

Deterministic Networking Working Group
Internet-Draft
Intended status: Informational
Expires: January 9, 2020

P. Liu
L. Geng
China Mobile
July 08, 2019

Dynamic Latency Guarantee
draft-liu-detnet-dynamic-latency-guarantee-01

Abstract

Aiming at the deterministic demand for network latency in future vertical industry applications, this document analyzes the existing latency control methods for data transmission, points out the possible shortcomings, and proposes some directions for optimizing the latency control method. .

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 January 9, 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

Internet-Draft

Deterministic Networking

July 2019

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
2.	Technologies of Latency Control	2
2.1.	IEEE 802.1Qav Forwarding and Queuing Enhancements for Time-Sensitive Streams	3
2.2.	IEEE 802.1Qbv Enhancements for Scheduled Traffic	3
2.3.	IEEE 802.1Qbu Frame Preemption	3
3.	Problems and Requirments	3
3.1.	Problems in Bounded Latency	4
3.2.	Requirments of Deterministic latency	4
4.	Solutions	4
5.	Conclusion	6
6.	Security Considerations	6
7.	IANA Considerations	6
8.	References	6
8.1.	Normative References	7
8.2.	Informative References	7
	Authors' Addresses	7

[1.](#) Introduction

New types of services such as AR/VR, V2X, industrial motion control, etc. have stringent requirements for latency and stability. In order to meet those requirements, some network technologies such as time-sensitive network, deterministic network, etc., have proposed corresponding technical means to provide network bearers with deterministic latency and packet loss rate to guarantee the service experience. TSN includes a set of standards developed by the IEEE 802.1 Working Group's. Deterministic network (DETNET) is based on the mechanism of TSN and committed to applying the method to the IP layer to provide more reliable and stable network transmission. This document will present some problems when applying TSN in DETNET, and try to propose reference methods to solve the corresponding problems.

[2.](#) Technologies of Latency Control

Based on time synchronization, TSN has a range of bounded latency technologies.

[2.1.](#) IEEE 802.1Qav Forwarding and Queuing Enhancements for Time-Sensitive Streams

IEEE 802.1Qav inherited from the AVB, including priority mapping algorithms and Credit-based Traffic Shaping algorithms. The priority mapping algorithms is to mapping the priority to 'traffic class', which represents whether the stream is time sensitive or not. Credit-based Traffic Shaping algorithms provide the method to allocate bandwidth of different streams.

[2.2.](#) IEEE 802.1Qbv Enhancements for Scheduled Traffic

In IEEE 802.1Qbv, the gate control list is created according to the actual stream and timescale. It contains the transmission sequence of all streams, and controls whether the data stream of each priority is sent at the current time or not. All streams will be transmitted strictly according to the current list. More Than This, IEEE 802.1Qbv also defines the guard band mechanism and spares part of the time to guarantee the transmission of high priority data frames at the beginning of the next time slice.

[2.3.](#) IEEE 802.1Qbu Frame Preemption

In the preemption mechanism, high-priority frames can interrupt the transmission of low-priority data frames unless low-priority data frames can no longer be fragmented. This standard fully guarantees the transmission delay of the highest priority data frame, and also reduces the guard band in IEEE 802.1Qbv to 127 bytes. The frame preemption mechanism changes the transmission rules of the ethernet frame and is used in conjunction with the IEEE 802.3Qbr .

In addition to these, there are also other standards to guarantee the sequence of receiving data streams, which are fine-grained traffic scheduling technology and the key technologies of TSN in bounded latency.

[3.](#) Problems and Requirments

DETNET refers to the bounded latency mechanism of TSN, so it needs to pay attention to some problems in the bounded latency mechanism. There are several standards refers to bounded latency. Users can decide whether to use a specific standard or not, which depends on the requirments of network and business. Some TSN testbeds have been established these years whose basic concept is realizing 802.1Qbv to ensure the deterministic transmission of time sensitive stream. Though it realized ignoring the interfere of background stream, the testbed was too simple. In fact, networking is complicated. There

will be more than two kind of streams being transmitted. So it is not that easily to apply those mechanisms on real networks.

