronization

It is usually hard to achieve the full time synchronization when the scale of networks become large. Some networks like mobile backhaul use frequency synchronization, such as SyncE, instead of the strict time synchronization. It is desired that the same deterministic performance in term of the bounded latency and jitter SHOULD be achieved when full time synchronization is not available, that is to say, when only partial synchronization (SyncE is one of the examples) is in use.

3.1.4. Provide Mechanisms not Requiring Synchronization

There can be a large number of traffic flows in a deterministic network and some of them are not periodic. Asynchronization based methods can meet the requirements of those traffic flows. Moreover, The mechanisms not requiring the time and/or frequency synchronization eliminate the hardware cost and difficulty at the network nodes. [[IEEE802.1Qcr](#)] conceptually uses per-flow based asynchronous shaper to achieve bounded latency. The effectiveness of the per-flow based asynchronous shaper can be proven through mathematical analysis. It can naturally tolerate the time variance, but it exhibits the concerns of per-flow state buffer management as shown in [[I-D.eckert-detnet-bounded-latency-problems](#)]. When it is in use, the requirement in [Section 3.3](#) SHOULD be carefully met.

3.2. Support Large Single-hop Propagation Latency

In some deterministic networks, a single hop distance is enough to generate large latency. The speed of optical transmission in fiber is 200 km/ms. Thus, the propagation delay of a single hop can be in the order of a few milliseconds. It is much greater than that of a LAN, and introduces impacts on queuing mechanisms, such as cyclic or time aware scheduling method. So the requirement for propagation

delays in end-to-end computations is needed, such as considering the propagation latency when setting the period in both time synchronization or frequency synchronization based methods, or setting the extra buffer in the asynchronization based methods, to meet the requirements of deterministic forwarding between the network nodes.

Here, we use an example to describe the influence of large single-hop propagation delay on cycle based methods, but not to specify any solution. For a cyclic based method, suppose a deterministic network wants to keep using the simple cycle mapping relationship, however the link distance between two nodes is longer. Moreover, a downstream node may have many upstream nodes with different link propagation delays (e.g., 9 us, 10 us, 11 us, 15 us and 20 us). In order to absorb the longest link propagation delay, the length of cycle must be set to more than 20 us. In [IEEE802.1Qch], propagation delay is part of the dead time imposed in a cycle, which impacts the bandwidth utilization. However, since packet's arrival time varies within the receiving cycle, larger cycle length means larger delay variance.

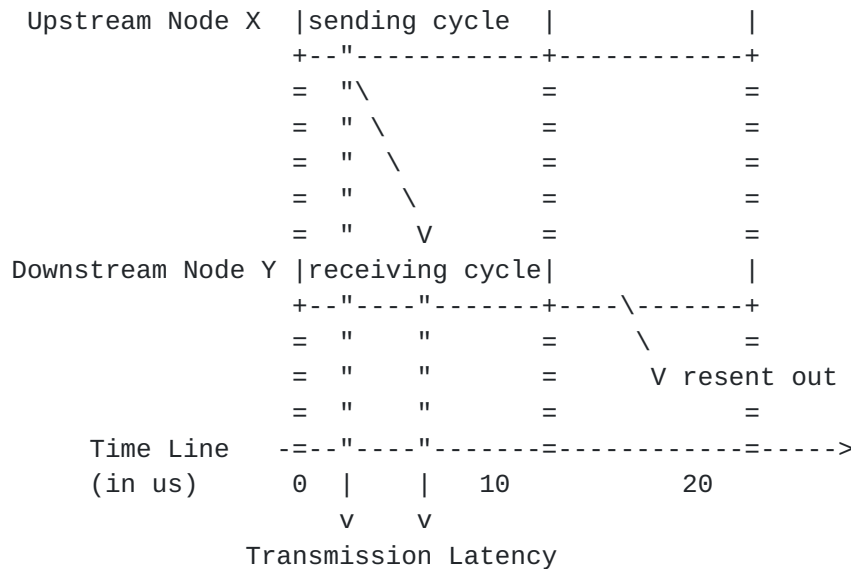


Figure 2: The influence of transmission latency on a cyclic method

Note that Figure 2 is just to show an example of latency caused by large single-hop propagation. CQF is not limited to 2 queues, instead using more than 2 queues can be an option to solve large link latency related concerns. [Multiple-CQF] describes this problem in more detail and also proposes a mechanism to address it.

