Workgroup: Network Working Group
Internet-Draft:
draft-ietf-teas-ietf-network-slices-08
Published: 6 March 2022
Intended Status: Informational
Expires: 7 September 2022
Authors: A. Farrel, Ed.
Old Dog Consulting
J. Drake, Ed.
Juniper Networks
R. Rokui
Ciena
S. Homma
NTT
K. Makhijani
Futurewei
L.M. Contreras
Telefonica
J. Tantsura
Microsoft

Framework for IETF Network Slices

Abstract

This document describes network slicing in the context of networks built from IETF technologies. It defines the term "IETF Network Slice" and establishes the general principles of network slicing in the IETF context.

The document discusses the general framework for requesting and operating IETF Network Slices, the characteristics of an IETF Network Slice, the necessary system components and interfaces, and how abstract requests can be mapped to more specific technologies. The document also discusses related considerations with monitoring and security.

This document also provides definitions of related terms to enable consistent usage in other IETF documents that describe or use aspects of IETF Network Slices.

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 7 September 2022.

Copyright Notice

Table of Contents

1. Introduction

A number of use cases benefit from network connections that, along with connectivity, provide assurance of meeting a specific set of objectives with respect to network resources use. This connectivity and resource commitment is referred to as a network slice and is expressed in terms of connectivity constructs (see Section 3) and service objectives (see Section 4). 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. 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 (known as Service Demarcation Points (SDPs) - see Section 2.1 and Section 4.2) over a shared underlay network. The term IETF Network Slice service is also introduced to describe the service requested by and provided to the service provider's customer.

Services that might benefit from IETF Network Slices include, but are not limited to:

- *5G services (e.g. eMBB, URLLC, mMTC)(See [TS23501])
- *Network wholesale services
- *Network infrastructure sharing among operators
- *NFV connectivity and Data Center Interconnect

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 with different requirements to coexist over a shared underlay network. A request for an IETF Network Slice service is agnostic to the technology in the underlying network so as to allow a customer to describe their network connectivity objectives in a common format, independent of the underlying technologies used.

This document also provides a framework for discussing IETF Network Slices. The framework is intended as a structure for discussing interfaces and technologies. It is not intended to specify a new set of concrete interfaces or technologies.

For example, virtual private networks (VPNs) have served the industry well as a means of providing different groups of users with logically isolated access to a common network. The common or base network that is used to support the VPNs is often referred to as an underlay network, and the VPN is often called an overlay network. An overlay network may, in turn, serve as an underlay network to support another overlay network.

Note that it is conceivable that extensions to IETF technologies are needed in order to fully support all the ideas that can be implemented with network slices. Evaluation of existing technologies, proposed extensions to existing protocols and interfaces, and the creation of new protocols or interfaces is outside the scope of this document.

1.1. Background

The concept of network slicing has gained traction driven largely by needs surfacing from 5G ([NGMN-NS-Concept], [TS23501], and [TS28530]). In [TS23501], a Network Slice is defined as "a logical network that provides specific network capabilities and network characteristics", and a Network Slice Instance is defined as "A set of Network Function instances and the required resources (e.g. compute, storage and networking resources) which form a deployed Network Slice." According to [TS28530], an end-to-end network slice consists of three major types of network segments: Radio Access Network (RAN), Transport Network (TN) and Core Network (CN). An IETF Network Slice provides the required connectivity between different entities in RAN and CN segments of an end-to-end network slice, with a specific performance commitment. For each end-to-end network slice, the topology and performance requirement on a customer's use of an IETF Network Slice can be very different, which requires the underlay network to have the capability of supporting multiple different IETF Network Slices.

While network slices are commonly discussed in the context of 5G, it is important to note that IETF Network Slices are a narrower concept with a broader usage profile, and focus primarily on particular network connectivity aspects. Other systems, including 5G deployments, may use IETF Network Slices as a component to create entire systems and concatenated constructs that match their needs, including end-to-end connectivity.

An IETF Network Slice could span multiple technologies and multiple administrative domains. Depending on the IETF Network Slice customer's requirements, an IETF Network Slice could be isolated from other, often concurrent IETF Network Slices in terms of data, control and management planes.

The customer expresses requirements for a particular IETF Network Slice service by specifying what is required rather than how the requirement is to be fulfilled. That is, the IETF Network Slice customer's view of an IETF Network Slice is an abstract one.

Thus, there is a need to create logical network structures with required characteristics. The customer of such a logical network can require a degree of isolation and performance that previously might not have been satisfied by overlay VPNs. Additionally, the IETF Network Slice customer might ask for some level of control of their virtual networks, e.g., to customize the service paths in a network slice.

This document specifies definitions and a framework for the provision of an IETF Network Slice service. Section 6 briefly indicates some candidate technologies for realizing IETF Network Slices.

2. Terms and Abbreviations

The following abbreviations are used in this document.

*NSC: Network Slice Controller

*SLA: Service Level Agreement

*SLI: Service Level Indicator

*SLO: Service Level Objective

The meaning of these abbreviations is defined in greater details in the remainder of this document.

2.1. Core Terminology

The following terms are presented here to give context. Other terminology is defined in the remainder of this document.

Customer: A customer is the requester of an IETF Network Slice service. Customers may request monitoring of SLOs. A customer may be an entity such as an enterprise network or a network operator, an individual working at such an entity, a private individual contracting for a service, or an application or software component. A customer may be an external party (classically a paying customer) or a division of a network operator that uses the service provided by another division of the same operator.

Other terms that have been applied to the customer role are "client" and "consumer".

Provider: A provider is the organization that delivers an IETF Network Slice service. A provider is the network operator that controls the network resources used to construct the network slice (that is, the network that is sliced). The provider's network maybe a physical network or may be a virtual network supplied by another service provider.

Customer Edge (CE): The customer device that provides connectivity to a service provider. Examples include routers, Ethernet switches, firewalls, 4G/5G RAN or Core nodes, application accelerators, server load balancers, HTTP header enrichment functions, and PEPs (Performance Enhancing Proxy). In some circumstances CEs are provided to the customer and managed by the provider.

Provider Edge (PE): The device within the provider network to which a CE is attached. A CE may be attached to multiple PEs, and multiple CEs may be attached to a given PE.

Attachment Circuit (AC): A channel connecting a CE and a PE over which packets that belong to an IETF Network Slice service are exchanged. An AC is, by definition, technology specific: that is, the AC defines how customer traffic is presented to the provider network. The customer and provider agree (through configuration) on which values in which combination of layer 2 and layer 3 header and payload fields within a packet identify to which {IETF Network Slice service, connectivity construct, and SLOs/SLEs} that packet is assigned. The customer and provider may agree on a per {IETF Network Slice service, connectivity construct, and SLOs/SLEs} basis to police or shape traffic on the AC in both the ingress (CE to PE) direction and egress (PE to CE) direction, This ensures that the traffic is within the capacity profile that is agreed in an IETF Network Slice service. Excess traffic is dropped by default, unless specific out-of-profile policies are agreed between the customer and the provider. As described in Section 4.2 the AC may be part of the IETF Network Slice service or may be external to it.

Service Demarcation Point (SDP): The point at which an IETF Network Slice service is delivered by a service provider to a customer. Depending on the service delivery model (see Section 4.2 this may be a CE or a PE, and could be a device, a software component, or in the case of network functions virtualization (for example), be an abstract function supported within the provider's network. Each SDP must have a unique identifier (e.g., an IP address or

MAC address) within a given IETF Network Slice service and may use the same identifier in multiple IETF Network Slice services.

An SDP may be abstracted as a Service Attachment Point (SAP) [I-D.ietf-opsawg-sap] for the purpose generalizing the concept across multiple service types and representing it in management and configuration systems.

Connectivity Construct: A set of SDPs together with a communication type that defines how traffic flows between the SDPs. An IETF Network Slice service is specified in terms of a set of SDPs, the associated connectivity constructs and the service objectives that the customer wishes to see fulfilled.

