This document is an individual contribution to the Time-Sensitive Networking Task Group of the IEEE 802.1 working group, by Tongtong Wang and Norman Finn, and is intended to further the progress of project P802.1DF. It is not an official draft of the Task Group or Working Group.

P802.1DF™/Dfw2
Draft Standard for Time-Sensitive Networking Profile for Service Provider Networks

Abstract: This standard defines profiles that select features, options, configurations, defaults, protocols, and procedures of bridges and end-stations defined in IEEE Std 802.1Q and IEEE Std 802.1CB that are necessary to provide Time-Sensitive Networking (TSN) quality of service features for non-fronthaul shared service provider networks. The standard also provides use cases, and informative guidance for network operators on how to configure their networks for those use cases.

Keywords: IEEE 802.1Q, Time-Sensitive Networking, service provider, profile, network calculus, network slicing, hard partitioning.

<< Editor's forward as used in other IEEE 802.1 documents, and IEEE boilerplate from the Word template, go here. >>

Contents

Draft Standard for Time-Sensitive Networking Profile for Service Provider Networks

1. Overview

1.1 Scope

This standard defines profiles that select features, options, configurations, defaults, protocols, and procedures of bridges and end-stations defined in IEEE Std 802.1Q and IEEE Std 802.1CB that are necessary to provide Time-Sensitive Networking (TSN) quality of service features for non-fronthaul shared service provider networks. The standard also provides use cases, and informative guidance for network operators on how to configure their networks for those use cases.

1.2 Purpose

This standard provides guidance for equipment vendors, designers, and operators of service provider networks that are shared by multiple users and applications, and that need the TSN Quality of Service (QoS) features offered by IEEE Std 802.1Q bridges. These networks have links with a very large bandwidth-delay product. The TSN features include dependable bandwidth and bounded latency.

1.3 Introduction

Service provider networks, also called carrier networks, provide connectivity between access node and content sources (usually in data centers) for multiple users and applications. While 5G new technologies come into market, URLLC (Ultra Reliable Low Latency Communication) applications (e.g. vertical applications / utility networks) bring on strict latency requirements over carrier networks.

As shown in Figure 1, a typical service provider topology is a layered ring network with sufficient redundant connections for better reliability and load balance. Usually user end stations are connected on the access ring network and multiple access rings could be linked to one aggregation ring that has larger bandwidth links. Backbone connections to the aggregation ring are not shown. A backbone layer is also possible in service provider networks, to further aggregate traffics and communicate with other service providers. Nevertheless, topology in Figure 1 is an example to show typical connections, not the constraints on all scenarios. All devices in service provider network are provider nodes, while a relay node can act as ingress and/or egress edge node. A service provider network in this project is in one TSN control domain, that means traffic shaping,

scheduling and buffering policy on provider nodes are coordinated such that end to end latency in service provider network can be guaranteed.

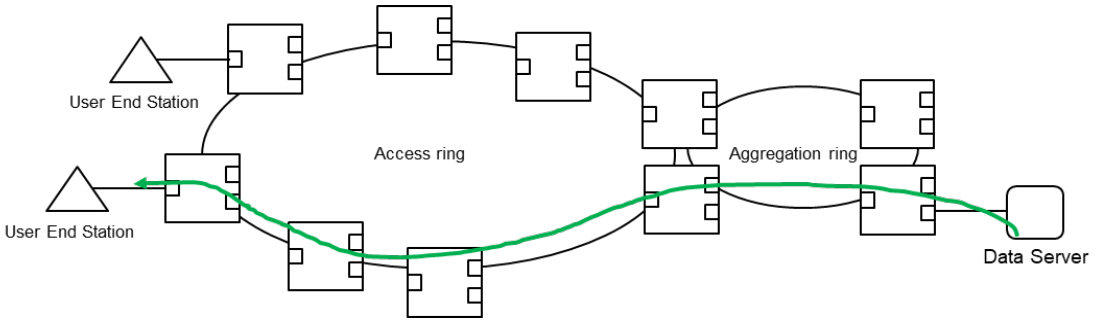

Figure 1 Example topology of service provider networks

To specify and explain the selection of features and options, this document:

a) Describes latency and packet loss requirements for critical applications in service provider networks (Clause 6).
b) Describes how the operation of bridges and bridged networks affects the quality of service provided by the carrier bridged network (Clause 7), provides details in the calculation of latency (Annex C), and the tradeoffs inherent in the use of TSN QoS techniques;
c) Specifies multiple profiles (Clause 8) that support the construction of bridged networks meeting latency requirements and jitter requirements.
d) Defines service provider network profile conformance requirements (Clause 5) for bridges and other network components meeting specific profile requirements, for end stations, and for time synchronization.
e) Describes how TSN techniques are used by Layer 3 devices such as routers, in conformance to the documents published by the Internet Engineering Task Force (IETF) Deterministic Networking (DetNet) Working Group (Clause 9).
f) Provides a Profile Conformance Statement (PCS, Annex A) to support clear detailed statements of equipment conformance to Service provider network profile requirements.
g) Provide basic knowledge on Network Calculus to assist network latency evaluation.

2. Normative references

The following referenced documents are indispensable for the application of this document (i.e., they must be understood and used, so each referenced document is cited in text and its relationship to this document is explained). For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments or corrigenda) applies.

IEEE Std 802, IEEE Standard for Local and Metropolitan Area Networks—Overview and Architecture.

