

Recommendation

ITU-T Y.3183 (01/2023)

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

Future networks

Framework for network slicing management assisted by machine learning leveraging quality of experience feedback from verticals

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

GLOBAL INFORMATION INFRASTRUCTURE

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

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

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

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.3183

Framework for network slicing management assisted by machine learning leveraging quality of experience feedback from verticals

Summary

Recommendation ITU-T Y.3183 provides a framework for machine learning-assisted network slicing management, leveraging vertical end users' feedback on quality of experience (QoE), which can help achieve the runtime optimization of user perceived performance. The overall architecture, components, workflow and related APIs of this framework are specified with respect to the high-level requirements identified. A use case is provided in Appendix I to show an application example of this framework. Example implementations of the key application programming interfaces (APIs) are also provided.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y.3183	2023-01-13	13	11.1002/1000/15238

Keywords

ML, network slice, QoE, QoS, vertical feedback.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

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 2023

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	2
	3.1 Terms defined elsewhere	2
	3.2 Terms defined in this Recommendation.....	3
4	Abbreviations and acronyms	3
5	Conventions	4
6	Introduction	5
7	High-level requirements	7
	7.1 General requirements.....	7
	7.2 Functional requirements	7
8	Framework and workflow.....	8
	8.1 Overall framework.....	8
	8.2 QoE feedback workflow	10
9	Components and related APIs	13
	9.1 OSA	13
	9.2 P&P controller	14
	9.3 Control plane	15
10	Security considerations.....	16
	Appendix I – eHealth use case	18
	Appendix II – Examples of P&P controller APIs	20
	II.1 Example of a continuous monitoring data stream API request	20
	II.2 Continuous monitoring data stream API response	20
	II.3 QoE feedback API.....	21
	Appendix III – Examples of control plane APIs.....	22
	III.1 Example usage of QoS control	22
	III.2 Example usage of NF config control.....	23
	III.3 Example usage of IPC control.....	24
	Bibliography.....	30

Recommendation ITU-T Y.3183

Framework for network slicing management assisted by machine learning leveraging quality of experience feedback from verticals

1 Scope

This Recommendation provides a framework for machine learning-assisted network slicing leveraging feedback from verticals to improve performance. It addresses the following subjects:

- Overview and requirements;
- Framework for machine learning-assisted network slice (NS) management leveraging input from verticals, focusing on QoE feedback;
- Components and APIs in the framework.

An eHealth use case is provided in Appendix I to show an application example of this framework.

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-T X.1601] Recommendation ITU-T X.1601 (2015), *Security framework for cloud computing*.
- [ITU-T Y.2701] Recommendation ITU-T Y.2701 (2007), *Security requirements for NGN release 1*.
- [ITU-T Y.3101] Recommendation ITU-T Y.3101 (2018), *Requirements of the IMT-2020 network*.
- [ITU-T Y.3104] Recommendation ITU-T Y.3104 (2018), *Architecture of the IMT-2020 network*.
- [ITU-T Y.3110] Recommendation ITU-T Y.3110 (2017), *IMT-2020 network management and orchestration requirements*.
- [ITU-T Y.3111] Recommendation ITU-T Y.3111 (2017), *IMT-2020 network management and orchestration framework*.
- [ITU-T Y.3170] Recommendation ITU-T Y.3170 (2018), *Requirements for machine learning-based quality of service assurance for the IMT-2020 network*.
- [ITU-T Y.3172] Recommendation ITU-T Y.3172 (2019), *Architectural framework for machine learning in future networks including IMT-2020*.
- [ITU-T Y.3174] Recommendation ITU-T Y.3174 (2020), *Framework for data handling to enable machine learning in future networks including IMT-2020*.
- [ITU-T Y.3177] Recommendation ITU-T Y.3177 (2021), *Architectural framework for artificial intelligence-based network automation for resource and fault management in future networks including IMT-2020*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 control plane [b-ITU-T Y.2011]: The set of functions that controls the operation of entities in the stratum or layer under consideration, plus the functions required to support this control.

3.1.2 data plane [b-ITU-T Y.2011]: The set of functions used to transfer data in the stratum or layer under consideration.

