

Workgroup:

Deterministic Networking Working Group

Internet-Draft:

draft-liu-detnet-large-scale-requirements-00

Published: 13 October 2021

Intended Status: Informational

Expires: 16 April 2022

Authors: P. Liu Z. Du Y. Li
 China Mobile China Mobile Huawei

Requirements of large scale deterministic network

Abstract

Aiming at the large scale deterministic network, this document specifies the technical and operational requirements when the different deterministic levels of applications co-exist and are transported over a wide area.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

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 16 April 2022.

Copyright Notice

Copyright (c) 2021 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- [1. Introduction](#)
- [2. Diversified application requirements and trial status](#)
 - [2.1. Different levels of application requirements](#)
 - [2.2. Examples in terms of large scale](#)
- [3. Technical requirements in large scale deterministic network](#)
 - [3.1. Tolerate time asynchrony](#)
 - [3.1.1. Support asynchronous clocks across domains](#)
 - [3.1.2. Tolerate clock jitter & wander within a clock synchronous domain](#)
 - [3.1.3. Provide mechanisms not requiring full time synchronization](#)
 - [3.1.4. Support asynchronization based methods](#)
 - [3.2. Support the large single-hop propagation latency](#)
 - [3.3. Accommodate the higher link speed](#)
 - [3.4. Be scalable](#)
 - [3.4.1. Be scalable to numerous network devices](#)
 - [3.4.2. Be scalable to massive traffic flows](#)
 - [3.5. Tolerate failures of links or nodes and topology changes](#)
 - [3.6. Support incremental device updates](#)
- [4. Summary of the proposed queuing mechanisms besides TSN and IntServ/GS](#)
- [5. Conclusion](#)
- [6. Security Considerations](#)
- [7. IANA Considerations](#)
- [8. Acknowledgements](#)
- [9. Normative References](#)
- [Authors' Addresses](#)

1. Introduction

Since the time sensitive network and deterministic network were proposed, the application use case has always been the hottest topic. It may originate from the industry, audio and video, and has more demand in the era of 5G and industrial Internet. As years of development, TSN has been used in several industries, and has enough public awareness of the industry for it's scope. DetNet also has done a lot of work and the standards are mature, and people concern more on how to guarantee the deterministic demand on Layer 3 network.

However, when to provide deterministic network services, network providers always face the problem of how to match application needs to the technology, so more work are needed for network service providers to successfully sell DetNet type services to customers. For example,

The service level objective definitions, considering absolute or relative latency and jitter bounds, flows types and physical network scale;

The suitable queuing mechanisms, considering more option of queuing mechanisms for different service level;

The deployment issues, considering how to integrate into existing networks, service, and controller-plane.

2. Diversified application requirements and trial status

2.1. Different levels of application requirements

[[RFC8578](#)] gives some requirements of industry, electricity, buildings, etc.. some of them clearly specify the requirements for latency and jitter, while some not for the jitter. Different types of users have different demand, just as network provider provide different network services for personal business or enterprise business, so as to the detnet service for different uses.

One kind has critical SLAs requirement, such as remote control or cloud 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 can't tolerate with any non-deterministic situation and can pay more on the network service.

Another kind has relatively lower levels of SLA requirement, such as cloud gaming, cloud VR and online meeting for "consumer" networks. Users of these applications hope to have a better network experience, but they can tolerate it to a certain extent if the network quality is not good sometime. So they are willing to spend more money for high-quality network services. In some aspects, because such services have no industry barriers and can tolerate exceeding the upper boundary of latency within a small probability, they have relatively lower requirements for the network and may be easier to deploy.

Different application needs are actually related to cost. For strict deterministic services, strict 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 equipment or share corresponding network resources.

