

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: January 4, 2018

H. Flinck  
C. Sartori  
A. Andrianov  
C. Mannweiler  
N. Sprecher  
Nokia  
July 3, 2017

Network Slicing Management and Orchestration  
draft-flinck-slicing-management-00

Abstract

Network Slicing is worked in multiple SDOs from different view points. As network slicing is an end-to-end topic, this draft proposes that network slices architecture [NS-Framework] aligns with the work done in NGMN, 3GPP and ETSI with relation to management and orchestration. The key aspect that this draft makes is the rational for role and need for Network Slice Management Function (NSMF) entity that operates above Network Virtualization Function Orchestrator and PNFs Management Functions. NSMF needs to support different abstractions of resources and to offer access to different management entities.

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 <http://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 January 4, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	2
1.1. Acronyms and Abbreviations . . . . .	5
2. Different levels of Network Slice Control exposure . . . . .	5
3. Network Slice Management Function (NSMF) . . . . .	6
4. IANA considerations . . . . .	8
5. Security considerations . . . . .	8
6. Acknowledgements . . . . .	8
7. Informative References . . . . .	8
Authors' Addresses . . . . .	9

## 1. Introduction

The purpose of this draft is to highlight the essential aspects of network slice management from 3GPP, NGMN and ETSI relevant for the network slices architecture as described in [NS-Framework] and to propose a minimal alignment between these works to ensure compatibility between them. NGMN documents "161010\_NGMN\_Network\_Slicing\_framework\_v1.0.8" [NGMN\_NS] and "5G Network and Service Management including Orchestration" [NGMN\_NSMN] define Network Slicing and how it relates to overall Service and Network Management architecture. The NGMN documents define as well the terminology adopted later by 3GPP and reflected in 3GPP [TR28.801]. In this paper, for sake of simplicity, only an "executive summary" of network slicing is given, while relying on both terminology and complete descriptions on the above mentioned documents.

Network Slicing provides multiple logical networks on top of a partially shared network infrastructure as described in [NS-Framework]. Each instance of a network slice represents an independent end-to-end network that allows deployment of different architectural flavors in parallel slices. These slices may be deployed and/or operated by the slice provider, or by the tenant who requested the slice.

A network slice can span across different administrative domains. NGMN Network slicing white paper [NS-Framework] defines various forward-looking business models engaging multiple administrative domains that may be envisioned in the industry. An administrative domain refers to the scope of jurisdiction of a provider. A provider may obtain services from 3rd parties (i.e. sub-providers) to enrich the services it provides to its end customers. A provider could also benefit from offering its spare capabilities or resources to other providers becoming itself a sub-provider. A network service can be a single user connectivity service, NaaS (Network as a Service) such as a service instance, a network slice instance or a subnetwork slice (note NGMN and 3GPP use a different terminology for what IETF netslices drafts call for "network slice segment") instance offering for a business vertical that utilizes forward-looking business models, or IaaS (Infra structure as a Service).

Depending on the use cases and type of services for which the end-to-end slice has been instantiated multiple levels of control may be exposed to the tenants by the slice provider. On the lowest level of the exposed control the network slice provider grants only access to use the slice and means to monitor its performance. At second level a control exposure is to allow tenants to change the configuration of the network functions associated to the tenant's network slice. At the highest level of control tenants can compose network slices and manage them with their own management system. These different levels of control exposure require that the network slice management must work on multiple levels of abstractions where highest level is at the Service Management & Orchestration (M&O) and lowest level at the network functions. The slice provider must be able to isolate these control functions of different tenants to match the "Slice Provider" - "Slice Consumer" -relationship.

A network slice instance can contain virtualized network functions as well as physical network functions. Virtualized network functions (VNF) are decoupled from physical network equipment by a virtualization layer. Both the lifecycle of the types of the network functions can span beyond the lifecycle of a Network Slice and they need their own life cycle management functions. The life cycle management of these two types of network functions differ. The environment in which VNFs are deployed is called Network Functions Virtualisation Infrastructure (NFVI) and is managed by Virtualised Infrastructure Manager (VIM) according to ETSI NFV-MANO [MANO] reference architecture. VNFs are instantiated by requests of NFV Orchestrator (NFVO). In the MANO architecture NFV Orchestrator (NFVO) uses VNF Managers for the lifecycle management of VNF instances and the VIM allocates the needed virtualized resources as requested by the NFVO into the NFVI. However, the same approach cannot be applied to network functions of dedicated hardware

(Physical Network Functions, PNF) as their resources are not controlled by NFVO nor VIMS. Network Functions (whether PNF or VNF) require their function specific management, as well as their resource management.

When adding support for the virtualized version of the PNFs their management systems will evolve to either extend their capability with an embedded VNF management functionality or will delegate their virtual resource management to an external VNF manager. In either case, the VNF management function interacts with the NFVO and the VIM through the MANO defined interfaces and provides the cloud resource FCAPS management for the network functions. Another key issue for provisioning of network slices is the identification, design, and management of network functions which can be shared by multiple end-to-end slices [Rost].

For Network slice management function (NSMF), which is a slice-dedicated function with slice-specific view on any FCAPS data and management procedures, such sharing or common usage should be transparent, i.e., the multiplexing of multiple network slices to a commonly used function/element is done by EMS/NMS. NSMF operates above NFVO and PNFs Management Functions in the Service M&O. In view of 3GPP as well as ETSI NFV, NSMF belongs to OSS/BSS. When a network slice contains PNFs the NSMF instructs the PNFs Management Functions to configure the physical network components to deliver the required slice characteristics.

This draft introduces the role of NSMF in the context of 3GPP [TR28.801], [TS28.530] and ETSI [MANO] work and reflects that back to netslices-architecture presented in [NS-Framework]. We argue that the NSMF is at the Service M&O level, even at a tenant. This is because of several reasons:

- o Need for exposing different levels of network slice control to the tenants.
- o Different life cycle management approaches for PNFs and VNFs. NSMF must have interfaces both to NFVO and to PNFs Management and is therefore above of the NFVO and PNF management and it should support service level abstractions.

Network slicing is end-to-end concept, thus including several network components, (Network Slice Subnetwork Functions according to 3GPP terminology). Often those components belong to different administrative domains (e.g. RAN, Core Network, Transport) and therefore the need for a higher level of abstraction. Transport network [ACTN] is a subnetwork slice in the 3GPP model and recursion can be applied to slices as well as to subnetwork slices.

### 1.1. Acronyms and Abbreviations

This document uses the following acronyms:

3GPP	3rd Generation Partnership Project
BSS/OSS	Business Support Systems/Operations Support Systems
EMS	Element Management System
ETSI	European Telecommunications Standards Institute
FCAPS	Fault, Configuration, Accounting, Performance, Security
IaaS	Infra structure as a Service
KPI	Key Performance Indicator
MANO	ETSI Management and Orchestration
LCM	Life Cycle Management
MNO	Mobile Network Operator
M&O	Management & Orchestration
NaaS	Network as a Service
NGMN	Next Generation Mobile Networks
NMS	Network Management System
NSMF	Network Slice Subnet Management Functions
NSSMF	Network Slice Management Function
NFVI	Network Functions Virtualisation Infrastructure
NVFO	Network Virtualization Function Orchestrator
PNF	Physical Network Function
RAN	Radio Access Network
SLA	Service Level Agreement
VIM	Virtualised Infrastructure Manager
VNF	Virtualised Network Function

### 2. Different levels of Network Slice Control exposure

Depending on the "Slice Provider" - "Slice Consumer" -relationship the Slice Provider can offer various levels of control to the Slice Consumers. Roughly speaking levels of control can be categorized onto follow cases:

1. Monitoring only. The Slice Provider offers only means to monitor the slice KPIs as agreed in the contract. Network slice configuration is chosen from a catalogue of readymade slice templates. Accesses via dashboard-like web service and/or north bound interfaces provided by the Slice Provider.
2. Limited control to Slice Consumer to perform design and composition of network slice. Slice Consumer can change configuration of deployed network functions and /or onboard own certified network functions into Slice Provider's repository using interfaces provided by the Slice Provider.

3. Extended Control. In this case the Slice Consumer deploys and operates the network slice using its own MANO stack and NMS. The Slice consumer has tight control over its own network functions and services while has limited control over MNO network functions.

Because of these varying levels of network slice control, the NSMF needs to support different abstractions of resources and to offer access to different management entities (e.g. PNFs management functions, NFV-MANO). Consequently, the logical place for NSMF function in the network slice management architecture is at the Service Management & Orchestration (M&O).

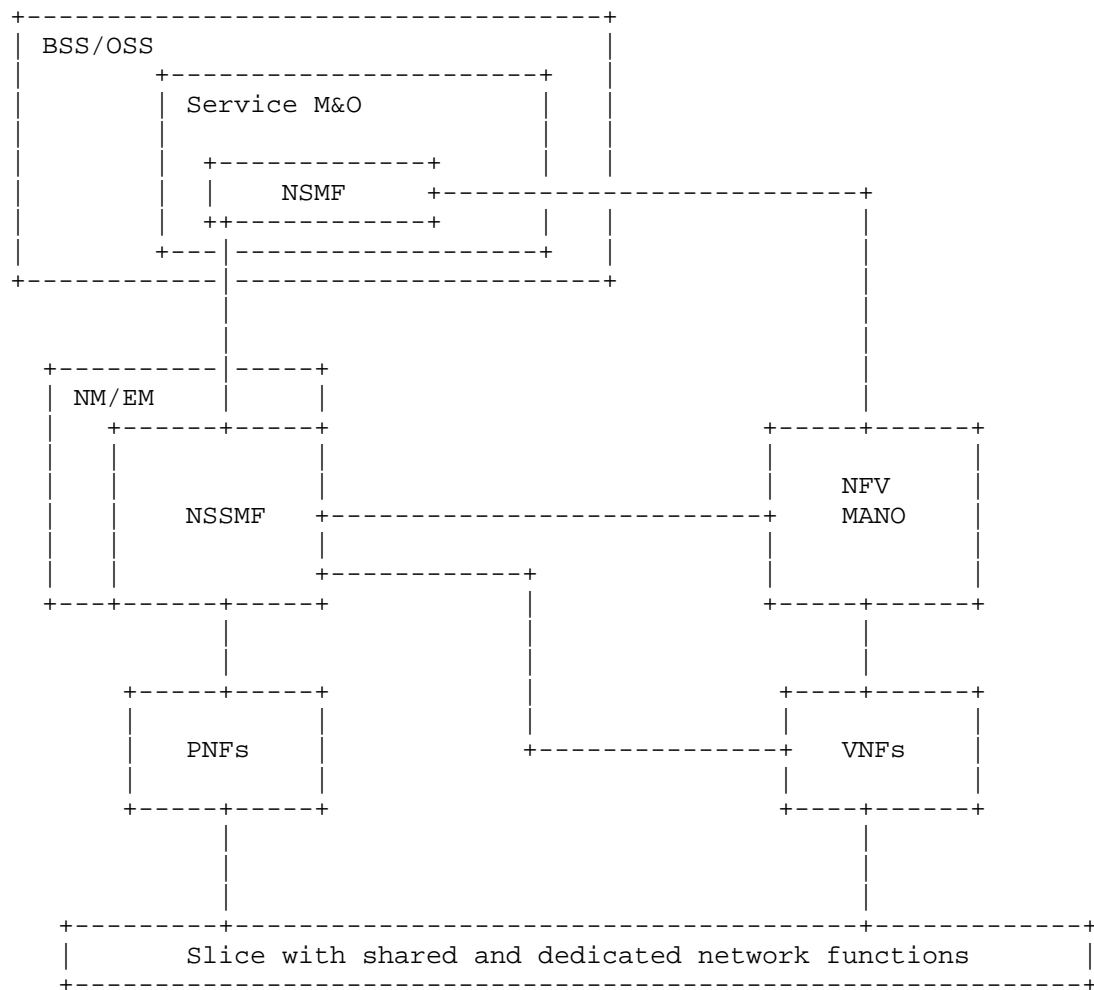
### 3. Network Slice Management Function (NSMF)

Network slicing concept of NGMN consists of 3 layers: Service Instance Layer, Network Slice Instance Layer, and Resource layer [NGMN\_NS]. The Service Instance Layer is managed by service orchestrator that is considered to be part of BSS/OSS according to the 3GPP view [TR28.801]. Network Slicing Instance Layer is a Business to Business service and may pass across multiple administrative domains. Network Slice Management Function resides at this layer and is consequently part of Service Orchestration and BSS/OSS.

The end-to-end network slice management (NSMF) can use different technology domains and their segments to create an end-to-end slice. It has full visibility and control to the end-to-end slice and its performance. It resides above the Network Slice Subnet Management Functions (NSSMF). It monitors slice specific FCAPS to maintain and to expose the overall SLAs of the end-to-end slices to the tenant.

NSMF interfaces domain specific Network Management and Element Management Systems through Network Slice Subnet Management Functions (NSSMF). In addition, NSMF also interfaces NFV-MANO to manage virtualization aspects (through "OS-Ma-nfvo"-interface).

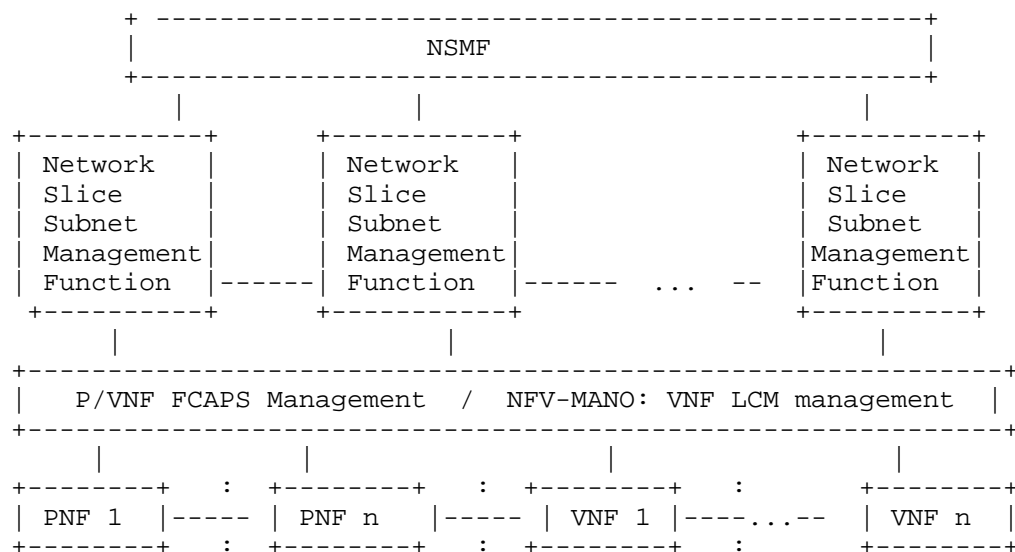
NSSMF manages Network Slice Subnet (3GPP defined management abstraction) composed of Network Functions (virtualized or not) and other Network Slice Subnets (recursion principle). NM/EM could play the role of NSSMF. For management of virtualization aspects (such as NS and VNF LCM) and TN, NSSMF interacts with NFV-MANO (through "Os-Ma-nfvo"-interface). The 3GPP defined Network Slice Subnets correspond to ETSI NFV defined NSs composed from either network functions and/or nested network slices (recursion principle).



Network Slice Management functional architecture.

Figure 1

Based on the above reasoning we propose to replace the "Figure 2: E2E Slice Orchestration"-figure of the section of Management and Orchestration of Network Slicing in [NS-Framework] with the following figure with the above stated reasoning.



Network Slice Management Function (Network Slice segment term corresponds roughly to Network Slice subnetwork term used by 3GPP/ NGMN)

Figure 2

#### 4. IANA considerations

This document makes no request of IANA.

#### 5. Security considerations

Each element and their interface of the proposed management architecture needs to address their security requirements.

#### 6. Acknowledgements

#### 7. Informative References

- [ACTN] Ceccarelli, D. and Lee, Y., "Framework for Abstraction and Control of Traffic Engineered Networks", draft-ietf-teas-actn-framework-06 (work in progress), June 2017.
- [MANO] ETSI, "ETSI GS NFV-MAN 001: Network Functions Virtualization (NFV); Management and Orchestration", 2014.



- [NGMN\_NS] NGMN Alliance, "Description of Network Slicing Concept", [https://www.ngmn.org/uploads/media/161010\\_NGMN\\_Network\\_Slicing\\_framework\\_v1.0.8.pdf](https://www.ngmn.org/uploads/media/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf) , 2016.
- [NGMN\_NSMN] NGMN Alliance, "5G Network and Service Management including Orchestration", <https://www.ngmn.org/publications/all-downloads/article/5g-network-and-service-management-including-orchestration.html> , 2017.
- [NS-Framework] Geng, L., Dong, J., Bryant, S., Makhijani, K., Galis, A., De Foy, X., and Kuklinsk, S., "Network Slicing Architecture", draft-geng-netslices-architecture-01 (work in progress), June 2017.
- [Rost] Rost, P., Mannweiler, C., Diomidis, M., Sartori, C., Sciancalepore, V., Sastry, N., Holland, O., Tayade, S., Han, B., Bega, D., Aziz, D., Bakker, H., and IEEE Communications Magazine, Volume: 55 Issue: 5,, "Network Slicing to Enable Scalability and Flexibility in 5G Mobile Networks", May 2017.
- [TR28.801] 3GPP, "Study on management and orchestration of network slicing for next generation network, Release 14)3GPP TR 28.801 V1.2.0", <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3091> , 2017.
- [TS28.530] 3GPP, "Management of network slicing in mobile networks; Concepts, use cases and requirements. Technical specification. Release 15. 3GPP TR 28.530.", <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3091> , 2017.