IEEE Std 802.1Q, IEEE Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks.

IEEE Std 802.1CB, IEEE Standard for Local and Metropolitan Area Networks—Frame Replication and Elimination for Reliability.

IEEE Std 802.1DC, IEEE Standard for Local and Metropolitan Area Networks—Quality of Service Provision by Network Systems.

IEEE Std 802.3, IEEE Standard for Ethernet.

IEEE Std 802.3br, IEEE Standard for Ethernet—Amendment 5: Specification and Management Parameters for Interspersing Express Traffic.

IETF RFC xx05, (draft-ietf-detnet-ip-over-tsn) DetNet Data Plane: IP over IEEE 802.1 Time Sensitive Networking (TSN).

IETF RFC xx07, (draft-ietf-detnet-mpls-over-tsn) DetNet Data Plane: MPLS over IEEE 802.1 Time Sensitive Networking (TSN).

IETF RFC xx09, (draft-ietf-detnet-tsn-vpn-over-mpls) DetNet Data Plane: IEEE 802.1 Time Sensitive Networking over MPLS.

3. Definitions

For the purposes of this document, the following terms and definitions apply. The *IEEE Standards Dictionary Online* should be consulted for terms not defined in this clause.[1]

This standard makes use of the following terms defined in IEEE Std 802:

- bridge
- end station
- Ethernet
- forwarding
- frame
- Local Area Network (LAN)

This standard makes use of the following terms defined in IEEE Std 802.1Q:

- bridged network
- latency
- port
- priority-tagged frame
- Stream
- traffic class
- untagged frame
- Virtual Local Area Network (VLAN)
- VLAN Bridge
- VLAN-tagged frame

The following terms are specific to this standard:

[1] *IEEE Standards Dictionary Online* is available at: http://dictionary.ieee.org. An IEEE Account is required for access to the dictionary, and one can be created at no charge on the dictionary sign-in page.

4. Abbreviations

ATS – Asynchronous Traffic Shaping

CBS – Credit Based Shaper

DetNet – IETF Deterministic Networking

GBR – Guarantee Bit Rate

IETF – Internet Engineering Task Force

MPLS – Multi-Protocol Label Switching

QoS – Quality of Service

RFC – Request for Comments

TAS – Time Aware Shaper

5. Conformance

A claim of conformance to this standard is a claim that the behavior of an implementation of a bridge (5.3, 5.4) or of an end station (5.5, 5.6) meets the mandatory requirements of this standard and may support options identified in this standard.

<< Contributors' Note: This profile will distinguish between an end station, on the one hand, and a router, label switch, or other network device defined by IETF, on the other, in order to link our standard to the relevant RFCs from the IETF DetNet Working Group. It is possible that this will result in a third requirement clause. >>

5.1 Requirements terminology

For consistency with existing IEEE and IEEE 802.1 standards, requirements placed upon conformant implementations of this standard are expressed using the following terminology:

h) *Shall* is used for mandatory requirements;
i) *May* is used to describe implementation or administrative choices ("may" means "is permitted to," and hence, "may" and "may not" mean precisely the same thing);
j) *Should* is used for recommended choices (the behaviors described by "should" and "should not" are both permissible but not equally desirable choices).

The Profile Conformance Statement (PCS) proformas (see Annex A) reflect the occurrences of the words "shall," "may," and "should" within the standard.

The standard avoids needless repetition and apparent duplication of its formal requirements by using *is*, *is not*, *are*, and *are not* for definitions and the logical consequences of conformant behavior. Behavior that is permitted but is neither always required nor directly controlled by an implementer or administrator, or whose conformance requirement is detailed elsewhere, is described by *can*. Behavior that never occurs in a conformant implementation or system of conformant implementations is described by *cannot*. The word *allow* is used as a replacement for the phrase "Support the ability for," and the word *capability* means "can be configured to."

5.2 Profile Conformance Statement (PCS)

The supplier of an implementation that is claimed to conform to this standard shall provide the information necessary to identify both the supplier and the implementation, and shall complete a copy of the PCS proforma provided in Annex A.

5.3 Bridge requirements

5.4 Bridge options

5.4.1 Ingress Bridge

5.4.2 Core Bridge

5.4.3 Egress Bridge

5.5 End station requirements

5.6 End station options

6. Service provider networks

Possible emerging applications on 5G carrier networks are discussed in 3GPP TS 23.501 [B2], and summarized into three types of services shown in Table 1. Bandwidth sensitive services have strict requirement on average bandwidth and loose constraints on latency, while connection services just require message delivery from time to time. A new type of service is the delay critical service, that is the subject of this standard.

Service Catalog	Examples	Packet delay budget	Packet loss rate	Default Max Data Burst
Bandwidth Sensitive Services (GBR)	Conservational Voice	100ms	10^{-2}	N/A
	Conversational Video (live streaming)	150ms	10^{-3}	N/A
	Real Time Gaming	50ms	10^{-3}	N/A
Connection Services (Non-GBR)	Buffered Streaming Video	300ms	10^{-6}	N/A
Latency Sensitive Services (Delay Critical)	Intelligent Transport Systems	30ms	10^{-5}	1354 bytes
	Smart Grid Tele-protection	5ms	10^{-5}	255 bytes

Table 1 Typical services in 5G carrier networks

6.1 Bandwidth sensitive services

Bandwidth sensitive services such as conversational voice usually have relaxed delay requirements over a carrier network. Currently, all such packets traverse IP DiffServ networks (Differentiated Services, [B8], [B9], [B10]) using QoS methods like strict priority, weighted round robin, etc. Since carrier networks usually have large bandwidth and are utilized in as balanced a manner as possible, traffic congestion rarely happens to high priority data streams. Bandwidth sensitive applications get satisfactory performance as long as adequate throughput and buffering capabilities are reserved and provided in time.

