teas
Internet-Draft
Intended status: Informational
Expires: August 26, 2021

R. Rokui
Nokia
S. Homma
NTT
K. Makhijani
Futurewei
LM. Contreras
Telefonica
J. Tantsura
Juniper Networks
February 22, 2021

Definition of IETF Network Slices
draft-ietf-teas-ietf-network-slice-definition-01

Abstract

This document provides a definition of the term "IETF Network Slice" for use within the IETF and specifically as a reference for other IETF documents that describe or use aspects of network slices.

The document also describes the characteristics of an IETF network slice, related terms and their meanings, and explains how IETF network slices can be used in combination with end-to-end network slices or independent of them.

Table of Contents

1. Introduction

A number of use cases benefit from network connections that along with the connectivity provide assurance of meeting a specific set of objectives wrt network resources use. In this document, as detailed in the subsequent sections, we refer to this connectivity and

resource commitment as an IETF Network Slice. Services that might benefit from the network slices include but not limited to:

- 5G services (e.g. eMBB, URLLC, mMTC)(See [TS.23.501-3GPP])
- Network wholesale services
- Network infrastructure sharing among operators
- NFV connectivity and Data Center Interconnect

The use cases are further described in [I-D.nsdt-teas-ns-framework].

This document defines the concept of IETF network slices that provide connectivity coupled with a set of specific commitments of network resources between a number of endpoints over a shared network infrastructure. Since the term network slice is rather generic, the qualifying term 'IETF' is used in this document to limit the scope of network slice to network technologies described and standardized by the IETF.

IETF network slices are created and managed within the scope of one or more network technologies (e.g., IP, MPLS, optical). They are intended to enable a diverse set of applications that have different requirements to coexist on the shared network infrastructure. A request for an IETF network slice is technology-agnostic so as to allow a consumer to describe their network connectivity objectives in a common format, independent of the underlying technologies used.

2. Terms and Abbreviations

The terms and abbreviations used in this document are listed below.

- NS: Network Slice
- NSC: Network Slice Controller
- NBI: NorthBound Interface
- SBI: SouthBound Interface
- SLI: Service Level Indicator
- SLO: Service Level Objective
- SLA: Service Level Agreement

The above terminology is defined in greater details in the remainder of this document.

3. Definition and Scope of IETF Network Slice

The definition of a network slice in IETF context is as follows:

An IETF network slice is a logical network topology connecting a number of endpoints using a set of shared or dedicated network resources that are used to satisfy specific Service Level Objectives (SLOs).

An IETF network slice combines the connectivity resource requirements and associated network behaviors such as bandwidth, latency, jitter, and network functions with other resource behaviors such as compute and storage availability. IETF network slices are independent of the underlying infrastructure connectivity and technologies used. This is to allow an IETF network slice consumer to describe their network connectivity and relevant objectives in a common format, independent of the underlying technologies used.

IETF network slices may be combined hierarchically, so that a network slice may itself be sliced. They may also be combined sequentially so that various different networks can each be sliced and the network slices placed into a sequence to provide an end-to-end service. This form of sequential combination is utilized in some services such as in 3GPP's 5G network [TS.23.501-3GPP].

An IETF network slice is technology-agnostic, and the means for IETF network slice realization can be chosen depending on several factors such as: service requirements, specifications or capabilities of underlying infrastructure. The structure and different characteristics of IETF network slices are described in the following sections.

Term "Slice" refers to a set of characteristics and behaviours that separate one type of user-traffic from another. IETF network slice assumes that an underlying network is capable of changing the configurations of the network devices on demand, through in-band signaling or via controller(s) and fulfilling all or some of SLOs to all of the traffic in the slice or to specific flows.

4. IETF Network Slice System Characteristics

The following subsections describe the characteristics of IETF network slices.

