

Workgroup: DetNet Working Group

Internet-Draft:

draft-ietf-detnet-dataplane-taxonomy-00

Published: 14 May 2024

Intended Status: Informational

Expires: 15 November 2024

Authors: J. Joung                      X. Geng      S. Peng  
         Sangmyung University      Huawei      ZTE Corporation  
         T. Eckert  
         Futurewei Technologies

## **Dataplane Enhancement Taxonomy**

### **Abstract**

This draft is to facilitate the understanding of the data plane enhancement solutions, which are suggested currently or can be suggested in the future, for deterministic networking. This draft provides criteria for classifying data plane solutions. Examples of each category are listed, along with reasons where necessary. Strengths and limitations of the categories are described. Suitability of the solutions for various services of deterministic networking are also briefly mentioned.

### **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 15 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](#)
- [2. Terminology](#)
  - [2.1. Terms Used in This Document](#)
  - [2.2. Abbreviations](#)
- [3. Conventions Used in This Document](#)
- [4. Taxonomy with Performance](#)
  - [4.1. Per Hop Dominant Factor for Latency Bound](#)
- [5. Taxonomy with Functional Characteristics](#)
  - [5.1. Periodicity](#)
  - [5.2. Network Synchronization](#)
  - [5.3. Traffic Granularity](#)
  - [5.4. Work Conserving](#)
  - [5.5. Target Transmission Time](#)
  - [5.6. Service Order](#)
- [6. IANA Considerations](#)
- [7. Security Considerations](#)
- [8. Acknowledgements](#)
- [9. Contributor](#)
- [10. References](#)
  - [10.1. Normative References](#)
  - [10.2. Informative References](#)
- [Authors' Addresses](#)

## 1. Introduction

This draft is to facilitate the understanding of the data plane enhancement solutions, which are suggested currently or can be suggested in the future, for deterministic networking.

An enhancement solution can be a combination of multiple data plane functional entities, such as regulators, queues, and schedulers. A solution can also include functional entities across network nodes, e.g. traffic enforcement or regulation functions at the edge. A regulator, or equivalently a shaper, is defined as a functional entity that makes the arrival process of a flow conform to a predefined process. A packet scheduler, or simply a scheduler, is a functional entity that determines when a packet is transmitted.

We use the term taxonomy as a synonym to the criteria for classifying the solutions accordingly. A category is a subset of solutions classified into a single group with a taxonomy. This draft provides

several taxonomies and the criteria for classifying data plane solutions. These taxonomies are orthogonal to each other.

Examples of the categories are listed, along with reasons where necessary. Strengths and limitations of the categories are described.

Suitability of the solutions for various services of deterministic networking are also briefly mentioned. The services can be classified according to the flow characteristics and the performance requirements. For example, Requirements for Reliable Wireless Industrial Services [[I-D.ietf-detnet-raw-industrial-req](#)] characterizes the services by the latency bound, the burst size, the burst transmission period, the number of nodes, etc. This document adopts this characterization rule, and classifies the services into one of tight/loose latency, large/small burst, periodic/non-periodic, and large/small scale services. For example, the display information service defined in Section 4.4. of [[I-D.ietf-detnet-raw-industrial-req](#)] is a loose latency, large burst, non-periodic, and small scale service.

The taxonomies described in this draft can be applied for the solutions of other standardization bodies, such as IEEE 802.1 TSN TG.

In this draft, the candidate solutions currently being proposed in DetNet WG are simply listed without any descriptions. The details of the solutions are intentionally omitted. Interested readers may refer to the corresponding drafts. When necessary, the solutions from IEEE TSN TG or existing popular ones are used as examples to better understand the taxonomy and the derived categories.

The mechanisms raised in the DetNet WG are not entirely new concepts but rather variations of existing mechanisms. These deliberate approaches aim to address the scalability requirements defined in [[I-D.ietf-detnet-scaling-requirements](#)] while ensuring a degree of continuity and compatibility with the current practices. The taxonomy in this draft reflects how new mechanisms extend existing ones to address scalability challenges.