<< Contributors' Note: Consider providing guidelines on bandwidth analysis over carrier networks, considering delay guaranteed bandwidth, rather than average bandwidth. >>

6.2 Latency sensitive services

Latency sensitive services put more stringent requirements on end-to-end latency over a carrier network; any packet arriving later than a certain deadline (the tolerable deadline) is regarded as a failure of packet delivery. An example of the latency sensitive service is the smart grid tele-protection application, which requires 5 ms end-to-end latency over carrier networks with 99.999% reliability [B2].

It is important to note the difference between minimum average latency and an assurance of a finite worst-case latency. In general, the best possible average latency is obtained using simple, strict priority-based forwarding. The average latency for packets of a highest-priority Stream can be very near the minimum latency. However, the worst-case latency for any given packet is generally many times larger than that minimum. Packets delivered late are equivalent to packets lost. The techniques called out in the present standard (see 7.3) are concerned with providing a finite worst-case latency. When this can be achieved, zero congestion loss is a pleasant byproduct.

Two main requirements in latency sensitive services are bounded latency and/or bounded jitter. Most of URLLC applications requires upper bound, e.g. smart grid tele-protection, while a small number of use cases require bounded jitter.

Different requirement leads to different TSN solutions, see more analysis in section 6.5.

<< Contributors' Note: Consider providing detailed latency evaluation method and compare multiple TSN techniques in section.6.5. >>

7. Quality of Service provision

7.1 Causes of packet loss

The most common causes of packet loss can be classified, for our purposes, as:

a) Congestion: Fluctuations in traffic load can result in dropping a packet due to lack of buffer space in some network node.
b) Equipment failure: The loss of a wire or a node in the network causes some number of packets to be lost until either the failure can be corrected, or following packets can be re-routed.
c) Electromagnetic Interference: Electromagnetic events can cause some number of packets to be lost, or received with checksum errors that cause them to be discarded.
d) Random: Random thermal or quantum mechanical events in physical interfaces or buffer memory can cause the loss of a packet, typically due to a checksum error.

Of these, congestion generally causes the most packet loss for best-effort traffic. The TSN techniques described in section 7.3 can completely eliminate congestion loss for Streams. Frame Replication and Elimination for Reliability (FRER) can provide a significant reduction in packet loss from the other three causes.

7.2 Causes of excessive latency

The most common causes of end-to-end delivery times exceeding the applications' requirements can be classified, for our purposes, as:

a) Physical causes: The number of hops along the path of a Stream through the network, the length of the links, the minimum forwarding time for each hop, and the speed of each link determine the minimum possible end-to-end delivery time for packets of a Stream. For most physical link technologies, this part of the latency is practically constant once the path of the Stream is determined.
b) Interference causes: Packets can arrive simultaneously in a network node from different input ports, but they have to be output serially, if relayed to the same output port. Therefore, multiple time-critical Streams can interfere with each other; some packets have to be buffered until they can be transmitted. This interference can cause gaps or bursts of packets relative to the average rate of a Stream, which can cause further interference downstream.

Latency due to physical causes can be improved by altering the physical topology or link characteristics, or by altering the path taken by the Stream through the network, and is outside the scope of this standard. Excessive latency or latency variation caused by interference can be successfully addressed by TSN techniques, so that congestion losses can be avoided. Section 7.3 will elaborate on each of the available TSN techniques.

<< Contributors' Note: Should we have a section somewhere talking about altering the route of one or more Streams, either to achieve a latency goal, or to open up resources for another Stream? >>

7.3 Providing the services

There are six basic data plane techniques for providing the services described in Clause 7:

a) **FRER:** Frame Replication and Elimination for Reliability. Data is serial-numbered, replicated, sent along multiple disjoint paths through the network, and recombined into a single stream without duplicates (7.3.1).

b) **Strict Priority:** Critical data is given the highest priority in the network, perhaps even higher than network control protocol traffic (e.g. network topology control) (7.3.2).
c) **CBS:** Credit-based shaper, defined as Enhancements for Time-Sensitive Streams in 8.6.8.2 of IEEE Std 802.1Q-2018, shapes the transmission of a traffic class (7.3.3).
d) **ATS:** Per-Stream shaping and policing, using IEEE Std 802.1Qcr Asynchronous Traffic Shaping (ATS) and/or IETF IntServ ([B3], [B4], [B5], [B6], [B7]), can apply a state machine to shape individual Streams at each hop (7.3.4).
e) **Scheduled Traffic:** Timed transmission windows control traffic classes' (queues') transmission opportunities, using 8.6.8.4 in IEEE Std 802.1Q-2018 (7.3.5).
f) **CQF:** Time-synchronized transmission windows, using Annex T in IEEE Std 802.1Q-2018 Cyclic Queuing and Forwarding, transmit bunches of packets in lock-step through the network (7.3.7).