4.1. Objectives for IETF Network Slices

An IETF network slice is defined in terms of several quantifiable characteristics or service level objectives (SLOs). SLOs along with terms Service Level Indicator (SLI) and Service Level Agreement (SLA) are used to define the performance of a service at different levels.

A Service Level Indicator (SLI) is a quantifiable measure of an aspect of the performance of a network. For example, it may be a measure of throughput in bits per second, or it may be a measure of latency in milliseconds.

A Service Level Objective (SLO) is a target value or range for the measurements returned by observation of an SLI. For example, an SLO may be expressed as "SLI <= target", or "lower bound <= SLI <= upper bound". A network slice is expressed in terms of the set of SLOs that are to be delivered for the different connections between endpoints.

A Service Level Agreement (SLA) is an explicit or implicit contract between the consumer of an IETF network slice and the provider of the slice. The SLA is expressed in terms of a set of SLOs and may include commercial terms as well as the consequences of missing/ violating the SLOs they contain.

Additional descriptions of IETF network slice attributes is covered in [I-D.contreras-teas-slice-nbi].

4.1.1. Service Level Objectives

SLOs define a set of network attributes and characteristics that describe an IETF network slice. SLOs do not describe 'how' the IETF network slices are implemented or realized in the underlying network layers. Instead, they are defined in terms of dimensions of operation (time, capacity, etc.), availability, and other attributes. An IETF network slice can have one or more SLOs associated with it. The SLOs are combined in an SLA. The SLOs are defined for sets of two or more endpoints and apply to specific directions of traffic flow. That is, they apply to specific source endpoints and specific connections between endpoints within the set of endpoints and connections in the network slice.

4.1.2. Minimal Set of SLOs

This document defines a minimal set of SLOs and later systems or standards could extend this set as per Section 4.1.3.

SLOs can be categorized in to 'Directly Measurable Objectives' or 'Indirectly Measurable Objectives'. Objectives such as guaranteed minimum bandwidth, guaranteed maximum latency, maximum permissible delay variation, maximum permissible packet loss rate, and availability are 'Directly Measurable Objectives'. While 'Indirectly Measurable Objectives' include security, geographical restrictions, maximum occupancy level objectives. The later standard might define other SLOs as needed.

Editor's Note TODO: replace Minimal set to most commonly used objectives to describe network behavior. Other directly or indirectly measurable objectives may be requested by that consumer of an IETF network slice.

The definition of these objectives are as follows:

Guaranteed Minimum Bandwidth

Minimum guaranteed bandwidth between two endpoints at any time. The bandwidth is measured in data rate units of bits per second and is measured unidirectionally.

Guaranteed Maximum Latency

Upper bound of network latency when transmitting between two endpoints. The latency is measured in terms of network characteristics (excluding application-level latency). [RFC2681] and [RFC7679] discuss round trip times and one-way metrics, respectively.

Maximum Permissible Delay Variation

Packet delay variation (PDV) as defined by [RFC3393], s the difference in the one-way delay between sequential packets in a flow. This SLO sets a maximum value PDV for packets between two endpoints.

Maximum permissible packet loss rate

The ratio of packets dropped to packets transmitted between two endpoints over a period of time. See [RFC7680]

Availability

The ratio of uptime to the sum of uptime and downtime, where uptime is the time the IETF network slice is available in accordance with the SLOs associated with it.

Security

An IETF network slice consumer may request that the network applies encryption or other security techniques to traffic flowing between endpoints.

Note that the use of security or the violation of this SLO is not directly observable by the IETF network slice consumer and cannot be measured as a quantifiable metric.

Also note that the objective may include request for encryption (e.g., [RFC4303]) between the two endpoints explicitly to meet architecture recommendations as in [TS33.210] or for compliance with [HIPAA] and/or [PCI].

Editor's Note: Please see more discussion on security in Section 10.

4.1.3. Other Objectives

Additional SLOs may be defined to provide additional description of the IETF network slice that a consumer requests.