NOTE – In the ITU-T IMT-2020 related standard documents, "user plane" is used preferentially to "data plane".

3.1.3 management [b-ITU-T Y.3100]: In the context of IMT-2020, the processes aiming at fulfilment, assurance and billing of services, network functions and resources in both physical and virtual infrastructure including compute, storage, and network resources.

3.1.4 network function [b-ITU-T Y.3100]: In the context of IMT-2020, a processing function in a network.

NOTE 1 – Network functions include but are not limited to network node functionalities, e.g., session management, mobility management and transport functions, whose functional behaviour and interfaces are defined.

NOTE 2 – Network functions can be implemented on a dedicated hardware or as virtualized software functions.

NOTE 3 – Network functions are not regarded as resources, but rather any network functions can be instantiated using the resources.

3.1.5 network infrastructure management provider [b-ITU-T Y.3103]: The network infrastructure management provider integrates infrastructures of multiple infrastructure providers to offer the combined resources to the NS management and orchestration provider.

3.1.6 network infrastructure provider [b-ITU-T Y.3103]: The network infrastructure provider is the owner, the provider and the manager of the network infrastructure.

3.1.7 network slice [b-ITU-T Y.3100]: A logical network that provides specific network capabilities and network characteristics.

NOTE 1 – Network slices enable the creation of customized networks to provide flexible solutions for different market scenarios which have diverse requirements, with respect to functionalities, performance and resource allocation.

NOTE 2 – A network slice may have the ability to expose its capabilities.

NOTE 3 – The behaviour of a network slice is realized via network slice instance(s).

3.1.8 network slice blueprint [b-ITU-T Y.3100]: A complete description of the structure, configuration and work flows on how to create and control a network slice instance during its life cycle.

NOTE – A network slice template can be used synonymously with a network slice blueprint.

3.1.9 network slice instance [b-ITU-T Y.3100]: An instance of network slice, which is created based on a network slice blueprint.

NOTE 1 – A network slice instance is composed of a set of managed run-time network functions, and physical/logical/virtual resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the service instance(s).

NOTE 2 – A network slice instance may also be shared across multiple service instances provided by the network operator. A network slice instance may be composed of none, one or more sub-network slice instances which may be shared with another network slice instance.

3.1.10 network slice management and orchestration provider [b-ITU-T Y.3103]: The NS management and orchestration provider is responsible for orchestrating NS(s) and managing the lifecycle of NS instance(s) based on NS blueprint(s) (a term defined in [b-ITU-T Y.3100] (see definition in clause 3.1.2)).

NOTE – The NS blueprints can be provided by third parties.

3.1.11 network slice provider [b-ITU-T Y.3103]: The NS provider is the owner of the NS instance(s) and provides the NS instance(s).

3.1.12 network slice service provider [b-ITU-T Y.3103]: The NS service provider is the user of the NS instance(s), and is responsible for providing services to its NS service users via the NS instance(s).

3.1.13 network slice service user [b-ITU-T Y.3103]: The NS service user uses the service(s) provided by the NS instance(s).

3.1.14 network sub-slice management and orchestration provider [b-ITU-T Y.3103]: The network sub-slice management and orchestration provider is responsible for orchestrating network sub-slice(s) and managing the lifecycle of the network sub-slice instance(s).

3.1.15 network sub-slice provider [b-ITU-T Y.3103]: The network sub-slice provider is the owner and provider of the network sub-slice instance(s).

3.1.16 orchestration [b-ITU-T Y.3100]: In the context of IMT-2020, the processes aiming at the automated arrangement, coordination, instantiation and use of network functions and resources for both physical and virtual infrastructures by optimization criteria.

3.1.17 quality of experience (QoE) [b-ITU-T P.10]: The degree of delight or annoyance of the user of an application or service.

3.1.18 quality of service (QoS) [b-ITU-T Q.1741.9]: The collective effect of service performances, which determine the degree of satisfaction of a user of a service. It is characterized by the combined aspects of performance factors applicable to all services, such as:

- service operability performance;
- service accessibility performance;
- service retainability performance;
- service integrity performance; and
- other factors specific to service.

3.1.19 role [b-ITU-T Y.3502]: A set of activities that serves a common purpose.