## Authors' Addresses

Hannu Flinck  
Nokia  
Espoo  
FI

Phone: +358504839522  
Email: [hannu.flinck@nokia.com](mailto:hannu.flinck@nokia.com)

Cinzia Sartori  
Nokia  
Munich  
DE

Phone: +491713008990  
Email: cinzia.sartori@nokia-bell-labs.com

Anatoly Andriannov  
Nokia  
Arlington Heights, IL  
US

Phone: +1-847-668-0394  
Email: anatoly.andrianov@nokia.com

Christian Mannweiler  
Nokia  
Munich  
DE

Phone: +491715581581  
Email: christian.mannweiler@nokia-bell-labs.com

Nurit Sprecher  
Nokia  
Hod HaSharon  
IL

Phone: +97297751229  
Email: nurit.sprecher@nokia.com

No Working Group  
Internet-Draft  
Intended Status: Informational  
Expires: January 3, 2018

A. Galis (editor)  
University College London  
et al.  
July 3, 2017

Network Slicing - Revised Problem Statement  
draft-galis-netslices-revised-problem-statement-01

Abstract

This document introduces Network Slicing problems and the motivation for new work areas. It represents an initial revision of the Network Slicing problem statement derived from the analysis of the technical gaps in IETF protocols ecosystem. It complements and brings together the efforts being carried out in several other IETF working groups to achieve certain aspects of Network Slicing functions and operations.

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 <http://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 January 2, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

## Table of Contents

1	Introduction	3
1.1	Network Slicing Value Characteristics	4
1.2	Network Slicing Work Scope	5
1.3	Notes	7
2.	Network Slicing - Selected Problems and Work Areas	7
2.1	Global Issues - Problems (GP)	7
2.1.1	Problem GI1: Uniform Reference Model (***)	7
2.1.2	Problem GI2: Requirements for operations and interactions (**)	8
2.1.3	Problem GI3: Slice Templates (***)	9
2.2	Network Slice Capabilities - Problems (NSC)	10
2.2.1	Problem NSC1: Guarantees for isolation (***)	10
2.2.2	Problem NSC2: Fulfilling diverse requirements (*)	10
2.2.3	Problem NSC3: Efficiency in slicing (*)	11
2.2.4	Problem NSC4: Slice Recursion (**)	11
2.2.5	Problem NSC5: Customized security mechanisms per slice (*)	12
2.2.6	Problem NSC6: flexibility and efficiency trade-offs (*)	12
2.2.7	Problem NSC7: Optimisation (**)	12
2.2.8	Problem NSC8: Monitoring and Discovery (**)	13
2.2.9	Problem NSC9: Capability exposure and APIs (***)	13
2.2.10	Problem NSC10: Programmability of Operations of Network Slices (**)	14
2.3	Network Slice Operations - Problems (NSO)	15
2.3.1	Problem NSO1: Slice life cycle management (***)	15
2.3.2	Problem NSO2: Autonomic slice management and operation (**)	15
2.3.3	Problem NSO3 - Slice stitching / composition (***)	16
2.3.4	Problem NSO4: E2E Network Orchestration (***)	17
2.3.5	Problem NSO5: Service Mapping in a single domain and Cross-Domain Coordination (***)	18
2.3.6	Problem NSO6: Roles in Network Slicing (*)	18
2.3.7	Problem NSO7: Efficient enablers and methods for integration (*)	19
2.4	Notes	20
3.	OAM Operation with Customized Granularity	21
4	Summary	22
5	Security Considerations	23
6	IANA Considerations	23
7	Acknowledgements	23
8	References	23
7.1	IETF References	23
7.2	Informative References	26

Authors' Addresses . . . . .	27
------------------------------	----

## 1 Introduction

Network slicing (NS) is an approach of flexible isolation and allocation of network resources and network functions for a network instance, providing high level of customization and quality service guarantee. It transform the networking perspective by abstracting, isolating, orchestrating, softwarizing, and separating logical network components from the underlying physical network supporting the introduction of new network architectures ([RFC1958], [RFC3439], [RFC3234]) and new service delivery [5G-ICN]. In general, a particular network slice consists of a union of subsets of (connectivity, storage, computing) resources & (Virtual) Network Functions & Service functions [RFC7665] at the data & control & management planes at a given time that are managed together to provide a logical networking infrastructure in support of a variety of services.

Network slicing enables at a given time the dynamic creation of multiple, parallel sub-networks of different features by flexible isolation of allocated to a slice network resources and network functions and providing high level of customization and quality guarantee.

The management plane allocates the grouping of network resources (whereby network resources can be physical, virtual or a combination thereof), it connects with the physical and virtual network and service functions ([SFC WG]) as appropriate, and it instantiates all of the network and service functions assigned to the slice. On the other hand, for slice operations, the slice control plane functionality that may be operated by slice tenant, takes over the control and governing of all the network resources, network functions, and service functions assigned to the slice. It (re-) configures them as appropriate and as per elasticity needs, in order to provide an end-to-end service. In particular, slice ingress routers are configured so that appropriate traffic is bound to the relevant slice.

Allocation of traffic flows to slices as may be based on simple rules (relying on a subset of the transport coordinate, DSCP/traffic class, or flow label), or may be a more sophisticated one (to be further defined) such as enabling new slice specific constructs in the data plane. Also, the flows to slice allocation rules that are specified for a slice can be changed dynamically, based on some events (e.g. triggered by a service request). The slice control plane is responsible for instructing the involved elements to honor such

needs.

Network operators can use NS to enable different services to receive different treatment and to allow the allocation and release of network resources according to the context and contention policy of the operators. Such an approach using NS would allow significant reduction of the operations expenditure. In addition, there is an enabling link between NS and softwarization. On one hand NS makes possible softwarization, programmability ([RFC7149]), and the innovation necessary to enrich the offered services. On the other hand Network softwarization techniques [IMT2020-2015], [IMT2020-2016] may be used to realise and manage [MANO-2014] network slicing. NS provides the means for the network operators to provide network programmable capabilities to both service providers and other market players without changing their physical infrastructure. NS enables the concurrent deployment of multiple logical, self-contained and independent, shared or partitioned networks on a common infrastructure. Slices may support dynamic multiple services, multi-tenancy, and the integration means for vertical market players (e.g. automotive industry, energy industry, healthcare industry, media and entertainment industry, etc.)

Please refer to [I-D.geng-netslices-architecture] for related terminologies and definitions.

### 1.1 Network Slicing Value Characteristics

As a differentiation from non-partition networks and those with simple partitions of connectivity resources (e.g. VPNs)/ Virtual Networks/Other abstractions of the data traffic layer, the following key value-added characteristics of Network Slicing and corresponding usage are identified:

- \* Network Slice is a dedicated network that is build on an infrastructure mainly composed of, but not limited to, connectivity, storage and computing.
- \* Each network slice has the ability to dynamically expose and possibly negotiate the parameters that characterize an NS.
- \* Each network slice will have its own operator that sees it as a complete network infrastructure (i.e. router instances, programmability, using any appropriate communication protocol, caches, provide dynamic placement of virtual network functions according to traffic patterns, to use its own controller, finally it can manage its network as its own).
- \* Network slicing support tenants that are strongly independent on infrastructure.
- \* A Network Slicing aware infrastructure allows operators to use part of the network resources to meet stringent resource

requirements

- \* Network slicing introduces an additional layer of abstraction by the creation of logically or physically isolated groups of network resources and network function/virtual network functions configurations separating its behavior from the underlying physical network.
- \* Network slicing covers the full life cycle of slices that are managed groups of infrastructure resources, network functions and services (e.g. the network slice components are: service instance, a network functions instance, resources, slice manager and capability exposure).
- \* Network slices are dynamically and non-disruptively reprovisioned
- \* Network slices will need to be self-managed by automated, autonomic and autonomous, systems in order to cope with dynamic requirements, such as scalability or extensibility of an infrastructure (organically growing/shrinking of resources to meet the size of their organizations)
- \* Network slices are configurable and programmable and they have the ability to expose their capabilities and characteristics. The slice protocols and functions are selected according to slice required features. The behaviour of the network slice realized via network slice instance(s).
- \* Network slices are concurrently deployed as multiple logical, self-contained and independent, partitioned network functions and resources on a common physical infrastructure.
- \* Network slicing supports dynamic multi-services, multi-tenancy and the means for backing vertical market players.
- \* Network slicing simplifies the provisioning of services manageability of networks and integration and operational challenges especially for supporting. communication services.
- \* Network slicing offers native service customization enabled by the selection and configuration of network functions for coordinating/orchestration and control of network resource.
- \* Network slicing considerably transforms the networking perspective by abstracting, isolating, orchestrating and separating logical network behaviors from the underlying physical network resources

## 1.2 Network Slicing Work Scope

The purpose of the NS work in IETF is to develop a set of protocols and/or protocol extensions that enable efficient slice creation, activation / deactivation, composition, elasticity, coordination/orchestration, management, isolation, guaranteed SLA, and safe and secure operations within a connectivity network or network cloud / data centre environment that assumes an IP and/or

MPLS-based underlay.

While there are isolated efforts being carried out in several IETF working groups Network WG [I-D.leeking-actn-problem-statement 03], TEAS WG [I-D.teas-actn-requirements-04], [I-D.dong-network-slicing-problem-statement], ANIMA WG [I-D.galis-anima-autonomic-slice-networking], [IETF-Slicing1], [IETF-Slicing2], [IETF-Slicing3], [IETF-Slicing4], [IETF-Slicing5], [IETF-Mobility], [IETF-Virtualization], [IETF-Coding], [IETF-Anchoring] to achieve certain aspects of network slice functions and operations, there is a clear need to look at the complete life-cycle management characteristics of Network Slicing solutions though the discussions based on the following architectural tenets:

- \* Underlay tenet: support for an IP/MPLS-based underlay data plane.
- \* Governance tenet: a logically centralized authority for network slices in a domain.
- \* Separation tenet: slices may be virtually or physically independent of each other and have an appropriate degree of isolation (note 1) from each other.
- \* Capability exposure tenet: each slice allows third parties to access via dedicated interfaces and /or APIs and /or programming methods information regarding services provided by the slice (e.g., connectivity information, mobility, autonomicity, etc.) within the limits set by the operator or the slice owner.

NS approaches that do not adhere to these tenets are explicitly outside of the scope of the proposed work at IETF.

In pursuit of the solutions described above, there is a need to document an architecture for network slicing within both wide area network and edge/central data center environments.

Elicitation of requirements (examples are [RFC2119], [RFC4364]) for both Network Slice control and management planes will be needed, Facilitating the selection, extension, and/or development of the protocols for each of the functional interfaces identified to support the architecture.

Additionally, documentation on the common use-cases for slice validation for 5G is needed, such as mission-critical ultra-low latency communication services; massive-connectivity machine communication services (e.g. smart metering, smart grid and sensor networks); extreme QoS; independent operations and management; independent cost and/or energy optimisation; independent multi-topology routing; multi-tenant operations; new network architecture enablement, etc.



The proposed NS work would be coordinated with other IETF WGs (e.g. TEAS WG, DETNET WG, ANIMA WG, SFC WG, NETCONF WG, SUPA WG, NVO3 WG, DMM WG, Routing Area WG (RTGWG), Network Management Research Group (NMRG) and NFV Research Group (NFVRG)) to ensure that the commonalities and differences in solutions are properly considered. Where suitable protocols, models or methods exist, they will be preferred over creating new ones.

### 1.3 Notes

- (1) This issue requires efficient interaction between an upper layer in the hierarchy and a lower layer for QoS guarantees and for most of the operations on slicing.

## 2. Network Slicing - Selected Problems and Work Areas

The goal of this proposed work is to develop one or more protocol specifications (or extensions to existing protocols) to address specific slicing problems that are not met by the existing tools. The following problems were selected according to the analysis of the technical gaps in IETF protocols ecosystem. Each problem is associated with one identified IETF gap (draft-qiangu-netslices-gap-analysis-01). In addition an initial priority level is attached to each problem. [(\*\*\*) high priority, (\*\*) medium priority and (\*) low priority]. The proposed WG charter would include at least the high priority problems.

### 2.1 Global Issues - Problems (GP)

#### 2.1.1 Problem GI1: Uniform Reference Model (\*\*\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Uniform Reference Model for Network Slicing (Architecture document):

- \* Description of all of the functional elements required for network slicing.
- \* Description of shared non-sliced network parts. Establishes the boundaries to the basic network slice operations (creation, management, exposure, consumption).
- \* Describes the minimum functional roles derived from basic network slice operations including infrastructure owner (creation, exposure, management), slice operator (exposure, management, consumption), slice customer (management, consumption). Describe the interactions between infrastructure owner <-> slice operator, slice operator <-> slice operator, slice operator <-> slice customer.

- \* Additionally, this working area will normalize nomenclature and definitions for Network Slicing.

Short explanation: A uniform definition and architecture of Network slicing is presented in the NS Architecture draft. A Network slice is a managed group of subsets of resources, network functions/network virtual functions at the data, control, management/orchestration planes and services at a given time. Network slice is programmable and has the ability to expose its capabilities. The behaviour of the network slice realized via network slice instance(s).

- (1) The Service Instance Component represents the end-user service or business services. An instance of an end-user service or a business service that is realized within or by a NS. Would be provided by the network operator or by 3rd parties.
- (2) A Network Slice Instance component
  - a. Represented by a set of network functions, virtual network functions and resources at a given time
  - b. Forms a complete instantiated logical network to meet certain network characteristics required by the Service Instance(s).
  - c. Provides network characteristics which are required by a Service Instance.
  - d. May also be shared across multiple Service Instances
- (3) Resources component - it includes: Physical, Logical & Virtual resources
  - a. Physical & Logical resources - An independently manageable partition of a physical resource, which inherits the same characteristics as the physical resource and whose capability is bound to the capability of the physical resource. It is dedicated to a Network Function or shared between a set of Network Functions.
  - b. Virtual resources - An abstraction of a physical or logical resource, which may have different characteristics from that resource, and whose capability may not be bound to the capability of that resource.
- (4) Slice Element Manager (SEM) and Capability exposure component
  - a. Slice Element Manager (SEM) is instantiated in each Network Slice and it manages all access permissions and all interaction between a Network Slice and external functions (i.e. other Network Slices, Orchestrators, etc). Each SEM converts requirements from orchestrator into virtual resources and manages

Consolidation and versification of the above definition is required. New protocols are needed for the creation, for discovery and for orchestrating network slicing.

#### 2.1.2 Problem GI2: Requirements for operations and interactions (\*\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Review common scenarios from the requirements for operations and interactions point of view. Describes the roles (owner, operator, user) which are played by entities with single /multiple entities playing different roles.

Short explanation: Review of the functional and non- functional NS requirements is needed to ensure that resource utilization is maximized and infrastructure costs are minimized as services will need to operate over a union of shared network infrastructures, as against the traditional monolithic model operated either as dedicated network or as an overlay.

### 2.1.3 Problem GI3: Slice Templates (\*\*\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Design the slices to different scenarios [ChinaCom-2009], [GENI-2009], [IMT2020-2016bis], [NGMN-2016], [NGS-3GPP-2016], [ONF-2016]); Outlines an appropriate slice template definition that may include capability exposure of managed partitions of network resources (i.e. connectivity ([CPP]), compute and storage resources), physical and/or virtual network and service functions that can act as an independent connectivity network and/or as a network cloud.

Short explanation: A network slice template based on uniform reference model would enable the creation of a network slice instance. A template defines an abstraction of the overall network resources and functions requirement for a particular network slice instance. Different templates can also be regarded as definitions of individual network slice types. Besides the reference model for network resources and functions, each template has a complete description of the structure, configuration and the plans/work flows for how a certain type of network slice instance should be instantiated and managed during its life cycle.

There should be a clear definition of the level of abstraction of the network slice template according to the arrangement and specification of network slice life cycle management system. A valid network slice instance profile created based on specific network slice template is going to be decomposed into configuration profiles to certain OAM domains for the purpose of network slice instance creation.

The creation of a specific network slice template strictly relies on the exposed network capabilities. The network slice life cycle

management system should not allow a template with parameters exceeding the capabilities to be created.

## 2.2 Network Slice Capabilities - Problems (NSC)

### 2.2.1 Problem NSC1: Guarantees for isolation (\*\*\*)

Related Identified IETF Gap: "Slicing specific extension on Isolation".

Four-dimensional efficient slice creation with guarantees for isolation in each of the Data /Control /Management /Service planes. Enablers for safe, secure and efficient multi-tenancy in slices.

Short explanation: Network slices MUST support multi-tenancy, ensuring that isolation and performance guarantees are provided at the data, control, management and service planes. This involves the following:

- \* A network slice SHOULD provide a guaranteed level of service, according to a negotiated SLA between the customer and the slice provider
- \* Slices MUST be isolated at service level (e.g., one slice must not impact on the level of service of the other slides, even if sharing resources).
- \* Slices MUST be isolated at data level, even if sharing resources. Security and privacy mechanisms should be in place to ensure this.
- \* A network slice SHOULD be provided with exclusive control and/or management interfaces (depending on the type of network slice), enabling the deployment of different logical network slices over shared resources.

### 2.2.2 Problem NSC2: Fulfilling diverse requirements (\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Methods to enable diverse requirements for NS including guarantee for the end-to-end QoS of service in a slice.

Short explanation: The main goal of fulfilling NS requirements is to ensure that service operators can utilize or benefit from Network Slicing through multi-tenancy, enabling different customized infrastructures for different group of services across different network segments and operating them independently. It includes the tasks that go into determining the needs or conditions to meet for NS systems, taking account of the possibly conflicting requirements of

the various stakeholders and a prioritisation of requirements. New protocols are needed for interoperability between diverse type of network slices which are fulfilling diverse requirements

#### 2.2.3 Problem NSC3: Efficiency in slicing (\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Specifying policies and methods to realize diverse requirements without re-engineering the infrastructure.

Short explanation: This item is deployment-specific and cannot be promised as a problem to be solved entirely by protocols. An underlying infrastructure will always be needed to be reengineered and maintained to support up-to-date technologies and emerging requirements (including instantiating new service functions or withdrawing service functions, adding new nodes to absorb for traffic, ...). It is a local decision to figure out whether many services will be bound to the same slice, how many slices are to be instantiated and so on. Exposing standard interfaces to capture requirements will help to rationale the use of resources and how the requirements are fulfilled, however it is a challenge to guarantee in an absolute manner that slicing allows "diverse requirements without re-engineering the infrastructure".

#### 2.2.4 Problem NSC4: Slice Recursion (\*\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Short explanation: Recursion is a property of some functional blocks: a larger functional block can be created by aggregating a number of a smaller functional block and interconnecting them with a specific topology. As such recursive network slice definition is defined as the ability to build a new network slice out of existing network slice (s). A certain resource or network function /virtual network function could scale recursively, meaning that a certain pattern could replace part of itself. This leads to a more elastic network slice definition, where a network slice template, describing the functionality, can be filled by a specific pattern or implementation, depending on the required performance, required QoS or available infrastructure.