If the IETF network slice consumer service is traffic aware, other traffic specific characteristics may be valuable including MTU, traffic-type (e.g., IPv4, IPv6, Ethernet or unstructured), or a higher-level behavior to process traffic according to user-application (which may be realized using network functions).

Maximal occupancy for an IETF network slice should be provided. Since it carries traffic for multiple flows between the two endpoints, the objectives should also say if they are for the entire connection, group of flows or on per flow basis. Maximal occupancy should specify the scale of the flows (i.e. maximum number of flows to be admitted) and optionally a maximum number of countable resource units, e.g IP or MAC addresses a slice might consume.

4.2. IETF Network Slice Endpoints

As noted in Section 3, an IETF network slice describes connectivity between multiple endpoints across the underlying network. These connectivity types are: point-to-point, point-to-multipoint, multipoint-to-point multipoint-to-point, or multipoint-to-multipoint.

Figure 1 shows an IETF network slice along with its NSEs.

The characteristics of IETF network slice endpoints (NSEs) are as follows:

- The IETF network slice endpoints (NSEs) are conceptual points of connection to IETF network slice. As such, they serve as the IETF network slice ingress/egress points.

- Each endpoint could map to a device, application or a network function. A non-exhaustive list of devices, applications or network functions might include but not limited to: routers, switches, firewalls, WAN, 4G/5G RAN nodes, 4G/5G Core nodes, application acceleration, Deep Packet Inspection (DPI), server load balancers, NAT44 [RFC3022], NAT64 [RFC6146], HTTP header enrichment functions, and TCP optimizers.

- An NSE should be identified by a unique ID in the context of an IETF network slice consumer.

- In addition to an identifier, each NSE should contain a subset of attributes such as IPv4/IPv6 addresses, encapsulation type (i.e., VLAN tag, MPLS Label etc.), interface/port numbers, node ID etc.

- A combination of NSE unique ID and NSE attributes defines an NSE in the context of the IETF network slice controller.

- During the realization of the IETF network slice, in addition to SLOs, all or subset of IETF NSE attributes will be utilized by IETF network slice controller (NSC) to find the optimal realization in the IETF network.

- Similarly to IETF network slices, the IETF network slice endpoints are logical entities that are mapped to services/tunnels/paths endpoints in IETF network slice during its initialization and realization.

Note that there are various IETF TE terms such as access points (AP) defined in [RFC8453], Termination Point (TP) defined in [RFC8345], and Link Termination Point (LTP) defined in [RFC8795] which are tightly coupled with TE network type and various realization techniques. At the time of realization of the IETF network slice, the NSE could be mapped to one or more of these based on the network slice realization technique in use.

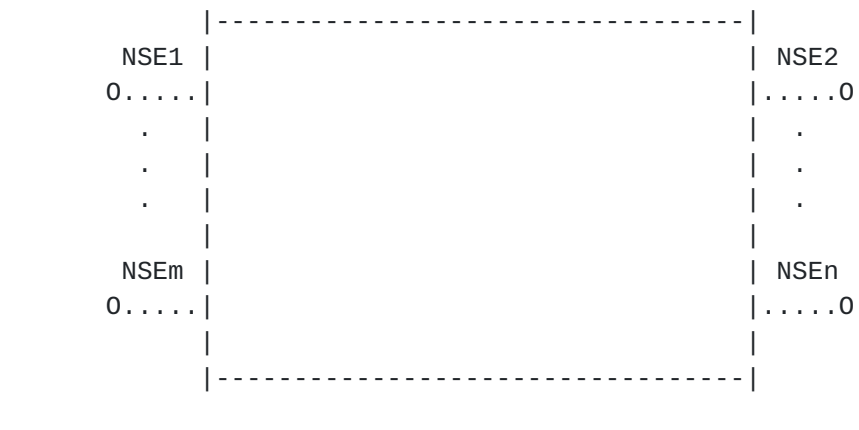


Figure 1: An IETF Network Slice Endpoints (NSE)

4.2.1. IETF Network Slice Connectivity Types

The IETF Network Slice connection types can be point to point (P2P), point to multipoint (P2MP), multi-point to point (MP2P), or multi-point to multi-point (MP2MP). They will requested by the higher level operation system.