3. IETF Network Slice Objectives

IETF Network Slices are created to meet specific requirements, typically expressed as bandwidth, latency, latency variation, and other desired or required characteristics. Creation of an IETF Network Slice is initiated by a management system or other application used to specify network-related conditions for particular traffic flows in response to an actual or logical IETF Network Sliceservice request.

Once created, these slices can be monitored, modified, deleted, and otherwise managed.

Applications and components will be able to use these IETF Network Slices to move packets between the specified end-points of the service in accordance with specified characteristics.

3.1. Definition and Scope of IETF Network Slice

An IETF Network Slice service enables connectivity between a set of Service Demarcation Points (SDPs) with specific Service Level Objectives (SLOs) and Service Level Expectations (SLEs) over a common underlay network.

An IETF Network Slice combines the connectivity resource requirements and associated network capabilities such as bandwidth, latency, jitter, and network functions with other resource behaviors such as compute and storage availability. The definition of an IETF Network Slice is independent of the connectivity and technologies used in the underlay network. This allows an IETF Network Slice service customer 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 [TS23501].

An IETF Network Slice service is agnostic to the technology of the underlying network, and its realization may be selected based upon multiple considerations including its service requirements and the capabilities of the underlay network.

The term "Slice" refers to a set of characteristics and behaviors that differntiate one type of user-traffic from another. An IETF Network Slice assumes that an underlay 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 the SLOs/SLEs to specific flows or to all of the traffic in the slice.

3.2. IETF Network Slice Service

A service provider delivers an IETF Network Slice service for a customer. The IETF Network Slice service is specified in terms of a set of SDPs, a set of one or more connectivity constructs between subsets of these SDPs, and a set of SLOs and SLEs for each SDP sending to each connectivity construct. A communication type (point-to-point (P2P), point-to-multipoint (P2MP), or any-to-any (A2A)) is specified for each connectivity construct. That is, in a given IETF Network Slice service there may be one or more connectivity constructs of the same or different type, each connectivity construct may be between a different subset of SDPs, for a given connectivity construct each sending SDP has its own set of SLOs and SLEs, and the SLOs and SLEs in each set may be different. Note that a service provider may decide how many connectivity constructs per IETF Network Slice service it wishes to support such that an IETF Network Slice service may be limited to one connectivity construct or may support many.

This approach results in the following possible connectivity constructs:

* For a P2P connectivity construct, there is one sending SDP and one receiving SDP. This construct is like a private wire or a tunnel. All traffic injected at the sending SDP is intended to be received by the receiving SDP. The SLOs and SLEs apply at the sender (and implicitly at the receiver).

* For a P2MP connectivity construct, there is only one sending SDP and more than one receiving SDP. This is like a P2MP tunnel or multi-access VLAN segment. All traffic from the sending SDP is

intended to be received by all the receiving SDPs. There is one set of SLOs and SLEs that applies at the sending SDP (and implicitly at all receiving SDPs).

*With an A2A connectivity construct, any sending SDP may send to any one receiving SDP or any set of receiving SDPs in the construct. There is an implicit level of routing in this connectivity construct that is not present in the other connectivity constructs because the provider's network must determine to which receiving SDPs to deliver each packet. This construct may be used to support P2P traffic between any pair of SDPs, or to support multicast or broadcast traffic from one SDP to a set of other SDPs. In the latter case, whether the service is delivered using multicast within the provider's network or using "ingress replication" or some other means is out of scope of the specification of the service. A service provider may choose to support A2A constructs, but to limit the traffic to unicast.

The SLOs/SLEs in an A2A connectivity construct apply to individual sending SDPs regardless of the receiving SDPs, and there is no linkage between sender and receiver in the specification of the connectivity construct. A sending SDP may be "disappointed" if the receiver is over-subscribed. If a customer wants to be more specific about different behaviors from one SDP to another SDP, they should use P2P connectivity constructs.

A customer traffic flow may be unicast or multicast, and various network realizations are possible:

*Unicast traffic may be mapped to a P2P connectivity construct for direct delivery, or to an A2A connectivity construct for the service provider to perform routing to the destination SDP. It would be unusual to use a P2MP connectivity construct to deliver unicast traffic because all receiving SDPs would get a copy, but this can still be done if the receivers are capable of dropping the unwanted traffic.

*A bidirectional unicast service can be constructed by specifying two P2P connectivity constructs. An additional SLE may specify fate-sharing in this case.

*Multicast traffic may be mapped to a set of P2P connectivity constructs, a single P2MP connectivity construct, or a mixture of P2P and P2MP connectivity constructs. Multicast may also be supported by an A2A connectivity construct. The choice clearly influences how and where traffic is replicated in the network. With a P2MP or A2A connectivity construct, it is the operator's choice whether to realize the construct with ingress replication,

multicast in the core, P2MP tunnels, or hub-and-spoke. This choice should not change how the customer perceives the service.

* The concept of a multipoint-to-point (MP2P) service can be realized with multiple P2P connectivity constructs. Note that, in this case, the egress may simultaneously receive traffic from all ingresses. The SLOs at the sending SDPs must be set with this in mind because the provider's network is not capable of coordinating the policing of traffic across multiple distinct source SDPs. It is assumed that the customer, requesting SLOs for the various P2P connectivity constructs, is aware of the capabilities of the receiving SDP. If the receiver receives more traffic than it can handle, it may drop some and introduce queuing delays.

* The concept of a multipoint-to-multipoint (MP2MP) service can best be realized using a set of P2MP connectivity constructs, but could be delivered over an A2A connectivity construct if each sender is using multicast. As with MP2P, the customer is assumed to be familiar with the capabilities of all receivers. A customer may wish to achieve an MP2MP service using a hub-and-spoke architecture where they control the hub: that is, the hub may be an SDP or an ancillary SDP (see Section 3.2.1) and the service may be achieved by using a set of P2P connectivity constructs to the hub, and a single P2MP connectivity construct from the hub.

From the above, it can be seen that the SLOs of the senders define the SLOs for the receivers on any connectivity construct. That is, and in particular, the network may be expected to handle the traffic volume from a sender to all destinations. This extends to all connectivity constructs in an IETF Network Slice service.

Note that the realization of an IETF Network Slice service does not need to map the connectivity constructs one-to-one onto underlying network constructs (such as tunnels, etc.). The service provided to the customer is distinct from how the provider decides to deliver that service.

If a CE has multiple attachment circuits to a PE within a given IETF Network Slice service and they are operating in single-active mode, then all traffic between the CE and its attached PEs transits a single attachment circuit; if they are operating in in all-active mode, then traffic between the CE and its attached PEs is distributed across all of the active attachment circuits.

A given sending SDP may be part of multiple connectivity constructs within a single IETF Network Slice service, and the SDP may have different SLOs and SLEs for each connectivity construct to which it is sending. Note that a given sending SDP's SLOs and SLEs for a

given connectivity construct apply between it and each of the receiving SDPs for that connectivity construct.

An IETF Network Slice service provider may freely make a deployment choice as to whether to offer a 1:1 relationship between IETF Network Slice service and connectivity construct, or to support multiple connectivity constructs in a single IETF Network Slice service. In the former case, the provider might need to deliver multiple IETF Network Slice services to achieve the function of the second case.

It should be noted that per Section 9 of [RFC4364] an IETF Network Slice service customer may actually provide IETF Network Slice services to other customers in a mode sometimes referred to as "carrier's carrier". In this case, the underlying IETF Network Slice service provider may be owned and operated by the same or a different provider network. As noted in Section 4.3, network slices may be composed hierarchically or serially.

Section 4.2 provides a description of endpoints in the context of IETF network slicing. These are known as Service Demarcation Points (SDPs). For a given IETF Network Slice service, the customer and provider agree, on a per-SDP basis which end of the attachment circuit provides the SDP (i.e., whether the attachment circuit is inside or outside the IETF Network Slice service). This determines whether the attachment circuit is subject to the set of SLOs and SLEs at the specific SDP.