3.3. Accommodate the Higher Link Speed

A deterministic network can use higher speed links, especially for its backbone. At the time of publication, deterministic mechanisms used in a local network is usually deployed in link speed of 10 Mbps or 1 Gbps, or possibly 10 Gbps. The data rates of 10G, 100G, 400G and even higher are commonly used in wide area networks. With the increasing of the data rate, the network scheduling cycle can be reduced if the same amount of the data is required to be sent in each cycle for each application. Or, more data can be sent if the network cycle time remains the same. For the former, it requires the more precise time control (e.g. cycle in the order of a few microseconds or sub- microseconds) for the input stream gate and the timed output buffer. For the latter, more buffer space is required, which imposes more complex buffer or queue management and larger memory consumption.

Another aspect to consider is the aggregation of the flows. In the deterministic network, the number of flows can be hundreds or tens of thousands. They can be aggregated into a small number of deterministic path or tunnels. It is practical to have a few flow-based or aggregated-flow based status in the local network. But in higher speed and larger scale networks, it is hardly feasible. If [[IEEE802.1Qcr](#)] is in use, it requires more buffers comparing to the other full/partial time synchronized mechanisms. Therefore, it requires optimizations to support higher link speeds.

3.4. Be Scalable to A Large Number of Flows and Tolerate High Utilization of Bandwidth

A deterministic network may have a large number of traffic flows. According to [[RFC2475](#)], per-flow state identification in the network becomes infeasible as flow count scales up. So, it is almost impossible to identify individual IP flows from a massive number of flows in the DetNet data plane or configure the per-flow state machine in each node. DetNet allows the leverage of the flow aggregation, while individual flows may join and exit the aggregated flow rapidly in the scaling network with large number of flows, which causes the dynamic in identification of the aggregated DetNet flow. The wildcards and value ranges used in the identification may have to change in order to ensure the aggregated flows have compatible deterministic characteristics.

For scaling network, proper provision at the control plane to accommodate such higher aggregation is required. Provisioning on the aggregated flows normally improve the scalability at the cost of coarse granularity of the incoming traffic filtering and protection. It is desirable that adding a flow to the network does not affect the

latency of the existing flows or require complex re-calculations, but should require as less work as possible. For instance, the filtering and policing configuration is performed only on ingress nodes and not on intermediate nodes. The [[IEEE802.1Qbv](#)] uses traffic classes to divide the flows, the number of which is usually 8. Therefore, the forwarding mechanism itself is not complex with a large number of flows or higher aggregation. However, when adding new flows, the Gate Control List may be changed, and the re-calculation is complex. There might be the method to simplify the calculation or configuration, which need more work to enhance it.

Meanwhile, the traffic that requires deterministic networking can significantly fill up the capacity of a link or the portion of the link which is dedicated to such traffic, for example, more than 75% and/or up to near 100% utilization. Usually, the resources required for DetNet are reserved, however, the over-provisioning of link capacity may not work in such cases. In order to guarantee deterministic latency and jitter in this environment, it is better to provide scalable queuing solutions to improve the bandwidth utilization based on the current methods, including the TSN and other published standards. For instance, when the bandwidth utilization is high, the guard band in each cycle in [[IEEE802.1Qch](#)] is a type of over-provisioning and can be improved with more scalable queuing additions.

[3.5. Tolerate Failures of Links or Nodes and Topology Changes](#)

Deterministic networks may involve more network devices, and the increase or decrease of network devices in deterministic networks is more frequent than that in LANs. A simple use case to understand is ultra-low-latency (public) 5G transport networks, which would require DetNet extend to every 5G base station. For some network operators, their networks may need to connect to ~100 K base stations (serving multiple mobile networks operators), and this number will only increase with 5G.