4.3. IETF Network Slice Composition

Operationally, an IETF network slice may be decomposed in two or more IETF network slices as specified below. Decomposed network slices are then independently realized and managed.

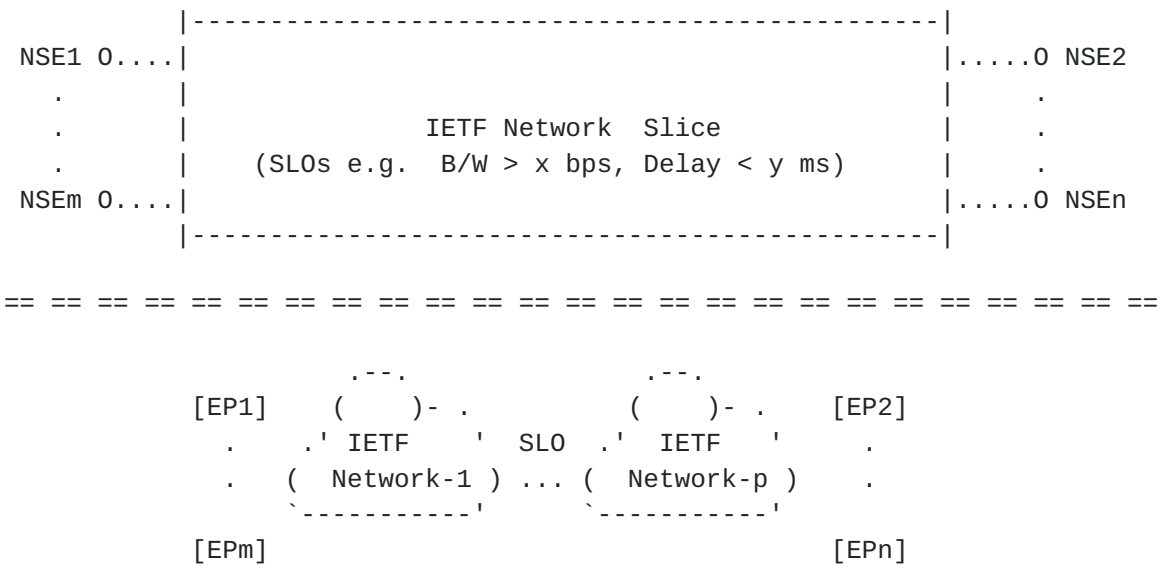
- Hierarchical (i.e., recursive) composition: An IETF network slice can be further sliced into other network slices. Recursive composition allows an IETF network slice at one layer to be used by the other layers. This type of multi-layer vertical IETF network slice associates resources at different layers.

- Sequential composition: Different IETF network slices can be placed into a sequence to provide an end-to-end service. In sequential composition, each IETF network slice would potentially support different dataplanes that need to be stitched together.

5. IETF Network Slice Structure

Editor's note: This content of this section merged with Relationship with E2E slice discussion.

An IETF network slice is a set of connections among various endpoints to form a logical network that meets the SLOs agreed upon.



Legendy
NSE: IETF Network Slice Endpoints
EP: Serivce/tunnels/path Endpoints used to realize the IETF Network Slice

Figure 2: IETF Network slice

Figure 2 illustrates a case where an IETF network slice provides connectivity between a set of IEFT network slice endpoints (NSE) pairs with specific SLOs (e.g. guaranteed minimum bandwidth of x bps and guaranteed delay of no more than y ms). The IETF network slice endpoints are mapped to the underlay IETF networks endpoints (EP). Also, the IETF network slice endpoints on the same IETF network slice may belong to the same or different address spaces.