For instance, Cycle Specified Queuing and Forwarding (CSQF) [[I-D.chen-detnet-sr-based-bounded-latency](#)], Tagged Cyclic Queuing and Forwarding (TCQF) [[I-D.eckert-detnet-tcqf](#)], IEEE 802.1Qdv Enhanced CQF (ECQF) are enhancements built upon the foundation of Cyclic Queuing and Forwarding (CQF). Similarly, Work Conserving Stateless Core Fair Queuing (C-SCORE) [[I-D.joung-detnet-stateless-fair-queuing](#)] is an extension of Fair Queuing (FQ). Timeslot Queuing and Forwarding (TQF) [[I-D.peng-detnet-packet-timeslot-mechanism](#)] is an extension of IEEE 802.1Qbv, also known as Time Aware Shaper (TAS). Earliest Deadline First (EDF) [[I-D.peng-detnet-deadline-based-forwarding](#)] proposed to DetNet WG is a variation of the well-known mechanism that

has the same name. Other well-known mechanisms that could provide bounded latency are also covered, for example Deficit Round Robin (DRR) and Asynchronous Traffic Shaping (ATS) [[IEEE 802.1Qcr](#)].

## **2. Terminology**

### **2.1. Terms Used in This Document**

### **2.2. Abbreviations**

## **3. 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.

## **4. Taxonomy with Performance**

Taxonomy based on the performance, such as E2E latency bounds and jitter bounds, is helpful to understand the solutions. The performance should be exhibited as a mathematical expression with the network and traffic parameters.

### **4.1. Per Hop Dominant Factor for Latency Bound**

One possible taxonomy would be based on the per hop dominant factor for the latency bound. The dominant factor is defined as the largest sum term in the expression, when the network and traffic conditions are the worst. The worst condition typically means high network utilization, large packet and burst sizes, and large number of hops. Any existing solution can be put into one of three categories.

Category 1 (Max Packet Length/Service Rate): FQ and its variations like C-SCORE fall into this category, where the latency bound is primarily influenced by the ratio of a flow's maximum packet size to its allocated service rate. This category emphasizes individual flow isolation. The consequence is that the variation of E2E latency bound for a flow is minimized with the other flows' join and leave. Therefore, this category performs well with dynamic flows. This category also fits well to services with large bursts, since the burst sizes of flows are not the dominant factor of the latency bound.

Category 2 (Sum of Max Packet Lengths/Capacity): Solutions like DRR belong here, where the dominant factor is the sum of maximum packet lengths of all DetNet flows over the total allocated bandwidth. This category typically has less implementation complexity than Category 1 but can impact individual flow isolation. The other flows' max packet

lengths affect the latency bound, which can be altered as flows join and leave.

Category 3 (Sum of Max Burst Sizes/Capacity): CQF, TAS, their variations (including CSQF, TCQF, ECQF, TQF), and EDF fall into this category. The key influence on latency here is the total burst sizes of all DetNet flows relative to the network capacity. This category prioritizes bounded latency guarantees but may require tighter burst control mechanisms. Once the burst is controlled, for example by an extremely strict regulation, into a packet length level, then this category may be indistinguishable with Category 2. This category fits well to the services for static flows with small bursts.

As an example, assuming the capacities and maximum packet lengths are identical in all the links along the path of a flow under observation, the E2E latency bound of the flow by FQ is given as the following [[STILIADIS-LRS](#)].

$$(B-L)/r + H(Lh/Rh + L/r), \quad (1)$$

where  $B$ ,  $L$ , and  $r$  are the maximum burst size of, the maximum packet length of, and the allocated service rate to the flow, respectively;  $H$  is the number of hops;  $Lh$  and  $Rh$  are the maximum packet length and the capacity of all the links.