[3.1.](#) Problems in Bounded Latency

Because of the complicated of real networks, there may be some situations that the preemptable data frame transmission delay is too large or cannot be transmitted. Thoes might occur when both Enhancements for Scheduled Traffic and Frame Preemption are enabled.

Except for the highest priority, the others may be preempted by the time slice to wait for transmission. In the actual scenario, the preemptable data frame is not necessarily a completely non-time sensitive frame, so it also need to guarantee the transmission of some preemptable frame. However, Under the current mechanism, there may be multiple preemption to cause a very large transmission delay or no transmission of preemptable frame, depending on the size of the express frame and the period of the timescale. In an actual scenario, a data frame with a Secondary high priority may also be a time-sensitive. If it cannot be transmitted or the transmission delay is large, the service cannot be operated.

[3.2.](#) Requirments of Deterministic latency

Deterministic network includes deterministic latency and deterministic packet loss. We need to think how to apply the bounded latency mechanism effectively. Before using the bounded latency mechanism, network manager needs to know enough about the network and applications. For example, which kind of stream is time sensitive?

How about the frame's transceiver frequency of thoes stream? How much bandwidth does it need? ... When you have a clear understanding of the real-time state of the network, you can configure a delay-limited algorithm for the network.

However, the transmission state of the network is not invariable. Some transfer table might make corresponding adjustments according to the current network situation. So the parameters that have been configured before should also be changed. More than this, the bounded latency mechanism also need a feedback system to receive current network status and adjust/reconfigure the network.

[4.](#) Solutions

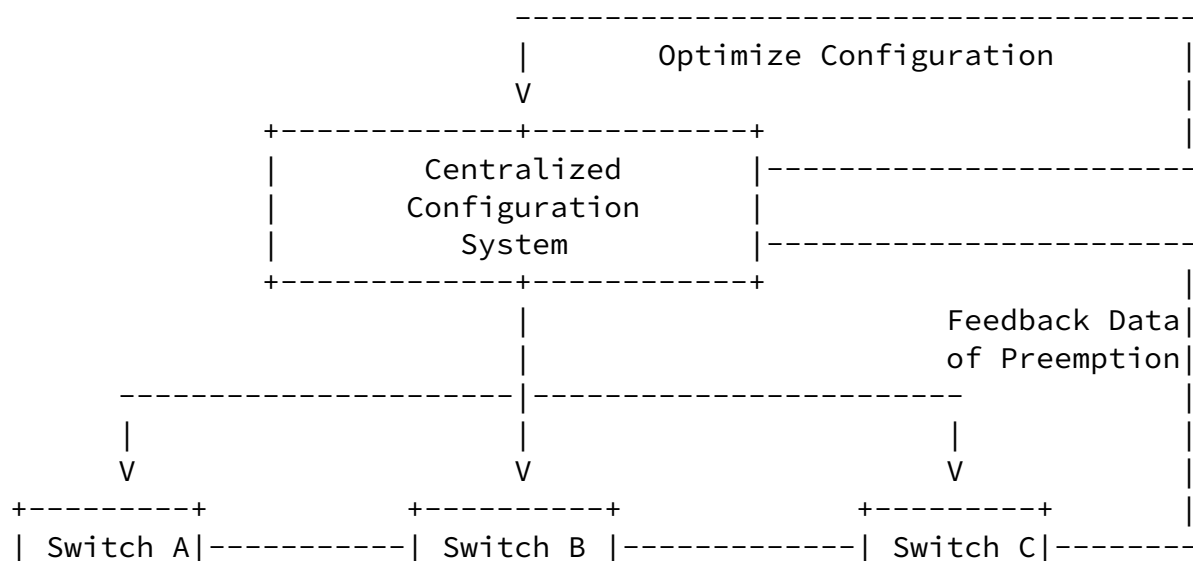
The implementation of the mechanism to guarantee latency requires sophisticated calculation, including timescale and gate control tist . When the stream in the network becomes diverse, it will consume a lot of computing resources to schedule each stream. Therefore, a single transmission rule may not be able to meet the problem of