IETF Network slice structure fits into a broader concept of end-to-end network slices. A network operator may be responsible for delivering services over a number of technologies (such as radio networks) and for providing specific and fine-grained services (such as CCTV feed or High definition realtime traffic data). That

operator may need to combine slices of various networks to produce an end-to-end network service. Each of these networks may include multiple physical or virtual nodes and may also provide network functions beyond simply carrying of technology-specific protocol data units.An end-to-end network slice is defined by the 3GPP as a complete logical network that provides a service in its entirety with a specific assurance to the consumer [TS.23.501-3GPP].

An end-to-end network slice may be composed from other network slices that include IETF network slices. This composition may include the hierarchical (or recursive) use of underlying network slices and the sequential (or stitched) combination of slices of different networks.

6. IETF Network Slice Stakeholders

An IETF network slice and its realization involves the following stakeholders and it is relevant to define them for consistent terminology.

Consumer: A consumer is the requester of an IETF network slice. Consumers may request monitoring of SLOs. A consumer may manage the IETF network slice service directly by interfacing with the IETF network slice controller or indirectly through an orchestrator.

Orchestrator: An orchestrator is an entity that composes different services, resource and network requirements. It interfaces with the IETF network slice controllers.

IETF Network Slice Controller (NSC): It realizes an IETF network lice in the underlying network, maintains and monitors the run-time state of resources and topologies associated with it. A well-defined interface is needed between different types of IETF network slice controllers and different types of orchestrators. An IETF network slice operator (or slice operator for short) manages one or more IETF network slices using the IETF network slice Controller(s).

Network Controller: is a form of network infrastructure controller that offers network resources to NSC to realize a particular network slice. These may be existing network controllers associated with one or more specific technologies that may be adapted to the function of realizing IETF network slices in a network.

7. IETF Network Slice Controller Interfaces

The interworking and interoperability among the different stakeholders to provide common means of provisioning, operating and monitoring the IETF network slices is enabled by the following communication interfaces (see Figure 3).

NSC Northbound Interface (NBI): The NSC Northbound Interface is an interface between a consumer's higher level operation system (e.g., a network slice orchestrator) and the NSC. It is a technology agnostic interface. The consumer can use this interface to communicate the requested characteristics and other requirements (i.e., the SLOs) for the IETF network slice, and the NSC can use the interface to report the operational state of an IETF network slice to the consumer.

NSC Southbound Interface (SBI): The NSC Southbound Interface is an interface between the NSC and network controllers. It is technology-specific and may be built around the many network models defined within the IETF.

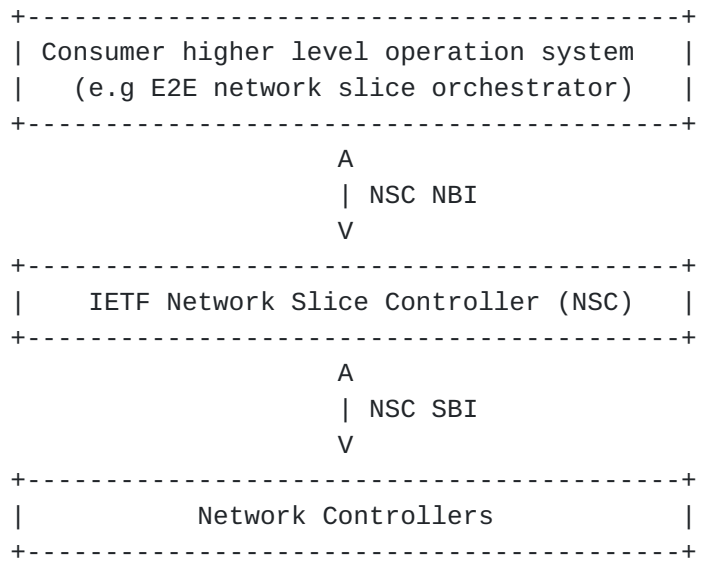


Figure 3: Interface of IETF Network Slice Controller

8. Realizing IETF Network Slice

Realization of IETF network slices is out of scope of this document. It is a mapping of the definition of the IETF network slice to the

underlying infrastructure and is necessarily technology-specific and achieved by the NSC over the SBI.