New protocols are needed for use of network slice template segmentation allowing a slicing hierarchy with parent - child relationships.

#### 2.2.5 Problem NSC5: Customized security mechanisms per slice (\*)

Related Identified IETF Gap: "Mechanisms for customized granularity OAM (Operations, Administration, and Maintenance)".

Short explanation: Customized securing mechanisms will be needed on a per slice basis. This may be provided by enabling dedicated service functions. For such cases, SFC techniques can be used here. Soliciting distinct SFs per slice can be provided with existing tools. I don't see a new problem out there.

This may be provided by configuring dedicated policies in a given security service function. In such case, I2NSF techniques can be used to interact with a given service function.

Traffic isolation may be needed for some services. Legacy tools can be used. I'm not sure if there is specific work specific to slicing other than making sure that appropriate flows are grafted to the appropriate slice and no data leaking between slices is to happen.

#### 2.2.6 Problem NSC6: flexibility and efficiency trade-offs (\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Methods and policies to manage the trade-offs between flexibility and efficiency in slicing.

Short explanation: Mechanisms SHOULD be in place to allow different levels of flexibility when providing network slices: from the ones that provided greater levels of flexibility in the provided resources and services that compound the slice, allowing to dynamically change/scale/migrate it over time within a negotiated range, to the ones that ensure the efficiency of the use of the resources at the cost of a smaller degree of flexibility.

#### 2.2.7 Problem NSC7: Optimisation (\*\*)

Related Identified IETF Gap: "non-overlay OAM solution".

Methods for network resources automatic selection for NS; global resource view formed; global energy view formed; Network Slice deployed based on global resource and energy efficiency; Mapping protocols.

Short explanation: NS optimization includes methods which enable that resources utilization is monitored and maximize, that infrastructure operational costs are minimized and that QoS are managed and

maximized at the time of creation of network slice instance and well as during NS operation.

#### 2.2.8 Problem NSC8: Monitoring and Discovery (\*\*)

Related Identified IETF Gap: "Mechanisms for dynamic discovery of service with function instances and their capabilities".

Monitoring status and behaviour of NS in a single and/or multi-domain environment; NS interconnection.

Short explanation: A Network slice is a managed group of subsets of resources, network functions / network virtual functions at the data, control, management/orchestration planes and services at a given time. Monitoring of slices interacts with and it is part of the NS Lifecycle management to aiming at reporting the performance of the running NS. As input, the Monitoring Subsystem receives the detailed service monitoring requests with references to resource allocation and Network functions instances in a NS. The Monitoring Subsystem is responsible for the monitoring continuously the state all 4 components of a NS (Service Instance component, Network Slice Instance component, Resources component). New protocols are needed for discovery and monitoring probes of all NS components and NS itself itself and for dynamic discovery of service with function instances and their capability.

#### 2.2.9 Problem NSC9: Capability exposure and APIs (\*\*\*)

Related Identified IETF Gap: "A detailed specification of Network Slicing Specification".

Capability exposure for NS; plus APIs for slices

Short explanation: To exploit the flexibility offered by network slices their users (customers, overlying operators) would need to know the features offered by both individual resources and complete slices. This means that there must be interfaces to deliver such information to the entity that needs it, but that will be also transitively delivered to the following chains of the slicing structure towards the final users.

To this sense, there are two specific interfaces that must be defined to address such function:

- \* The bottom-up interface, offered by underlying resource providers to resource consumers (operators) of any layer.
- \* The top-down interface, offered by overlying operators to lower level providers.

On the one hand, the first interface will, obviously, enable slice operators to access the slices owned by underlying providers and manage the resources they have been assigned in them. On the other hand, the second interface will enable lower layers to know details about the resources managed by overlying operators and the requirements they impose to the overlying network slices.

In this respect, both interfaces will emphasize the relation among the original resources, as well as the links from them to the resulting resources. This forms the main key of their management operations.

#### 2.2.10 Problem NSC10: Programmability of Operations of Network Slices (\*\*)

Related Identified IETF Gap: "Mechanisms for customized granularity OAM (Operations, Administration, and Maintenance)".

Short explanation: Network slice operations consist of all operations related to life cycle management of a slice and its optimized operation. Slice instance lifecycle management includes all operations related to slice instance creation, activation, update and deactivation. All these operations are automated and driven by appropriate policies. A slice instance is created according to a slice template and related policies. A unique identifier is assigned to each slice after its creation and a list of active slice instances are stored in slice repository. Several slice types are predefined which describe their functions as access, core, transport, data center and edge cloud slices. As example to each slice instance a Slice Priority parameter is assigned which describe the way of handling of slice degradation in case of lack of resources that can be allocated to slices. The parameter is also used in emergency situations in which there is a need to release resources from existing slices and to allocate them to newly created slices that are used for emergency situation handling.

The end-to-end slice can be a composition of per administrative or technological domain slices that are created according to their local templates. The process of slice creation can be recursive. The slice level are split between slice operator and slice tenant. The slice tenant obtains information about slices related KPIs and is expressing his reconfiguration wills as intents (high level policies), which are implemented in an automated manner by slice control and management planes. The slice operator is responsible for slice lifecycle and slice FCAPS handling. During operations of slice the slice resources are allocated in a dynamic way in order to provide required performance but in an economical way.



Each network slice exhibits following features: protection (note 2), elasticity (note 3), extensibility (note 4) and safety (note 5).

## 2.3 Network Slice Operations - Problems (NSO)

### 2.3.1 Problem NSO1: Slice life cycle management (\*\*\*)

Related Identified IETF Gap: "non-overlay OAM solution".

Slice life cycle management including creation, activation / deactivation, protection (note 2), elasticity (note 3), extensibility (note 4), safety (note 5), sizing and scalability of the slicing model per network and per network cloud: slices in access, core and transport networks; slices in data centres, slices in edge clouds.

Short explanation: Network slicing enables the operator to create logically partitioned networks at a given time customized to provide optimized services for different market scenarios. These scenarios demand diverse requirements in terms of service characteristics, required customized network and virtual network functionality (at the data, control, management planes), required network resources, performance, isolation, elasticity and QoS issues. A network slice is created only with the necessary network functions and network resources at a given time. They are gathered from a complete set of resources and network /virtual network functions and orchestrated for the particular services and purposes.

New protocols are needed for realising full Slice life cycle management at two distinct levels:

- (1) "network slice life-cycle management level" (i.e. the series of state of functional activities through which a network slice passes: creation, operation, deletion) and
- (2) "network slice instances level" (activated network slice level).

Functions for creating and managing network slice instances and the functions instantiated in the network slice instance are mapped to respective framework level.

### 2.3.2 Problem NSO2: Autonomic slice management and operation (\*\*)

Related Identified IETF Gap: "non-overlay OAM solution".

It includes self-configuration, self-composition, self-monitoring, self-optimisation, self-elasticity are carried as part of the slice protocols.

Short explanation: Network slice is a dynamic entity which lifecycle and operations should be automated. There are 3 main reasons of this automation:

- (1) There is an expectation that the number of slice instances can be huge what rises the problem of management scalability if it is performed in a classical, manual way.
- (2) Network slice instances can be created on demand by the end-users or verticals. They may play a role of slice instance administrator (making some reconfigurations or monitoring slice performance. It is however not expected that such administrator will have required experience and tools related to slice instance management. They can express some high-level requests that has to be translated into low level operations
- (3) Multiple network slice instances have to share common resources in economical way but with preserving required performance indicators. The problem of allocation of resources between slices combined with real-time optimization of slice operations can be only solved by continuous monitoring of slice performance and making continuous adaptations of resources allocated to them.

The mentioned reasons call for autonomic management which part should be autonomic management of each slice and autonomic management of resources that are allocated to slices. The autonomic operations at the slice instance level comprise of self-configuration, self-composition, efficient self-monitoring, self-healing and real-time optimization (self-optimisation) as a part of the autonomic management framework.

### 2.3.3 Problem NS03 - Slice stitching / composition (\*\*\*)

Related Identified IETF Gap: "non-overlay OAM solution".

Having enablers and methods for efficient stitching /composition/ decomposition of slices:

- \* vertically (service + management + control planes) and/or
- \* horizontally (between different domains part of access, core, edge segments) and /or
- \* vertically + horizontally.

Short explanation: Slice stitching. The network slice has to provide end-to-end communication and services. In some cases such end-to-end network slice instance can be created directly but in multi-domain environment the end-to slice will be a composition of slices of different domains in the each network segments (i.e. access, core, edge, transport, etc.). In such a case the domain slices will be created and maintained using domain specific slice templates and use

domain specific operations and all the domain slicing will be stitched together horizontally. The operation is supported by appropriate descriptions of domain slices, exchange of slice related policies between domains. Slice stitching operations are supported by uniform slice descriptors and appropriate matching of them. Each slices has appropriate set of mechanisms (slice border control functions) that support horizontal stitching of slices.

The vertical stitching of slices is an operation that modifies functionality of existing slice by adding and merging of functions of another slice (i.e. enhancing control plane properties be functions defined in another slice template). In general the vertical stitching of slices is used to enrich slice services.

Slices will be recursively used as components in software architectures. This means that several slices will be able to be used together to build a "composite network service" that inherits the capabilities of the original slices. The recursive property means that both slices and derived composite services can be again "split" into pieces to form new slices. The straight result of this aspect is that complex services are highly simplified by "stitching" slices and/or part of them, achieving the actual complexity by exploiting layering, which is the de-facto standard composition capability typically mapped into network architectures, but also by exploiting the abstraction levels offered by network service composability.

However, to hide such complexity and thus achieve the intended abstraction, the network architecture (and slicing reference model) must include, adopt, and promote the deployment of the necessary mechanisms and functions that support slice stitching and network service composability.

#### 2.3.4 Problem NSO4: E2E Network Orchestration (\*\*\*)

Related Identified IETF Gap: "non-overlay OAM solution (Operations, Administration, and Maintenance) solution".

End-to-end network segments orchestration of slices ([GUERZONI-2016], [KARL-2016]).

Short explanation: Network service composition has demonstrated to be highly beneficial for both operators and final users [GRAMMATIKOU-2012]. It allows the formation of large number of different services, which will be specialized to the particular needs of a user or a specific situation. However, the current network architecture is far for being ideal to implement such function.

One of the keys of network slicing is the flexibility it adds to the

network and the resulting "de-ossification" of network resources. Thus, this environment is much more optimal to allow the proliferation of network service composition, but it means that some sort of specific requirements must be pushed towards the architecture that supports the general slicing.

First, a proper composable network service model needs network resources to be compatible, regardless of the domain to which they pertain. Then they must be homogeneously described, so a user can actually understand their individual capabilities and "draw" the service they want to build by combining them. Finally, the resources living among separated network slices must be "connectable" to each other. This means that they must cross the domain of their providers/owners in order to reach their destination.

New protocols are needed for full end to end orchestration between the layers, from the IP layer and up.

#### 2.3.5 Problem NS05: Service Mapping in a single domain and Cross-Domain Coordination (\*\*\*)

Related Identified IETF Gap: "Companion YANG data model for network slicing - single domain and Cross-Domain Coordination".

Having dynamic and Automatic Mapping of Services to slices; YANG models for slices.

Short Description: The main goal of the service mapping framework is to enable on-demand processing anywhere in the physically distributed network, with dynamic and fine granular service (re-)provisioning, which can hide significant part of the resource management complexity from service providers and users, hence allowing them to focus on service and application innovation. It include a slice-aware YANG information model based on necessary connectivity, storage, compute resources, network functions, capabilities exposed and service elements in a single domain as well as Cross-Domain Coordination. As such the service mapping participates in management of the network slices.

#### 2.3.6 Problem NS06: Roles in Network Slicing (\*)

Related Identified IETF Gap: "Detailed specification of Network Slicing Specification".

Enablers and methods for the above mentioned capabilities and operations from different viewpoints on slices (note 6).

Short explanation: Several viewpoints emerge from the global and

wide interaction among the Network Infrastructure Owner (NIO), Network Slice Provider (NSP), and Network Slice Tenant(NST), and they must be treated by network slicing to ensure their use cases are correctly covered. They are:

- NIO <=> NSP:
  - + NIO offers the physical infrastructure to NSP, and NSP creates and manages the "slice" of network resources.
  - + NSP interacts vertically to request and instantiate (embed) composite network services onto the underlying physical infrastructures.
  - + NSP can possibly act as NIO.
- NSP <=> NST:
  - + NSP offers the individual objects/resources obtained after slicing the physical infrastructure to the NST.
  - + NST requests to the NSP the necessary CRUD (Create, Retrieve, Update, Delete) operations on its own Network Slices.
- NSP <=> NSP:
  - + Allows inter-provider tasks (e.g. migration of resources or whole slices among providers.
  - + Organizes the interoperability levels among Network Slices managed by different providers.
  - + Facilitates the recursive slicing, so a new NSP slices the resources offered by other NSP.
- NIO <=> NIO:
  - + Horizontal communication between owners to coordinate the required interactions among physical infrastructure resources, and/or the migration of whole slices among different NIOs.
  - + It may be common for NIO to provide network infrastructures to NSP in an old-fashion way where no network slicing is concerned.

However, a NIO may become a double role of NIO+NSP once it provides NSaaS.

Any NSP can become a NST if it uses specific network slice instance for a particular service, or it purchase NSaaS from another NSP.

#### 2.3.7 Problem NS07: Efficient enablers and methods for integration (\*)

Related Identified IETF Gap: "non-overlay OAM solution".

Efficient enablers and methods for integration of above capabilities

and operations.

Short explanation: In order to enable the above required capabilities and operation for network slicing, well defined reference points among the involved actors and entities are required, as well as the proper interface definitions ensuring interoperability of all involved pieces. Some examples of the required reference points/interfaces include:

Customer/Vertical (user of the slice) - Network Slice Provider. The user of the slice SHOULD be able to specify the characteristics of the slice and provide it in a suitable/understandable format to the NSP. A proper information model is needed to convey the customer slice requirements. And the model might need to support different levels of abstraction, to support different use cases.

Network Slice Provider - Network Slice Provider / Network Slice Operator / Network Service Provider. The slice provider MUST be able to request resources to compose a slice to other slice providers, slice operator or service operators. The interface needs to support recursiveness and different levels of abstraction (the request might involve resources or services).

Inter-domain interactions at different levels. Another way of composing a slice is by interaction of players at the same level (peering, instead of recursive), by delegating the request to other providers/operators. This type of interaction can take place at different levels (resource, network service, etc), and therefore would impose different requirements. In all cases, security issues are key due to the inter-operator nature.

## 2.4 Notes

- (1) Protection refers to the related mechanisms so that events within one slice, such as congestion, do not have a negative impact on another slice.
- (2) Elasticity refers to the mechanisms and triggers for the growth/shrinkage of network resources, and/or network and service functions.
- (3) Extensibility refers to the ability to expand a NS with additional functionality and/or characteristics, or through the modification of existing functionality/characteristics, while minimizing impact to existing functions.
- (4) Safety refers to the conditions of being protected against

different types and the consequences of failure, error harm or any other event, which could be considered non-desirable.

- (5) Multiple viewpoints on slices: I) viewpoint of the slice's owner towards user: from this viewpoint a slice is defined as a means to "split" physical or virtual infrastructure elements to "service" smaller portions. This action would be recursively done from the owner of the initial and physical infrastructure element to the users. II) viewpoint of from the user towards the physical infrastructure owner. From this viewpoint a slice is viewed just as a set of resources that must be managed (requests to a provider, listed, changed, returned to the provider, etc.). This viewpoint emphasizes those issues that would be used in the SLA definition of a slice.

### 3. OAM Operation with Customized Granularity

In accordance with [RFC6291], OAM is used to denote the following:

- \* Operations: refer to activities that are undertaken to keep the network and the services it deliver up and running. It includes monitoring the underlying resources and identifying problems.
- \* Administration: refer to activities to keep track of resources within the network and how they are used.
- \* Maintenance: refer to activities to facilitate repairs and upgrades. Maintenance also involves corrective and preventive measures to make the managed network run more effectively, e.g., adjusting configuration and parameters.

As per [RFC6291], NetSlices provisioning operations are not considered as part of OAM. Provisioning operations are discussed in other sections. Maintaining automatically-provisioned slices within a network raises the following requirements:

- \* Ability to run OAM activities on a provider's customized granularity level. In other words, ability to run OAM activities at any level of granularity that a service provider see fit. In particular:
  - An operator must be able to execute OAM tasks on a per slice basis.
  - These tasks can cover the "whole" slice within a domain or portion of that slice (for troubleshooting purposes, for example).
  - For example, OAM tasks can consist in tracing resources that are bound to a given slice, tracing resources that are invoked when forwarding a given flow bound to a given

- network slice, assessing whether flow isolation characteristics are in conformance with the NS Resource Specification, or assessing the compliance of the allocated slice resource against flow/customer requirements.
- An operator must be able to enable differentiated failure detect and repair features for a specific/subset of network. For example, a given slice may require fast detect and repair mechanisms (e.g., as a function of the nature of the traffic (pattern) forwarded through the NS), while others may not be engineered with such means.
- When a given slice is shared among multiple services/customers, an operator must be able to execute (per-slice) OAM tasks for a particular service or customer.
- \* Ability to automatically discover the underlying service functions and the slices they are involved in or they belong to.
- \* Ability to dynamically discover the set of NetSlices that are enabled within a network. Such dynamic discovery capability facilitates the detection of any mismatch between the view maintained by the control plane and the actual network configuration. When mismatches are detected, corrective actions must be undertaken accordingly.

#### 4 Summary

The following is a summary of the selected higher priority problems based on previous analysis in this document:

(I) Identified IETF Gap: "A detailed specification of Network Slicing Specification"; Requirement: Network Slicing Specification.

- \* Problem GI1: Uniform Reference Model
- \* Problem GI3: Slice Templates
- \* Problem NSC9: Capability exposure and APIs

(II) Identified IETF Gap: "A companion YANG data model for Network Slicing"; Requirement: Network Slicing Specification.

- \* Problem NSO5: Service Mapping in a single domain and Cross-Domain Coordination.

(III) Identified IETF Gap: "Slicing specific extension on Isolation (Performance Guarantee and Isolation-PGI"; Requirement: Performance Guarantee and Isolation.