<< Contributors' Note: Paternoster and multi-layer CQF would also be useful. Shall we include them? >>

There are other potential data plane techniques that are **not** addressed by the present standard:

g) Frame preemption: The transmission of a frame is interrupted in order to transmit a frame requiring lower latency, then transmission of the original frame is resumed from the point of interruption. Frame preemption is defined in 6.7.2 in IEEE Std 802.1Q-2018 and clause 99 in IEEE Std 802.3-2018. Preemption is not considered in this standard, because it is useful primarily on low-speed links.
h) Cut-through forwarding: The transmission of a frame is initiated before the last bit of the frame has been received and the checksum examined for errors. This document assumes store-and-forward operation. Cut-through forwarding is most useful on low-speed links.
i) Dedicated devices or resources: A network is constructed for the exclusive use of one or a small number of critical applications, and this network is physically isolated from other networks. This is a technique is often used for critical traffic, today; avoiding such duplication is a goal of this standard.
j) Congestion detection: Packets belonging to a flow that is experiencing congestion are marked, and messages eventually sent towards the sender, to cause the flow to slow down. Such flows are certainly of interest to service providers, but are not addressed in this standard. We assume, here, that the applications generating Streams cannot slow down the real-time physical world to accommodate the network's current load.
k) Congestion avoidance: New Streams are routed over less-congested network paths, or existing Streams are re-routed. Congestion avoidance can be useful in a TSN service provider network; the technique is used, today. But congestion avoidance is not specifically a TSN technique, and so is not addressed in this standard.

7.3.1 Frame Replication and Elimination for Reliability

Frame Replication and Elimination for Reliability (FRER) is described in IEEE Std 802.1CB. See clause 7 in IEEE Std 802.1CB-2017 for an overview of the technique. The packet replication and packet elimination functions described in 3.2.2.2 in IETF RFC 8655 [B14] are a generalization of IEEE Std 802.1CB FRER.

FRER is aimed at packet loss due to equipment failure (point b) in 7.1) or electromagnetic interference (point c) in 7.1).[2] Packets in a Stream are sequence numbered, then replicated, and flow along two (or more) paths to a point nearer the destination. At that point, the paths are combined into a single Stream, again. A multicast Stream can be combined a multiple points, for multiple receivers. The combining can be done in a network node, or in the receiving end station. When combined, packet sequence numbers are compared continuously, and the duplicates eliminated. For FRER to be useful, the cost of losing packets while the network recovers from a failure must be greater than the cost of doubling the bandwidth used by the critical, replicated, Streams.

[2] In rare environments, where congestion loss is eliminated (by other methods), and equipment failure and interference are very rare, FRER can improve packet loss due to random events (point d) in 7.1).

7.3.2 Strict priority

Strict priority is described in section 8.6.8.1 and Annex L of IEEE Std 802.1Q-2018. The strict priority techniques orders two or more queues on an output port in a monotonic sequence. When the link is ready to transmit a frame, the highest-priority queue that is not empty is chosen, and one frame is transmitted from that queue. If there is enough traffic in the highest-priority queue or queues to fill the output link's bandwidth, then frames in the lower-priority queues are never transmitted. Every bridge conformant to IEEE Std 802.1Q, that has more than one queue per output port, is required to implement strict priority.

If the total bandwidth of the Streams is a small fraction of the bandwidth of any link over which they pass, then Streams can be assigned to the highest priority queue that is used for data (see below for control implications). More than one level of Streams can be identified and assigned different priority levels. This is how IEEE Std 802.1CM, Time-Sensitive Networking for Fronthaul [B1], provides bounded low latency and zero congestion loss for Common Public Radio Interface (CPRI) fronthaul networks.

In order to determine the worst-case latency delivered by strict priority, as well as the amount of buffer space required to ensure zero congestion loss, the characteristics of the Streams must be known in detail. Obviously, latency and loss cannot be controlled if the volume of highest-priority traffic traversing a link exceeds the bandwidth of that link in the long term. IEEE Std 802.1CM provides an analysis of latency and buffer requirements for Streams whose transmissions are coordinated using synchronized clocks. If transmissions are not coordinated, then relatively low bandwidth utilization rates (5% or 10%) can result in significant worst-case delays and buffer requirements due to the coincidence of random events, and the difficulty of computing the worst case becomes more difficult than for coordinated transmissions.

In best-effort networks, the industry best practice has been to reserve the highest-priority one or two queues for network control traffic, e.g. network error detection, recovery, network management, and topology control. This way, an uncontrolled source of high-priority data cannot starve the control traffic, and thus prevent the network from maintaining its basic ability to move data over non-congested links. However, network control traffic is not always easy to characterize, and can, on occasion, be excessive. Therefore, in order to obtain the necessary quality of service for time-critical Streams, it may be necessary to give some Streams the very highest priority, and to install metering and policing functions to ensure that malfunctioning high-priority senders cannot totally block control traffic.

Strict priority is always in use on any port with more than one class of service queue, no matter what other techniques are employed. Each of the other techniques, when applied to a queue, selects whether that queue is or is not eligible to compete for a chance to transmit a frame. If more than one queue is enabled and has a frame to transmit, then the highest-priority queue wins, and a frame from that queue is selected for transmission. Thus, any combination of the techniques in this section 7.3 can be used on the same port.