The realization can be achieved in a form of either physical or logical connectivity through VPNs (see, for example, [I-D.ietf-teas-enhanced-vpn], a variety of tunneling technologies such as Segment Routing, MPLS, etc. Accordingly, endpoints may be realized as physical or logical service or network functions.

9. Isolation in IETF Network Slices

An IETF network slice consumer may request, that the IETF Network Slice delivered to them is isolated from any other network slices of services delivered to any other consumers. It is expected that the changes to the other network slices of services do not have any negative impact on the delivery of the IETF network slice.

9.1. Isolation as a Service Requirement

Isolation may be an important requirement of IETF network slices for some critical services. A consumer may express this request as an SLO.

This requirement can be met by simple conformance with other SLOs. For example, traffic congestion (interference from other services) might impact on the latency experienced by an IETF network slice. Thus, in this example, conformance to a latency SLO would be the primary requirement for delivery of the IETF network slice service, and isolation from other services might be only a means to that end.

It should be noted that some aspects of isolation may be measurable by a consumer who have the information about the traffic on a number of IETF network slices or other services.

9.2. Isolation in IETF Network Slice Realization

Delivery of isolation is achieved in the realization of IETF network slices, with existing, in-development, and potential new technologies in IETF. It depends on how a network operator decides to operate their network and deliver services.

Isolation may be achieved in the underlying network by various forms of resource partitioning ranging from dedicated allocation of resources for a specific IETF network slice, to sharing or resources with safeguards. For example, traffic separation between different IETF network slices may be achieved using VPN technologies, such as L3VPN, L2VPN, EVPN, etc. Interference avoidance may be achieved by network capacity planning, allocating dedicated network resources,

traffic policing or shaping, prioritizing in using shared network resources, etc. Finally, service continuity may be ensured by reserving backup paths for critical traffic, dedicating specific network resources for a selected number of network slices, etc.

10. Security Considerations

This document specifies terminology and has no direct effect on the security of implementations or deployments. In this section, a few of the security aspects are identified.

- Conformance to security constraints: Specific security requests from consumer defined IETF network slices will be mapped to their realization in the unerlay networks. It will be required by underlay networks to have capabilities to conform to consumer's requests as some aspects of security may be expressed in SLOs.

- IETF network slice controller authentication: Unerlying networks need to be protected against the attacks from an adversary NSC as they can destablize overall network operations. It is particularly critical since an IETF network slice may span across different networks, therefore, IETF NSC should have strong authentication with each those networks. Futhermore, both SBI and NBI need to be secured.

- Specific isolation criteria: The nature of conformance to isolation requests means that it should not be possible to attack an IETF network slice service by varying the traffic on other services or slices carried by the same underlay network. In general, isolation is expected to strengthen the IETF network slice security.

- Data Integrity of an IETF network slice: A consumer wanting to secure their data and keep it private will be responsible for applying appropriate security measures to their traffic and not depending on the network operator that provides the IETF network slice. It is expected that for data integrity, a consumer is responsible for end-to-end encryption of its own traffic.

Note: see NGMN document [NGMN_SEC] on 5G network slice security for discussion relevant to this section.

11. IANA Considerations

This memo includes no request to IANA.

12. Acknowledgment

The entire TEAS NS design team and everyone participating in those discussion has contributed to this draft. Particularly, Eric Gray, Xufeng Liu, Jie Dong, Adrian Farrel, and Jari Arkko for a thorough review among other contributions.

13. Informative References

[HIPAA] HHS, "Health Insurance Portability and Accountability Act - The Security Rule", February 2003, <https://www.hhs.gov/hipaa/for-professionals/security/index.html>.