- \* Problem NSC1: Guarantees for isolation.



(IV) Identified IETF Gap: "Mechanisms for dynamic discovery of service with function instances and their capabilities"; Requirement: Network Slicing OAM.

- \* Problem NSC8: Monitoring and Discovery

(V) Identified IETF Gap: "non-overlay OAM (Operations, Administration, and Maintenance) solution"; Requirement: Network Slicing OAM.

- \* Problem NSO1: Slice life cycle management

- \* Problem NSO4: E2E Network Orchestration

(VI) Identified IETF Gap: "Mechanisms for customized granularity OAM (Operations, Administration, and Maintenance)"; Requirement: Network Slicing OAM.

- \* Problem NSO2: Autonomic slice management and operation

- \* Problem NSO3: Slice stitching / composition

## 5 Security Considerations

Security will be a major part of the design of network slicing.

## 6 IANA Considerations

This document requests no IANA actions.

## 7 Acknowledgements

Thanks to Sheng Jiang (Huawei Technologies), Kevin Smith (Vodafone), Satoru Matsushima (SoftBank), Christian Jacquenet (Orange), Mohamed Boucadair (Orange) for reviewing this draft. Thanks to Stuart Clayman (UCL) for creating the nroff markup for this document.

## 8 References

### 7.1 IETF References

[I-D.dong-network-slicing-problem-statement] Dong, J. and S. Bryant, "Problem Statement of Network Slicing in IP/MPLS Networks", draft-dong-network-slicing-problem-statement-00 (work in progress), October 2016.

[I-D.galis-anima-autonomic-slice-networking] Galis, A., Makhijani, K., and D. Yu, "Autonomic Slice Networking-Requirements

and Reference Model", draft-galis-anima-autonomic-slice-networking-01 (work in progress), October 2016.

[RFC7665] Halpern, J., Pignataro, C., "Service Function Chaining (SFC) Architecture", <https://tools.ietf.org/html/rfc7665>, October 2015.

[I-D.leeking-actn-problem-statement 03] Ceccarelli, D., Lee, Y., "Framework for Abstraction and Control of Traffic Engineered Networks", draft-leeking-actn-problem-statement-03 (work in progress), September 2014.

[I-D.teas-actn-requirements-04] Lee, Y., Dhody, D., Belotti, S., Pithewan, K., Ceccarelli, D., "Requirements for Abstraction and Control of TE Networks", draft-ietf-teas-actn-requirements-04.txt, January 2017.

[IETF-Slicing1] "Presentations - Network Slicing meeting at IETF 97 of 15th November 2016", n.d., <[https://www.dropbox.com/s/ax2ofdwygjema8z/0-Network%20Slicing%20Side%20Meeting%20Introduction\\_IETF97.pdf](https://www.dropbox.com/s/ax2ofdwygjema8z/0-Network%20Slicing%20Side%20Meeting%20Introduction_IETF97.pdf)>.

[IETF-Slicing2] "Presentations - Network Slicing meeting at IETF 97 of 15th November 2016", n.d., <[https://www.dropbox.com/s/k2or6sd0ddzrc6c/1-Network%20Slicing%20Problem%20Statement\\_IETF97.pdf](https://www.dropbox.com/s/k2or6sd0ddzrc6c/1-Network%20Slicing%20Problem%20Statement_IETF97.pdf)>.

[IETF-Slicing3] "Presentations - Network Slicing meeting at IETF 97 of 15th November 2016", n.d., <[https://www.dropbox.com/s/g8zvfvbrtkysjs1/2-Autonomic%20Slice%20Networking\\_IETF97.pdf](https://www.dropbox.com/s/g8zvfvbrtkysjs1/2-Autonomic%20Slice%20Networking_IETF97.pdf)>.

[IETF-Slicing4] "Presentations - Network Slicing meeting at IETF 97 of 15th November 2016", n.d., <[https://www.dropbox.com/s/d3rk4pjpeg552ilv/3-Architecture%20for%20delivering%20multicast%20mobility%20services%20using%20network%20slicing\\_IETF97.pdf](https://www.dropbox.com/s/d3rk4pjpeg552ilv/3-Architecture%20for%20delivering%20multicast%20mobility%20services%20using%20network%20slicing_IETF97.pdf)>.

[IETF-Slicing5] "Presentations - Network Slicing meeting at IETF 97 of 15th November 2016", n.d., <[https://www.dropbox.com/s/e3isnlbxwwhaw8g/4-ACTN%20and%20network%20slicing\\_IETF97.pdf](https://www.dropbox.com/s/e3isnlbxwwhaw8g/4-ACTN%20and%20network%20slicing_IETF97.pdf)>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <<http://www.rfc-editor.org/info/rfc4364>>.
- [RFC1958] Carpenter, B., "Architectural Principles of the Internet", RFC 1958, <<https://www.ietf.org/rfc/rfc1958.txt>>.
- [RFC3439] Bush, R., Meyer, D., "Some Internet Architectural Guidelines and Philosophy", RFC3439, <<https://www.ietf.org/rfc/rfc3439.txt>>.
- [RFC3234] Carpenter, B., Brim S., "Middleboxes: Taxonomy and Issues", RFC3439, <<https://tools.ietf.org/html/rfc3234>>.
- [RFC7149] Boucadair, M., Jacquenet, C. , " Software-Defined Networking: A Perspective from within a Service Provider Environment", RFC 7149, March 2014 <<https://tools.ietf.org/html/rfc7149>>.
- [SFG WG] "Service Function Chaining WG" <<https://datatracker.ietf.org/doc/charter-ietf-sfc/>>.
- [CPP] Boucadair M., Jacquenet, C., Wang, N., "IP Connectivity Provisioning Profile (CPP)" <https://tools.ietf.org/html/rfc7297>
- [IETF-Mobility] Truong-Xuan Do, Young-Han Kim, "Architecture for delivering multicast mobility services using network slicing" 2016-10-31<[draft-xuan-dmm-multicast-mobility-slicing-00.txt](#)>
- [IETF-Virtualization] Carlos Bernardos, Akbar Rahman, Juan Zuniga, Luis Contreras, Pedro Aranda, " Network Virtualization Research Challenges" 2016-10-31<[draft-irtf-nfvrg-gaps-network-virtualization-03.txt](#)>
- [IETF-Coding] M.A. Vazquez-Castro, Tan Do-Duy, Paresh Saxena, Magnus Vikstrom, "Network Coding Function Virtualization" 2016-11-14 <[draft-vazquez-nfvrg-netcod-function-virtualization-00.txt](#)>
- [IETF-Anchoring] Anthony Chan, Xinpeng Wei, Jong-Hyouk Lee, Seil Jeon, Alexandre Petrescu, Fred Templin "Distributed Mobility Anchoring" 2016-12-15 <[draft-ietf-dmm-distributed-mobility-anchoring-03.txt,.pdf](#)>
- [RFC6291] L. Andersson, H. van Helvoort, R. Bonica, D. Romascanu, S. Mansfield "Guidelines for the Use of the "OAM" Acronym in

the IETF" - June 2011 <https://tools.ietf.org/html/rfc6291>

## 7.2 Informative References

- [ChinaCom-2009] "A. Galis et al - Management and Service-aware Networking Architectures (MANA) for Future Internet - Invited paper IEEE 2009 Fourth International Conference on Communications and Networking in China (ChinaCom09) 26-28 August 2009, Xi'an, China", n.d., <<http://www.chinacom.org/2009/index.html>>.
- [GENI-2009] "GENI Key Concepts - Global Environment for Network Innovations (GENI)", n.d., <<http://groups.geni.net/geni/wiki/GENIConcepts>>.
- [GUERZONI-2016] "Guerzoni, R., Vaishnavi, I., Perez-Caparrros, D., Galis, A., et al Analysis of End-to-End Multi Domain Management and Orchestration Frameworks for Software Defined Infrastructures - an Architectural Survey", June 2016, <[onlinelibrary.eilex.com/10.1002/ett.3084/pdf](http://onlinelibrary.eilex.com/10.1002/ett.3084/pdf)>.
- [IMT2020-2015] "Report on Gap Analysis", ITU-T FG IMT2020, December 2015, <<http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx>>.
- [IMT2020-2016] "Draft Technical Report Application of network softwarization to IMT-2020 (O-041)", ITU-T FG IMT2020, December 2016, <<http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx>>.
- [IMT2020-2016bis] "Draft Terms and definitions for IMT-2020 in ITU-T (O-040)", ITU-T FG IMT2020, December 2016, <<http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx>>.
- [KARL-2016] "Karl, H., Peuster, M, Galis, A., et al DevOps for Network Function Virtualization - An Architectural Approach", July 2016, <<http://onlinelibrary.wiley.com/doi/10.1002/ett.3084/full>>.
- [MANO-2014] "Network Functions Virtualisation (NFV); Management and Orchestration v1.1.1.", ETSI European Telecommunications Standards Institute., December 2014, <[http://www.etsi.org/deliver/etsi\\_gs/NFV-MAN/001\\_099/001/01.01.01\\_60/gs\\_nfv-man001v010101p.pdf](http://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_nfv-man001v010101p.pdf)>.
- [NGMN-2016] "Hedmar, P., Mschner, K., et al - Description of Network

Slicing Concept", NGMN Alliance NGS-3GPP-2016, January 2016, <[https://www.nmn.org/uploads/media/160113\\_Network\\_Slicing\\_v1\\_0.pdf](https://www.nmn.org/uploads/media/160113_Network_Slicing_v1_0.pdf)>.

[NGS-3GPP-2016] "Study on Architecture for Next Generation System - latest version v1.0.2", September 2016, <[http://www.3gpp.org/ftp/tsg\\_sa/WG2\\_Arch/Latest\\_SA2\\_Specs/Latest\\_draft\\_S2\\_Specs](http://www.3gpp.org/ftp/tsg_sa/WG2_Arch/Latest_SA2_Specs/Latest_draft_S2_Specs)>.

[ONF-2016] Paul, M, Schallen, S., Betts, M., Hood, D., Shirazipor, M., Lopes, D., Kaippallimalit, J., - Open Network Foundation document "Applying SDN Architecture to 5G Slicing", Open Network Foundation, April 2016, <[https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/Applying\\_SDN\\_Architecture\\_to\\_5G\\_Slicing\\_TR-526.pdf](https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/Applying_SDN_Architecture_to_5G_Slicing_TR-526.pdf)>.

[5G-ICN] Ravi Ravindran, Asit Chakraborti, Syed Obaid Amin, Aytac Azgin, G.Q.Wang, "5G-ICN: Delivering ICN Services in 5G using Network Slicing", IEEE Communication Magazine, May, 2017.

[GRAMMATIKOU-2012] Grammatikou, M; Marinos, C; Martinez-Julia, P; Jofre, J; Gheorghiu, S; et al. Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA); Athens: 1-5. Athens: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). (2012)

[GAL] A. Galis, Chih-Lin I" Towards 5G Network Slicing - Motivation and Challenges" IEEE 5G Tech Focus, Volume 1, Number 1, March 2017 - <http://5g.ieee.org/tech-focus/march-2017#networkslicing>

[GAPS] Gap Analysis for Network Slicing draft-qiang-netslices-gap-analysis-01

[NS UseCases] Network Slicing Use Cases: Network Customization for different services draft-makhijani-netslices-usecase-customization-03

[NS ARCH] Network Slicing Architecture draft-geng-netslices-architecture-02

Authors' Addresses

Alex Galis  
University College London  
Email: a.galis@ucl.ac.uk

Slawomir Kuklinski  
Orange  
Email: slawomir.kuklinski@orange.com

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

Liang Geng  
China Mobile  
Email: gengliang@chinamobile.com

Kiran Makhijani  
Huawei Technologies  
Email: kiran.makhijani@huawei.com

Hannu Flinck  
Nokia  
Email: hannu.flinck@nokia-bell-labs.com

Ravi Ravindran  
Huawei Technologies  
Email: ravi.ravindran@huawei.com

Luis Miguel Contreras Murillo  
Telefonica  
Email: luismiguel.contrerasmurillo@telefonica.com

Stewart Bryant  
Huawei Technologies  
Email: stewart.bryant@gmail.com

Pedro Martinez-Julia  
National Institute of Information and Communications Technology  
(NICT)  
Email: pedro@nict.go.jp

Susan Hares  
Huawei Technologies  
Email: shares@ndzh.com

Carlos Jesus Bernardos Cano  
University Carlos III Madrid  
Email: cjbc@it.uc3m.es



Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: January 4, 2018

L. Geng  
China Mobile  
J. Dong  
S. Bryant  
K. Makhijani  
Huawei Technologies  
A. Galis  
University College London  
X. de Foy  
InterDigital Inc.  
S. Kuklinsk  
Orange  
July 3, 2017

Network Slicing Architecture  
draft-geng-netslices-architecture-02

Abstract

This document defines the overall architecture of network slicing. Based on the general architecture, basic concepts of network slicing and examples of network slicing instances are introduced for clarification purposes. Some architectural considerations about the data plane, control plane, management and orchestration of network slicing are described to give a general view of network slicing implementation principles.

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 <http://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 January 4, 2018.



## Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
1.1. Requirements Language . . . . .	4
1.2. Terminology . . . . .	4
2. Demand for Network Slicing . . . . .	6
2.1. Guaranteed Service Performance . . . . .	7
2.2. End-to-end Customization . . . . .	7
2.3. Network Slicing as a Service . . . . .	7
3. Network Slicing Architecture . . . . .	8
3.1. Requirements . . . . .	8
3.2. High-Level Functional Components . . . . .	8
3.2.1. Service Component . . . . .	11
3.2.2. Network Slicing Management and Orchestration . . . . .	11
3.2.3. Resource Component . . . . .	14
3.3. Network Slicing Capabilities . . . . .	15
3.3.1. Reclusiveness . . . . .	15
3.3.2. Protection . . . . .	15
3.3.3. Elasticity . . . . .	16
3.3.4. Extensibility . . . . .	16
3.3.5. Safety . . . . .	16
3.3.6. Isolation . . . . .	16
3.4. Network Slices Capability Exposure . . . . .	16
4. Data Plane of Network Slicing . . . . .	17
4.1. Propagation of Guarantees . . . . .	17
4.2. The Underlying Physical Layer . . . . .	17
4.3. Hard vs Soft Slicing in the Data-plane . . . . .	18
4.4. The Role of Deterministic Networking . . . . .	18
4.5. The Role of VPNs . . . . .	19
4.6. Dynamic Reprovisioning . . . . .	19
4.7. Non-IP Data Plane . . . . .	19
5. Control Plane of Network Slicing . . . . .	19
5.1. NS Infrastructure Control Plane . . . . .	20

5.2.	NS Infrastructure Control Operations and Protocols . . .	20
5.3.	Programmability of the NS Infrastructure Control Plane .	21
5.4.	Intra-Slice Control Plane . . . . .	21
6.	Management Plane of Network Slicing . . . . .	22
6.1.	Network Slice Creation - Reservation / Release Messages Flow . . . . .	22
6.2.	Self- Management Operations . . . . .	23
6.3.	Programmability of the Management Plane . . . . .	24
6.4.	Management plane slicing protocols . . . . .	24
7.	Service Functions and Mappings . . . . .	24
8.	OAM and Telemetry . . . . .	24
9.	IANA Considerations . . . . .	25
10.	Security Considerations . . . . .	25
11.	Acknowledgements . . . . .	25
12.	References . . . . .	25
12.1.	Normative References . . . . .	25
12.2.	Informative References . . . . .	26
	Authors' Addresses . . . . .	26

## 1. Introduction

The Internet has always been designed to support a variety of services. The emerging 5G market is expected to bring this diversity of services to a new level [NS\_WP]. Typical examples of new bandwidth-hungry services enabled by 5G include high definition (HD) video, virtual reality (VR) and augmented reality (AR). The high bandwidth requirement of these services is not particularly challenging thanks to the continuing advancing technologies. However, the guarantee of high bandwidth performance of these services based-on a spontaneous on-demand pattern is fairly challenging. Moreover, providing high bandwidth with strict packet loss tolerances and high mobility is also difficult for the current networks which are commonly designed for best effort purposes.

Given that most Internet protocols are designed to comply with a best effort, or enhanced best effort paradigm, it is inevitable that the network will suffer from performance degradation in case of congestion. Recent work on deterministic networking (DetNet) [I-D.finn-detnet-architecture] aims to improve this situation by providing a ceiling on latency for a particular traffic flow, which significantly improves packet error rate for specific DetNet services. This pioneering work gives a great example that new approaches are investigated to make the Internet aware of certain performance requirement other than the bandwidth.

Taking a look at the network infrastructure, service provider used to build dedicated network and resources for services requiring guaranteed performance. This is simply not cost-effective, neither

is it flexible. The emergence of virtualization and VPN technologies make it possible to set up logically isolated computing and network instances from shared infrastructures. This can be used dedicatedly by specific services for improved performances. However, many questions are still to be answered as different technologies in various domains need to be combined to build network slices, which may require the separation of different resources and various types of performance guarantees.

### 1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

### 1.2. Terminology

#### I. Networking Servicing Terms

**Service** - A piece of software that performs one or more functions and provides one or more APIs to applications or other services of the same or different layers to make use of said functions and returns one or more results. Services can be combined with other services, or called in a certain serialized manner, to create a new service.

**Service Instance** - An instance of an end-user service or a business service that is realized within or by a network slice. Each service is represented by a service instance. Services and service instances would be provided by the network operator or by third parties.

**Administrative domain** - A collection of systems and networks operated by a single organization or administrative authority. Infrastructure domain is an administrative domain that provides virtualized infrastructure resources such as compute, network, and storage, or a composition of those resources via a service abstraction to another Administrative Domain, and is responsible for the management and orchestration of those resources.

#### II. Network Resource Terms

**Resource** - A physical or virtual (network, compute, storage) component available within a system. Resources can be very simple or fine-grained (e.g., a port or a queue) or complex, comprised of multiple resources (e.g., a network device).

**Logical Resource** - An independently manageable partition of a physical resource, which inherits the same characteristics as the

physical resource and whose capability is bound to the capability of the physical resource.

Virtual Resource - An abstraction of a physical or logical resource, which may have different characteristics from that resource, and whose capability may not be bound to the capability of that resource.