On the one hand, the numerous devices may lead to more network link failures. Path switching or re-convergence of routing will cause high latency of packet loss and retransmission, which is usually in seconds before the network becomes stable again. As the added delay magnitudes involved are too large to use jitter compensation techniques. It is necessary to support certain mechanisms to adapt to failures of links or nodes and topology changes.

On the other hand, the change of the number of devices may affect the implementation and adjustment of deterministic network mechanism, such as the topology discovery, queuing mechanism and packet replication and elimination. For instance, fully disjoint paths when

implementing the Packet Replication, Elimination, and Ordering Functions (PREOF) gives a better chance of survival when one of the nodes or links in the path fails. At the same time, it brings the challenges of finding paths with similar distance and/or number of hops so that there is enough buffer space to absorb the latency difference caused by different paths when the scale is large. So, a method is expected to support flexible planning of multiple paths and provide a solid foundation for the implementation of PREOF.

3.6. Prevent Flow Fluctuation

More kinds of DetNet traffic flows described in [Section 3.4](#) will cause more dynamic joining or leaving of the flows, which will further cause more flow fluctuation as well as more unpredictability of the DetNet flows. Such as:

- * Various and massive traffic flows of different applications in scaling networks easily cause more bursty traffic.
- * There will be more aggregation nodes which receive the flows from more upstream nodes adding the nondeterministic delay of the packet treatment.
- * The bursts of flows can be accumulated as the flows traverse, join, and separate over hops. Once one of the nodes makes a minor error in packet treatment, it will have a cumulative effect on downstream nodes.
- * Loops formed in a network topology increase the maximum bursts of flows exponentially [[ANDREWS](#)][[BOUILLARD](#)][[THOMAS](#)].
- * The node and link failures are more common in a large network ([Section 3.5](#)) which requires dynamic traffic steering to an alternate path, it will also easily cause the flow fluctuation.

Noting that the non-DetNet flows are also massive and may have potential impact on the scalability of the DetNet flows, for instance, causing the high utilization of the bandwidth and suppressing the possibility of more resource reservation and the traffic steering of DetNet flows. However, it is assumed that there will be the strategy in the ingress to deal with the non-DetNet flows and prevent the real-time influence on the DetNet flows.

It is required to support bursty traffic and some methods to decrease the micro-burst. So the pre-planned, ingress traffic conditioning, scalable queuing, and enhanced capacity of buffer are required to accommodate the flow fluctuation, and the time required for network reconfiguration to reflect such changes is required to be controlled, e.g., less than a specified amount of time.

3.7. Be Scalable to a Large Number of Hops with Complex Topology

Scaling networks often results in situations where an end-to-end flow involves a large number of hops, e.g., 15 or more. The network topology can also be complex, including star, ring, mesh, and their combinations, and can possibly be hierarchical. It is required to support networks with such various types of topologies and large hop counts.

Delivering DetNet Quality of Service (QoS) in large and complex networks requires end-to-end bounds on both latency and jitter, as discussed in [section 3.1 of \[RFC8655\]](#). Achievable end-to-end latency bounds necessarily depend on the number of hops for a flow. In the best case, the added queuing mechanism latency for DetNet QoS is bounded by a fixed constant per hop maximum value so that the resulting end-to-end latency bounds are a linear function of the number of hops in addition to the inherent forwarding latency of the nodes involved. In contrast, it is possible to achieve fixed constant end-to-end jitter bounds that are independent of the number of hops. Such fixed constant jitter bounds are strongly preferable to jitter bounds that increase with an increasing number of hops.

DetNet QoS requires resource allocation in advance (e.g., of link bandwidth and node buffer resources), as discussed in [section 3.2.1 of \[RFC8655\]](#). The complexity of resource allocation processing may range from linear (e.g. allocating resources for each hop via a path-based resource reservation protocol such as RSVP [\[RFC2205\]](#)) to potentially exponential (e.g., if solving a complex graph optimization problem is required). This resource allocation complexity does not directly affect achievable end-to-end latency and jitter bounds, but it does surface in other areas such as the amount of computation and elapsed time required to admit a new flow to a DetNet network without disrupting the DetNet QoS provided to already admitted flows.