Critical latency requirements:

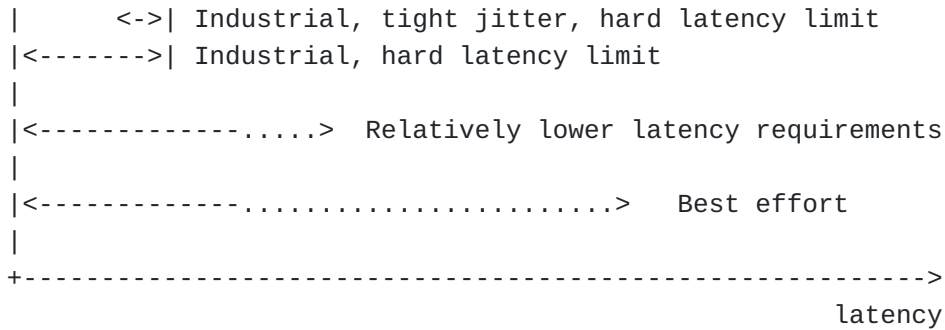


Figure 1: Figure 1: Different levels of application requirements

2.2. Examples in terms of large scale

Ahead of the formulation of standards, some trials have been carried out to verify 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) was deployed, which is a network built for new network technology's trial. This trial spanned 3000km and has 13-hops devices, the jitter is controlled within 100us.

In order to verify the remote control on Deterministic IP, which required that the latency should be controlled within 4ms and jitter should be controlled within 20us. A trial cooperated with Baosteel spanned 600km 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.

In order to realize multi flows synchronization on inter provincial network in an exhibition, Emergen proposed the requirements that two flows of video and VR were sent from province A, and arrived at province B together, so the 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 SLA.

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.

3. Technical requirements in large scale deterministic network

Due to the different kinds of application requirements in large scale network, the corresponding technique requirements should be considered.

3.1. Tolerate time asynchrony

3.1.1. Support asynchronous clocks across domains

A large scale network may span over multiple networks with one or more administrators. 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[[TSN-Qbv](#)][[TSN-Qch](#)][[TSN-Qav](#)]. However, different TSN islands may have different clocks which are not synchronized as shown in Figure 2, where the time difference of two TSN domain is D. DetNet needs to connect these two TSN domains together and provide end-to-end deterministic latency service. The mechanism adopted by DetNet should be able to support the interaction across time domains by putting extra buffer space at the ingress of a new domain or increase the dead time as a guard band, or using some timing compensation mechanism. This document does not intend to list all the potential ways.

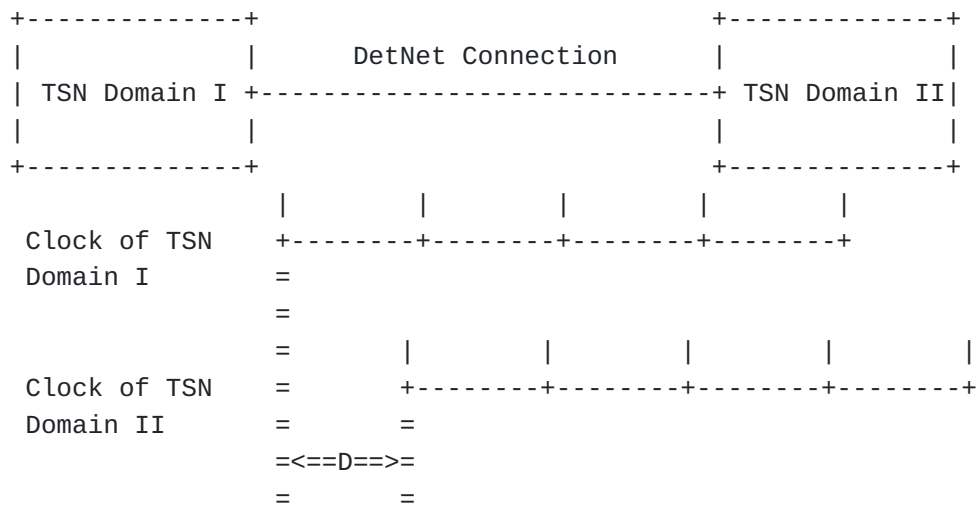


Figure 2: Figure 2: TSN islands interconnecting

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](#)], 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 large scale networks should be able to recover or absorb such time variance within a domain and across multiple domains.