Network Function (NF) - A processing function in a network. It includes but is not limited to network nodes functionality, e.g. session management, mobility management, switching, routing functions, which has defined functional behaviour and interfaces. Network functions can be implemented as a network node on a dedicated hardware or as a virtualized software functions. Data, Control, Management, Orchestration planes functions are Network Functions.

Virtual Network Function (VNF) - A network function whose functional software is decoupled from hardware. One or more virtual machines running different software and processes on top of industry-standard high-volume servers, switches and storage, or cloud computing infrastructure, and capable of implementing network functions traditionally implemented via custom hardware appliances and middle-boxes (e.g. router, NAT, firewall, load balancer, etc.)

Network Element - A network element is defined as a manageable logical entity uniting one or more network devices. This allows distributed devices to be managed in a unified way using one management system. It means also a facility or equipment used in the provision of a communication service. Such term also includes features, functions, and capabilities that are provided by means of such facility or equipment, including subscriber numbers, databases, signalling systems, and information sufficient for billing and collection or used in the transmission, routing, or other provision of a telecommunications service.

### III. Network Slicing Terms used in this draft

Resource Slice - A grouping of physical or virtual (network, compute, storage) resources. It inherits the characteristics of the resources which are also bound to the capability of the resource. A resource slice could be one of the components of Network Slice, however on its own does not represent fully a Network Slice.

Network Slice - A Network slice is a managed group of subsets of resources, network functions / network virtual functions at the data, control, management/orchestration planes and services at a given time. Network slice is programmable and has the ability to expose its capabilities. The behaviour of the network slice realized via network slice instance(s).

End-to-end Network Slice - A cross-domain network slice which may consist of access network (fixed or cellular), transport network, (mobile) core network and etc. End-to-end network slice can be customized according to the requirements of network slice tenants

Network Slice Instance - An activated network slice. It is created based on network template. A set of managed run-time network functions, and resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the service instance(s). It provides the network characteristics that are required by a service instance. A network slice instance may also be shared across multiple service instances provided by the network operator.

Network Slice Provider - A network slicing provider, typically a telecommunication service provider, is the owner or tenant of the network infrastructures from which network slices are created. The network slicing provider takes the responsibilities of managing and orchestrating corresponding resources that the network slicing consists of.

Network Slice Terminal - A terminal that is network-slice-aware, typically subscribed to the service which is hosted within a network slice instance. A network slice terminal may be capable of subscribing to multiple network slice instance simultaneously.

Network Slice Tenant - A network slice tenant is the user of specific NSIs, with which specific services can be provided to end customers. Network slice tenants can make requests of the creation of new network slice instances. Certain level of management capability should be exposed to network slice tenant from network slice service provider.

Network Slice Repository - A repository that in each domain consists of a list of active Network Slices with their identifiers and description. This description defines also the rules that have to be fulfilled in order to access a slice. Network Slice Repository is updated by slice orchestrator. In case of recursive slicing the Network Slice Repository keeps information about all slices that compose a higher level slice but such slice has its own identifier and descriptors.

## 2. Demand for Network Slicing

It is expected that a diversity of new services will emerge in both mobile/5G and fixed networks.[I-D.qin-netslices-use-cases] describes many of the differentiated services (e.g. smart home, industrial control, remote healthcare, Vehicle-to-Everything (V2X) etc.) and

their relevance to the Network Slicing. These use cases are typical examples of service verticals requiring features beyond connectivity such as uRLL, high-bandwidth, and isolation.

### 2.1. Guaranteed Service Performance

One of the most challenging requirements for future network is to provide guaranteed performance for varieties of new services whilst maintaining the economies of scale that accrue through resource sharing. It has been foreseen that the requirements of different services would be diversified and complex.

Network slicing can deal with these challenges by mapping the performance requirements to physically or logically dedicated resources.

### 2.2. End-to-end Customization

Customization is another significant feature of future services. Many vertical industries are expected to offer customization capabilities as a service to both internal manufacturing processes and specific end users. Meanwhile, these customized services need to be deployed with short time-to-market. The network needs to adapt to this challenge since customers may frequently adjust and refine their customization requirements.

There is ongoing work such as network orchestration, software defined networks and network function virtualization that aims to address this problem. In principle, these new technologies share a common request for the network to provide the ability to provide agile resource allocation.

### 2.3. Network Slicing as a Service

It is anticipated that the operation of 5G and future networks will involve new business models. Given that the network is more flexible, elastic, modularized and customized, the shared network infrastructure can be sliced and offered as a service to the customer. For instance, dedicated, isolated, end-to-end network resources with a customized topology can be provided as a network slice service to the tenant of this network slice. The tenants are allowed to have a certain level of provisioning of their network slices.

### 3. Network Slicing Architecture

This section introduces the general system architecture of network slicing.

#### 3.1. Requirements

To meet the diversified Quality of Experience (QoE) demands of different vertical industries, the gap analysis document has identified the following requirements:

- o Req.1 Network Slicing Resource Specification
- o Req.2 Cross-Network Segment; Cross-Domain Negotiation
- o Req.3 Guaranteed Slice Performance and Isolation
- o Req.4 Slice Discovery and Identification
- o Req.5 NS Domain-Abstraction
- o Req.6 OAM Operations with Customized Granularity

In the following sections, these requirements will be addressed and associated with different aspects of the Network Slicing architecture.

#### 3.2. High-Level Functional Components

End-to-end network slice is a broad area and comprises of several functional components. In the context of distribution of role and responsibilities, a network slice consists of the following components as shown in Figure 1. It can be seen that two network slice instances are created from the shared network infrastructures. In principle, the network slicing subnets (NS Subnets) represent any general physical and logical network resources for demonstration purposes. The two network slice instances created share the computing, connectivity and storage resources, whether they are in physical or virtual forms.

It is fundamental to network slicing that slices may be created, the topology and/or its resources modified, and that the slices may be decommissioned in a timely manner with minimum work by the network slicing provider or the customer. This is not however unique to network slicing, it is a goal of modern classical networks to be able to do this.

The descriptions of functional components are introduced in the following sections.



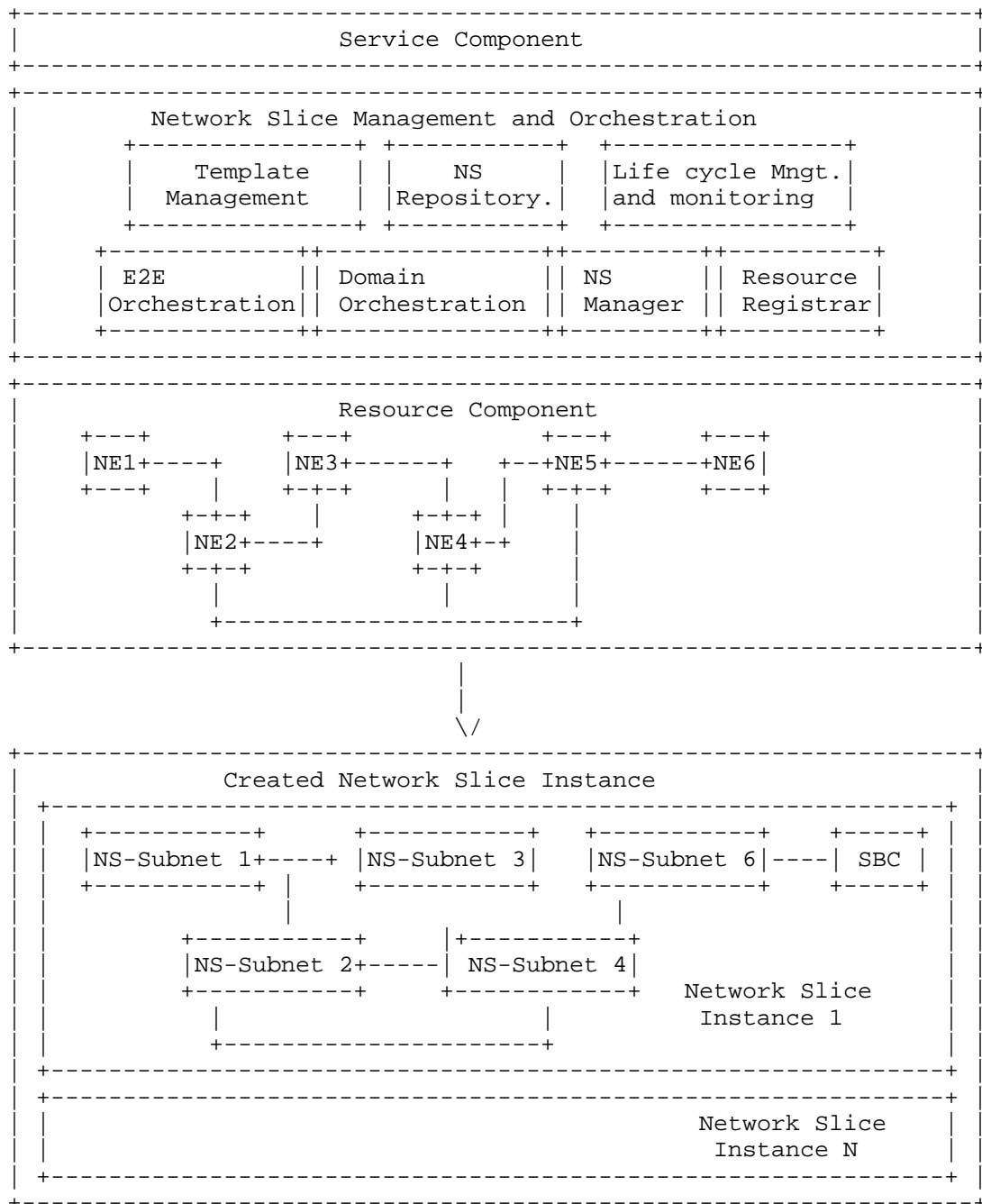


Figure 1: Network Slicing Architecture

### 3.2.1. Service Component

A service represents an end-user's business logic. It is realized within or by Network slice instance. A service may demand a set of network resources and attributes in form of a network slice. A service is either mapped to a network slice instance or an ordered chain of network slice instances.

### 3.2.2. Network Slicing Management and Orchestration

As seen in Figure 1, The management and orchestration layer of network slicing system consist of the following functional components.

#### 1. Template Management

A network slice template consists of complete description of the structure, configuration and the plans/work flows for how to instantiate and control the network slice instance during its life cycle.

#### 2.NS Repository

To provide mechanism that will allow the end-user selection and attachment to a slice instance or, if required, to multiple slice instances at the same time, a NS repository (or repositories) is needed, in which there are stored slices with the description of their properties and access rules.

The service component should have an access to such repository in order to check if the required slice exists. If such a slice doesn't exist a matching procedure should allow an attachment of the service to a slice which properties are the most similar ones to the requested slice (under certain policies agreement between network slice provider and tenant). Optionally the service may trigger the deployment of a new slice. During the attachment of the service component to a slice the slice data forwarding mechanisms are configured in way that will redirect a selected part of the end-user traffic to the slice.

#### 3.Life cycle management and monitoring

Network slicing enables the operator to create logically partitioned networks at a given time customized to provide optimized services for different market scenarios. These scenarios demand diverse requirements in terms of service characteristics, required customized network and virtual network functionality (at the data, control, management planes), required network resources, performance,

isolation, elasticity and QoS issues. A network slice is created only with the necessary network functions and network resources at a given time. They are gathered from a complete set of resources and network /virtual network functions and orchestrated for the particular services and purposes.

A network slice is a dynamic entity therefore its lifecycle has to be managed. The network slice lifecycle management is (creation, update, deletion) is managed by the network slice orchestrator. The slice orchestrator according to requests that can be send by the orchestrator operator, 3rd parties or even by the end-users creates a new slice instance that is based on slice template that is stored in slice template repository however it takes into account slice operator (owner) preferences (policies).

#### 4.E2E Orchestration

This section describes E2E Slices Orchestration and its functionality. Orchestration refers to the system functions in a domain that automate and autonomically co-ordination of network functions in slices autonomically coordinate the slices lifecycle and all the components that are part of the slice (i.e. Service Instances, Network Slice Instances, Resources, Capabilities exposure) to ensure an optimized allocation of the necessary resources across the network. The main functionality of E2E slice orchestration may include the following aspects.

- (1) Coordinate a number of interrelated resources, often distributed across a number of subordinate domains, and to assure transactional integrity as part of the process.
- (2) Autonomically control of slice life cycle management, including concatenation of slices in each segment of the infrastructure including the data plane, the control plane, and the management plane.
- (3) Autonomically coordinate and trigger of slice elasticity and placement of logical resources in slices.
- (4) Coordinates and (re)-configure logical resources in the slice by taking over the control of all the virtualized network functions assigned to the slice.

It is the continuous process of allocating resources to satisfy contending demands in an optimal manner. The idea of optimization would include at least prioritized SLA commitments , and factors such as customer endpoint location, geographic or topological proximity, delay, aggregate or fine-grained load, monetary cost, fate- sharing

or affinity. The word continuing incorporates recognition that the environment and the service demands constantly change over the course of time, so that orchestration is a continuous, multi-dimensional optimization feedback loop. The E2E slice orchestration should have the following characteristics.

- o It protects the infrastructure from instabilities and side effects due to the presence of many slice components running in parallel.
- o It ensures the proper triggering sequence of slice functionality and their stable operation.
- o It defines conditions/constraints under which service components will be activated, taking into account operator service and network requirements (inclusive of optimize the use of the available network; compute resources and avoid situations that can lead to sub-par performance and even unstable and oscillatory behaviors).

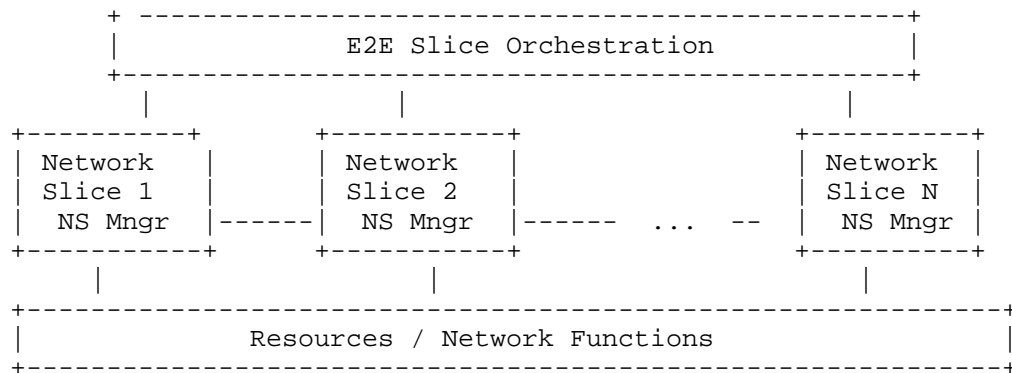


Figure 2: E2E Slice Orchestration

## 5. Domain Orchestration

Another value that the network slicing brings is fast, automated and dynamic deployment services in end-to-end manner, even in heterogeneous environment. In order to achieve that goal the problem of providing a slice that spans multiple domains has to be solved. There two possible solutions. The first one lies on appropriate allocation of resources in each domain (i.e. creation of the resource slice instance), their aggregation and using of a single orchestrator in order to deploy a slice. Another possibility is to use per domain orchestrators with domain specific template and provide the chaining of domain slices in order to obtain the end-to-end slice. In such a case the orchestration is hierarchical one, i.e. the domain

orchestration is driven by a high level orchestrator that interacts with orchestrators of all domains that are involved in end-to-end slice instance creation. The slice that is composed of multiple domain level slices requires specific mechanisms for inter-slice operations like topology information exchange and/or appropriate protocol conversion/adaptation.

The approach may lead to recursive slicing (or sub-slicing) in which higher level slice instances are composed of lower level ones. The creation of end-to-end slice composed of several slices may require specific description of such slice and changes of functions of domain slices. For example the traffic redirection can be implemented only in this domain slice which is an ingress slice.

## 6. NS Manager

NS Manager management entity for a specific network slice instance. it manages all access permissions and all interaction between a Network Slice and external functions (i.e. other Network Slices, Orchestrators, etc). Each NS Manager maps requirements from orchestrator into network resources and manages these resources of a specific network slice instance.

Allow 3rd parties to access via APIs information regarding services provided by the slice (e.g. connectivity information, QoS, mobility, autonomicity, etc.)

Allow dynamical customization of the network characteristics for different diverse use cases within the limits set of functions by the operator. Network slice enables the operator to create networks customized to provide flexible solutions for different market scenarios, which have diverse requirements, with respect to the functionality, performance and resource separation.

It includes a description of the structure (and contained components) and configuration of the slice instance.

## 7. Resource Registration

Resource registration component manages the exposed capability of the network infrastructure. Details description is TBD.

### 3.2.3. Resource Component

Resource component includes physical, logical and virtual resources (defined in Section 2). An abstraction of resources is required in order to consistently map the requirements such as latency, reliability, band-width. Resource component may need interfaces with

elements in network slice functional component as well as NS manager for that purpose of discovering capabilities.

### 3.3. Network Slicing Capabilities

#### 3.3.1. Reclusiveness

Recursion is a property of some functional blocks: a larger functional block can be created by aggregating a number of a smaller functional block and interconnecting them with a specific topology. As such one could summarize the concept of recursive network slice definition as the ability to build a new network slice out of existing network slice (s). A certain resource or network function /virtual network function could scale recursively, meaning that a certain pattern could replace part of itself. This leads to a more elastic network slice definition, where a network slice template, describing the functionality, can be filled by a specific pattern or implementation, depending on the required performance, required QoS or available infrastructure. If a certain part of a network slice can be replaced by different patterns, this can offer some advantages:

- o Each pattern might have its own capabilities in terms of performance. Depending on the required workload, a network function /virtual network function might be replaced by a pattern able to process at higher performance. Similarly, a service or network function /virtual network function can be decomposed so it can be deployed on the available infrastructure.
- o From an orchestrating point of view, above way of using recursive network slice templates, can be beneficial for the placement algorithm used by the orchestrator. The success rate, solution quality and/or runtime of such an embedding algorithm benefits from information on both possible scaling or decomposition topologies and available infrastructure.
- o Enabling methods for network slice template segmentation allowing a slicing hierarchy with parent - child relationships.

#### 3.3.2. Protection

Protection refers to the related capability and mechanisms so that events within one network slice, such as congestion, do not have a negative impact on another slice.

### 3.3.3. Elasticity

Elasticity refers to the capability, mechanisms and triggers for the growth /shrinkage of network resources, and/or network and service functions in an Network Slice as function of service needs.