3.2.1. Ancillary SDPs

It may be the case that the set of SDPs needs to be supplemented with additional senders or receivers. An additional sender could be, for example, an IPTV or DNS server either within the provider's network or attached to it, while an extra receiver could be, for example, a node reachable via the Internet. This is modelled as a set of ancillary SDPs which supplement the other SDPs in one or more connectivity constructs, or which have their own connectivity constructs. Note that an ancillary SDP can either have a resolvable address, e.g., an IP address or MAC address, or the SDP may be a placeholder, e.g., IPTV or DNS server, which is resolved within the provider's network when the IETF Network Slice service is instantiated.

4. IETF Network Slice System Characteristics

The following subsections describe the characteristics of IETF Network Slices in addition to the list of SDPs, the connectivity constructs, and the technology of the ACs.

4.1. Objectives for IETF Network Slices

An IETF Network Slice service is defined in terms of quantifiable characteristics known as Service Level Objectives (SLOs) and unquantifiable characteristics known as Service Level Expectations (SLEs). SLOs are expressed in terms Service Level Indicators (SLIs), and together with the SLEs form the contractual agreement between service customer and service provider known as a Service Level Agreement (SLA).

The terms are defined as follows:

- 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 customer can determine whether the provider is meeting the SLOs by performing measurements on the traffic.

- A Service Level Expectation (SLE) is an expression of an unmeasurable service-related request that a customer of an IETF Network Slice makes of the provider. An SLE is distinct from an SLO because the customer may have little or no way of determining whether the SLE is being met, but they still contract with the provider for a service that meets the expectation.

- A Service Level Agreement (SLA) is an explicit or implicit contract between the customer of an IETF Network Slice service and the provider of the slice. The SLA is expressed in terms of a set of SLOs and SLEs that are to be applied for a given connectivity construct between a sending SDP and the set of receiving SDPs, and may describe the extent to which divergence from individual SLOs and SLEs can be tolerated, and commercial terms as well as any consequences for violating these SLOs and SLEs.

4.1.1. Service Level Objectives

SLOs define a set of measurable network attributes and characteristics that describe an IETF Network Slice service. SLOs do not describe how an IETF Network Slice service is 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 service may include multiple connection constructs that associate sets of endpoints (SDPs). SLOs apply to a given connectivity construct and apply to a specific direction of traffic flow. That is, they apply to a specific sending SDP and the connection to specific receiving SDPs.

The SLOs are combined with Service Level Expectations in an SLA.

4.1.1.1. Some Common SLOs

SLOs can be described as 'Directly Measurable Objectives': they are always measurable. See Section 4.1.2 for the description of Service Level Expectations which are unmeasurable service-related requests sometimes known as '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'. Future specifications (such as IETF Network Slice service YANG models) may precisely define these SLOs, and other SLOs may be introduced as described in Section 4.1.1.2.

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). [RFC7679] discusses one-way metrics.

Maximum Permissible Delay Variation: Packet delay variation (PDV) as defined by [RFC3393], is 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 connectivity construct is available in accordance with all of the SLOs associated with it.

4.1.1.2. Other Service Level Objectives

Additional SLOs may be defined to provide additional description of the IETF Network Slice service that a customer requests. These would be specified in further documents.

If the IETF Network Slice 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).

4.1.2. Service Level Expectations

SLEs define a set of network attributes and characteristics that describe an IETF Network Slice service, but which are not directly measurable by the customer. Even though the delivery of an SLE cannot usually be determined by the customer, the SLEs form an important part of the contract between customer and provider.

Quite often, an SLE will imply some details of how an IETF Network Slice service is realized by the provider, although most aspects of the implementation in the underlying network layers remain a free choice for the provider.

SLEs may be seen as aspirational on the part of the customer, and they are expressed as behaviors that the provider is expected to apply to the network resources used to deliver the IETF Network Slice service. The SLEs are combined with SLOs in an SLA.

An IETF Network Slice service may include multiple connection constructs that associate sets of endpoints (SDPs). SLEs apply to a given connectivity construct and apply to specific directions of traffic flow. That is, they apply to a specific sending SDP and the connection to specific receiving SDPs. However, being more general in nature than SLOs, SLEs may commonly be applied to all connection constructs in an IETF Network Slice service.

4.1.2.1. Some Common SLEs

SLEs can be described as 'Indirectly Measurable Objectives': they are not generally directly measurable by the customer.

Security, geographic restrictions, maximum occupancy level, and isolation are example SLEs as follows.

Security: A customer may request that the provider applies encryption or other security techniques to traffic flowing between SDPs of a connectivity construct within an IETF Network Slice service. For example, the customer could request that only

network links that have MACsec [MACsec] enabled are used to realize the connectivity construct.

This SLE may include a request for encryption (e.g., [RFC4303]) between the two SDPs explicitly to meet the architectural recommendations in [TS33.210] or for compliance with [HIPAA] or [PCI].

Whether or not the provider has met this SLE is generally not directly observable by the customer and cannot be measured as a quantifiable metric.

Please see further discussion on security in Section 9.

Geographic Restrictions: A customer may request that certain geographic limits are applied to how the provider routes traffic for the IETF Network Slice service. For example, the customer may have a preference that its traffic does not pass through a particular country for political or security reasons.

Whether or not the provider has met this SLE is generally not directly observable by the customer and cannot be measured as a quantifiable metric.

Maximal Occupancy Level: The maximal occupancy level specifies the number of flows to be admitted and optionally a maximum number of countable resource units (e.g., IP or MAC addresses) an IETF Network Slice service can consume. Since an IETF Network Slice service may include multiple connection constructs, this SLE should also say whether it applies for the entire IETF Network Slice service, for group of connections, or on a per connection basis.

Again, a customer may not be able to fully determine whether this SLE is being met by the provider.

Isolation: As described in Section 7, a customer may request that its traffic within its IETF Network Slice service is isolated from the effects of other network services supported by the same provider. That is, if another service exceeds capacity or has a burst of traffic, the customer's IETF Network Slice service should remain unaffected and there should be no noticeable change to the quality of traffic delivered.

In general, a customer cannot tell whether a service provider is meeting this SLE. They cannot tell whether the variation of an SLI is because of changes in the underlying network or because of interference from other services carried by the network. If the service varies within the allowed bounds of the SLOs, there may be no noticeable indication that this SLE has been violated.

Diversity:

A customer may request that traffic on the connection between one set of SDPs should use different network resources from the traffic between another set of SDPs. This might be done to enhance the availability of the connectivity constructs within an IETF Network Slice service.

While availability is a measurable objective (see Section 4.1.1.1) this SLE requests a finer grade of control and is not directly measurable (although the customer might become suspicious if two connections fail at the same time).

4.2. IETF Network Slice Service Demarcation Points

As noted in Section 3.1, an IETF Network Slice is a logical network topology connecting a number of endpoints. Section 3.2 goes on to describe how the IETF Network Slice service is composed of a set of one or more connectivity constructs that describe connectivity between the Service Demarcation Points (SDPs) across the underlying network.

The characteristics of IETF Network Slice (SDPs are as follows.

* SDPs are conceptual points of connection to an IETF Network Slice. As such, they serve as the IETF Network Slice ingress/egress points.

* Each SDP maps to a device, application, or a network function, such as (but not limited to) routers, switches, interfaces/ports, firewalls, WAN, 4G/5G RAN nodes, 4G/5G Core nodes, application accelerators, server load balancers, NAT44 [RFC3022], NAT64 [RFC6146], HTTP header enrichment functions, and Performance Enhancing Proxies (PEPs) [RFC3135].

* An SDP is identified by a unique identifier in the context of an IETF Network Slice customer.

* Each SDP is associated with a set of provider-scope identifiers such as IP addresses, encapsulation-specific identifiers (e.g., VLAN tag, MPLS Label), interface/port numbers, node ID, etc.

* SDPs are mapped to endpoints of services/tunnels/paths within the IETF Network Slice during its initialization and realization.

  - A combination of the SDP identifier and SDP provider-network-scope identifiers define an SDP in the context of the Network Slice Controller (NSC) (see Section 5.3).

  - The NSC will use the SDP provider-network-scope identifiers as part of the process of realizing the IETF Network Slice.