3.1.20 user plane (UP) [b-ITU-T Y.1714]: Refers to the set of traffic forwarding components through which traffic flows.

NOTE – "User plane" is also referred to as "transport plane" in other ITU-T Recommendations.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

API Application Programming Interface

CN	Core Network
CPS	Control Plane Service
CPSR	Control Plane Service Register
DPP	Data Plane Programmability
ECA	which Event and Condition what Action
E2E	End-to-End
HATEOAS	Hypertext As The Engine Of Application State
IMSI	International Mobile Subscriber Identity
IMT-2020	International Mobile Telecommunications-2020
IPC	InterPoint of presence Connection
MANO	Management and Orchestration
MEC	Mobile/Multiaccess Edge Computing
ML	Machine Learning
NF	Network Function
NFV	Network Function Virtualization
NS	Network Slice
NSI	Network Slice Instance
NSp	Network Slice providers
NSsp	Network Slice service providers
NSsu	Network Slice service users
OSA	One-Stop Access
P&P	Plug and Play
PoP	Point of Presence
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
REST	REpresentational State Transfer
SDN	Software Defined Networking
SLA	Service-Level Agreement
VNF	Virtual Network Function
WAN	Wide Area Network

5 Conventions

In this Recommendation:

The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted, if conformance to this Recommendation is to be claimed.

The keywords "is recommended" indicate a specification which is recommended but which is not absolutely required. Thus, this specification need not be present to claim conformance.

The keywords "can optionally" indicate an optional requirement which 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/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with this Recommendation.

6 Introduction

Network slicing is a primary approach in future networks including IMT-2020 and beyond to meet the diverging service requirements from diverse business verticals by creating multiple logical networks with the required quality of service (QoS). Network slicing connects virtual and/or physical network functions (NFs) [b-ITU-T Y.3100] with the required underlying resources through chaining to create the intended service. Truly end-to-end network slicing needs to create properly isolated logical networks across all the physical network segments all the way from source to destination, including the radio access network (RAN), mobile/Multiaccess edge computing (MEC), backhaul/transport network, core network (CN) and wide area network (WAN), regardless of the heterogeneity of the underlying infrastructures. Moreover, there are various aspects to be addressed to achieve advanced network slicing, including guaranteed QoS and customized network services, among others. Guaranteed QoS can be achieved through QoS control of the various network segments via radio resource allocation specific to various and often heterogeneous wireless/mobile systems and data plane programmability (DPP) for diverse wired data path technologies. Customized network services are created through virtual network function (VNF) or more generally NF configuration and chaining across the interested points of presence (PoPs). Therefore, a set of NS control services is necessitated.

On top of these network slicing capabilities, for optimized network slicing operation, machine learning (ML) can be used to enhance network slicing performance. Figure 1 illustrates this vision.