### 3.3.4. Extensibility

Extensibility refers to the capability and ability to expand a network slice with additional functionality and/or characteristics, or through the modification of existing network function / virtual network function while minimizing impact to existing functions.

### 3.3.5. Safety

Safety refers to the conditions in within one network slice of being protected against different types and the consequences of failure, error harm or any other event, which could be considered non-desirable in an other network slice.

### 3.3.6. Isolation

Efficient slice creation is expected to guarantee the isolation and non interference between network slices in the Data /Control /Management planes as well as safety and security for multi-tenancy in slices.

## 3.4. Network Slices Capability Exposure

An important value of network slicing is the capability of a slice to be tightly coupled with services, i.e. the slice instance can be designed that way that it support a specific service or limited number of services only, but not all of them in the same slice. The property means that not only the slice data plane operations are properly tuned, but also the control plane can be designed according to the requirements of slice specific services. In general it is possible that a single slice instance may support a single service only, however it is more scalable to provide more than a single service per slice. Such approach has important implications. First of all in order to add services to a slice each slice should expose its functions to services/applications. Moreover the service lifecycle management is different than slice lifecycle management. This is similar to the existing networks, however in opposite to them the deployment of a new service may lead to important reconfiguration of a slice to which the service is attached (the slice is programmable what means that we are going beyond the API approach - the services templates are melted with the slice template). The goal is to have tightly coupled services with networks and providing joint

optimization of networks and services at the level that is impossible to achieve in present, hardware based solutions.

#### 4. Data Plane of Network Slicing

In the network slicing architecture, the data plane in the edge and core of the network will likely be one or more of the standard IETF data planes: IPv4/IPv6, MPLS or Pseudo-wires (PW). This section assumes that the IETF protocol stack exists as-is, and describes the performance consideration in different layers of the data plane.

##### 4.1. Propagation of Guarantees

Guarantees of delay start at the physical layer and propagate up the stack layer by layer. Any layer can add delay, and can take various steps to minimize the impact of delay on its layer, but no layer can reduce the delay introduced by a lower layer.

Guarantees of loss and jitter can, by contrast be upheld or improved at any layer of the protocol stack, but usually at a cost of increased delay. Where delay is a constrain as it is in some 5G applications the option of trading delay for better loss or jitter characteristics is not an option. In these circumstances it is critical that the quality characteristics start at the physical layer and be maintained at each layer of the protocol stack.

##### 4.2. The Underlying Physical Layer

A point to point dedicated physical channel provides the delay, jitter and loss characteristics limited only by the media itself. This does not fulfill the need for rapid reconfiguration of the network to provision new services.

To address the need to provision a slice of the data-plane one approach that can be deployed is to time-slice access to the physical service. Ignoring many of the classic TDM offering as being too slow, a number of technologies are available that might be applied including OTN and FlexE. Whilst the provisioning of the channel provided by underlays such as FlexE and the interconnection of FlexE channels is within the scope of this architecture the operation of the underlay is outside its scope.

The logical sub-division of a physical channel be that a single channel with the full bandwidth available or a channel multiplexed at the physical layer such as is provided by FlexE we will consider in the following section.



#### 4.3. Hard vs Soft Slicing in the Data-plane

Hard slicing refers to the provision of resources in such a way that they are dedicated to a specific NSI. Data-plane resources are provided in the data-plane through the allocation of a lambda, through the allocation of a time domain multiplexed resource such as a FlexE channel or through a service such as an MPLS hard-pipe. Note that although hard-pipes can be used to allocate dedicated, non-shared resources to an NSI, the using of allocation is bandwidth, which can result in more "lumpiness" in the physical channel that would not be present with a true physical layer multiplexing scheme.

Soft slicing refers to the provision of resources in such a way that whilst the slices are separated such that they cannot statically interfere with each other (one cannot receive the others packets or observe or interfere with the other's storage), they can interact dynamically (one may find the other is sending a packet just when it wants to, or the other may be using CPU cycles just when the other needs to process some information), which means they may compete for some particular resource at some specific time. Soft slicing is achieved through logically multiplexing the data-plane over a physical channel include various types of tunnel (IP or MPLS) or various types of pseudo-wire (again IP or MPLS). Although the design of deterministic networking techniques helps, it is not possible to achieve the same degree of isolation with these techniques as it is possible to achieve with pure physical layer multiplexing techniques. However where such techniques provide sufficient isolation their use leads to a network design that may be deployed on existing equipment designs and which can make unused bandwidth available to best effort traffic.

#### 4.4. The Role of Deterministic Networking

Deterministic networking is a technology under development in the IETF that aims to both minimize congestion loss and set an upper bound on per hop latency. It allows a packet layer to emulate the behaviour of a fully partitioned underlay such might be provided through some physical layer multiplexing system such as FlexE.

Deterministic networking works by policing the ingress rate of a flow to an agreed maximum and then scheduling the transmission time of each flow to reduce the "lumpiness" and hence the possible buildup of queues and hence congestion loss.

Whilst deterministic networking is not as perfect as physical layer multiplexing in terms of latency minimization, because the scheduling is hop by hop and not end to end meaning that at each hop a packet has to wait for the transmission slot allocated to its flow, it has

the advantage that it is able to allocate slots not needed by the allocated traffic to best effort traffic. This reallocation of the unused transmission slots to background traffic significantly improves the efficiency of the network by amortizing the cost between the scheduled high priority users and the best effort users.

#### 4.5. The Role of VPNs

VPNs are considered candidate technologies for network slicing. The existing VPN technologies mainly focus on the isolation of forwarding tables between different tenants and provide a virtual topology for the connectivity between different sites of a tenant. The VPN layer and the underlying network resources are usually loosely coupled, and statistical multiplexing is adopted to improve network utilization.

Although VPNs have been widely used to provide enterprise services in service provide networks, it is unclear that whether VPNs along with existing underlying tunnel technologies can meet the performance and isolation requirements of critical services in the vertical industries.

#### 4.6. Dynamic Reprovisioning

A requirement of the network slicing system is that it can be dynamically and non-disruptively reprovisioned. That is not an unusual requirement of a modern network. However the frequency of reprovisioning with network slicing will be relatively high, such that it in many cases it is not possible to hide any disruption during a "quiet" time.

Physical multiplexing methods such as FlexE have the ability to seamlessly reprovision multiplex slots. At the network layer techniques such as make-before-break, segment routing, and loop-free-convergence can be used to provide uninterrupted operation during a topology change.

#### 4.7. Non-IP Data Plane

Non-IP data plane in support of Information Centric Networking (ICN), some of the IoT services and other similar requirements will be added in a future version.

### 5. Control Plane of Network Slicing

There are two control plane systems that need to be considered. The first is the control plane of the slicing infrastructure itself (NS Infrastructure Control Plane), the second is the control plane of an individual slice (Intra-Slice Control Plane).

### 5.1. NS Infrastructure Control Plane

The NS infrastructure control plane receives the instruction of creating a network slice with particular requirements from the orchestration layer. It then creates the network slice by allocating a set of network resources in the corresponding network infrastructure. This set of network resources is associated with the network slice during this operation.

The NS infrastructure control plane is also responsible, with the support of the orchestration layer, for dynamically adjusting the network according to slice change requests (e.g. from slice tenants), and to changes in network infrastructure. As it is critical to meet the service requirements of a network slice independently from activity and changes occurred in other network slices or in infrastructure, appropriate service assurance mechanisms should be deployed in the network. The control plane, with the support of the orchestration layer, **MUST** be able to react within a pre-determined (possibly system-specific) time to any network events, such as resource addition and failure. The orchestration layer **SHOULD** be involved, directly or indirectly, to take reactive decisions, e.g. to re-route a flow, to ensure that other network slices are not affected. Indirect involvement includes, for example, reactive programming by the orchestration layer to address foreseeable events or cases where connection to the orchestration layer is lost.

The NS infrastructure control plane can be implemented as an extension of the Virtual Infrastructure Manager (VIM), in cases where the NFV-MANO architecture is used for the management and control architecture of the system. Especially, the VNF Manager is considered part of the management plane and not control plane. From technology standpoint, NS infrastructure control plane can be an extension of Cloud infrastructure technology (e.g. OpenStack), which itself can integrate SDN technology for network control. This logically centralized control can be supplemented or replaced with distributed control protocols, that can provide some benefits in scenarios which require fast reaction, robustness and efficient information distribution. A hybrid architecture is anticipated, where distributed protocols complement and simplify a centralized control system.

### 5.2. NS Infrastructure Control Operations and Protocols

The following operations should be supported. Different control protocols can be used to control different types of resources. Multiple control protocols can be supported simultaneously.

- o Setting up or tearing down network function instances within a slice. Set, increase or decrease compute capacity of NFs.
  - \* Control protocols can be based on openstack APIs and other Cloud infrastructure control protocols.
- o Setting up, tearing down, increase or decrease capacity of connectivity between network function instances within a slice, e.g. as L2-L3 virtual network or software function chain.
  - \* Control protocols can include NVO3 control protocol, SFC control protocol and NetConf.
- o Reservation/release of traffic flows within a slice, possibly with associated QoS and routing requirements.
  - \* Control protocols can include DETNET, MPLS-TE, etc.
  - \* Interconnect slices or slice flows, including across domains
  - \* Control protocols are TBD.

### 5.3. Programmability of the NS Infrastructure Control Plane

The NS Control Plane exposes a Northbound API, typically for use by the orchestration layer. A higher-than-physical representation level of abstraction can be used, enabling the manipulation of a logical network, that is translated down to physical resource manipulation by the NS infrastructure control plane. The level of this abstraction and of its associated logical network is TBD. Programmability should include programming reactions to events, which reduces the dynamic involvement of the orchestration layer, and therefore reaction time to events.

### 5.4. Intra-Slice Control Plane

Intra-slice control plane maintains proper connectivity and networking characteristics within the slice. A full range of existing control plane technologies needs to be permissible. Intra-slice control plane technologies can include existing IGP protocols (such as IS-IS or OSPF), BGP, overlay control (such as NVO3 or SFC). Some slices may be controlled by their own SDN controllers. Intra-slice control plane can span across multiple domains (since NS infrastructure control deals with slice interconnection).

## 6. Management Plane of Network Slicing

It is expected that the management and orchestration layer would use state of the art management technologies to support short time-to-market, and help the operators to build an open ecosystem for new services in vertical industries. In multi-tenant environment the slice tenants can trigger the creation of slice instances for them by interacting with the E2E Orchestrator. After the creation of the slice the slice tenant is able to monitor slice KPIs (performance, faults) and send slice reconfiguration requests to E2E Orchestrator.

The basic functional architecture of management and orchestration layer of network slicing system has been discussed in section 3. This section further introduces some essential characteristics.

### 6.1. Network Slice Creation - Reservation / Release Messages Flow

The establishment of Network slices is both business-driven (i.e. slices are in support for different types and service characteristics and business cases) and technology-driven as network slice is a grouping of physical or virtual resources (network, compute, storage) and a grouping network functions and virtual network functions (at the data, control and management planes) which can act as a sub network at a given time. A network slice can accommodate service components and network functions (physical or virtual) in all network segments: access, core and edge / enterprise networks.

The management plane creates the grouping of network resources (physical, virtual or a combination thereof), it connects with the physical and virtual network and service functions and it instantiates all of the network and service functions assigned to the slice.

Once a network slice is created, the slice control plane takes over the control, slice operations and governing of all the network resources, network functions, and service functions assigned to the slice. It (re-) configures them as appropriate and as per elasticity needs, in order to provide an end-to-end service. In particular, ingress routers are configured so that appropriate traffic is bound to the relevant slice. Identification means for the traffic may be simple (relying on a subset of the transport coordinate, DSCP/traffic class, or flow label), or identification may be a more sophisticated one. Also, the traffic capacity that is specified for a slice can be changed dynamically, based on some events (e.g. triggered by a service request). The slice control plane is responsible for instructing the involved elements to guarantee such needs.

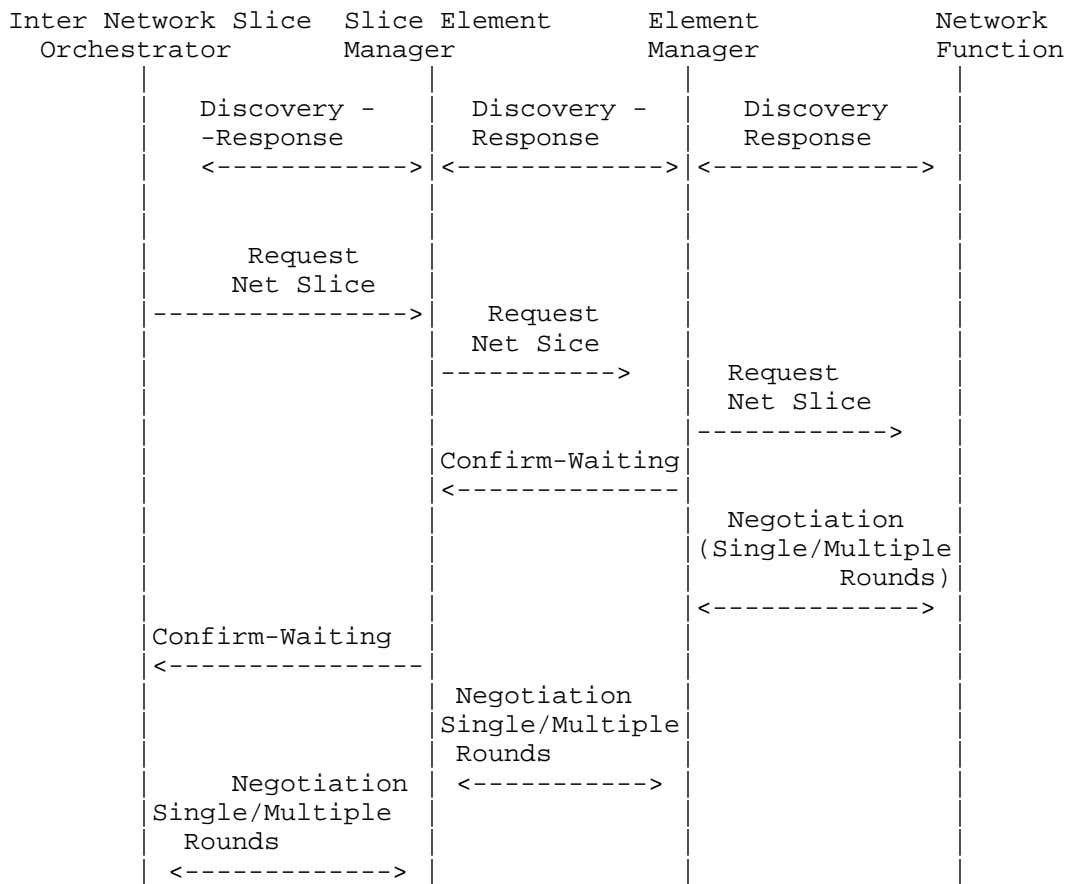


Figure 3: Network Slice Reservation / Release Messages Flow

## 6.2. Self- Management Operations

Self-management operations are focused on self-optimization and self-healing of network slice instances (including intra-slice functions management), network slice instance services and resources that are used for all slice instances. All these operations are combined with efficient and economical monitoring and reconfigurations at appropriate level. In order to make the management scalable and environment aware the management architecture is composed of many functional entities that follows the feedback loop management paradigm (aka autonomic management). The self-management functions may realize different goals and have to be coordinated according to slice instance and infrastructure operator policies. The self-management deals with dynamic (1) allocation of resources to slice instances in a economical way that provides required slice instances

performance, (2) self-optimization and self-healing of slice instances during their deployment (lifecycle management) and operations (3) self-optimization and self-healing of services of each slice instance. Their lifecycle, that is typically different than slice instance lifecycle should also be managed in the autonomous way. Despite the self-managed functions may have different goals and involved entities the slice instance self-management should be coordinated with self-management of their services and self-management of resources (inter-slice operations) should be aligned with in-slice self-management operations. In the implementation the self-management functionality is split between NS manager (that is a part of slice template) and slice orchestrator in case of slice management and between service specific management and NS manager in case of services that use a specific slice.

### 6.3. Programmability of the Management Plane

The Management Plane is composed of multiple functional entities and is responsible for resource, slice instance and slice service management. In case of slice instances and services their management comes as a part of appropriate slice or service template respectively. That way slice or service related management functions are instantiated for each slice and/or service. The Management Plane may expose a set of APIs which can be used by additional management services that are added independently on service or slice instance lifecycle. Using these APIs and allocation additional resource the slice or service operator can add advanced and new management functions. That way the Management Plane programmability is provided.

### 6.4. Management plane slicing protocols

At this stage it is too early to define protocols (IMHO). We have to define the management architecture first with functional entities and reference points/interfaces. Having them we could define which protocol(s) we want to use for each of them. Maybe we can mention some protocols but generally they should be a part of separate specification.

## 7. Service Functions and Mappings

## 8. OAM and Telemetry

OAM and telemetry to instrument the system need to be provided for each NSI so that the NSI provider can monitor the health of the NSI and so that the NSI owner can independently verify the health of their NSI.

Running OAM on the NSI from the perspective of its owner can be undertaken by the owner using the native tools for the NSI network type. For example if the NSI is IP, tools like ICMP [RFC792], ICMPv6 [RFC4443], or IPFIX [RFC7011] can be used. Similarly the native OAM tools for MPLS and Ethernet can be used. If the NSI provides a partial emulation of the network type that limits the ability to operate such native instrumentation tools, then this needs to be made clear to the NSI owner.

Similarly running OAM on the underlay will also use the native tools for the network type providing the underlay. Care must be taken that any OAM run by the NS provider does not impinge on the operation of the NSI, and SHOULD be undetectable in the NSI.

Telemetry will need to be provided to both the NS provider and the NSI owner. Telemetry of the underlay will use the NS providers pub-sub system of choice.

Telemetry of the NSI may be provided purely by the NSI owner installing a telemetry collection system. However significant efficiencies may be realised by if the NS provider exports relevant telemetry to the NSI owner's pub-sub system. Where this is done, consideration must be given to the security of the measurement and export system so to no information is leaked between NSIs.

## 9. IANA Considerations

This document makes no request of IANA.

## 10. Security Considerations

Each layer of the system has its own security requirements.