For a given IETF Network Slice service, the IETF Network Slice customer and provider agree where the endpoint (i.e., the service demarcation point) is located. This determines what resources at the edge of the network form part of the IETF Network Slice and are subject to the set of SLOs and SLEs for a specific endpoint.

Figure 1 shows different potential scopes of an IETF Network Slice that are consistent with the different SDP locations. For the purpose of this discussion and without loss of generality, the figure shows customer edge (CE) and provider edge (PE) nodes connected by attachment circuits (ACs). Notes after the figure give some explanations.

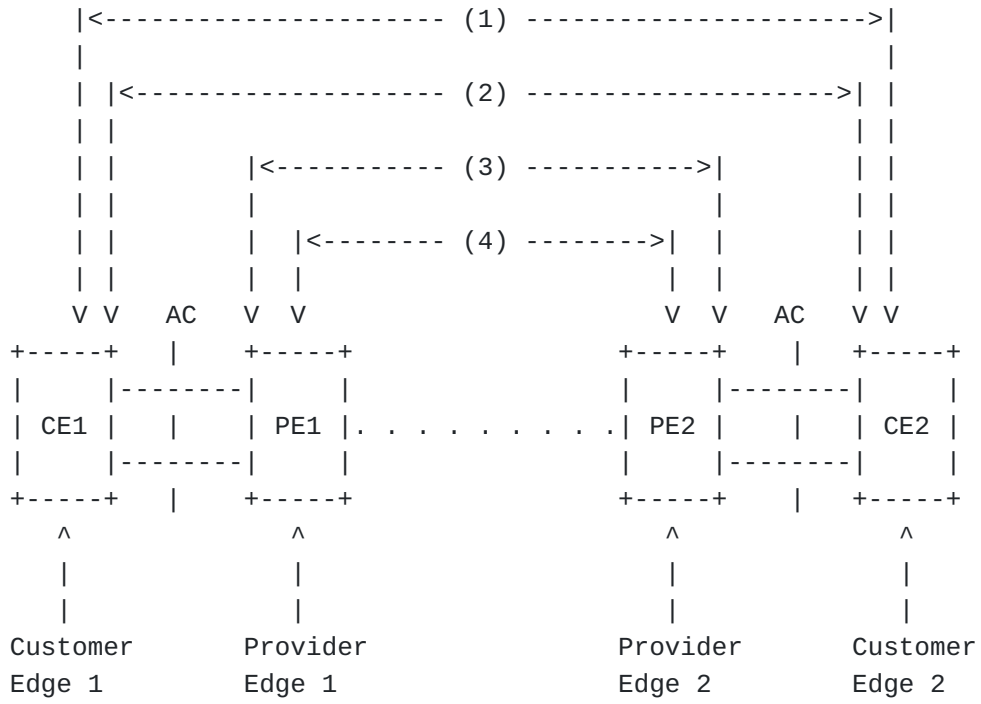

Figure 1: Positioning IETF Service Demarcation Points

Explanatory notes for Figure 1 are as follows:

1. If the CE is operated by the IETF Network Slice service provider, then the edge of the IETF Network Slice may be within the CE. In this case the slicing process may utilize resources from within the CE such as buffers and queues on the outgoing interfaces.

2. The IETF Network Slice may be extended as far as the CE, to include the AC, but not to include any part of the CE. In this case, the CE may be operated by the customer or the provider.

Slicing the resources on the AC may require the use of traffic tagging (such as through Ethernet VLAN tags) or may require traffic policing at the AC link ends.

3. In another model, the SDPs of the IETF Network Slice are the customer-facing ports on the PEs. This case can be managed in a way that is similar to a port-based VPN: each port (AC) or virtual port (e.g., VLAN tag) identifies the IETF Network Slice and maps to an IETF Network Slice SDP.

4. Finally, the SDP may be within the PE. In this mode, the PE classifies the traffic coming from the AC according to information (such as the source and destination IP addresses, payload protocol and port numbers, etc.) in order to place it onto an IETF Network Slice.

The choice of which of these options to apply is entirely up to the network operator. It may limit or enable the provisioning of particular managed services and the operator will want to consider how they want to manage CEs and what control they wish to offer the customer over AC resources.

Note that Figure 1 shows a symmetrical positioning of SDPs, but this decision can be taken on a per-SDP basis through agreement between the customer and provider.

In practice, it may be necessary to map traffic not only onto an IETF Network Slice, but also onto a specific connectivity construct if the IETF Network Slice supports more than one with a source at the specific SDP. The mechanism used will be one of the mechanisms described above, dependent on how the SDP is realized.

Finally, note (as described in Section 2.1) that an SDP is an abstract endpoint of an IETF Network Slice service and as such may be a device, interface, or software component and may, in the case of network functions virtualization (for example), be an abstract function supported within the provider's network.

4.3. IETF Network Slice Decomposition

Operationally, an IETF Network Slice may be composed of two or more IETF Network Slices as specified below. Decomposed network slices are 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. Framework

A number of IETF Network Slice services will typically be provided over a shared underlying network infrastructure. Each IETF Network Slice consists of both the overlay connectivity and a specific set of dedicated network resources and/or functions allocated in a shared underlay network to satisfy the needs of the IETF Network Slice customer. In at least some examples of underlying network technologies, the integration between the overlay and various underlay resources is needed to ensure the guaranteed performance requested for different IETF Network Slices.

5.1. IETF Network Slice Stakeholders

An IETF Network Slice and its realization involves the following stakeholders. The IETF Network Slice customer and IETF Network Slice provider (see Section 2.1) are also stakeholders.

Orchestrator: An orchestrator is an entity that composes different services, resource, and network requirements. It interfaces with the IETF NSC when composing a complex service such as an end-to-end network slice.

IETF Network Slice Controller (NSC): The NSC realizes an IETF Network Slice in the underlying network, and maintains and monitors the run-time state of resources and topologies associated with it. A well-defined interface is needed to support interworking between different NSC implementations and different orchestrator implementations.

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

5.2. Expressing Connectivity Intents

An IETF Network Slice customer communicates with the NSC using the IETF Network Slice Service Interface.

An IETF Network Slice customer may be a network operator who, in turn, use the IETF Network Slice to provide a service for another IETF Network Slice customer.

Using the IETF Network Slice Service Interface, a customer expresses requirements for a particular slice by specifying what is required rather than how that is to be achieved. That is, the customer's view of a slice is an abstract one. Customers normally have limited (or no) visibility into the provider network's actual topology and resource availability information.

This should be true even if both the customer and provider are associated with a single administrative domain, in order to reduce the potential for adverse interactions between IETF Network Slice customers and other users of the underlay network infrastructure.

The benefits of this model can include the following.

- Security: The underlay network components are less exposed to attack because the underlay network (or network operator) does not need to expose network details (topology, capacity, etc.) to the IETF Network Slice customers.

- Layered Implementation: The underlay network comprises network elements that belong to a different layer network than customer applications. Network information (advertisements, protocols, etc.) that a customer cannot interpret or respond to is not exposed to the customer. (Note - a customer should not use network information not exposed via the IETF Network Slice Service Interface, even if that information is available.)

- Scalability: Customers do not need to know any information beyond that which is exposed via the IETF Network Slice Service Interface.

The general issues of abstraction in a TE network are described more fully in [RFC7926].

This framework document does not assume any particular technology layer at which IETF Network Slices operate. A number of layers (including virtual L2, Ethernet or, IP connectivity) could be employed.

Data models and interfaces are needed to set up IETF Network Slices, and specific interfaces may have capabilities that allow creation of slices within specific technology layers.

Layered virtual connections are comprehensively discussed in other IETF documents. See, for instance, GMPLS-based networks [RFC5212] and [RFC4397], or Abstraction and Control of TE Networks (ACTN) [RFC8453] and [RFC8454]. The principles and mechanisms associated with layered networking are applicable to IETF Network Slices.

There are several IETF-defined mechanisms for expressing the need for a desired logical network. The IETF Network Slice Service Interface carries data either in a protocol-defined format, or in a formalism associated with a modeling language.

For instance:

* The Path Computation Element (PCE) Communication Protocol (PCEP) [RFC5440] and GMPLS User-Network Interface (UNI) using RSVP-TE [RFC4208] use a TLV-based binary encoding to transmit data.