With Strict priority schedulers, theoretically high priority traffic suffers no interference from low priority data. Only same class traffic multiplexing is considered, service curve can be modeled as:

$$\beta\ (t) = (C - \sum \text{FlowRate}_{\text{samePri}})^{*}\ (t - (\sum \text{burst}_{\text{samePri}} + \text{MaxPacketLength}_{\text{lowPri}})/(C - \sum \text{FlowRate}_{\text{samePri}}));$$

Latency from this scheduler is T + burst/R, extended as follows,

$$\text{Strict priority scheduler delay} = (\sum \text{burst}_{\text{samePri}} + \text{MaxPacketLength}_{\text{lowPri}} + \text{burst})/\ (C - \sum \text{FlowRate}_{\text{samePri}});$$

Low latency bound is achievable when the burst size of high priority traffic is constrained. When the number of flows or the burst size of high priority traffics rises, the latency bound deteriorates quickly.

7.3.3 Credit-Based Shaper

Section 8.6.8.2 and Annex L of IEEE 802.1Q-2018 describe Forwarding and Queuing for Time-Sensitive Streams (FQTSS). FQTSS can apply a credit-based shaper (CBS) function to any of up to seven of the eight classes of service (queues) supported by an IEEE Std 802.1Q bridge output port. CBS throttles the transmission of a class of Streams to the sum of the bandwidth of those Streams, at every hop along the path through the network. This throttling can prevent momentary bursts of critical data, caused by recent interference upstream, from combining to overflow the buffer capacity of a downstream node, and thus cause congestion loss (point a) in 7.1).

As each Stream is added or deleted from the network, the CBS shapers' bandwidth parameters along the path of the Stream have to be adjusted. If none of the queues along the path carry any other Streams, then the calculation of the worst-case latency for a new Stream and the buffer requirements to avoid congestion loss are relatively straightforward. In many applications, however, the limited number of shapers available on an output port precludes dedicating a shaper to a single Streams; multiple Streams have to share a class of service, and hence a shaper. In this case, all Streams that share a queue with the new or deleted Stream are affected, and their latency and buffer requirements recalculated. These changes, in turn, can trigger further recalculations. Alternatively, the calculation can be carried to a point that it is clear that either all Streams' requirements can be met, or they cannot.

$Delay_a = L_0/C + L_a/idleslope_a$;

$Delay_b = (L_0 + L_a)/C + (L_0 * Idleslope_a/(C - Idleslopea))$;

L_a, L_0 are maximum packet length for SR Class A and best effort class defined in FQTSS specification. Also according to FQTSS specification, there is no constrained delay traffic considered.

<< Contributors' Note: This will be further expanded. >>

IEEE Std 802.1Q provides for eight traffic classes. Assuming that a given network port has more than that number of Streams passing through a port, then it is highly likely that any given class supports more than one Stream. Interference between Streams sharing a traffic class can, in some cases, cause congestion loss. CBS has been found, in practice, to provide better packet loss characteristics than other common methods, e.g. Weighted Fair Queuing, but latency calculations become very difficult as the number of shared traffic classes increases. It is, however, relatively cheap to implement, because it has one shaper per traffic class, instead of one per Stream.

7.3.4 Asynchronous Traffic Shaping

Asynchronous Traffic Shaping (ATS) is described in clauses 8.6.11 and 49 of IEEE Std 802.1Qcr-2020[3]. With ATS, each Stream can be assigned its own state machine for regulating the flow of that Stream through each network node. When ATS is implemented, a network management system can compute the worst-case latency for any packet belonging to a Stream, and can compute the amount of per-Stream and/or per-class buffering needed at every hop along the path in order to assure that packet loss due to congestion (point a) in 7.1) is mathematically impossible. Because ATS can be configured to give two Streams sharing the same queue two different shapers, much of the interference between Streams that is commonly present in the FQTSS/CBS (7.3.3) is avoided, and the calculations can be performed more easily. The worst-case latency and buffer requirements for any given Stream are also better than with FQTSS/CBS, for the same reason. This is the result of, and comes at the cost of, the additional shapers required in each network node.

[3] IEEE Std 802.1Qcr will be incorporated into the next edition of IEEE Std 802.1Q following the 2019 edition of that standard.

7.3.5 Scheduled Traffic

Enhancements for Scheduled Traffic are described in clauses 8.6.3 and 37 of IEEE Std 802.1Q-2018. This feature attaches a timed binary gate to each of the (up to) eight class of service queues on an output port. A rotating schedule is established by management action. This schedule repeats at a fixed, rational (a ratio of two integers) number of nanoseconds. During each cycle, a fixed order of events are executed at fixed integer numbers of nanoseconds from the start of the cycle. Each event sets each of the eight queues' gates either to contend or not to contend for transmission opportunities on the port. Such scheduling can be used, for example, to give a queue serving a set of critical Streams exclusive access to the port for regular, brief periods.

7.3.6 Per-Stream filtering and policing (PSFP)

The allocation of resources to a Stream or a class of Streams is predicated on the assumption that the transmitter of a Stream will not exceed the contract made with the network at the time the resources are allocated. This means, at the least, that the contracted frame size and number of frames per measurement interval will not be exceeded. If a Stream exceeds its contract to a significant degree, then at the very least, its latency and congestion loss requirements will not be met, at at worst, it will interfere with other Streams, and cause their requirements to also be violated.

Per-Stream Filtering and Policing (PSFP) is described in clause 8.6.5.2.1 of IEEE Std 802.1Q-2018. This feature can be configured to monitor a specific Stream or class of Streams, and mark or discard frames in excess of the contract. This can isolate the consequences of misbehaving devices.