## 11. Acknowledgements

## 12. References

### 12.1. Normative References

[I-D.finn-detnet-architecture]  
Finn, N. and P. Thubert, "Deterministic Networking Architecture", draft-finn-detnet-architecture-08 (work in progress), August 2016.



[I-D.qin-netslices-use-cases]

Qin, J., kiran.makhijani@huawei.com, k., Dong, J., Qiang, L., and S. Peng, "Network Slicing Use Cases: Network Customization for Different Services", draft-qin-netslices-use-cases-00 (work in progress), March 2017.

## 12.2. Informative References

[NS\_WP] China Mobile Communication Corporation, Huawei Technologies Co. Deutsche Telekom AG, Volkswagen, "5G Service-Guaranteed Network Slicing White Paper", 2016, <<http://labs.chinamobile.com/pdf/5GService-GuaranteedNetworkSlicingWhitePaper.pdf>>.

## Authors' Addresses

Liang Geng  
China Mobile  
Beijing  
China

Email: gengliang@chinamobile.com

Jie Dong  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095

Email: jie.dong@huawei.com

Stewart Bryant  
Huawei Technologies  
U.K.

Email: stewart.bryant@gmail.com

Kiran Makhijani  
Huawei Technologies  
2890 Central Expressway  
Santa Clara CA 95050

Email: kiran.makhijani@huawei.com

Alex Galis  
University College London  
London  
U.K.

Email: a.galis@ucl.ac.uk

Xavier de Foy  
InterDigital Inc.  
1000 Sherbrooke West  
Montreal  
Canada

Email: Xavier.Defoy@InterDigital.com

Slawomir Kuklinski  
Orange

Email: slawomir.kuklinski@gmail.com

TEAS Working Group  
Internet-Draft  
Intended status: Informational  
Expires: October 2, 2021

D. King  
Old Dog Consulting  
J. Drake  
Juniper Networks  
H. Zheng  
Huawei Technologies  
A. Farrel  
Old Dog Consulting  
March 31, 2021

Applicability of Abstraction and Control of Traffic Engineered Networks  
(ACTN) to Network Slicing  
draft-king-teas-applicability-actn-slicing-10

Abstract

Network abstraction is a technique that can be applied to a network domain. It utilizes a set of policies to select network resources and obtain a view of potential connectivity across the network.

Network slicing is an approach to network operations that builds on the concept of network abstraction to provide programmability, flexibility, and modularity. It may use techniques such as Software Defined Networking (SDN) and Network Function Virtualization (NFV) to create multiple logical or virtual networks, each tailored for a set of services that share the same set of requirements.

Abstraction and Control of Traffic Engineered Networks (ACTN) is described in RFC 8453. It defines an SDN-based architecture that relies on the concept of network and service abstraction to detach network and service control from the underlying data plane.

This document outlines the applicability of ACTN to network slicing in a Traffic Engineering (TE) network that utilizes IETF technology. It also identifies the features of network slicing not currently within the scope of ACTN, and indicates where ACTN might be extended.

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 October 2, 2021.

#### Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

#### Table of Contents

1. Introduction . . . . .	3
1.1. Terminology . . . . .	4
2. Requirements for Network Slicing . . . . .	5
2.1. Resource Slicing . . . . .	6
2.2. Network Virtualization . . . . .	6
2.3. Service Isolation . . . . .	6
2.4. Control and Orchestration . . . . .	7
3. Abstraction and Control of Traffic Engineered (TE) Networks (ACTN) . . . . .	7
3.1. ACTN Virtual Network as a Network Slice . . . . .	8
3.2. ACTN Virtual Network for Network Slice Aggregation . . . . .	9
3.3. Management Components for ACTN and Network Slicing . . . . .	9
3.4. Examples of ACTN Delivering Types of Network Slices . . . . .	10
3.4.1. ACTN Used for Virtual Private Line . . . . .	10
3.4.2. ACTN Used for VPN Delivery Model . . . . .	12
3.4.3. ACTN Used to Deliver a Virtual Consumer Network . . . . .	13
4. YANG Models . . . . .	15
4.1. Network Slice Service Mapping from TE to ACTN VN Models . . . . .	15
4.2. Interfaces and Yang Models . . . . .	16
4.3. ACTN VN Telemetry . . . . .	17
5. IANA Considerations . . . . .	18
6. Security Considerations . . . . .	18
7. Acknowledgements . . . . .	19
8. Contributors . . . . .	19

9. Informative References . . . . .	19
Authors' Addresses . . . . .	22

## 1. Introduction

The principles of network resource separation are not new. For years, the concept of separated overlay and logical (virtual) networking has existed, allowing multiple services to be deployed over a single physical network comprised of single or multiple layers. However, several key differences exist that differentiate overlay and virtual networking from network slicing.

A network slice is a virtual (that is, logical) network with its own network topology and a set of network resources that are used to provide connectivity that conforms to a specific Service Level Agreement (SLA) or set of Service Level Objectives (SLOs). The network resources used to realize a network slice belong to the network that is sliced. The resources may be assigned and dedicated to an individual slice, or they may be shared with other slices enabling different degrees of service guarantee and providing different levels of isolation between the traffic in each slice.

[I-D.ietf-teas-ietf-network-slice-definition] provides a number of useful definitions for network slicing in the context of IETF network technologies. In particular, that document defines the term "IETF network slice" to be the generic network slice concept applied to a network that uses IETF technologies. An IETF network slice could span multiple technologies (such as IP, MPLS, or optical) and multiple administrative domains. The logical network that is an IETF network slice may be kept separate from other concurrent logical networks each with independent control and management: each can be created or modified on demand. Since this document is focused entirely on IETF technologies, it uses the term "network slice" as a more concise expression. Further discussion on the topic of IETF network slices can be found in [I-D.ietf-teas-ietf-network-slice-framework].

At one end of the spectrum, a virtual private wire or a virtual private network (VPN) may be used to build a network slice. In these cases, the network slices do not require the service provider to isolate network resources for the provision of the service - the service is "virtual".

At the other end of the spectrum there may be a detailed description of a complex service that will meet the needs of a set of applications with connectivity and service function requirements that may include compute resource, storage capability, and access to content. Such a service may be requested dynamically (that is,

instantiated when an application needs it, and released when the application no longer needs it), and modified as the needs of the application change. This type of service is called an enhanced VPN and is described in more detail in [I-D.ietf-teas-enhanced-vpn]. It is often based on Traffic Engineering (TE) constructs in the underlay network.

Abstraction and Control of TE Networks (ACTN) [RFC8453] is a framework that facilitates the abstraction of underlying network resources to higher-layer applications and that allows network operators to create virtual networks for their customers through the abstraction of the operators' network resources.

As noted in [I-D.ietf-teas-ietf-network-slice-framework], ACTN is a toolset capable of delivering network slice functionality. This document outlines the application of ACTN and associated enabling technologies to provide network slicing in a network that utilizes IETF technologies such as IP, MPLS, or GMPLS. It describes how the ACTN functional components can be used to support model-driven partitioning of resources into variable-sized bandwidth units to facilitate network sharing and virtualization. Furthermore, the use of model-based interfaces to dynamically request the instantiation of virtual networks can be extended to encompass requesting and instantiation of specific service functions (which may be both physical or virtual), and to partition network resources such as compute resource, storage capability, and access to content. Finally, this document highlights how the ACTN approach might be extended to address the requirements of network slicing where the underlying network is TE-capable.

### 1.1. Terminology

As far as is possible, this document re-uses terminology from [I-D.ietf-teas-ietf-network-slice-definition], [I-D.ietf-teas-enhanced-vpn] and [I-D.ietf-teas-ietf-network-slice-framework]. The terms defined below are give context and meaning for use in this document only and do not force wider applicability. As other work matures, it is hoped that the terminology will converge.

**Service Provider:** A server network or collection of server networks. The persons or organization responsible for operating such networks.

**Consumer:** As defined in [I-D.ietf-teas-ietf-network-slice-definition], a consumer is the component or entity that requests and uses a network slice. This may be any application, client network, or customer of a service

provider. In the ACTN framework [RFC8453] the consumer of a network service is termed a 'customer' because it will often be the case that a VPN consumer is a customer of the operator of the core network that delivers the service. In the context of a network slice, the consumer may well be a customer, but might also be a client network of the service provider (which could also be an internal organization of the service provider), or an application that engineers traffic in the network.

**Service Functions (SFs):** Components that provide specific functions within a network. SFs are often combined in a specific sequence called a service function chain to deliver services [RFC7665].

**Resource:** Any feature including connectivity, bufferage, compute, storage, and content delivery that forms part of or can be accessed through a network. Resources may be shared between users, applications, and clients, or they may be dedicated for use by a unique consumer.

**Infrastructure Resources:** The hardware and software for hosting and connecting SFs. These resources may include computing hardware, storage capacity, network resources (e.g., links and switching/routing devices enabling network connectivity), and physical assets for radio access.

**Service Level Agreement (SLA):** Per [I-D.ietf-teas-ietf-network-slice-definition], an SLA is an explicit or implicit contract between the consumer of a network slice and the provider of the slice. The SLA is expressed in terms of a set of Service Level Objectives (SLOs) and may include commercial terms as well as the consequences of violating the SLOs. The SLA describes the quality with which features and functions are to be delivered. It may include measures of bandwidth, latency, and jitter; the types of service (such as firewalls or billing) to be provided; the location, nature, and quantities of services (such as the amount and location of compute resources and the accelerators required).

**Network Slice Service:** An agreement between a consumer and a service provider to deliver network resources according to a specific service level agreement.

## 2. Requirements for Network Slicing

According to [I-D.ietf-teas-ietf-network-slice-framework] the consumer expresses requirements for a particular IETF network slice by specifying what is required rather than how the requirement is to be fulfilled. That is, the IETF network slice consumer's view of a IETF network slice is an abstract one.

The concept of network slicing is a key capability to serve consumers with a wide variety of different service needs expressed as SLOs in term of latency, reliability, capacity, and service function specific capabilities.

This section outlines the key capabilities required to realize network slicing in a TE-enabled IETF technology network.

### 2.1. Resource Slicing

Network resources need to be allocated and dedicated for use by a specific network slice, or they may be shared among multiple slices. This allows a flexible approach that can deliver a range of services by partitioning (that is, slicing) the available network resources to make them available to meet the consumer's SLA.

### 2.2. Network Virtualization

Network virtualization enables the creation of multiple virtual networks that are operationally decoupled from the underlying physical network, and are run on top of it. Slicing enables the creation of virtual networks as consumer services.

### 2.3. Service Isolation

A consumer may request, through their SLA, that changes to the other services delivered by the service provider do not have any negative impact on the delivery of the service. This quality is referred to as "isolation" [I-D.ietf-teas-ietf-network-slice-definition] [I-D.ietf-teas-enhanced-vpn].

Delivery of such service isolation may be achieved in the underlying network by various forms of resource partitioning ranging from dedicated allocation of resources for a specific slice, to sharing or resources with safeguards.

Although multiple network slices may utilize resources from a single underlying network, isolation should be understood in terms of the following three categorisations.

- o Performance isolation requires that service delivery for one network slice does not adversely impact congestion or performance levels of other slices.
- o Security isolation means that attacks or faults occurring in one slice do not impact on other slices. Moreover, the security functions supporting each slice must operate independently so that an attack or misconfiguration of security in one slice will not



prevent proper security function in the other slices. Further, privacy concerns require that traffic from one slice is not delivered to an end point in another slice, and that it should not be possible to determine the nature or characteristics of a slice from any external point.

- o Management isolation means that each slice must be independently viewed, utilized, and managed as a separate network. Furthermore, it should be possible to prevent the operator of one slice from being able to control, view, or detect any aspect of any other network slice.

#### 2.4. Control and Orchestration

Orchestration combines and coordinates multiple control methods to provide a single mechanism to operate one or more networks to deliver services. In a network slicing environment, an orchestrator is needed to coordinate disparate processes and resources for creating, managing, and deploying the network slicing service. Two aspects of orchestration are required:

- o Multi-domain Orchestration: Managing connectivity to set up a network slice across multiple administrative domains.
- o End-to-end Orchestration: Combining resources for an end-to-end service (e.g., underlay connectivity with firewalling, and guaranteed bandwidth with minimum delay).

### 3. Abstraction and Control of Traffic Engineered (TE) Networks (ACTN)

ACTN facilitates end-to-end connectivity and provide virtual connectivity services (such as virtual links and virtual networks) to the user. The ACTN framework [RFC8453] introduces three functional components and two interfaces:

- o Customer Network Controller (CNC)
- o Multi-domain Service Coordinator (MDSC)
- o Provisioning Network Controller (PNC)
- o CNC-MDSC Interface (CMI)
- o MDSC-PNC Interface (MPI)

RFC 8453 also highlights how:

- o Abstraction of the underlying network resources is provided to higher-layer applications and consumers.
- o Virtualization is achieved by selecting resources according to criteria derived from the details and requirements of the consumer, application, or service.
- o Creation of a virtualized environment is performed to allow operators to view and control multi-domain networks as a single virtualized network.
- o A network is presented to a consumer as a single virtual network via open and programmable interfaces.

The ACTN managed infrastructure consists of traffic engineered network resources. The concept of traffic engineering is broad: it describes the planning and operation of networks using a method of reserving and partitioning of network resources in order to facilitate traffic delivery across a network (see [I-D.ietf-teas-rfc3272bis] for more details). In the context of ACTN, traffic engineering network resources may include:

- o Statistical packet bandwidth.
- o Physical forwarding plane sources, such as wavelengths and time slots.
- o Forwarding and cross-connect capabilities.

The ACTN network is "sliced" with consumers each being given a different partial and abstracted topology view of the physical underlay network.

### 3.1. ACTN Virtual Network as a Network Slice

To support multiple consumers, each with its own view of and control of a virtual network constructed using a server network, a service provider needs to partition the server network resources to create network slices assigned to each consumer.

An ACTN Virtual Network (VN) is a consumer view of a slice of the ACTN-managed infrastructure. It is a network slice that is presented to the consumer by the ACTN provider as a set of abstracted resources. See [I-D.ietf-teas-actn-vn-yang] for a detailed description of ACTN VNs and an overview of how various different types of YANG model are applicable to the ACTN framework.

Depending on the agreement between consumer and provider, various VN operations are possible:

- o **Network Slice Creation:** A VN could be pre-configured and created through static configuration or through dynamic request and negotiation between consumer and service provider. The VN must meet the network slice requirements specified in the SLA to satisfy the consumer's objectives.
- o **Network Slice Operations:** The VN may be modified and deleted based on consumer requests. The consumer can further act upon the VN to manage the consumer's traffic flows across the network slice.
- o **Network Slice View:** The VN topology is viewed from the consumer's perspective. This may be the entire VN topology or a collection of tunnels that are expressed as consumer end points, access links, intra domain paths and inter-domain links.

[RFC8454] describes a set of functional primitives that support these different ACTN VN operations.

### 3.2. ACTN Virtual Network for Network Slice Aggregation

Scaling considerations for IETF network slicing are an important consideration. If the service provider must manage and maintain network state for every network slice then this will quickly limit the number of customer services that can be supported.

The importance of network slice aggregation is discussed in [I-D.ietf-teas-enhanced-vpn] and further in [I-D.dong-teas-enhanced-vpn-vtn-scalability]. That work notes the importance of aggregating network slices into groups of similar slices before realizing those aggregates in the network.

The same consideration applies to ACTN VNs. But fortunately, ACTN VNs may be arranged hierarchically by recursing the MDSCs so that one VN is realised over another VN. This allows the VNs presented to the customer to be aggregated before they are instantiated in the physical network.

### 3.3. Management Components for ACTN and Network Slicing

The ACTN management components (CNC, MDSC, and PNC) and interfaces (CMI and MPI) are introduced in Section 3 and described in detail in [RFC8453]. The management components for network slicing are described in [I-D.ietf-teas-ietf-network-slice-framework] and are known as the consumer orchestration system, the IETF network slice controller (NSC), and the network controller. The network slicing

management components are separated by the network slice controller northbound interface (NSC NBI) and the network slice controller southbound interface (NSC SBI).

[I-D.ietf-teas-ietf-network-slice-framework] describes the mapping between network slicing management components and ACTN management components. This is presented visually in Figure 1 and provides a useful reference for understanding the material in Section 3.4 and Section 4.

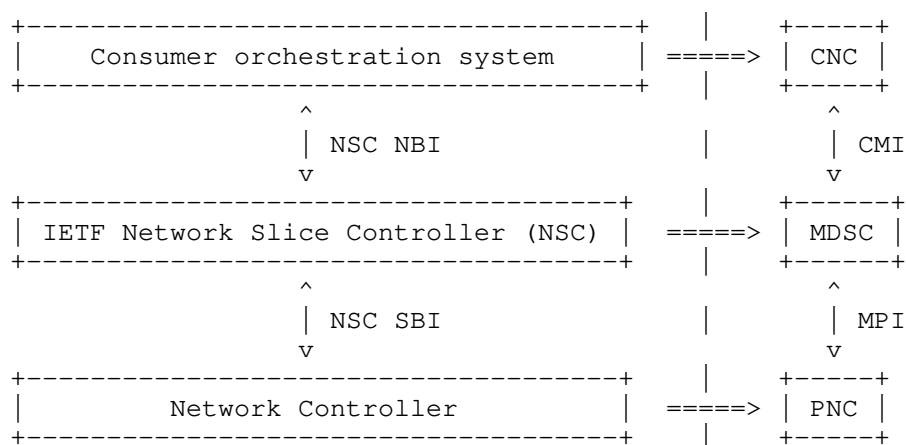


Figure 1: Mapping Between IETF Network Slice and ACTN Components

### 3.4. Examples of ACTN Delivering Types of Network Slices

The examples that follow build on the ACTN framework to provide control, management, and orchestration for the network slice life-cycle. These network slices utilize common physical infrastructure, and meet specific service-level requirements.

Three examples are shown. Each uses ACTN to achieve a different network slicing scenario. All three scenarios can be scaled up in capacity or be subject to topology changes as well as changes of consumer requirements.