3.1.3. Provide mechanisms not requiring full time synchronization

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

3.1.4. Support asynchronization based methods

There are large amount of traffic flows in large scale network and some of them are acyclic. Asynchronization based methods can meet the requirements of those traffic flow. Moreover, The mechanisms not requiring the time and/or frequency synchronization eliminate the hardware cost and difficulty at the network nodes. [[TSN-Qcr](#)] conceptually uses per-flow based asynchronous shaper to achieve bounded latency. The formula proof shows its effectiveness. 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 subsection 3.3 should be carefully met.

3.2. Support the large single-hop propagation latency

In a large scale network, a single hop distance is enough to generate a larger latency. The speed of optical transmission in fiber is 200km/ms. Thus the propagation delay of a single hop can be in the order of low number of msec. It is much great than a LAN, and introduces impacts on queuing mechanisms, such as cyclic or time aware scheduling method.

For cyclic based method, suppose a large scale 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 each 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, then the length of cycle must be set to at least 20 us. However since packet's arrival time varies within the receiving cycle, larger cycle length means larger delay variance.

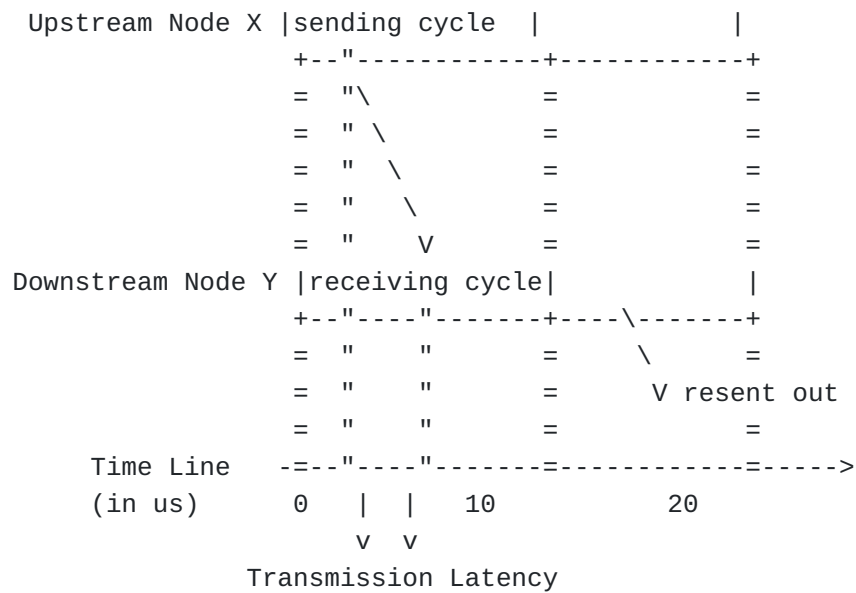


Figure 3: Figure 3: The influence of transmission latency on cyclic method

3.3. Accommodate the higher link speed

The large scale network normally uses the higher link speed, especially for its backbone. Current deterministic mechanisms used in the local network is usually deployed in link speed of 10Mbps or 1Gbps and possibly 10Gbps; and even higher is 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 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 low number of usec or sub-usec) 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 large scale network, the number of flows can be hundreds or tens of thousands. They can be aggregated into a few number of deterministic path or tunnels. It is practical to have a few flow-based or aggregated-flow based status in a local network. But in higher speed and larger scale network, it is hardly feasible. If TSN-ATS[] is in use, it requires more number of buffers comparing to the other full/partial time synchronized mechanisms. Therefore it requires optimizations to support higher link speed.

3.4. Be scalable

Comparing to a LAN, large scale network may has more network devices and traffic flows, and there is a greater possibility of adding or removing network devices and traffic flows. The deterministic latency forwarding mechanisms must scale to networks of significant size with numerous network devices and massive traffic flows.