In this example, the term  $(Lh/Rh + L/r)$  can be seen as the per hop latency, because the max burst size,  $B$ , appears only once. The service rate of a flow,  $r$ , is likely to be much less than the link capacity,  $Rh$ , while the maximum lengths  $L$  and  $Lh$  would not differ too much. Therefore, the dominant factor here is  $L/r$ .

The dominant factor determines the level of flow isolation, as well as the level of E2E latency bound value.

## 5. Taxonomy with Functional Characteristics

Taxonomy based on the functional characteristics is the key to understanding the solutions. The taxonomy listed in this section is orthogonal to each other, if not stated explicitly.

### 5.1. Periodicity

If a solution transmits packets in a periodic pattern, in which a packet is assigned to a time slot based on a predefined rule and a set of consecutive time slots repeated periodically, then the solution is periodic. Otherwise, the solution is non-periodic.

The set of consecutive time slots are called a period. Note that here we use the term period to avoid confusion with the term cycle used in CQF, which is equivalent to the time slot defined in this draft.

According to the above definition, IEEE 802.1Qbv TAS is a periodic solution. A finite Gate Control List (GCL) of TAS contains multiple gate control entries. Each entry represents a time slot with an assigned set of flows. A set of consecutive time slots forming a GCL is repeated periodically. Time slots can be overlapped with each other, as in ECQF.

TAS based solutions and CQF based solutions belong to periodic solutions, for example CSQF, TCQF, ECQF, TQF and so on.

Periodic solutions may fit well to periodic services, and vice versa.

## 5.2. Network Synchronization

According to whether network synchronization is required, a solution can be classified as either phase synchronous, frequency synchronous, or asynchronous.

Phase synchronous solutions require network nodes to be both phase and frequency synchronized. These solutions can be called strictly synchronous. TAS and CQF are in this category.

Frequency synchronous solutions require network nodes to be only frequency synchronized. Such nodes are often called syntonized. CQF variations and TAS variations are in this category, for example CSQF, TCQF, ECQF, TQF and so on.

Asynchronous solutions may also require loose phase and frequency synchronizations, for example ATS and EDF.

In non-synchronized networks, it has been shown that ignoring the timing inaccuracies can lead to network instability due to unbounded delay in per-flow or interleaved regulators [[THOMAS-Sync](#)]. However, the level of synchronization required is not high. The problem can be solved by adjusting the regulator parameters conservatively, even when loosely synchronized clocks are used. Thus, the solutions that require regulators such as ATS are categorized into asynchronous solutions.

The criteria to distinguish between synchronous and asynchronous solutions should be the level of required synchronization precision. One indicator suitable to such criteria would be the allowable Maximum Time Interval Error (MTIE). MTIE is usually calculated as the difference between the largest and smallest time differences in the ensemble of measurements. With this definition, a device that has an arbitrarily large and constant time difference with the standard reference has an MTIE value of 0, because MTIE is a measure of the evolution of the time difference, not the magnitude of the time difference itself. In this respect, the MTIE statistic is really a

measure of the frequency offset between the device under test and the standard reference.

Therefore, the allowable MTIE value can be applied equivalently, for the precision level evaluation, to both phase synchronous and frequency synchronous solutions.

In a distributed system, typical MTIE can be managed within nano second level. However, the exact value of the allowable MTIE as an indicator for synchronous solutions is for further study. It is expected to be within tens of nanoseconds.

Note that the taxonomy of network synchronization is closely related to the taxonomy of periodicity. However, these two can be used independently of each other.

### **5.3. Traffic Granularity**

This draft categorizes data plane solutions based on the granularity of their traffic control target, which refers to the size and specificity of the traffic entity they handle. Three granularity levels exist.

Flow level: Each packet is controlled based on its specific flow, which can be identified usually by the 5-tuple. Examples include FQ and its variations such as C-SCORE, which offer precise service differentiation but require potentially complex implementation.