* The Network Configuration Protocol (NETCONF) [RFC6241] and RESTCONF Protocol [RFC8040] use XML and JSON encoding.

* gRPC/GNMI [I-D.openconfig-rtgwg-gnmi-spec] uses a binary encoded programmable interface. ProtoBufs can be used to model gRPC and GNMI data.

* For data modeling, YANG ([RFC6020] and [RFC7950]) may be used to model configuration and other data for NETCONF, RESTCONF, and GNMI, among others.

While several generic formats and data models for specific purposes exist, it is expected that IETF Network Slice management may require enhancement or augmentation of existing data models. Further, it is possible that mechanisms will be needed to determine the feasibility of service requests before they are actually made.

5.3. IETF Network Slice Controller (NSC)

The IETF NSC takes abstract requests for IETF Network Slices and implements them using a suitable underlying technology. An IETF NSC is the key component for control and management of the IETF Network Slice. It provides the creation/modification/deletion, monitoring and optimization of IETF Network Slices in a multi-domain, a multi-technology and multi-vendor environment.

The main task of the IETF NSC is to map abstract IETF Network Slice requirements to concrete technologies and establish required connectivity ensuring that resources are allocated to the IETF Network Slice as necessary.

The IETF Network Slice Service Interface is used for communicating details of an IETF Network Slice (configuration, selected policies, operational state, etc.), as well as information about status and performance of the IETF Network Slice. The details for this IETF Network Slice Service Interface are not in scope for this document.

The controller provides the following functions.

* Provides an IETF Network Slice Service Interface for creation/modification/deletion of the IETF Network Slices that is agnostic to the technology of the underlying network. The API exposed by this interface communicates the Service Demarcation Points of the IETF Network Slice, IETF Network Slice SLO/SLE parameters (and possibly monitoring thresholds), applicable input selection (filtering) and various policies, and provides a way to monitor the slice.

* Determines an abstract topology connecting the SDPs of the IETF Network Slice that meets criteria specified via the IETF Network Slice Service Interface. The NSC also retains information about the mapping of this abstract topology to underlying components of the IETF Network Slice as necessary to monitor IETF Network Slice status and performance.

* Provides "Mapping Functions" for the realization of IETF Network Slices. In other words, it will use the mapping functions that:

  - map IETF Network Slice Service Interface requests that are agnostic to the technology of the underlying network to technology-specific network configuration interfaces.

  - map filtering/selection information as necessary to entities in the underlay network.

* The controller collects telemetry data (e.g., OAM results, statistics, states, etc.) via a network configuration interface for all elements in the abstract topology used to realize the IETF Network Slice.

* Evaluates the current performance against IETF Network Slice SLO parameters using the telemetry data from the underlying realization of an IETF Network Slice (i.e., services/paths/tunnels). Exposes this performance to the IETF Network Slice customer via the IETF Network Slice Service Interface. The IETF Network Slice Service Interface may also include the capability to provide notifications if the IETF Network Slice performance reaches threshold values defined by the IETF Network Slice customer.

5.3.1. 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 2).

IETF Network Slice Service Interface: The IETF Network Slice Service Interface is an interface between a customer's higher level operation system (e.g., a network slice orchestrator or a customer network management system) and the NSC. It is agnostic to the technology of the underlying network. The customer can use this interface to communicate the requested characteristics and other requirements for the IETF Network Slice, and the NSC can use the interface to report the operational state of an IETF Network Slice to the customer.

Network Configuration Interface: The Network Configuration Interface is an interface between the NSC and network controllers. It is technology-specific and may be built around the many network models already defined within the IETF.

These interfaces can be considered in the context of the Service Model and Network Model described in [RFC8309] and, together with the Device Configuration Interface used by the Network Controllers, provides a consistent view of service delivery and realization.

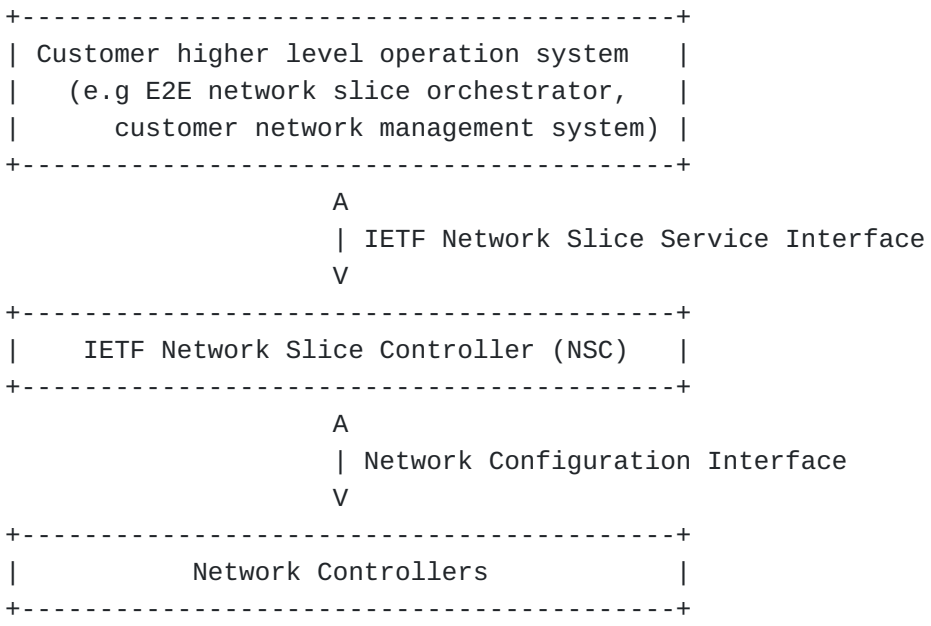

Figure 2: Interfaces of the IETF Network Slice Controller

5.3.1.1. IETF Network Slice Service Interface

The IETF Network Slice Controller provides an IETF Network Slice Service Interface that allows customers to request and monitor IETF Network Slices. Customers operate on abstract IETF Network Slices, with details related to their realization hidden.

The IETF Network Slice Service Interface is also independent of the type of network functions or services that need to be connected, i.e., it is independent of any specific storage, software, protocol, or platform used to realize physical or virtual network connectivity or functions in support of IETF Network Slices.

The IETF Network Slice Service Interface uses protocol mechanisms and information passed over those mechanisms to convey desired attributes for IETF Network Slices and their status. The information is expected to be represented as a well-defined data model, and should include at least SDP and connectivity information, SLO/SLE specification, and status information.

5.3.2. Management Architecture

The management architecture described in Figure 2 may be further decomposed as shown in Figure 3. This should also be seen in the context of the component architecture shown in Figure 5 and corresponds to the architecture in [RFC8309].