3.4.1. Be scalable to numerous network devices

The increase or decrease of network devices in large-scale networks is more frequent than that in LAN. 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 . 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 network may need to connect to ~100 K base stations (serving multiple mobile networks operators'), and this number will only increase with 5G.

3.4.2. Be scalable to massive traffic flows

It is almost impossible to identify individual IP flow at the Detnet data plane because of the large overhead and resource reservation for massive number of flows. Detnet allows the leverage of the flow aggregation. With the large scaling of the network, proper provision at the control plane to accommodate such higher aggregation is required. Individual flow may join and exit the aggregated flow rapidly which causes the dynamic in identification of the aggregated Detnet flow. The wildcards, value range used in the identification may have to change in order to ensure the aggregated flows have compatible deterministic characteristics. If each ultra-low-latency slice or MNO is treated as a separate deterministic latency traffic flow (or tunnel), then even if each base station has a limited number of ultra-low latency slices or MNOs (e.g. ~10), there will still be a lot of, ~1M, deterministic latency traffic flows on one network simultaneously.

3.5. Tolerate failures of links or nodes and topology changes

Network link failures are more common in large-scale networks. Path switching or re-convergence of routing will cause high latency of packet loss and retransmission, which is usually in seconds before the network is stable again. It is necessary to support certain mechanisms to adapt to failures of links or nodes and topology changes.

The change of path or topology poses a higher challenge to packet replication and elimination. The full disjoint paths when

implementing PREOF gives the better survival chance when one of the nodes 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.

3.6. Support incremental device updates

Do more shaping work on edge devices, so as to reduce the task of intermediate devices, which can be an advantage of deterministic network compared with the dedicated network. Since some applications that requires relatively loose levels of SLA, it will be acceptable for those applications to tolerate a deterministic low probability to exceed the upper boundary of latency. For those applications, some simple solutions that may be realized by update and configure the ingress and egress devices or part of network devices are expected. When the devices or traffic flows change, it can be realized through simple configuration. Meanwhile, the critical SLA of some applications, can be achieved by adding the existing or other new mechanisms and updating more devices.

4. Summary of the proposed queuing mechanisms besides TSN and IntServ/GS

There are some proposed queuing mechanisms beside TSN and IntServ/ Guaranteed service, which are not included in draft-ietf-detnet-bounded-latency.

[\[I-D.dang-queuing-with-multiple-cyclic-buffers\]](#) and [\[I-D.qiang-detnet-large-scale-detnet\]](#) are based on frequency synchronization and multiple cyclic buffers, and can be proved to provide the bounded latency and jitter. They use the flow aggregation and the Scalability is also good.

[\[I-D.du-detnet-layer3-low-latency\]](#) proposes a method to decrease the micro-burst based on a adjustable buffer. Though it can't prove a strict bounded latency, and the levels of deterministic is medium, it doesn't need the synchronization and have a good scalability, and can be easier deployed.

[\[I-D.stein-srtsn\]](#) is to encapsulate the time stamp in the packet, based on which can adjust forwarding behavior. The scalability is a driving force behind this draft, and the determinism is statistical in theory.

[\[I-D.shi-quic-dtp\]](#) is also based on the time stamp, which is a layer 4 solution. It's listed there to show that the latency is more important than before of the application requirements, and there is also queuing mechanism besides Layer 3 solution.

	Mechanisms	Levels of determinacy	Synchroni- zation	Cost	Scalabi- lity	Flow aggregation
1	draft-dang -queuing-with -multiple- cyclic-buffers /draft-qiang -detnet-large- scale-detnet	high	yes	high	good	yes
2	draft-du- detnet-layer3 -low-latency	medium	no	high	good	yes
3	draft-stein- srtsn	statistical determinism	yes	high	good	??
4	draft-shi- quic-dtp	low	yes	low	good	no

Figure 4: Proposed queuing mechanisms besides TSN and IntServ/GS

5. Conclusion

This draft specifies the technical requirements when ensuring the deterministic features in the large scale networks. Some of the proposed queueing mechanisms are analyzed and the authors of the document think those proposals give reasonably insights to enhancement the current queueing mechanisms to meet the deterministic requirements of the large scale networks.