Flow aggregate level: Flows are grouped by shared characteristics like traffic specification, service requirement, or routing path. This coarser level simplifies control but may offer less precise differentiation. Examples include interleaved regulators in ATS.

Class level: Flows are further grouped by similar service requirements, regardless of specific path or traffic details. This coarsest level simplifies control and accommodates traffic fluctuations but provides the least individual flow differentiation. Typically, time or time based information could be used for classification, such as in EDF, CQF and its variations.

For each level solution, packets within the same traffic entity receive the same treatment. For example, if a solution is flow aggregate level, then the packets within the same flow aggregate are treated identically, regardless of the flows they belong to.

There are cases in which a single solution consists of multiple functional entities that treat packets according to multiple traffic entities of different granularities. In such cases, it is defined that the functional entity with the coarsest granularity is dominant, thus the whole solution belongs to the coarsest granularity category.

For example, ATS consists of interleaved regulators (IRs) and a strict priority scheduler. An IR has a queue dedicated to a flow aggregate having the same class and the same input port. The regulation function itself is based on a flow. According to the definition above, IR is a flow aggregate level solution. On the other hand, the strict priority scheduler in ATS is class-based. Therefore, ATS as a whole is class level.

A finer granularity level solution has a benefit of a more accurate service differentiation among flows. Its limit is the larger implementation complexity. It fits to services with flows having various independent latency bound values.

Periodic solutions can further be categorized based on the traffic granularity. A time slot can be assigned per flow, per flow aggregate, or per class.

Note that TAS in 802.1Qbv is a scheduling mechanism defined in an output port with eight queues. The queues are controlled by GCL and its gate control entries. Each queue can serve a class. In an entry, queues can be either open or closed. Thus, TAS can be seen as a class level solution. However, in many cases TAS is understood as a scheduling mechanism, where the number of queues are not limited to 8. There could be a natural extension, such as TQF, which enables Qbv to allocate one queue to each flow or a flow aggregate.

Finer granularity periodic solutions have more strengths in jitter control. They also fit services with many periodic flows of independent period values.

#### **5.4. Work Conserving**

A work conserving solution never idle when there is a packet to send [[Fedorova](#)].

A non-work conserving solution can idle even if there is a packet to send in the queue.

A solution can be a combination of multiple data plane functional entities, and each functional entity has its own attribute of work conserving or non-work conserving. A solution is non-work conserving, as long as any of the functional entities included in the solution has the non-work conserving attribute.

FIFO, round robin schedulers, FQ and its variations like C-SCORE are examples of the work conserving solutions. TAS, CQF, ATS, and their variations are non-work conserving solutions, for example CSQF, TCQF, ECQF, TQF and so on. EDF can be operated either as work conserving or non-work conserving.



Work conserving solutions have strengths in terms of average delay. They usually show smaller observed maximum latencies than the theoretical latency bound expressions suggest. They also benefit from the statistical multiplexing gain without any wasted capacity, thus more room for best effort traffic.

Non-work conserving solutions have strengths to avoid burst accumulation and are also beneficial for jitter control. The burst size of a flow can be kept similar or the same with the initial burst size. Therefore, the buffer size necessary typically is less than those in work conserving solutions. This further makes the latency evaluation process simple.

### **5.5. Target Transmission Time**

Data plane solutions can be categorized as "on-time" or "in-time" based on how closely they adhere to predefined target transmission times for packets.

On-time solutions strive to transmit packets as close as possible to their target times without ever exceeding them. This ensures tight control over both latency and jitter, but it can sometimes lead to higher average latency.

In-time solutions allow more flexibility, transmitting packets without a specified target transmission time. FQ and its variations are in-time solutions.

ATS, which includes the interleaved regulator, is an in-time solution. A regulator determines an eligible time for a packet to be transmitted. Packets are always transmitted at or later than their eligible times. An eligible time is not a target transmission time. Note that ATS is a non-work conserving but in-time solution.