Different queuing mechanisms exhibit different properties across achievable end-to-end jitter bounds, achievable end-to-end latency bounds and complexity. All three of these areas are considerations in evaluation and selection of scalable DetNet queuing mechanisms.

3.8. Support Multi-Mechanisms in Single Domain and Multi-Domains

There will be a large number of flows as described in [Section 3.4](#), and the flows may have different levels of demand for DetNet service. [\[RFC8578\]](#) provides various use cases and their requirements in the areas of industry, electricity, buildings, etc. Some of them clearly specify the requirements for latency and jitter, while some others do not for the jitter. Different types of users have different demands, just as a network provider provides different network services for personal business or enterprise business.

One kind has critical SLA requirement, such as remote control or cloud Programmable Logic Controller (PLC) of manufacturing and differential protection of electricity. If these services exceed the boundaries of latency and jitter, it will bring property losses and security risks, so they cannot tolerate with any non-deterministic situation and can pay more on the network service. Another kind has relatively loose levels of SLA requirement, such as cloud gaming, cloud VR and online meeting for "consumer" networks. The users of these applications hope to have a better network experience, but they can tolerate it to a certain extent. For instance, exceeding the upper boundary of latency within a small probability is acceptable. Those different applications expect different kind of solutions, which are related to the cost more or less.

The different SLA demands need different DetNet technologies as well as the multi-domain demand in scaling networks, which results in the requirements for the diverse queuing mechanisms. For strict deterministic services, strict queuing technologies need to be used, and all network devices may need to be upgraded. For non-strict deterministic services, it may only be necessary to upgrade some network devices (maybe edge nodes) or share corresponding network resources. Moreover, as queuing mechanisms development, it is also desirable to provide the queuing solutions with multiple levels of deterministic capabilities and schedule the reasonable resources to achieve the optimization of network resource utilization. Those different queuing technologies may be used in:

- * the same network which requires the some of the equipment in the same network providing multiple queuing technologies. The operator can select the service type/level accordingly.
- * the multiple domain network which support different queuing technologies while needs the coordination with each other.
- * different network implementations intended for different application domains individually, where there is no additional requirements for the coordination.

Critical latency requirements:

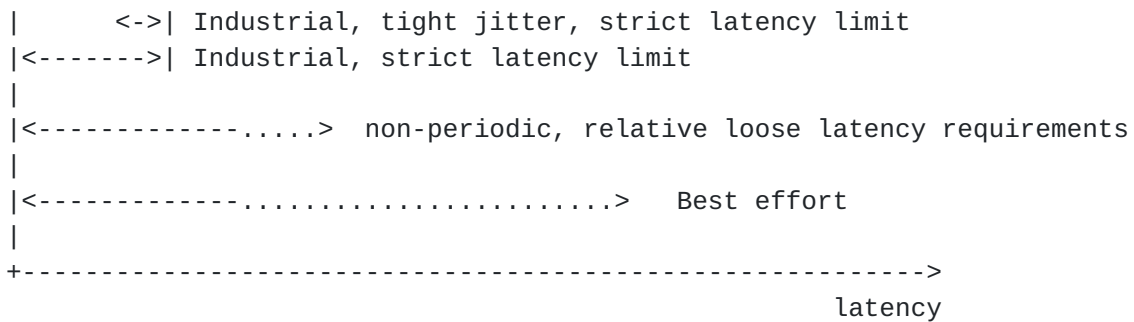


Figure 3: Different levels of application requirements

3.8.1. Support Provisioning of Multiple Mechanisms

It is required to provide diversified deterministic service for various applications in a deterministic network and to support the corresponding diversified mechanisms (possibly at multiple DetNet QoS levels). For instance, different queuing mechanisms can provide different levels of latency, jitter and other guarantees, and there may be situations where a network device provides multiple queuing mechanisms at the same time. For example, a network aggregation device may use the mechanisms specified in [IEEE802.1Qbv] and [IEEE802.1Qcr], and other mechanisms to forward traffic to different paths at the same time. By providing a variety of queuing mechanisms to meet diversified deterministic service requirements, compared with small scale environment, this demand is particularly prominent in large-scale networks. For instance, there may be more than eight queues or sub-queues to support more complicated queuing mechanisms comparing with the eight traffic classes in TSN enabled networks.