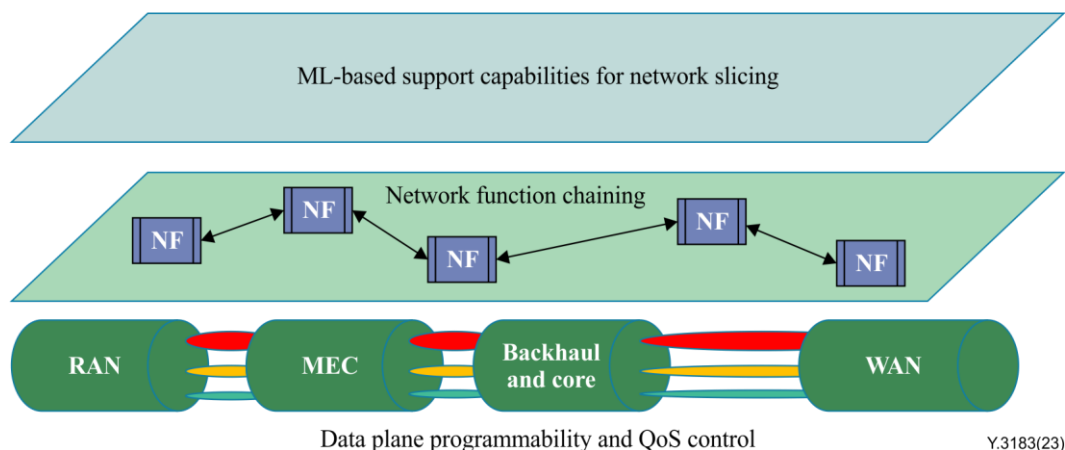


Figure 1 – ML-based support for network slicing

Furthermore, it is highly desirable that verticals be allowed to customize their NS-based services and provide as inputs the customization requirements in not only the service provisioning phase but also the runtime phase. Upon receiving such vertical feedback, the management system of a future network should apply the feedback to modify the service, and the corresponding end-to-end (E2E) NS for that vertical. Depending on the level of the vertical's feedback, the management system may need to convert quality of experience (QoE) requirements into QoS constraints and/or service-level agreement (SLA) terms. Subsequently, these QoS/SLA constraints are employed to influence the modification of the NS instance and/or inform other related operations to support the fulfilment of the requirements in terms of QoS/SLA constraints. For autonomous network control and

management, the above steps are required to be automated, enabled by ML, automation and other related technologies.

Such quality-aware NS control is a challenging task that requires a proper framework and tools in place to be executed efficiently. Of particular importance is the quality perceived by the network slice service users (NSSu) [b-ITU-T Y.3103], especially the verticals, since it has a correlation with their degree of satisfaction with the deployed service. As such, it becomes essential to capture the verticals' perspective on the runtime behaviour of the deployed NS and its influence on the services deployed on top. Network slice service providers (NSSp) [b-ITU-T Y.3103] will usually have contractual relationships with multiple types of verticals (e.g., automotive industry, power distribution, eHealth), and thus they will require a mechanism for gauging the satisfaction of the vertical customers, i.e., their QoE with the deployed NS.

The NSSp cooperate with the network slice providers (NSp) [b-ITU-T Y.3103] that create the corresponding NS in order to provide the required services to the NSSu.

Moreover, an inclusive stakeholder interface is expected for a common approach for various stakeholders to support all their interactions with the network slicing system via different projections of the information accessible for different purposes, based on the role of the stakeholder who consumes or influences the information and even operations. This will allow the stakeholders to utilize wireless technologies of various generations such as IMT-2020 and beyond, support dynamicity and adaptability with respect to the varied and distributed demands of the served verticals, allow for the expression of a multitude of performance, security and other requirements, enable various stakeholders to be able to participate as providers, and allow verticals to have certain level of control over their NSs for applying the principles of their application logic.

Furthermore, the delivery of heterogeneous services customized for a large set of vertical businesses with diverging requirements (e.g., in terms of QoS/SLA and QoE) requires an effective and common framework able to keep verticals in the loop of services and NSs provisioning, from design to operation. However, current offerings from service providers and network operators are still based on a very limited involvement of verticals and consumers in the runtime control and management of the offered services. This gap needs to be filled along the development of network slicing and the stakeholder interface, by allowing a vertical to have extended control capabilities of their NS-based services in a controlled manner. In practice, dynamic exposure control is the key enabler for this evolution of services and NS control exposure towards a paradigm where verticals play an active role in the runtime control and management of their NSs.

This Recommendation describes a framework that enables vertical QoE-aware NS control empowered by ML technologies. This framework allows QoE feedback from vertical businesses to facilitate a network slicing system to optimize the runtime slicing and thus the corresponding verticals' NS-based services. Such vertical feedback will then be explored to inform configuration actions or other actuations to be executed so that the NS belonging to that vertical is modified accordingly to meet the dynamic requirements of the vertical. The overall framework is presented with a set of essential components described including a stakeholder interface to allow the vertical feedback (and other stakeholders to access the system), dynamic exposure control to customize the vertical's view and control with continuous monitoring capabilities, and an overlay control plane on top of IMT-2020 and beyond control plane and data plane to enforce the actions/actuations. A set of associated application programming interfaces (APIs) for monitoring or control are described (with implementation details included in Appendices II and III). An eHealth use case is used to illustrate the key processes in this framework (more details are presented in Appendix I). Finally, it is noted that there is no limitation of the framework to other scenarios of applicability beyond network slicing.