TAS, CQF, and their variations are on-time solutions. A time slot of TAS, within which a packet should be transmitted, can be seen as the target interval. EDF can be operated either as in-time or on-time.

The on-time/in-time taxonomy here is about the scheduling decision, which determines when a packet is transmitted. It is not about the consequence of the scheduling, whether the jitter bound is also guaranteed or not.

On-time solutions typically control the jitter as well as latency, but suffer from larger average latency. In-time solutions have limitations on controlling jitter. In-time solutions may have to handle the jitter with additional mechanisms.

## 5.6. Service Order

Data plane solutions prioritize packets from different flows using various decision rules, categorized as follows.

Rate-based: Packets are ordered based on the allocated service rate of their flows or flow aggregates. Examples include FQ and its variations like C-SCORE, and DRR.

Time-based: Packets are prioritized based on their allowed delay or deadline. Examples are CQF, TAS, their variations, and EDF.

Arrival-based: Packets are served in the order they arrive. FIFO is an example.

Priority-based: Packets are ordered based on assigned priorities.

A solution can determine the service order of the packets from different flows, based on a rule which considers the rate allocated to a flow or a flow aggregate, the delay a packet is allowed, the packet arrival time, or the packet priority. A rule may also be constructed with a combination of these characteristics. Note that the service order within a flow cannot be altered, thus is already decided. We focus only on the service order among packets from different flows.

According to its primary service order decision rule, a solution can be categorized into either rate-based, time-based, arrival-based, or priority-based. Any solution can also use the packet arrival time as a secondary decision rule.

Strict priority scheduler uses primarily the priority of a packet. It also uses the arrival times among packets of the same priority. In this case it is categorized as priority-based.

ATS has IRs and a strict priority scheduler. The service order among packets at an IR is arrival-based. The order among packets from different input ports are decided at the strict priority scheduler. Thus, ATS is priority-based.

Rate-based solutions have a simple admission condition check process that is dependent only on the service rates of flows. They benefit from the "pay burst only once" property, by which the maximum burst size of a flow contributes to the E2E latency bound only once, without being multiplied by the hop count. Rate-based solutions typically fit well to services with large burst and large scale services, without a need for overprovisioning, or additional burst control mechanisms.

Time-based solutions have strengths in precise delay control for packets or flows. The services with tight latency, small burst, and small scale services may fit this category.

Priority-based and arrival-based solutions benefit from the implementation simplicity. The latency and jitter differentiation among flows can be coarse, however. The services with loose latency, small burst, and non-periodic services may fit this category.

## 6. IANA Considerations

There might be matters that require IANA considerations associated with metadata. If necessary, relevant text will be added in a later version.

## 7. Security Considerations

This section will be described later.

## 8. Acknowledgements

## 9. Contributor

## 10. References

### 10.1. Normative References

[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>>.

[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>>.

### 10.2. Informative References

[Fedorova] Fedorova, A., Seltzer, M., and M.D. Smith, "A non-work-conserving operating system scheduler for SMT processors", In Proceedings of the Workshop on the Interaction between Operating Systems and Computer Architecture, vol. 33, p. 10-17, June 2006.

[I-D.chen-detnet-sr-based-bounded-latency] Chen, M., Geng, X., Li, Z., Joung, J., and J. Ryoo, "Segment Routing (SR) Based Bounded Latency", Work in Progress, Internet-Draft, draft-chen-detnet-sr-based-bounded-latency-03, 7 July 2023, <<https://datatracker.ietf.org/doc/html/draft-chen-detnet-sr-based-bounded-latency-03>>.

