

Recommendation

ITU-T Y.3129 (04/2024)

SERIES Y: Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities

Future networks

Requirements and framework for stateless fair queuing in large scale networks including IMT-2020 and beyond

Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities

GLOBAL INFORMATION INFRASTRUCTURE	Y.100-Y.999
General	Y.100-Y.199
Services, applications and middleware	Y.200-Y.299
Network aspects	Y.300-Y.399
Interfaces and protocols	Y.400-Y.499
Numbering, addressing and naming	Y.500-Y.599
Operation, administration and maintenance	Y.600-Y.699
Security	Y.700-Y.799
Performances	Y.800-Y.899
INTERNET PROTOCOL ASPECTS	Y.1000-Y.1999
General	Y.1000-Y.1099
Services and applications	Y.1100-Y.1199
Architecture, access, network capabilities and resource management	Y.1200-Y.1299
Transport	Y.1300-Y.1399
Interworking	Y.1400-Y.1499
Quality of service and network performance	Y.1500-Y.1599
Signalling	Y.1600-Y.1699
Operation, administration and maintenance	Y.1700-Y.1799
Charging	Y.1800-Y.1899
IPTV over NGN	Y.1900-Y.1999
NEXT GENERATION NETWORKS	Y.2000-Y.2999
Frameworks and functional architecture models	Y.2000-Y.2099
Quality of Service and performance	Y.2100-Y.2199
Service aspects: Service capabilities and service architecture	Y.2200-Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250-Y.2299
Enhancements to NGN	Y.2300-Y.2399
Network management	Y.2400-Y.2499
Computing power networks	Y.2500-Y.2599
Packet-based Networks	Y.2600-Y.2699
Security	Y.2700-Y.2799
Generalized mobility	Y.2800-Y.2899
Carrier grade open environment	Y.2900-Y.2999
FUTURE NETWORKS	Y.3000-Y.3499
CLOUD COMPUTING	Y.3500-Y.3599
BIG DATA	Y.3600-Y.3799
QUANTUM KEY DISTRIBUTION NETWORKS	Y.3800-Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	Y.4000-Y.4999
General	Y.4000-Y.4049
Definitions and terminologies	Y.4050-Y.4099
Requirements and use cases	Y.4100-Y.4249
Infrastructure, connectivity and networks	Y.4250-Y.4399
Frameworks, architectures and protocols	Y.4400-Y.4549
Services, applications, computation and data processing	Y.4550-Y.4699
Management, control and performance	Y.4700-Y.4799
Identification and security	Y.4800-Y.4899
Evaluation and assessment	Y.4900-Y.4999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T Y.3129

Requirements and framework for stateless fair queuing in large scale networks including IMT-2020 and beyond

Summary

Recommendation ITU-T Y.3129 specifies the requirements and framework for deterministic networking with a set of work conserving packet schedulers that guarantees end-to-end latency bounds to flows. The schedulers in core nodes do not need to maintain flow states. Instead, the entrance node of a flow marks an ideal service completion time according to a fluid model, called finish time (FT), of a packet in the packet header. The subsequent core nodes update FT by adding a delay factor, which is a function of the flow and upstream nodes. The packets in the queue of the scheduler are served in the ascending order of FT. This mechanism is called stateless fair queuing. The result is that flows are isolated from each other almost perfectly. The latency bound of a flow depends only on the flow's intrinsic parameters, except the maximum packet length among other flows sharing each output link with the flow.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T Y.3129	2024-04-29	13	11.1002/1000/15868

Keywords

Deterministic networking, fair queuing, finish time, flow isolation, latency guarantee, scheduler.

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2024

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

	Page
1 Scope.....	1
2 References.....	1
3 Definitions	1
3.1 Terms defined elsewhere	1
3.2 Terms defined in this Recommendation.....	1
4 Abbreviations and acronyms	1
5 Conventions	2
6 Introduction.....	2
7 Requirements	4
7.1 Requirements for FT generation and update	4
7.2 Requirements for FT properties in core nodes	5
8 Framework.....	5
8.1 General description.....	5
8.2 Selection of delay factor for latency guarantee	6
8.3 Network configuration for latency guarantee.....	6
8.4 Role of entrance node for generation and update of FT	6
8.5 Role of core node for update of FT	7
8.6 Mitigation of the complexity of entrance node	7
8.7 Compensation of time difference between nodes.....	7
9 Considerations for IMT-2020 networks	8
10 Security considerations	8
Appendix I – Mathematical representation of the requirements and framework	9
Appendix II – Delay factor function examples and their performances	11
II.1 Simulation environment	11
II.2 Delay factor for E2E latency bound	14
II.3 Delay factors without explicit E2E latency bound value	16
II.4 Implementation with FIFO queue per flow type	18
Appendix III – An example architecture for IMT-2020 networks.....	20
Bibliography.....	21

Recommendation ITU-T Y.3129

Requirements and framework for stateless fair queuing in large scale networks including IMT-2020 and beyond

1 Scope

This Recommendation specifies the requirements and framework for a scalable solution of latency guarantee and flow isolation for deterministic networking in large scale networks including IMT-2020 and beyond, as follows:

- the requirements to derive work conserving stateless fair queuing (FQ);
- the framework to achieve end-to-end (E2E) latency guarantee and flow isolation with stateless FQ.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-R M.1645] Recommendation ITU-R M.1645 (2003), *Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 large scale network [b-ITU-T Y.3120]: A network or a set of interconnected networks, with diameter of 16 or larger, in which the numbers of flows and nodes are proportional to the diameter of the network.