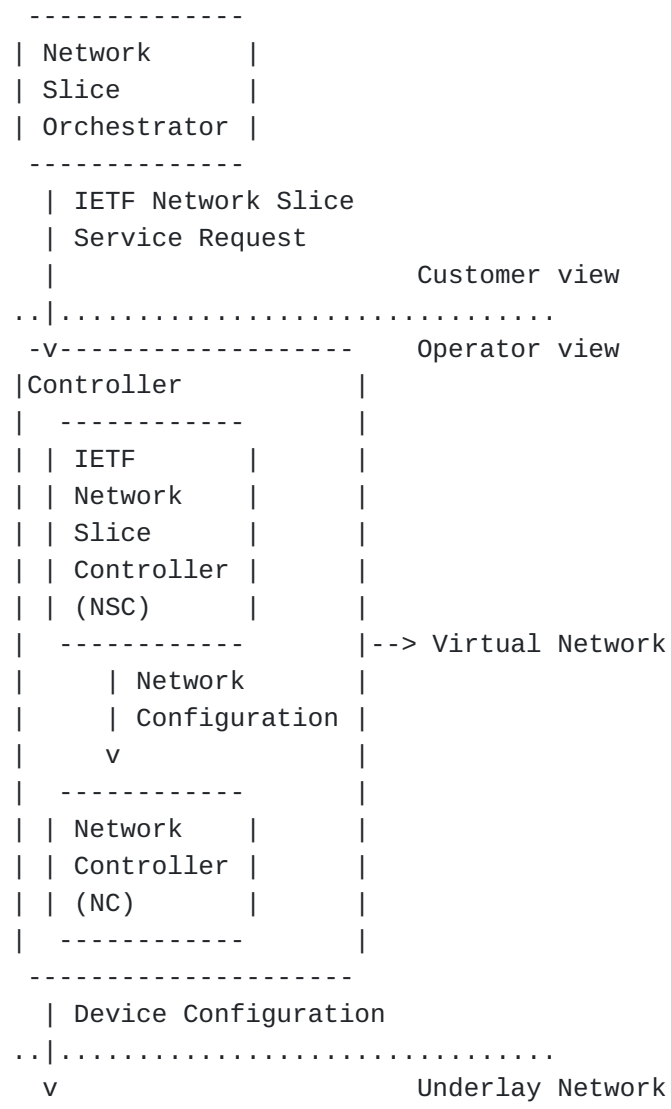

Figure 3: Interface of IETF Network Slice Management Architecture

5.4. IETF Network Slice Structure

An IETF Network Slice is a set of connection constructs between various SDPs to form a logical network that meets the SLOs agreed upon.

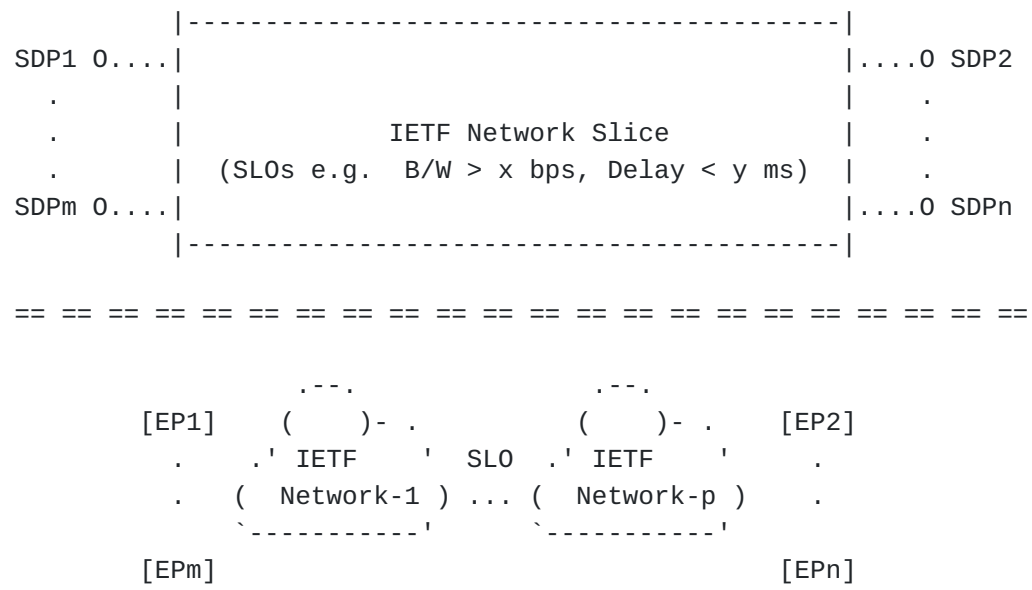

Figure 4: IETF Network Slice

Figure 4 illustrates this by showing a case where an IETF Network Slice provides connectivity between a set of SDP pairs (SDP1 to SDP2, ..., SDPm to SDPn) in a set of P2P connectivity constructs each 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 service/tunnel/path Endpoints (EP1 ro EPn) in the underlay network. Also, the SDPs in the same IETF Network Slice may belong to the same or different address spaces.

6. Realizing IETF Network Slices

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 Network Configuration Interface. However, this section provides an overview of the components and processes involved in realizing an IETF Network Slice.

The realization can be achieved in a form of either physical or logical connectivity using VPNs, virtual networks (VNs), or a variety of tunneling technologies such as Segment Routing, MPLS, etc. Accordingly, SDPs may be realized as physical or logical service or network functions.

6.1. Architecture to Realize IETF Network Slices

The architecture described in this section is deliberately at a high level. It is not intended to be prescriptive: implementations and technical solutions may vary freely. However, this approach provides a common framework that other documents may reference in order to facilitate a shared understanding of the work.

Figure 5 shows the architectural components of a network managed to provide IETF Network Slices. The customer's view is of individual IETF Network Slices with their SDPs, and connectivity constructs. Requests for IETF Network Slices are delivered to the NSC.

The figure shows, without loss of generality, the CEs, ACs, and PEs, that exist in the network. The SDPs are not shown and can be placed in any of the ways described in Section 4.2.

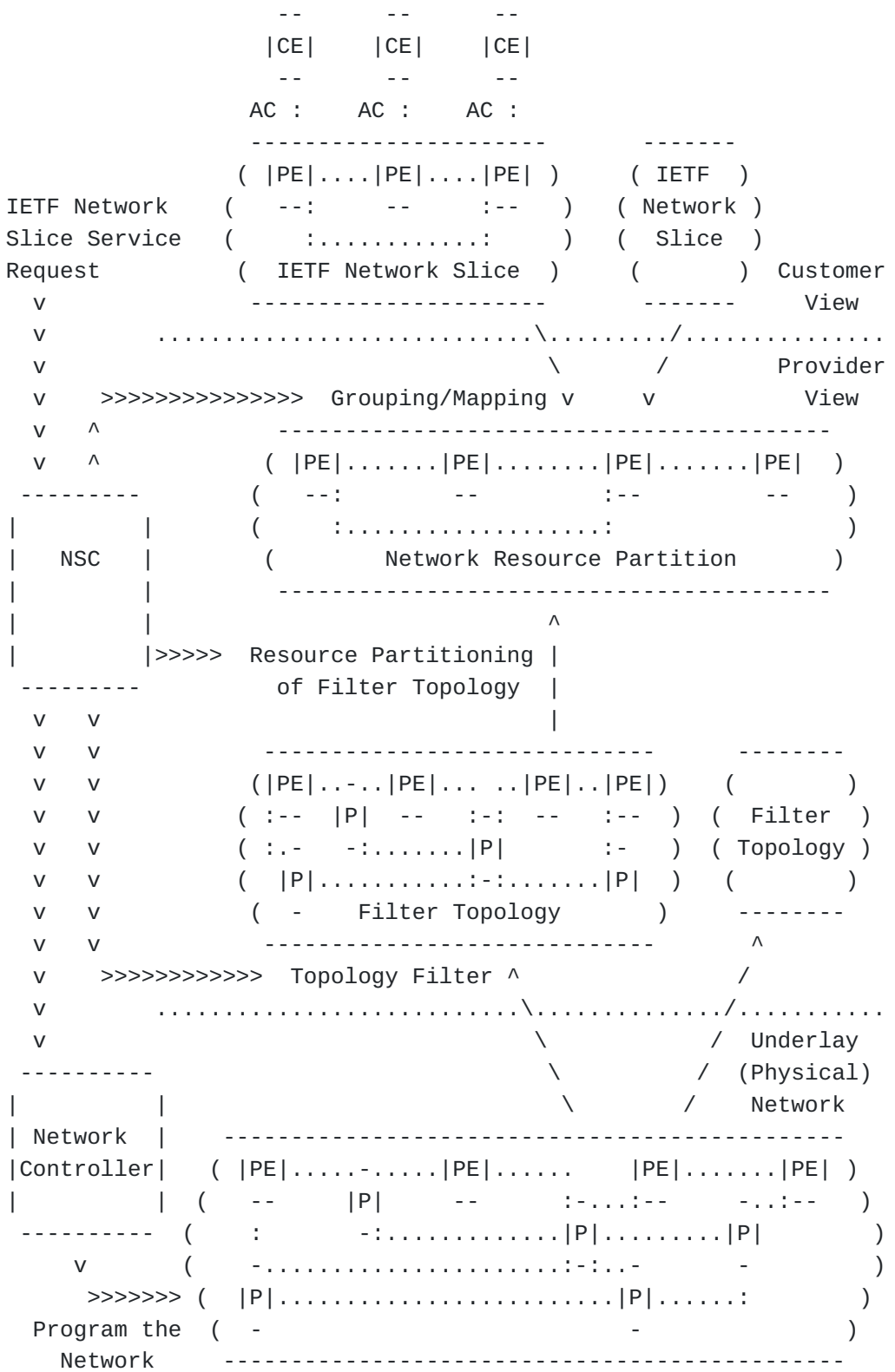

Figure 5: Architecture of an IETF Network Slice

The network itself (at the bottom of the figure) comprises an underlay network. This could be a physical network, but may be a virtual network. The underlay network is provisioned through network controllers that may utilize device controllers [RFC8309].