In a service provider environment, Streams may originate from different organizations, with each of which the service provider has contractual obligations for network performance. Typically, a provider applies PSFP at least to the points where Streams enter the provider's network. Whether PSFP is applied within a network depends on an assessment of the chances of internal equipment failure or misconfiguration, versus the cost of implementing and configuring PSFP internally, versus the chances of mishandling of Streams due to mistakes caused by failure or misconfiguration of the additional PSFP, itself.

7.3.7 Cyclic Queuing and Forwarding

Cyclic Queuing and Forwarding (CQF) uses a set of features that are included in IEEE Std 802.1Q for a number of techniques, including FRER (7.3.1) and ATS (7.3.4), in a combination that is described in Annex T of IEEE Std 802.1Q-2018. When employing CQF according to Annex T, a pair of queues on each port are devoted to the use of CQF. Both are assigned to a pair of classes of service or IEEE Std 802.1Q priority value. Using the scheduled traffic feature (7.3.5), the buffers are enabled for output alternately, on a regular cycle time *T*. The timed input gate feature of PSFP (7.3.6) is used at each input port to assign frames belonging to CQF Streams to alternate class of service values, and thus to alternate output queues. All of the input gates and all of the output gates used for CQF in the network switch at the same moment, within the accuracy of their synchronized clocks. At each port, at any given moment, one queue is filling and the other is enabled for transmitting. In this manner, data flows through the network hop by hop, taking two cycle times *T* per hop to reach its destination.

The disadvantages of CQF, compared to other TSN techniques, are that clock synchronization is required, and that it is not always possible to pick a suitable value of *T*, if the network is required to accommodate a wide range of Stream characteristics or link speeds. On the other hand, the calculation of worst-case latency is trivial, Streams do not affect the latency or buffer requirements of other Streams when added or deleted, and no per-Stream resources such as shapers need be implemented, allocated, or configured, making the addition and deletion of Streams easy and fast.

<< Contributors' Note: Paternoster and multi-layer CQF could both be of interest to service providers. Should anything be said? Do we need new projects? >>

8. Profiles

8.1 Introduction

Currently three main types of requirements are considered in TSN service provider networks, namely bounded latency, bounded jitter and high reliability. The following sections will discuss how to use TSN techniques in carrier networks to satisfy these requirements.

IEEE 802.1 TSN standard group provides multiple queueing and forwarding methods as listed in section 7.3 to achieve bounded latency, which can be categorized into two groups. One is working with synchronized clocking, also called as time-triggered methods, and the other group does not rely on global clocking and works in an event-triggered way.

<< Contributors' Note: Time-triggered group TSN techniques include …, while event triggered group include ... >>

Event-triggered methods such as strict priority, FQTSS and asynchronous traffic shaping, usually get lower average latency in light loaded networks, while its latency bound increase non-linearly when bandwidth utilization is high. This kind of approach is most suitable for VoIP similar communications, with small bandwidth consumption and strict latency requirement.

Time-triggered methods like Scheduled traffic and Cyclic Queueing and Forwarding have larger average latency since it holds packets until dedicate time gate is open on each hop, even no packets are transmitting on output port. On the other hand, jitter control performance is better with time-triggered methods.

With either time-triggered or event-triggered methods, resource reservation protocol or controlling is necessary along the traffic path to ensure bandwidth or buffering is available.

<< Contributors' Note: maybe refer to DetNet control plane for more info. >>

<< Contributors' Note: Three application examples likely to be extended in details,
- Industrial internet needs bounded latency and high reliability
- Smart grid need bounded low latency and high reliability
- Network slicing needs >>

The bridges of a service provider TSN network shall meet the bridge requirements (5.3) and each link is a full duplex point-to-point link. TSN service provider bridged network is designed, configured and operated to address the criteria specified in profiles (8.2, 8.3, 8.4).

A TSN service provider bridged network is designed, configured, and operated such that time critical data traffic does not exceed the required bandwidth during normal operation. Specifically, the data rate of each link of the TSN service provider bridged network is big enough to forward the desired time critical data traffic within required latency limitation. For example, if a bridge port aggregates multiple time critical data flows, its transmission rate is greater than the sum of bandwidth required by the received time critical data traffic under corresponding latency constraints.

Note: delay guarantee bandwidth is the bandwidth allocated on a port or a scheduler to meet latency requirement, which is different with average data rate or peak data rate of user traffic. A low rate application with tight latency requirement needs more bandwidth than average bandwidth consumption. For example, to transmit a 1500B data burst in 10us needs 1.2G bps bandwidth, although this data burst happens every 1ms with an average rate of 12M bps.

8.2 Latency bound guarantee profile

Since event trigger methods are more robust and scalable in large scale network, two options in event trigger approach are recommended to achieve bounded latency requirement described in Clause 6.5. Strict priority shall be supported in all bridges in carrier network with latency constraints, and weighted round robin is optional.

<< Contributors' Note: not sure WRR(weighted round robin) is in scope for IEEE TSN standard. >>

8.2.1 Shaping for time critical traffics

Service rate and burst size greatly affect latency performance, user traffic specification is determined before data packet transmission. An optional traffic shaping function is recommended on ingress provider edge node to ensure user traffic is in line with expectations.