Accordingly, the configuration for multiple queuing mechanisms can be complicated in deterministic networks and MUST support the unified or simplified scheduling and management of multiple queuing mechanisms. For example, in the distributed scenario with no controller, the related information of the queuing mechanisms could be advertised among the domain, including the types and related algorithms, queue forwarding capability with different levels of latency and jitter guarantees, etc. In the centralized scenario, the queuing mechanisms and other information could be reported to the controller to build a deterministic network resource topology pool for path calculation. In addition, for refined management of forward resources and providing resource assurance for deterministic forwarding when establishing/deleting sessions, it is required to establish unified mechanisms on quantification and measurement of resources associated with interfaces and queues.

3.8.2. Support Mechanisms Switchover Crossing Multi-domains

In deterministic networks, end-to-end service may cross multiple network domains, which may adopt a variety of different queuing mechanisms within each domain, or may have different link speeds ([Section 3.3](#)) for the same queuing mechanism.

Both of the two cases may generate additional end-to-end latency, jitter and packet loss, because the different queuing mechanisms and link speeds provide different scheduling granularities or phases between the domains. For the different queuing mechanisms switchover, such as from time synchronous mechanism [[IEEE802.1Qbv](#)] to asynchronous mechanism [[IEEE802.1Qcr](#)], a collaboration mechanism crossing multi-domains MUST be considered. For the different link speed switchover, such as from 1Gbps to 10Gbps (or reverse), the quantification of deterministic forwarding resources (such as time slots) of the queuing mechanism MUST match the link speed.

The device used to connect domains requires a flexible queue stitching function, such as increasing the buffer of the inter-domain devices to provide enough adjustment space for the flow to cross different queuing mechanisms, the jitter compression to reduce the coupling between the queuing mechanisms of two domains, or an additional traffic shaping solution to make the traffic smooth, so that for the same scale of service flows, they can be forwarded successfully based on different queuing mechanisms and/or the links with different speeds in multiple domains.

4. Data Plane Enhancement Requirements

According to [[RFC8938](#)], the DetNet data plane can provide or carry two metadata in MPLS and IP data planes: Flow-ID and sequence number. The Flow-ID could be used for identification of the DetNet flow or aggregate flow, and the sequence number could be used for PREOF for each DetNet flow. The Flow-ID is used by both the service and forwarding sub-layers, but the sequence number is only used by the service layer. Metadata can also be used for OAM indications and instrumentation of DetNet data plane operation.

Generally speaking, more data plane metadata and related processing SHOULD be supported in the deterministic networks. By introducing IPv6 Extension Headers [[RFC8200](#)] and Segment Routing over IPv6 [[RFC8986](#)], native IPv6 data plane should be supported with the IPv6-specific enhancement. This section lists the data plane enhancement requirements based on but not limited to the technical requirements in [Section 3](#), describing how to use IP and/or MPLS, and related OAM, to support a data plane method of flow identification and packet treatment over Layer 3. There might be some enhancement

of the control plane to meet the requirements in [Section 3](#), which is out of scope of this document and expected to be discussed and added in or other individual documents.

[\[I-D.ietf-detnet-controller-plane-framework\]](#)

4.1. Support Aggregated Flow Identification

Current IPv6 aggregated flow identification is generally based on 5 or 6 tuples, IP prefixes, or wildcards as indicated in [\[RFC8938\]](#). However, in deterministic networks the number of individual flows can be huge, and they may randomly join and leave the aggregated flow at each hop as described in [Section 3](#). Such behaviours lead to the difficulty in identifying aggregated flows by relying on the prefixes or wildcards.