[I-D.contreras-teas-slice-nbi] Contreras, L., Homma, S., and J. Ordonez-Lucena, "Considerations for defining a Transport Slice NBI", draft-contreras-teas-slice-nbi-01 (work in progress), March 2020.

[I-D.ietf-teas-enhanced-vpn] Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A Framework for Enhanced Virtual Private Networks (VPN+) Services", draft-ietf-teas-enhanced-vpn-05 (work in progress), February 2020.

[I-D.ietf-teas-yang-te-topo] Liu, X., Bryskin, I., Beeram, V., Saad, T., Shah, H., and O. Dios, "YANG Data Model for Traffic Engineering (TE) Topologies", draft-ietf-teas-yang-te-topo-22 (work in progress), June 2019.

[I-D.nsdt-teas-ns-framework] Gray, E. and J. Drake, "Framework for Transport Network Slices", draft-nsdt-teas-ns-framework-02 (work in progress), March 2020.

[NGMN_SEC] NGMN Alliance, "NGMN 5G Security - Network Slicing", April 2016, <https://www.ngmn.org/wp-content/uploads/Publications/2016/160429_NGMN_5G_Security_Network_Slicing_v1_0.pdf>.

[PCI] PCI Security Standards Council, "PCI DSS", May 2018, <https://www.pcisecuritystandards.org>.

[RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, DOI 10.17487/RFC2681, September 1999, <https://www.rfc-editor.org/info/rfc2681>.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, DOI 10.17487/RFC3022, January 2001, <https://www.rfc-editor.org/info/rfc3022>.

[RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, DOI 10.17487/RFC3393, November 2002, <https://www.rfc-editor.org/info/rfc3393>.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, DOI 10.17487/RFC4303, December 2005, <https://www.rfc-editor.org/info/rfc4303>.

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, DOI 10.17487/RFC6146, April 2011, <https://www.rfc-editor.org/info/rfc6146>.

[RFC7679] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Delay Metric for IP Performance Metrics (IPPM)", STD 81, RFC 7679, DOI 10.17487/RFC7679, January 2016, <https://www.rfc-editor.org/info/rfc7679>.

[RFC7680] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Loss Metric for IP Performance Metrics (IPPM)", STD 82, RFC 7680, DOI 10.17487/RFC7680, January 2016, <https://www.rfc-editor.org/info/rfc7680>.

[RFC8345] Clemm, A., Medved, J., Varga, R., Bahadur, N., Ananthakrishnan, H., and X. Liu, "A YANG Data Model for Network Topologies", RFC 8345, DOI 10.17487/RFC8345, March 2018, <https://www.rfc-editor.org/info/rfc8345>.

[RFC8453] Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for Abstraction and Control of TE Networks (ACTN)", RFC 8453, DOI 10.17487/RFC8453, August 2018, <https://www.rfc-editor.org/info/rfc8453>.

[TS.23.501-3GPP] 3rd Generation Partnership Project (3GPP), "3GPP TS 23.501 (V16.2.0): System Architecture for the 5G System (5GS); Stage 2 (Release 16)", September 2019, <http://www.3gpp.org/ftp//Specs/archive/23_series/23.501/23501-g20.zip>.

[TS33.210] 3GPP, "3G security; Network Domain Security (NDS); IP network layer security (Release 14).", December 2016, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2279>.

Authors' Addresses

Reza Rokui
Nokia
Canada

Email: reza.rokui@nokia.com

Shunsuke Homma
NTT
Japan

Email: shunsuke.homma.ietf@gmail.com

Kiran Makhijani
Futurewei
USA

Email: kiranm@futurewei.com

Luis M. Contreras
Telefonica
Spain

Email: luismiguel.contrerasmurillo@telefonica.com

Jeff Tantsura
Juniper Networks

Email: jefftant.ietf@gmail.com