6. Security Considerations

TBD.

7. IANA Considerations

TBD.

8. Acknowledgements

Thanks to Toerless Eckert, Yaakov Stein for helpful suggestion.
Thanks to Liang Geng, Peter Willis, Shunsuke Homma and Li Qiang for their previous work.

9. Normative References

[Fast-Ethernet-MII-clock]

"Fast Ethernet MII clock".

[G.8262] "G.8262 : Timing characteristics of a synchronous Ethernet equipment slave clock", 2009.

[G.8273] "G.8273: Framework of phase and time clocks".

[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, 22 February 2021, <<https://www.ietf.org/archive/id/draft-dang-queuing-with-multiple-cyclic-buffers-00.txt>>.

[I-D.du-detnet-layer3-low-latency] Du, Z. and P. Liu, "Micro-burst Decreasing in Layer3 Network for Low-Latency Traffic", Work in Progress, Internet-Draft, draft-du-detnet-layer3-low-latency-03, 11 July 2021, <<https://www.ietf.org/archive/id/draft-du-detnet-layer3-low-latency-03.txt>>.

[I-D.eckert-detnet-bounded-latency-problems] Eckert, T. and S. Bryant, "Problems with existing DetNet bounded latency queuing mechanisms", Work in Progress, Internet-Draft, draft-eckert-detnet-bounded-latency-problems-00, 12 July 2021, <<https://www.ietf.org/archive/id/draft-eckert-detnet-bounded-latency-problems-00.txt>>.

[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, 7 July 2019, <<https://www.ietf.org/archive/id/draft-geng-detnet-requirements-bounded-latency-03.txt>>.

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

Qiang, L., Geng, X., Liu, B., Eckert, T., Geng, L., and G. Li, "Large-Scale Deterministic IP Network", Work in Progress, Internet-Draft, draft-qiang-detnet-large-scale-detnet-05, 2 September 2019, <<https://www.ietf.org/archive/id/draft-qiang-detnet-large-scale-detnet-05.txt>>.

[I-D.shi-quick-dtp] Cui, Y., Liu, Z., Shi, H., Zhang, J., Zheng, K., and W. Wang, "Deadline-aware Transport Protocol", Work in Progress, Internet-Draft, draft-shi-quick-dtp-04, 25 July 2021, <<https://www.ietf.org/archive/id/draft-shi-quick-dtp-04.txt>>.

[I-D.stein-srtsn] Stein, Y. (., "Segment Routed Time Sensitive Networking", Work in Progress, Internet-Draft, draft-

stein-srtsn-01, 29 August 2021, <<https://www.ietf.org/archive/id/draft-stein-srtsn-01.txt>>.

- [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>>.
- [RFC8578] Grossman, E., Ed., "Deterministic Networking Use Cases", RFC 8578, DOI 10.17487/RFC8578, May 2019, <<https://www.rfc-editor.org/info/rfc8578>>.
- [TSN-Qav] Group, I. T. N. (. T., "802.1Qav - Forwarding and Queuing Enhancements for Time-Sensitive Streams", 2009.
- [TSN-Qbv] Group, I. T. N. (. T., "802.1Qbv - Enhancements for Scheduled Traffic", 2015.
- [TSN-Qch] Group, I. T. N. (. T., "P802.1Qch - Cyclic Queuing and Forwarding", 2015.
- [TSN-Qcr] IEEE, "P802.1Qcr - Bridges and Bridged Networks Amendment: Asynchronous Traffic Shaping", 2020.

Authors' Addresses

Peng Liu
China Mobile
Beijing
100053
China

Email: liupengyiy@chinamobile.com

Zongpeng Du
China Mobile
Beijing
100053
China

Email: duzongpeng@chinamobile.com

Yizhou Li
Huawei
Nanjing
210012
China

Email: liyizhou@huawei.com