The underlay network may optionally be filtered or customized by the network operator to produce a number of network topologies that we call Filter Topologies. Customization is just a way of selecting specific resources (e.g., nodes and links) from the underlay network according to their capabilities and connectivity in the underlay network. These actions are configuration options or operator policies. The resulting topologies can be used as candidates to host IETF Network Slices and provide a useful way for the network operator to know in advance that all of the resources they are using to plan an IETF Network Slice would be able to meet specific SLOs and SLEs. The creation of a Filter Topology could be an offline planning activity or could be performed dynamically as new demands arise. The use of Filter Topologies is entirely optional in the architecture, and IETF Network Slices could be hosted directly on the underlay network.

Recall that an IETF Network Slice is a service requested by / provided for the customer. The IETF Network Slice service is expressed in terms of one or more connectivity constructs. An implementation or operator is free to limit the number of connectivity constructs in a slice to exactly one. Each connectivity construct is associated within the IETF Network Slice service request with a set of SLOs and SLEs. The set of SLOs and SLEs does not need to be the same for every connectivity construct in the slice, but an implementation or operator is free to require that all connectivity constructs in a slice have the same set of SLOs and SLEs.

One or more connectivity constructs from one or more slices are mapped to a set of network resources called a Network Resource Partition (NRP). A single connectivity construct is mapped to only one NRP (that is, the relationship is many to one). An NRP may be chosen to support a specific connectivity construct because of its ability to support a specific set of SLOs and SLEs, or its ability to support particular connectivity types, or for any administrative or operational reason. An implementation or operator is free to map each connectivity construct to a separate NRP, although there may be scaling implications depending on the solution implemented. Thus, the connectivity constructs from one slice may be mapped to one or more NRPs. By implication from the above, an implementation or operator is free to map all the connectivity constructs in a slice to a single NRP, and to not share that NRP with connectivity constructs from another slice.

An NRP is simply a collection of resources identified in the underlay network. Thus, the NRP is a scoped view of a topology and may be considered as a topology in its own right. The process of determining the NRP may be made easier if the underlay network topology is first filtered into a Filter Topology in order to be aware of the subset of network resources that are suitable for specific NRPs, but this is optional.

The steps described here can be applied in a variety of orders according to implementation and deployment preferences. Furthermore, the steps may be iterative so that the components are continually refined and modified as network conditions change and as service requests are received or relinquished, and even the underlay network could be extended if necessary to meet the customers' demands.

6.2. Procedures to Realize IETF Network Slices

There are a number of different technologies that can be used in the underlay, including physical connections, MPLS, time-sensitive networking (TSN), Flex-E, etc.

An IETF Network Slice can be realized in a network, using specific underlying technology or technologies. The creation of a new IETF Network Slice will be realized with following steps:

* The NSC exposes the network slicing capabilities that it offers for the network it manages.

* The customer may issue a request to determine whether a specific IETF Network Slice could be supported by the network. The NSC may respond indicating a simple yes or no, and may supplement a negative response with information about what it could support were the customer to change some requirements.

* The customer requests an IETF Network Slice. The NSC may respond that the slice has or has not been created, and may supplement a negative response with information about what it could support were the customer to change some requirements.

* When processing a customer request for an IETF Network Slice, the NSC maps the request to the network capabilities and applies provider policies before creating or supplementing the NRP.

Regardless of how IETF Network Slice is realized in the network (i.e., using tunnels of different types), the definition of the IETF Network Slice service does not change at all. The only difference is how the slice is realized. The following sections briefly introduce how some existing architectural approaches can be applied to realize IETF Network Slices.

6.3. Applicability of ACTN to IETF Network Slices

Abstraction and Control of TE Networks (ACTN - [RFC8453]) is a management architecture and toolkit used to create virtual networks (VNs) on top of a TE underlay network. The VNs can be presented to customers for them to operate as private networks.

In many ways, the function of ACTN is similar to IETF network slicing. Customer requests for connectivity-based overlay services are mapped to dedicated or shared resources in the underlay network in a way that meets customer guarantees for service level objectives and for separation from other customers' traffic. [RFC8453] describes the function of ACTN as collecting resources to establish a logically dedicated virtual network over one or more TE networks. Thus, in the case of a TE-enabled underlay network, the ACTN VN can be used as a basis to realize IETF network slicing.

While the ACTN framework is a generic VN framework that can be used for VN services beyond the IETF Network Slice, it also a suitable basis for delivering and realizing IETF Network Slices.

Further discussion of the applicability of ACTN to IETF Network Slices including a discussion of the relevant YANG models can be found in [I-D.ietf-teas-applicability-actn-slicing].

6.4. Applicability of Enhanced VPNs to IETF Network Slices

An enhanced VPN (VPN+) is designed to support the needs of new applications, particularly applications that are associated with 5G services, by utilizing an approach that is based on existing VPN and TE technologies and adds characteristics that specific services require over and above VPNs as they have previously been specified.

An enhanced VPN can be used to provide enhanced connectivity services between customer sites and can be used to create the infrastructure to underpin a network slicing service.

It is envisaged that enhanced VPNs will be delivered using a combination of existing, modified, and new networking technologies.

[I-D.ietf-teas-enhanced-vpn] describes the framework for Enhanced Virtual Private Network (VPN+) services.

6.5. Network Slicing and Aggregation in IP/MPLS Networks

Network slicing provides the ability to partition a physical network into multiple isolated logical networks of varying sizes, structures, and functions so that each slice can be dedicated to specific services or customers.

Many approaches are currently being worked on to support IETF Network Slices in IP and MPLS networks with or without the use of Segment Routing. Most of these approaches utilize a way of marking packets so that network nodes can apply specific routing and forwarding behaviors to packets that belong to different IETF Network Slices. Different mechanisms for marking packets have been proposed (including using MPLS labels and Segment Routing segment IDs) and those mechanisms are agnostic to the path control technology used within the underlay network.

These approaches are also sensitive to the scaling concerns of supporting a large number of IETF Network Slices within a single IP or MPLS network, and so offer ways to aggregate the connectivity constructs of slices (or whole slices) so that the packet markings indicate an aggregate or grouping where all of the packets are subject to the same routing and forwarding behavior.

At this stage, it is inappropriate to mention any of these proposed solutions that are currently work in progress and not yet adopted as IETF work.

7. Isolation in IETF Network Slices

7.1. Isolation as a Service Requirement

An IETF Network Slice customer may request that the IETF Network Slice delivered to them is such that changes to other IETF Network Slices or to other services do not have any negative impact on the delivery of the IETF Network Slice. The IETF Network Slice customer may specify the degree to which their IETF Network Slice is unaffected by changes in the provider network or by the behavior of other IETF Network Slice customers. The customer may express this via an SLE it agrees with the provider. This concept is termed 'isolation'.

In general, a customer cannot tell whether a service provider is meeting an isolation SLE. If the service varies such that an SLO is breached then the customer will become aware of the problem, and if the service varies within the allowed bounds of the SLOs, there may be no noticeable indication that this SLE has been violated.

7.2. Isolation in IETF Network Slice Realization

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 of 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 IETF Network Slices.

8. Management Considerations

IETF Network Slice realization needs to be instrumented in order to track how it is working, and it might be necessary to modify the IETF Network Slice as requirements change. Dynamic reconfiguration might be needed.