**[I-D.eckert-detnet-tcqf]**

Eckert, T. T., Li, Y., Bryant, S., Malis, A. G., Ryoo, J., Liu, P., Li, G., Ren, S., and F. Yang, "Deterministic Networking (DetNet) Data Plane - Tagged Cyclic Queuing and Forwarding (TCQF) for bounded latency with low jitter in large scale DetNets", Work in Progress, Internet-Draft, draft-eckert-detnet-tcqf-05, 5 January 2024, <<https://datatracker.ietf.org/doc/html/draft-eckert-detnet-tcqf-05>>.

**[I-D.ietf-detnet-raw-industrial-req]** Sofia, R. C., Mendes, P., Bernardos, C. J., and E. Schooler, "Requirements for Reliable Wireless Industrial Services", Work in Progress, Internet-Draft, draft-ietf-detnet-raw-industrial-req-00, 19 January 2024, <<https://datatracker.ietf.org/doc/html/draft-ietf-detnet-raw-industrial-req-00>>.

**[I-D.ietf-detnet-scaling-requirements]**

Liu, P., Li, Y., Eckert, T. T., Xiong, Q., Ryoo, J., zhushiyin, and X. Geng, "Requirements for Scaling Deterministic Networks", Work in Progress, Internet-Draft, draft-ietf-detnet-scaling-requirements-05, 20 November 2023, <<https://datatracker.ietf.org/doc/html/draft-ietf-detnet-scaling-requirements-05>>.

**[I-D.joung-detnet-stateless-fair-queuing]** Joung, J., Ryoo, J., Cheung, T., Li, Y., and P. Liu, "Latency Guarantee with Stateless Fair Queuing", Work in Progress, Internet-Draft, draft-joung-detnet-stateless-fair-queuing-02, 29 February 2024, <<https://datatracker.ietf.org/doc/html/draft-joung-detnet-stateless-fair-queuing-02>>.

**[I-D.peng-detnet-deadline-based-forwarding]**

Peng, S., Du, Z., Basu, K., cheng, Yang, D., and C. Liu, "Deadline Based Deterministic Forwarding", Work in Progress, Internet-Draft, draft-peng-detnet-deadline-based-forwarding-09, 1 March 2024, <<https://datatracker.ietf.org/doc/html/draft-peng-detnet-deadline-based-forwarding-09>>.

**[I-D.peng-detnet-packet-timeslot-mechanism]**

Peng, S., Liu, P., Basu, K., Liu, A., Yang, D., and G. Peng, "Timeslot Queueing and Forwarding Mechanism", Work in Progress, Internet-Draft, draft-peng-detnet-packet-timeslot-mechanism-06, 4 March 2024, <<https://>

[datatracker.ietf.org/doc/html/draft-peng-detnet-packet-timeslot-mechanism-06](https://datatracker.ietf.org/doc/html/draft-peng-detnet-packet-timeslot-mechanism-06)>.

[IEEE\_802.1Qcr] 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>>.

[STILIADIS-LRS] Stiliadis, D. and A. Anujan, "Latency-rate servers: A general model for analysis of traffic scheduling algorithms", IEEE/ACM Trans. Networking, vol. 6, no. 5, pp. 611-624, 1998.

[THOMAS-Sync] Thomas, L. and J-Y. Le Boudec, "On Time Synchronization Issues in Time-Sensitive Networks with Regulators and Nonideal Clocks", Proceedings of the ACM on Measurement and Analysis of Computing Systems, vol. 4, no. 2, pp. 1-41, 2020.

#### Authors' Addresses

Jinoo Joung  
Sangmyung University

Email: [jjoung@smu.ac.kr](mailto:jjoung@smu.ac.kr)

Xuesong Geng  
Huawei

Email: [gengxuesong@huawei.com](mailto:gengxuesong@huawei.com)

Shaofu Peng  
ZTE Corporation

Email: [peng.shaofu@zte.com.cn](mailto:peng.shaofu@zte.com.cn)

Toerless Eckert  
Futurewei Technologies

Email: [tte@cs.fau.de](mailto:tte@cs.fau.de)