<< Contributors' Note: in service provider networks, end users usually do not know exact traffic specifications defined in 802.1Q-2018 TSEPC, parameters like average rate, peak rate, and max burst size are measured and user traffics could be re-shaped and enforced into traffic specification agreement. >>

8.2.2 Profile of time critical traffics

Streams for time critical application are served with high priority while other best effort traffics transmit in lower priority queues.

8.2.3 Meeting latency target

<< Contributors' Note: Give examples on SP and WRR, with Pros and Cons on average delay and worst case delay. >>

8.3 Jitter guarantee profile

Isolation is a widely used but not clearly defined term. From user experience point of view, only delay, jitter and packet loss ratio is observable. Harder isolation usually means less interference from other user traffics, thus results in smaller delay variation. Jitter guarantee profile is to provide bounded jitter over service provider networks.

<< Contributors' Note: Two options are considered in this profile, one is TAS/CQF ideas with proper time window planning; the other is put playout buffer on egress provider edge node. A high voltage power protection application example is helpful for understanding. >>

8.4 Reliability Guarantee profile

A typical way to provide high reliability, as described as low packet ratio, is discussed in IEEE Std 802.1CB-2017.

<< Contributors' Note: Check related discussion in DetNet. Here may only needs to remind needs for independent data paths for duplicated frames. An industrial internet example will be helpful in understanding. >>

9. Interface with DetNet

9.1 Introduction

The Deterministic Networking Working Group (DetNet, https://datatracker.ietf.org/wg/detnet/documents/) of the Internet Engineering Task Force (IETF) has worked closely with the IEEE 802.1 Time-Sensitive Networking Task Group, by means of common participation by individuals, to generate documents that provide very similar services, for IETF routers and label switches, that TSN provides for bridged LANs. Participants have endeavored to make the TSN and DetNet documents consistent and compatible. To date, DetNet has published a number of RFCs. The one most relevant to the present IEEE standard include:

a) RFC 8557, Deterministic Networking Problem Statement;
a) RFC 8578, Deterministic Networking Use Cases;
b) RFC 8655, Deterministic Networking Architecture;
c) RFC xx01, (draft-ietf-detnet-data-plane-framework) DetNet Data Plane Framework;
d) RFC xx02, (draft-ietf-detnet-flow-information-model) DetNet Flow Information Model;
e) RFC xx03, (draft-ietf-detnet-ip) DetNet Data Plane: IP;
f) RFC xx04, (draft-ietf-detnet-ip-over-mpls) DetNet Data Plane: IP over MPLS;
g) RFC xx05, (draft-ietf-detnet-ip-over-tsn) DetNet Data Plane: IP over IEEE 802.1 Time Sensitive Networking (TSN);
h) RFC xx06, (draft-ietf-detnet-mpls) DetNet Data Plane: MPLS;
i) RFC xx07, (draft-ietf-detnet-mpls-over-tsn) DetNet Data Plane: MPLS over IEEE 802.1 Time Sensitive Networking (TSN);
j) RFC xx08, (draft-ietf-detnet-mpls-over-udp-ip) DetNet Data Plane: MPLS over UDP/IP;
k) RFC xx09, (draft-ietf-detnet-tsn-vpn-over-mpls) DetNet Data Plane: IEEE 802.1 Time Sensitive Networking over MPLS;
l) RFC xx10, (draft-ietf-detnet-yang) Deterministic Networking (DetNet) Configuration YANG Model;

<< Contributors' Note: We expect the referenced IETF drafts, above, to achieve RFC status by the time the present draft standard is published. We expect only RFCs to be referenced in the published IEEE standard. >>

Section 10 of RFC 8578, use cases, gives the particular example of applying DetNet to provide network slicing capability for a 5G bearer network. (See RFC 8578 for the definitions of these terms.)

<< Contributors' Note: original ideas in this section is to consider how Layer 2 reservation protocol interwork with Layer 3 reservation protocols. Probably will delete if it is not clear to users >>

9.2 Data plane

If a network compliant to the present standard is intended to transport DetNet traffic, or if traffic in a compliant network is to be transported over an IP or MPLS network, then it shall conform to the relevant IETF standards, including RFC xx05, RFC xx07, and/or RFC xx09.

9.3 Control plane

<< Contributors' Note: At this writing, the IETF DetNet Working Group has not made sufficient progress on the control plane (e.g. resource reservation and fixed path establishment) for the present draft to make normative references. In the opinion of the author, the issue is more narrowing down choices than designing new protocols. It is possible that there is a need to augment

some IETF protocol(s) to support the Paternoster algorithm, but that algorithm has not been standardized in either IEEE or IETF, yet. There is also the possibility of implementing RAP in a router or label switch. This has not been sufficiently explored to determine whether it is a viable idea or not. As a consequence, this section will likely point the reader to the DetNet Working Group for further information. >>

Annex A

(informative)

PCS Proforma—TSN for Service Provider Networks Profiles

.

Annex B

(informative)

Bibliography

Bibliographical references are resources that provide additional or helpful material but do not need to be understood or used to implement this standard. Reference to these resources is made for informational use only.

[B1] IEEE Std 802.1CM-2018, "IEEE Standard for Local and metropolitan area networks—Time-Sensitive Networking for Fronthaul".

[B2] 3GPP TS 23.501, "System Architecture for the 5G System" V16.3.0.

[B3] IETF RFC 1633, "Integrated Services in the Internet Architecture: an Overview".

[B4] IETF RFC 2211, "Specification of the Controlled-Load Network Element Service".