The various management interfaces and components are discussed in Section 5.

9. 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 customer-defined IETF Network Slices will be mapped to their realization in the underlay networks. Underlay networks will require capabilities to conform to customer's requests as some aspects of security may be expressed in SLEs.

IETF NSC authentication: Underlying networks need to be protected against the attacks from an adversary NSC as this could destabilize overall network operations. An IETF Network Slice may span across different networks, therefore, the NSC should have strong authentication with each of these networks. Furthermore, both the IETF Network Slice Service Interface and the Network Configuration Interface 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 customer 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 customer is responsible for end-to-end encryption of its own traffic.

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

IETF Network Slices might use underlying virtualized networking. All types of virtual networking require special consideration to be given to the separation of traffic between distinct virtual networks, as well as some degree of protection from effects of traffic use of underlying network (and other) resources from other virtual networks sharing those resources.

For example, if a service requires a specific upper bound of latency, then that service can be degraded by added delay in transmission of service packets caused by the activities of another service or application using the same resources.

Similarly, in a network with virtual functions, noticeably impeding access to a function used by another IETF Network Slice (for instance, compute resources) can be just as service-degrading as delaying physical transmission of associated packet in the network.

While an IETF Network Slice might include encryption and other security features as part of the service, customers might be well advised to take responsibility for their own security needs, possibly by encrypting traffic before hand-off to a service provider.

10. Privacy Considerations

Privacy of IETF Network Slice service customers must be preserved. It should not be possible for one IETF Network Slice customer to discover the presence of other customers, nor should sites that are members of one IETF Network Slice be visible outside the context of that IETF Network Slice.

In this sense, it is of paramount importance that the system use the privacy protection mechanism defined for the specific underlying technologies that support the slice, including in particular those mechanisms designed to preclude acquiring identifying information associated with any IETF Network Slice customer.

11. IANA Considerations

This document makes no requests for IANA action.

12. 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.ietf-opsawg-sap]
Boucadair, M., Dios, O. G. D., Barguil, S., Wu, Q., and V. Lopez, "A Network YANG Model for Service Attachment Points (SAPs)", Work in Progress, Internet-Draft, draft-ietf-opsawg-sap-02, 22 February 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-opsawg-sap-02>.

[I-D.ietf-teas-applicability-actn-slicing] King, D., Drake, J., Zheng, H., and A. Farrel, "Applicability of Abstraction and Control of Traffic Engineered Networks (ACTN) to Network Slicing", Work in Progress, Internet-Draft, draft-ietf-teas-applicability-actn-slicing-00, 21 September 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-applicability-actn-slicing-00>.

[I-D.ietf-teas-enhanced-vpn] Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A Framework for Enhanced Virtual Private Network (VPN+) Services", Work in Progress, Internet-Draft, draft-ietf-teas-enhanced-vpn-09, 25 October 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-enhanced-vpn-09>.

[I-D.openconfig-rtgwg-gnmi-spec]
Shakir, R., Shaikh, A., Borman, P., Hines, M., Lebsack, C., and C. Morrow, "gRPC Network Management Interface (gNMI)", Work in Progress, Internet-Draft, draft-openconfig-rtgwg-gnmi-spec-01, 5 March 2018, <https://datatracker.ietf.org/doc/html/draft-openconfig-rtgwg-gnmi-spec-01>.

[MACsec] IEEE, "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Security", 2018, <https://1.ieee802.org/security/802-1ae>.

[NGMN-NS-Concept] NGMN Alliance, "Description of Network Slicing Concept", https://www.ngmn.org/uploads/media/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf , 2016.

[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>.

[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>.

[RFC3135] Border, J., Kojo, M., Griner, J., Montenegro, G., and Z. Shelby, "Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations", RFC 3135, DOI 10.17487/RFC3135, June 2001, <https://www.rfc-editor.org/info/rfc3135>.

[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>.

[RFC4208] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, DOI 10.17487/RFC4208, October 2005, <https://www.rfc-editor.org/info/rfc4208>.

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

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4397] Bryskin, I. and A. Farrel, "A Lexicography for the Interpretation of Generalized Multiprotocol Label Switching (GMPLS) Terminology within the Context of the ITU-T's Automatically Switched Optical Network (ASON) Architecture", RFC 4397, DOI 10.17487/RFC4397, February 2006, <https://www.rfc-editor.org/info/rfc4397>.

[RFC5212] Shiomoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, DOI 10.17487/RFC5212, July 2008, <https://www.rfc-editor.org/info/rfc5212>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/info/rfc5440>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020,

DOI 10.17487/RFC6020, October 2010, <https://www.rfc-editor.org/info/rfc6020>.

[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>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[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>.

[RFC7926] Farrel, A., Ed., Drake, J., Bitar, N., Swallow, G., Ceccarelli, D., and X. Zhang, "Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks", BCP 206, RFC 7926, DOI 10.17487/RFC7926, July 2016, <https://www.rfc-editor.org/info/rfc7926>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/info/rfc7950>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[RFC8309] Wu, Q., Liu, W., and A. Farrel, "Service Models Explained", RFC 8309, DOI 10.17487/RFC8309, January 2018, <https://www.rfc-editor.org/info/rfc8309>.

[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>.

[RFC8454] Lee, Y., Belotti, S., Dhody, D., Ceccarelli, D., and B. Yoon, "Information Model for Abstraction and Control of TE Networks (ACTN)", RFC 8454, DOI 10.17487/RFC8454,

September 2018, <https://www.rfc-editor.org/info/rfc8454>.

[TS23501] 3GPP, "System architecture for the 5G System (5GS)", 3GPP TS 23.501, 2019.

[TS28530] 3GPP, "Management and orchestration; Concepts, use cases and requirements", 3GPP TS 28.530, 2019.

[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>.

Acknowledgments

The entire TEAS Network Slicing design team and everyone participating in related discussions has contributed to this document. Some text fragments in the document have been copied from the [I-D.ietf-teas-enhanced-vpn], for which we are grateful.

Significant contributions to this document were gratefully received from the contributing authors listed in the "Contributors" section. In addition we would like to also thank those others who have attended one or more of the design team meetings, including the following people not listed elsewhere:

*Aihua Guo

*Bo Wu

*Greg Mirsky

*Lou Berger

*Rakesh Gandhi

*Ran Chen

*Sergio Belotti

*Stewart Bryant

*Tomonobu Niwa

*Xuesong Geng

Further useful comments were received from Daniele Ceccarelli, Uma Chunduri, Pavan Beeram, Tarek Saad, Kenichi Ogaki, Oscar Gonzalez de

Dios, Xiaobing Niu, Dan Voyer, Igor Bryskin, Luay Jalil, Joel Halpern, John Scudder, and John Mullooly.

This work is partially supported by the European Commission under Horizon 2020 grant agreement number 101015857 Secured autonomic traffic management for a Tera of SDN flows (Teraflow).

Contributors

The following authors contributed significantly to this document:

Eric Gray (The original editor of the foundation documents)
Independent
Email: ewgray@graiymage.com

Jari Arkko
Ericsson
Email: jari.arkko@piuha.net

Mohamed Boucadair
Orange
Email: mohamed.boucadair@orange.com

Dhruv Dhody
Huawei, India
Email: dhruv.ietf@gmail.com

Jie Dong
Huawei
Email: jie.dong@huawei.com

Xufeng Liu
Volta Networks
Email: xufeng.liu.ietf@gmail.com

Authors' Addresses

Adrian Farrel (editor)
Old Dog Consulting
United Kingdom

Email: adrian@olddog.co.uk

John Drake (editor)
Juniper Networks
United States of America

Email: jdrake@juniper.net

Reza Rokui
Ciena

Email: rrokui@ciena.com

Shunsuke Homma
NTT
Japan

Email: shunsuke.homma.ietf@gmail.com

Kiran Makhijani
Futurewei
United States of America

Email: kiranm@futurewei.com

Luis M. Contreras
Telefonica
Spain

Email: luismiguel.contrerasmurillo@telefonica.com

Jeff Tantsura
Microsoft Inc.

Email: jefftant.ietf@gmail.com