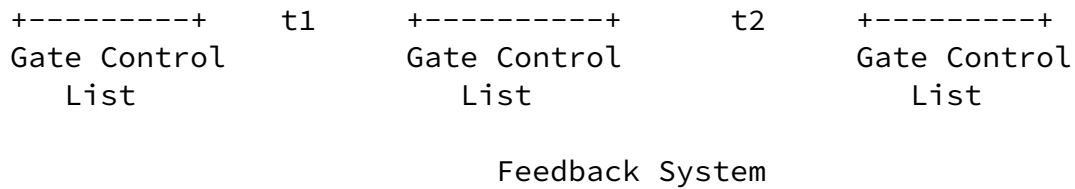
multiple streams' transmission. Worst of all, the gate control list is not properly calculated, the network may not transmit or failure.

Dynamic latency guarantee is a way of thinking based on the latency guarantee of the whole network. that is, to dynamically adjust the priority through the current network condition and the transmission of data stream, and a feedback system is needed to optimize the system. One of the reasons for this situation is that the prediction or mastery of the transmission of frames in the network is not accurate, so a feedback system is needed to tell the network to centrally configure the system. So it could help to optimize the gate control list to avoid the frequent occurring of this problems. The most basic case is that once there are multiple preemption occured, the switch need to report it to the Centralized Configuration System. It represent that there might be some unjustified configurations need to be reconfiguration. For example, distribute more bandwidth to the corresponding traffic class.

It should be noted that all devices in the network share the same gate control list. However, due to the difference in time of the transmission path, it is necessary to keep all devices in the network "asynchronous" to execute the gate control list. For example, when

the data frame is received by the device A, it is queued to be transmitted first in the currently divided time slice. When the frame is received by the device B, the time t_1 has elapsed. So the gate control list of device B needs to perform the time difference of t_1 with the A device, which can ensure that this frame arrives at every device with a first-transmitting in current time slice.





5. Conclusion

This draft described the existing mechanism of bounded latency and point out some problems when using them. It also proposed some reference methods to solve them. In the process of network evolution, there might also be more problems need to be noticed and discuss. For example, it also needs to consider whether the bounded latency mechanism of layer 2 can guarantee the deterministic processing of whole stack. There may be that deterministic forwarding mechanism is used in Layer 2, but due to the TCP/IP or other protocol in higher layer, data packets can not be processed in deterministic order in the queue, which leads to the uncertainty of latency.

6. Security Considerations

TBD.

7. IANA Considerations

TBD.

8. References

8.1. Normative References

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

Finn, N., Boudec, J., Mohammadpour, E., Zhang, J., Varga, B., and J. Farkas, "DetNet Bounded Latency", [draft-finn-detnet-bounded-latency-04](#) (work in progress), June 2019.

[I-D.ietf-detnet-architecture]

Finn, N., Thubert, P., Varga, B., and J. Farkas,
"Deterministic Networking Architecture", [draft-ietf-detnet-architecture-13](#) (work in progress), May 2019.

[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>>.

[8.2.](#) Informative References

[IEEE802.1Qav]
IEEE, "Forwarding and Queuing Enhancements for Time-Sensitive Streams (IEEE 802.1Qav)", 2009.

[IEEE802.1Qbu]
IEEE, "Frame Preemption", 2015.

[IEEE802.1Qch]
IEEE, "Cyclic Queuing and Forwarding", 2015.

[IEEE802.1Qbv]
IEEE, "Enhancements for Scheduled Traffic", 2016.

Authors' Addresses

Peng Liu
China Mobile
Beijing 100053
China

Email: liupengyjy@chinamobile.com

Liang Geng
China Mobile
Beijing 100053
China

Email: gengliang@chinamobile.com