[B5] IETF RFC 2212, "Specification of Guaranteed Quality of Service".

[B6] IETF RFC 2215, "General Characterization Parameters for Integrated Service Network Elements".

[B7] IETF RFC 2205, "Resource ReSerVation Protocol (RSVP) ".

[B8] IETF RFC 2474, Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <https://www.rfc-editor.org/info/rfc2474>.

[B9] IETF RFC 2475, Black, D., Wang, Z., Carlson, M., Weiss, W., Davies, E., and Blake, S., "An Architecture for Differentiated Services", RFC2475, DOI 10.17487/RFC2475, December 1998, <https://rfc-editor.org/rfc/rfc2475.txt>.

[B10] IETF RFC 4594, "Configuration Guidelines for DiffServ Service Classes".

[B11] IETF RFC 8100, "Diffserv-Interconnection Classes and Practice".

[B12] IETF RFC 8578, "Deterministic Networking Use Cases".

[B13] IETF RFC 8557, "Deterministic Networking Problem Statement".

[B14] IETF RFC 8655, "Deterministic Networking Architecture".

[B15] IETF RFC xx01, (draft-ietf-detnet-data-plane-framework) "DetNet Data Plane Framework".

[B16] IETF RFC xx02, (draft-ietf-detnet-flow-information-model) "DetNet Flow Information Model".

[B17] IETF RFC xx03, (draft-ietf-detnet-ip) "DetNet Data Plane: IP".

[B18] IETF RFC xx04, (draft-ietf-detnet-ip-over-mpls) "DetNet Data Plane: IP over MPLS".

[B19] IETF RFC xx06, (draft-ietf-detnet-mpls) "DetNet Data Plane: MPLS".

[B20] IETF RFC xx08, (draft-ietf-detnet-mpls-over-udp-ip) "DetNet Data Plane: MPLS over UDP/IP".

[B21] IETF RFC xx10, (draft-ietf-detnet-yang) "Deterministic Networking (DetNet) Configuration YANG Model".

<< Contributors' Note: We expect the referenced IETF drafts, above, to achieve RFC status by the time the present draft standard is published. We expect only RFCs to be referenced in the published IEEE standard. >>

[B22] "Network calculus: a theory of deterministic queuing systems for the internet", 2001, <https://arxiv.org/abs/1804.10608/>.

[B23] “Improved Credit Bounds for the Credit-Based Shaper in Time-Sensitive Networking” , 2019, E Mohammadpour, et al.

[B24] “communication networking, An analytical approach”, 2004, Anurag Kumar et al.

[B25] “Timing Analysis of AVB Traffic in TSN Networks Using Network Calculus", 2018 Luxi Zhao et al.

Annex C

(informative)

A concept for network calculus

<< Contributors' Note: Basis of Network Calculus will be introduced briefly. Also considering re-visit latency evaluation for existing TSN techniques, like CBS, TAS, etc. Probably leads to maintenance for 802.1Q-2018 with update on latency analysis >>

Latency analysis based on Network Calculus (Informative)

<< Contributors' Note: This clause may set an example on how to use profiles defined in this standard to setup a network to satisfy a certain use cases, such as smart grid or Cloud VR applications. >>

<< Contributors' Note: briefly introduce Network Calculus methodology with examples. Illustrate how to use Network calculus to analyze delay on single node and cascaded networks.>>

Network calculus theory emerged during 1990s as a latency evaluation theory for quality of service analysis of packet switching networks, it is originally focus on performance analysis for IntServ model over IP network. Data arrivals at a networked system are modelled by upper envelope functions. Minimum service guarantees that are provided by systems, such as a router, a scheduler, or a link, are characterized by service curves. Based on these concepts, network calculus offers convolution forms that enable worst case performance bounds evaluation including backlog and delay. Any number of bridged system in series can be transformed into a single equivalent system by convolution operation and obtain end-to-end performance.

C.1 Arrival curves

Streams can be described by arrival functions F(t) that are given as the cumulated number of bits seen in an interval[0,t]. Arrival curves are defined to give an upper bound on the arrival functions, where

$\alpha(t2-t1) = F(t2) - F(t1)$;

Token bucket based arrival curve is usually featured like in equation, $\alpha(t) = b + rt$, where b is burst size, r is data rate;

<< Contributors' Note: diagram of token bucket arrival curves will be helpful in this section. >>

C.2 Service curves

The service offered by the scheduler on an output port can be characterized by a minimum service curve, denoted by $\beta(t)$. A common service curve is described as rate-latency equation that includes a rate R and a wait time T, $\beta(t) = \max(0,R*(t-T)^{)}$.

Service curves for legacy Qos methods such as Priority Queuing (PQ) , Generalized Processor Sharing (GPS) and Weighted Fair Queuing (WFQ) are studied and proposed in multiple academia papers [B22]; TSN schedulers service curves are also under discussion and proposed in [B23] [B24] [B25]; In addition, aggregated scheduling networks resources shall be provisioned on an aggregate basis.

A.3 Upper bound of queueing delay

Queueing delay bound can be easily computed by comparing arrival curve and service curve in a queuing system, as the following equation shows,

Delay bound = T + b/R; where T and R are service curve parameters, and b is from arrival curve.

Detailed queueing delay for specific schedulers are provided in Section 6.5.

<< Contributors' Note: Consider a separate section to talk about aggregating mode. >>

Annex Z

(informative)

Committee Issues