3.1.2 relay node [b-ITU-T Y.3113]: A node supporting relay functionality that acts as an intermediary node, through which other nodes can pass their traffic (e.g., router, switch or gateway).

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 IMT-2020: Systems, system components, and related technologies that provide far more enhanced capabilities than those described in [ITU-R M.1645].

NOTE 1 – Based on [b-ITU-R M.2083].

NOTE 2 – [ITU-R M.1645] defines the framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000 for the radio access network.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AN Access Network

ATS	Asynchronous Traffic Shaping
CN	Core Network
DiffServ	Differentiated Service
DN	Data Network
DRR	Deficit Round Robin
E2E	End to End
FIFO	First-In First-Out
FQ	Fair Queuing
FT	Finish Time
GBR	Guaranteed Bit Rate
IMT-2020	International Mobile Telecommunications-2020
PGPS	Packetized Generalized Processor Sharing
QoS	Quality of Service
TSpec	Traffic Specification
UE	User Equipment
VC	Virtual Clock
WRR	Weighted Round Robin

5 Conventions

In this Recommendation:

The phrase "is required" indicates a requirement that must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The phrase "is recommended" indicates a requirement that is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

The phrase "can optionally" indicates an optional requirement that is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator or service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with this Recommendation.

6 Introduction

There are emerging applications that require both latency and jitter bounds in large-scale networks [b-ITU-T Y.Supp.66]. One of the key mechanisms of latency and jitter performance of a network is packet scheduling. This Recommendation specifies a scheduling mechanism that can completely isolate one flow from others that share a link. An ideal flow isolation scheduling mechanism would provide service as if there are no other flows in an imaginary link whose capacity is equal to the allocated service rate to the flow. In this case, the latency upper bound is a function of parameters that depend only on the flow. Latency guarantee and flow isolation are closely related. The degree of flow isolation is inversely proportional to the worst latency. There are mainly three groups of flow isolation mechanism.

- The first group includes the FQ mechanism, which limits interference between flows to the degree of maximum packet size. Packetized generalized processor sharing (PGPS) and

weighted FQ are representative examples of this mechanism [b-Parekh]. The key characteristic is to record the service status of the flow in real time to provide the assigned amount of service. The worst delay at each node is proportional to the maximum packet length divided by the service rate of the flow. However, for this purpose, the complexity of managing and recording a large amount of state information for each flow is a problem, so it is not usually applied in practice. Note that a node in this Recommendation represents an output port module of a relay node.

- The second group includes round-robin based mechanisms that do not maintain detailed individual states reflecting the characteristics of the flow. Such mechanisms only record the static requirements of the flow and service history, and determine future services. This group includes weighted round robin (WRR) and deficit round robin (DRR). The worst delay at each node is proportional to the sum of the maximum packet sizes of all the flows divided by the link capacity.
- The third group includes priority-based mechanisms, which store flows in a queue grouped by priority and simply service the queue in order of priority. The mechanism suggested by differentiated services (DiffServs) is a representative example of the third group [b-IETF RFC 2475]. The worst delay at each node is determined by the sum of the maximum bursts of the flows divided by the link capacity. The maximum burst of a flow means the total amount of data allowed for the flow to send simultaneously. In general, it represents a bundle of a large number of packets, and therefore it is considerably larger than the packet length. The worst delay proportional to their sum may be at an unacceptable level in some cases. Moreover, in a network to which one of these mechanisms is applied, if there is a cycle in the network topology, the worst delay is amplified as it passes through the nodes. There may be a situation where the worst delay cannot be guaranteed depending on the level of network utilization [b-Thomas].

The highest level of flow isolation by FQ is desirable. A consequence of the isolation is that the latency bound guaranteed to a flow upon its admission, alters minimally as other flows join or leave. The FQ mechanisms are preferable, but too complex, as stated in the following.

Generalized processor sharing suggested a paradigm for a fair service for flow as a fluid, and PGPS, which implemented it in a packet environment, played a pioneering role in this type of packet-based scheduler [b-Parekh]. PGPS determines the service order of packets in ascending order of finish time (FT) derived by Equation 1.

$$F(p) = \max \{F(p-1), V(A(p))\} + L(p)/r \quad (1)$$

where p and $p-1$ are the p th and $(p-1)$ th packet of a flow, $F(p)$ is the FT, $A(p)$ is the arrival time, $L(p)$ is the length of the packet p , and r is the service rate allocated to the flow. $V(t)$ is called a virtual time function and is a value obtained by multiplying the actual time t by the sum of r of flows being served and the ratio of link capacity. Virtual time represents the progress of the system. If the backlogged flows almost fill the link capacity, then it increases slowly. If there is only a handful of backlogged flows, then the virtual time increases at a higher rate, which prevents an unfair situation in which flows that entered late have relatively small FTs and thus receive services earlier for a considerable duration of time, compared to existing flows.

$F(p)$, the FT of p , represents the packet service FT of p in an ideal fluid system. In a real system, the packets are served in the increasing order of the FT. The FT can be calculated the moment the packet arrives at the node, so it can be recorded and used in the node in the form of meta-data of the packet before it is stored in the buffer. In general, there is a queue for each flow, the corresponding queue is managed by first-in first-out (FIFO), and the scheduler is implemented in the form of servicing the queue with the smallest completion time by looking at the head of queue of the queue for each flow. Alternatively, it is possible to put all the packets in one queue and insert them in the middle of the queue according to the value of the FT. The FT is the expected service completion time when the corresponding packet is fairly serviced. The point of Equation 1 is that, in the worst case, when all