In addition, the deterministic services may demand different deterministic QoS requirements according to different levels of application requirements. The flow identification with service-level aggregation should be supported. The flow identification is also used to quickly push a packet to a suitable queue. In scaling networks, there are large amount of flows requiring deterministic latency service and normal forwarding service. Explicit flow identification makes it easier to quickly distinguish the DetNet flows without requiring the longest match rule on multiple tuples in IP data plane. Therefore, explicit aggregated flow identification SHOULD be supported.

4.2. Support Information used by Functions Ensuring Deterministic Latency

According to [Section 3.1](#), the deterministic network should support synchronized or asynchronous queuing mechanisms. Different queuing mechanisms require different information to be defined as the DetNet-specific metadata to help the functions of ensuring deterministic latency, including regulation, queue management, etc. For instance, the data plane needs to support the identification of cycle for cyclic queuing and forwarding or the latency related information for time based queuing in order to enable the applicability of the respective queuing and/or scheduling mechanisms in the large scale network.

When different queuing mechanisms are deployed at a network node, metadata used for each queuing mechanism should be provided at the same time. When multiple metadata carried in one packet, metadata should be self-described and extensible to tolerate variable number of metadata. Meanwhile, extra data will cause extra processing, referring to fixed or variable length lookups, total number of read/write access to the packet header, etc. So the data plane processing

efficiency also needs to be considered when ensuring deterministic latency, but the specific method or solution is out of scope of this document.

This document does not specify what metadata and what format to be carried in data plane. The solution document should be specific enough on why and how the information carried as data plane meta data works in conjunction with the queuing or other functions and how it helps the deterministic network deployment.

5. Conclusion

This document specifies the technical requirements for scaling deterministic networks and the corresponding data plane enhancement requirements. Some of the proposed queuing mechanisms and trials are cited, and the authors of the document think those proposals give reasonably sound insights to enhance the current queuing mechanisms to meet the requirements of scaling deterministic networks.

6. Security Considerations

When DetNet flows span multiple domains and require multi domain collaboration, security guarantee needs to be strengthened. More considerations will be described later.

7. IANA Considerations

This section will be described later.

8. Acknowledgements

The authors would like to thank Daivd Black, Jino0 Joung, Lou Berger, Balazs Varga, Fan Yang, Tianran Zhou, Yaakov Stein for helpful suggestions. The authors also would like to thank Liang Geng, Peter Willis, Shunsuke Homma and Li Qiang for their previous works.

9. Contributors

The following people have substantially contributed to this document:

Zongpeng Du
China Mobile
EMail: duzongpeng@chinamobile.com

Lei Zhou
New H3C Technologies
Email: zhou.leih@h3c.com

10. Informative References

- [ANDREWS] M, A., "Instability of FIFO in the permanent sessions model at arbitrarily small network loads. ACM Trans. Algorithms, vol. 5, no. 3, pp. 1-29, doi:10.1145/1541885.1541894.", July 2009.
- [BOUILLARD]
Corronc, B. A. B. M. A. E. L., "Deterministic network calculus: From theory to practical implementation. in Networks and Telecommunications. Hoboken, NJ, USA: Wiley, doi: 10.1002/9781119440284.", January 2018.
- [Fast-Ethernet-MII-clock]
"Fast Ethernet MII clock".
- [G.8262] International Telecommunication Union, "Timing characteristics of a synchronous equipment slave clock", ITU-T Recommendation G.8262, November 2018.
- [G.8273] International Telecommunication Union, "Framework of phase and time clocks", ITU-T Recommendation G.8273, March 2018.
- [I-D.dang-queuing-with-multiple-cyclic-buffers]
Liu, B. and J. Dang, "A Queuing Mechanism with Multiple Cyclic Buffers", Work in Progress, Internet-Draft, [draft-dang-queuing-with-multiple-cyclic-buffers-00](https://datatracker.ietf.org/doc/html/draft-dang-queuing-with-multiple-cyclic-buffers-00), 22 February 2021, <<https://datatracker.ietf.org/doc/html/draft-dang-queuing-with-multiple-cyclic-buffers-00>>.
- [I-D.eckert-detnet-bounded-latency-problems]
Eckert, T. T. and S. Bryant, "Problems with existing DetNet bounded latency queuing mechanisms", Work in Progress, Internet-Draft, [draft-eckert-detnet-bounded-latency-problems-00](https://datatracker.ietf.org/doc/html/draft-eckert-detnet-bounded-latency-problems-00), 12 July 2021, <<https://datatracker.ietf.org/doc/html/draft-eckert-detnet-bounded-latency-problems-00>>.
- [I-D.geng-detnet-requirements-bounded-latency]
Geng, L., Willis, P., Homma, S., and L. Qiang, "Requirements of Layer 3 Deterministic Latency Service", Work in Progress, Internet-Draft, [draft-geng-detnet-requirements-bounded-latency-03](https://datatracker.ietf.org/doc/html/draft-geng-detnet-requirements-bounded-latency-03), 7 July 2019, <<https://datatracker.ietf.org/doc/html/draft-geng-detnet-requirements-bounded-latency-03>>.

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

Malis, A. G., Geng, X., Chen, M., Qin, F., Varga, B., and C. J. Bernardos, "Deterministic Networking (DetNet) Controller Plane Framework", Work in Progress, Internet-Draft, [draft-ietf-detnet-controller-plane-framework-05](https://datatracker.ietf.org/doc/html/draft-ietf-detnet-controller-plane-framework-05), 26 September 2023, <<https://datatracker.ietf.org/doc/html/draft-ietf-detnet-controller-plane-framework-05>>.

[I-D.qiang-detnet-large-scale-detnet]

Qiang, L., Geng, X., Liu, B., Eckert, T. T., Geng, L., and G. Li, "Large-Scale Deterministic IP Network", Work in Progress, Internet-Draft, [draft-qiang-detnet-large-scale-detnet-05](https://datatracker.ietf.org/doc/html/draft-qiang-detnet-large-scale-detnet-05), 2 September 2019, <<https://datatracker.ietf.org/doc/html/draft-qiang-detnet-large-scale-detnet-05>>.

[IEEE802.1Qav]

IEEE, "IEEE Standard for Local and metropolitan area networks -- Virtual Bridged Local Area Networks - Amendment 12: Forwarding and Queuing Enhancements for Time-Sensitive Streams", IEEE 802.1Qav-2009, DOI 10.1109/IEEESTD.2010.8684664, 5 January 2010, <<https://doi.org/10.1109/IEEESTD.2010.8684664>>.

[IEEE802.1Qbv]

IEEE, "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks - Amendment 25: Enhancements for Scheduled Traffic", IEEE 802.1Qbv-2015, DOI 10.1109/IEEESTD.2016.8613095, 18 March 2016, <<https://doi.org/10.1109/IEEESTD.2016.8613095>>.

[IEEE802.1Qch]

IEEE, "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks - Amendment 29: Cyclic Queuing and Forwarding", IEEE 802.1Qch-2017, DOI 10.1109/IEEESTD.2017.7961303, 28 June 2017, <<https://doi.org/10.1109/IEEESTD.2017.7961303>>.

[IEEE802.1Qcr]

IEEE, "IEEE Standard for Local and Metropolitan Area Networks -- Bridges and Bridged Networks - Amendment 34: Asynchronous Traffic Shaping", IEEE 802.1Qcr-2020, DOI 10.1109/IEEESTD.2020.9253013, 6 November 2020, <<https://doi.org/10.1109/IEEESTD.2020.9253013>>.

[IEEE802.1Qdv]

IEEE, "IEEE Standard for Local and metropolitan area networks -- Enhancements to Cyclic Queuing and Forwarding", IEEE 802.1Qdv-2023, 12 July 2023.

[IEEE802.1TSN]

IEEE Standards Association, "IEEE 802.1 Time-Sensitive Networking Task Group",
<<https://www.ieee802.org/1/pages/tsn.html>>.

[Multiple-CQF]

IEEE, "Multiple Cyclic Queuing and Forwarding."
<https://www.ieee802.org/1/files/public/docs2021/new-finn-multiple-CQF-0921-v02.pdf>".

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

[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>>.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", [RFC 2475](#), DOI 10.17487/RFC2475, December 1998,
<<https://www.rfc-editor.org/info/rfc2475>>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, [RFC 8200](#), DOI 10.17487/RFC8200, July 2017,
<<https://www.rfc-editor.org/info/rfc8200>>.

[RFC8578] Grossman, E., Ed., "Deterministic Networking Use Cases", [RFC 8578](#), DOI 10.17487/RFC8578, May 2019,
<<https://www.rfc-editor.org/info/rfc8578>>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", [RFC 8655](#), DOI 10.17487/RFC8655, October 2019,
<<https://www.rfc-editor.org/info/rfc8655>>.

- [RFC8938] Varga, B., Ed., Farkas, J., Berger, L., Malis, A., and S. Bryant, "Deterministic Networking (DetNet) Data Plane Framework", [RFC 8938](#), DOI 10.17487/RFC8938, November 2020, <<https://www.rfc-editor.org/info/rfc8938>>.
- [RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", [RFC 8986](#), DOI 10.17487/RFC8986, February 2021, <<https://www.rfc-editor.org/info/rfc8986>>.
- [THOMAS] Mifdaoui, T. L. L. B. J. A. A., "On cyclic dependencies and regulators in time-sensitive networks. IEEE Real-Time Syst. Symp. (RTSS), York, U.K., pp. 299-311.", December 2019.

Appendix A. Examples of Scaling Deterministic Network Trials

Some trials have been carried out to verify the concept of large-scale deterministic networks.

In order to verify the deterministic technology of large-scale networks, a trial of Deterministic IP on China Environment for Network Innovations (CENI), which is a network built for new network technology trial, was deployed. A network with a distance of 3,000 km over 13 hops was tested, and the jitter was controlled within 100us.

In order to verify the remote control on Deterministic IP, which required that the latency should be controlled within 4 ms and jitter should be controlled within 20 us. A trial cooperated with Baosteel spanned 600 km was deployed. Baosteel is a Chinese steel company and put forward this demand. Both of the first and second trials are based on a frequency synchronization solution. The mechanism details could be found in . [[I-D.dang-queuing-with-multiple-cyclic-buffers](#)][[I-D.qiang-detnet-large-scale-detnet](#)].

In order to realize multi flows synchronization on an inter-provincial network in an exhibition, Emergen proposed the requirement that two flows of video and virtual reality (VR) were sent from province A, and arrived at province B together, so people can see the synchronization of video collected by camera and the VR model. This requirement was proposed to facilitate the virtual industry product deployment. Due to time and other problems, it was realized by the edge network device for a relatively lower levels of service level agreement (SLA).

Teaming up with a smart factory operator, network operators, equipment companies, and universities, ETRI demonstrated an ultra-low latency, high-reliability 5G wired and wireless network-based remote industrial Internet of Things (IIoT) service by connecting a control center and a smart factory through three different operators' networks at a distance of 280 km. In this trail, it was demonstrated that real-time remote smart manufacturing service is possible by making round-trip delay below 3 ms within a smart factory and below 10 ms between remote 5G industrial devices. In the future, the team plans to examine feasibility of large-scale deterministic networking by connecting smart factories in Gyeongsan, South Korea and Oulu, Finland.

These trials show that both operators and enterprise users begin to put forward requirements for the certainty of large-scale networks, but the implementation technologies are not exactly the same.

Authors' Addresses

Peng Liu
China Mobile
Beijing
100053
China
Email: liupengyjy@chinamobile.com

Yizhou Li
Huawei
Nanjing
210012
China
Email: liyizhou@huawei.com

Toerless Eckert
Futurewei Technologies USA
Santa Clara, 95014
United States of America
Email: tte@cs.fau.de

Quan Xiong
ZTE Corporation
Wuhan
430223
China
Email: xiong.quan@zte.com.cn

Jeong-dong Ryoo
ETRI
Daejeon
34129
South Korea
Email: ryoo@etri.re.kr

Shiyin Zhu
New H3C Technologies
Beijing
100094
China
Email: zhushiyin@h3c.com

Xuesong Geng
Huawei
Beijing
100095
China
Email: gengxuesong@huawei.com