#### 3.4.1. ACTN Used for Virtual Private Line

In the example shown in Figure 2, ACTN provides virtual connections between multiple consumer locations (sites accessed through Customer Edge nodes - CEs). The service is requested by the consumer (via

CNC-A) and delivered as a Virtual Private Line (VPL) service. The benefits of this model include:

- o Automated: the service set-up and operation is managed by the network provider.
- o Virtual: the private line connectivity is provided from Site A to Site C (VPL1) and from Site B to Site C (VPL2) across the ACTN-managed physical network.
- o Agile: on-demand adjustments to the connectivity and bandwidth are available according to the consumer's requests.

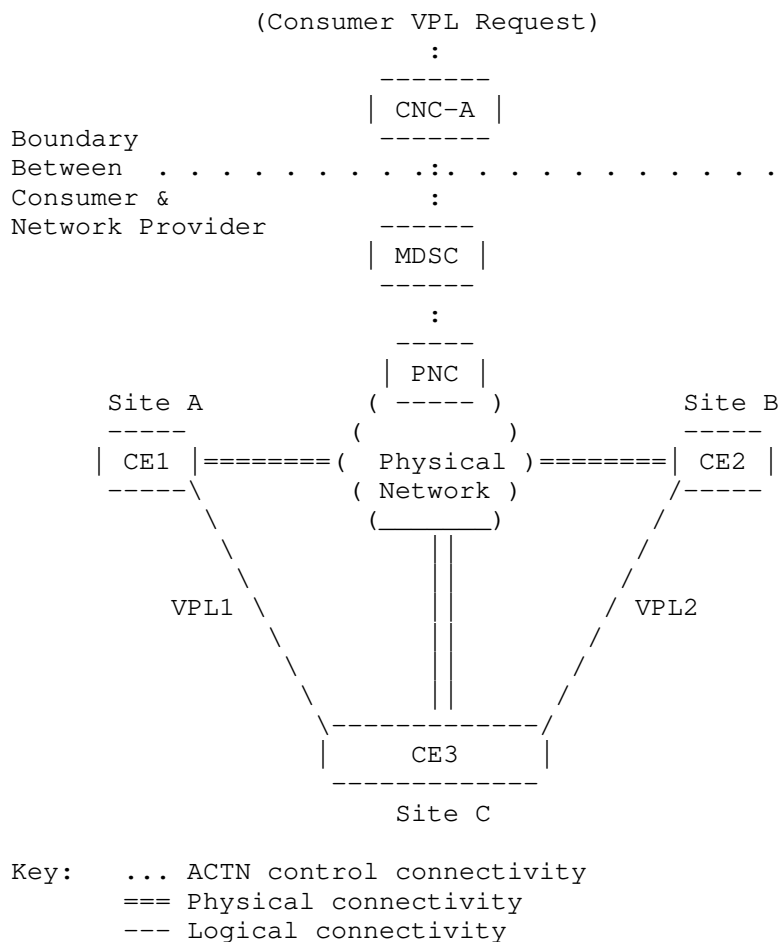


Figure 2: Virtual Private Line Model

### 3.4.2. ACTN Used for VPN Delivery Model

In the example shown in Figure 3, ACTN provides VPN connectivity between two sites across three physical networks. The requirements for the VPN are expressed by the users of the two sites who are the consumers. Their requests are directed to the CNC, and the CNC interacts with the network provider's MDSC. The benefits of this model include:

- o Provides edge-to-edge VPN multi-access connectivity.

- o Most of the function is managed by the network provider, with some flexibility delegated to the consumer-managed CNC.

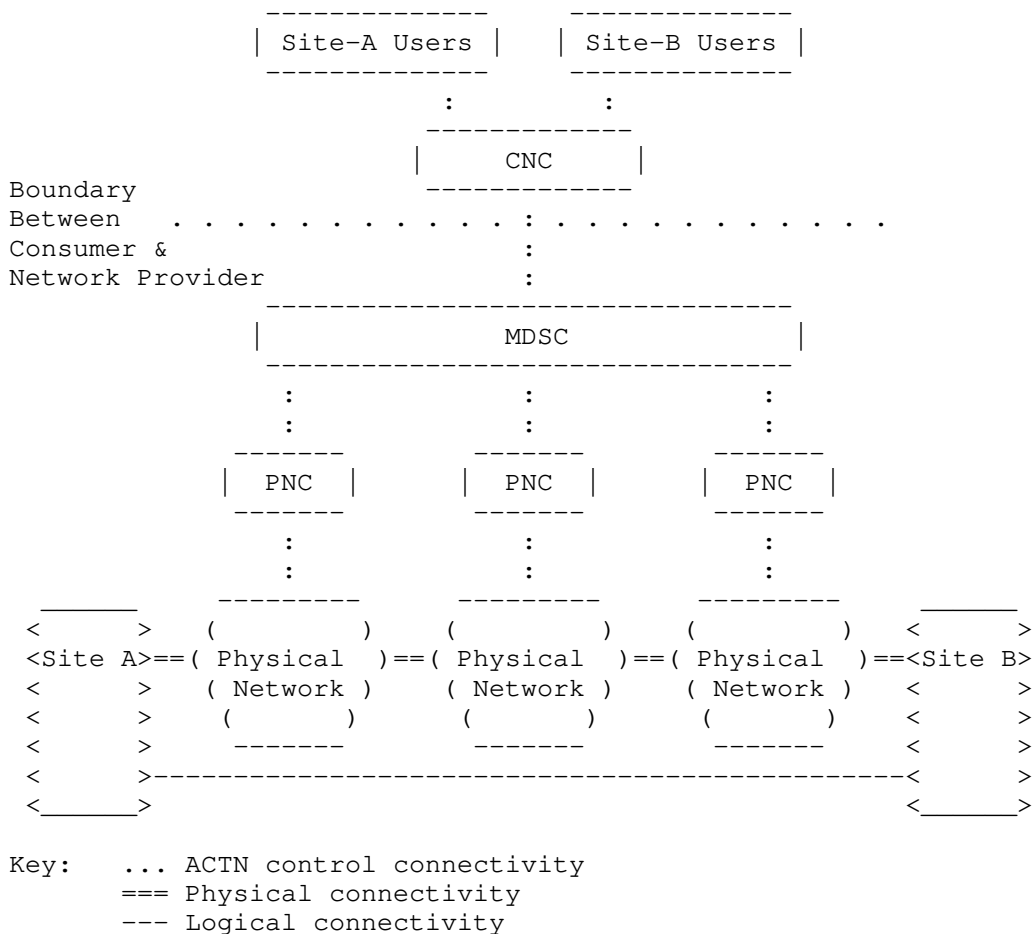


Figure 3: VPN Model

### 3.4.3. ACTN Used to Deliver a Virtual Consumer Network

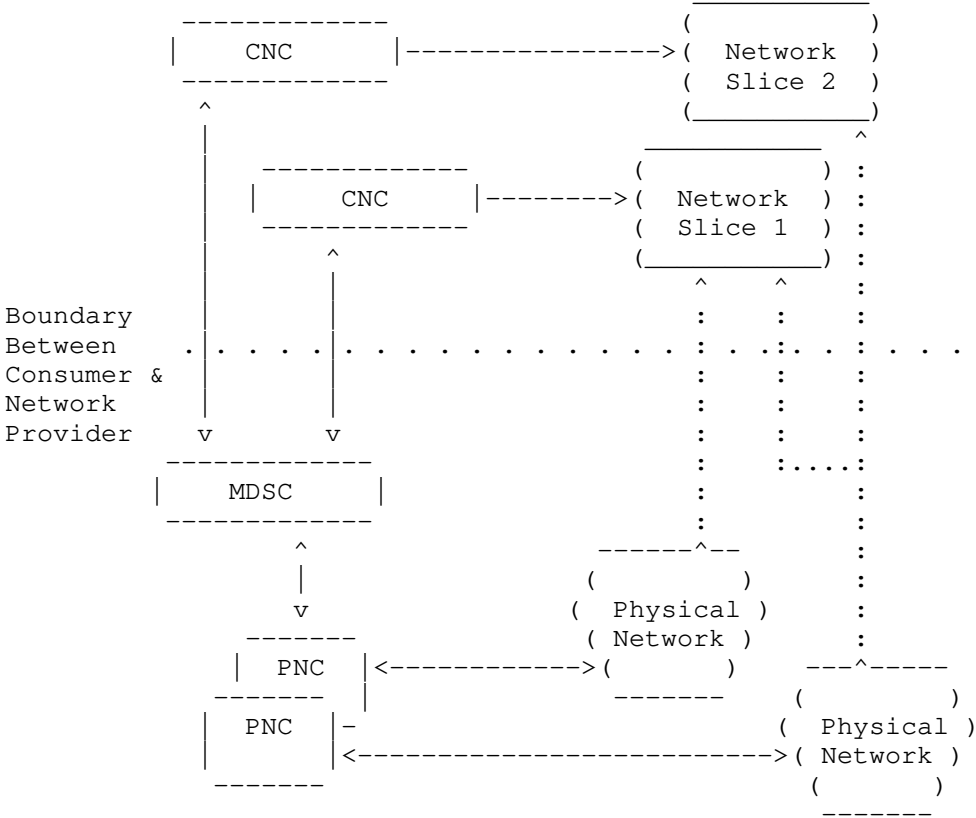
In the example shown in Figure 4, ACTN provides a virtual network to the consumer. This virtual network is managed by the consumer. The figure shows two virtual networks (Network Slice 1 and Network Slice 2) each created for a different consumer under the care of a different CNC. There are two physical networks controlled by separate PNCs. Network Slice 2 is built using resources from just

one physical network, while Network Slice 1 is constructed from resources from both physical networks.

The benefits of this model include:

- o The MDSC provides the topology to the consumer so that the consumer can control their network slice to fit their needs.
- o Applications can interact with their assigned network slices directly. The consumer may implement their own network control methods and traffic prioritization, and manage their own addressing schemes.
- o Consumers may further slice their virtual networks so that this becomes a recursive model.
- o Service isolation can be provided through selection of physical networking resources through a combination of efforts of the MSDC and PNC.
- o The network slice may include nodes with specific capabilities. These can be delivered as Physical Network Functions (PNFs) or Virtual Network Functions (VNFs).





Key: --- ACTN control connection  
... Virtualization/abstraction through slicing

Figure 4: Network Slicing

4. YANG Models

4.1. Network Slice Service Mapping from TE to ACTN VN Models

The role of the TE-service mapping model [I-D.ietf-teas-te-service-mapping-yang] is to create a binding relationship across a Layer 3 Service Model (L3SM) [RFC8299], Layer 2 Service Model (L2SM) [RFC8466], and TE Tunnel model [I-D.ietf-teas-yang-te], via the generic ACTN Virtual Network (VN) model [I-D.ietf-teas-actn-vn-yang].

The ACTN VN model is a generic virtual network service model that allows consumers to specify a VN (i.e., network slice) that meets the consumer's service objectives with various constraints on how the service is delivered.

The TE-service mapping model [I-D.ietf-teas-te-service-mapping-yang] is used to bind the L3SM with TE-specific parameters. This binding facilitates seamless service operation and enables visibility of the underlay TE network. The TE-service model developed in that document can also be extended to support other services including L2SM, and the Layer 1 Connectivity Service Model (L1CSM) [I-D.ietf-ccamp-llcsm-yang] L1CSM network service models.

Figure 5 shows the relationship between the models discussed above.

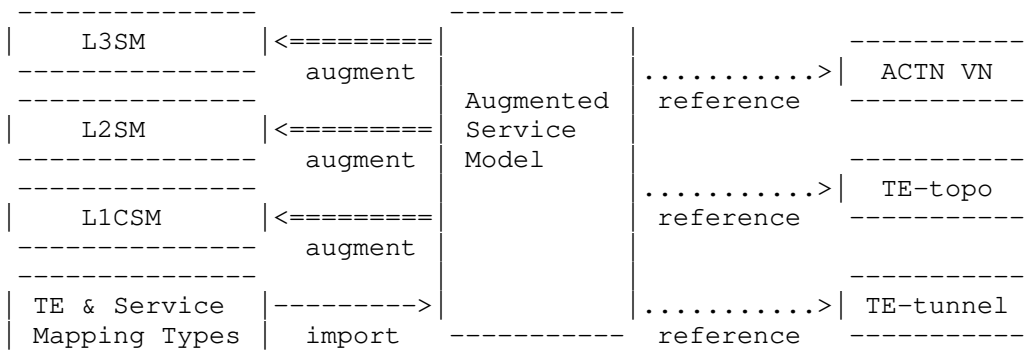


Figure 5: TE-Service Mapping

#### 4.2. Interfaces and Yang Models

Figure 6 shows the three ACTN components and two ACTN interfaces as listed in Section 3. The figure also shows the Southbound Interface (SBI) between the PNC and the devices in the physical network. That interface might be used to install state on every device in the network, or might instruct a "head-end" node if a control plane is used within the physical network. In the context of [RFC8309], the SBI uses one or more device configuration models.

The figure also shows the Network Slice Service Interface. This interface allows a consumer of a service to make requests for delivery of the service, and it facilitates the consumer modifying and monitoring the service. In the context of [RFC8309], this

"northbound interface (NBI)" is a customer service interface and uses a service model.

When an ACTN system is used to manage the delivery of network slices, a network slice resource model is needed. This model will be used for instantiation, operation, and monitoring of network and function resource slices. The YANG model defined in [I-D.wd-teas-transport-slice-yang] provides a suitable basis for requesting, controlling, and deleting, network slices.

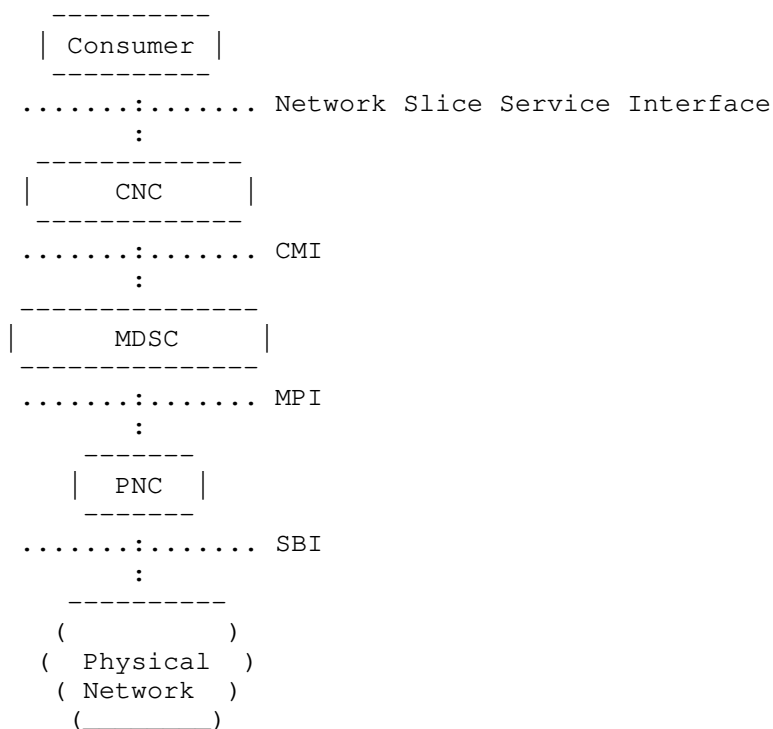


Figure 6: The Yang Interfaces in Context

#### 4.3. ACTN VN Telemetry

The ACTN VN KPI telemetry model [I-D.ietf-teas-actn-pm-telemetry-autonomics] provides a way for a consumer to define performance monitoring relevant for its VN/network slice via the NETCONF subscription mechanisms [RFC8639], [RFC8640], or using the equivalent mechanisms in RESTCONF [RFC8641], [RFC8650].

Key characteristics of [I-D.ietf-teas-actn-pm-telemetry-autonomics] include:

- o An ability to provide scalable VN-level telemetry aggregation based on consumer subscription model for key performance parameters defined by the consumer.
- o An ability to facilitate proactive re-optimization and reconfiguration of VNs/network slices based on network autonomic traffic engineering scaling configuration mechanism.

## 5. IANA Considerations

This document makes no requests for action by IANA.

## 6. Security Considerations

Network slicing involves the control of network resources in order to meet the service requirements of consumers. In some deployment models, the consumer is able to directly request modification in the behaviour of resources owned and operated by a service provider. Such changes could significantly affect the service provider's ability to provide services to other consumers. Furthermore, the resources allocated for or consumed by a consumer will normally be billable by the service provider.

Therefore, it is crucial that the mechanisms used in any network slicing system allow for authentication of requests, security of those requests, and tracking of resource allocations.

It should also be noted that while the partitioning or slicing of resources is virtual, as mentioned in Section 2.3 the consumers expect and require that there is no risk of leakage of data from one slice to another, no transfer of knowledge of the structure or even existence of other slices, and that changes to one slice (under the control of one consumer) should not have detrimental effects on the operation of other slices (whether under control of different or the same consumers) beyond the limits allowed within the SLA. Thus, slices are assumed to be private and to provide the appearance of genuine physical connectivity.

Some service providers may offer secure network slices as a service. Such services may claim to include edge-to-edge encryption for the consumer's traffic. However, a consumer should take full responsibility for the privacy and integrity of their traffic and should carefully consider using their own edge-to-edge encryption.

ACTN operates using the NETCONF [RFC6241] or RESTCONF [RFC8040] protocols and assumes the security characteristics of those protocols. Deployment models for ACTN should fully explore the authentication and other security aspects before networks start to carry live traffic.

## 7. Acknowledgements

Thanks to Qin Wu, Andy Jones, Ramon Casellas, Gert Grammel, and Kiran Makhijani for their insight and useful discussions about network slicing.

## 8. Contributors

The following people contributed text to this document.

Young Lee  
Email: [younglee.tx@gmail.com](mailto:younglee.tx@gmail.com)

Mohamed Boucadair  
Email: [mohamed.boucadair@orange.com](mailto:mohamed.boucadair@orange.com)

Sergio Belotti  
Email: [sergio.belotti@nokia.com](mailto:sergio.belotti@nokia.com)

Daniele Ceccarelli  
Email: [daniele.ceccarelli@ericsson.com](mailto:daniele.ceccarelli@ericsson.com)

## 9. Informative References

- [I-D.dong-teas-enhanced-vpn-vtn-scalability]  
Dong, J., Li, Z., Qin, F., and G. Yang, "Scalability Considerations for Enhanced VPN (VPN+)", draft-dong-teas-enhanced-vpn-vtn-scalability-01 (work in progress), November 2020.
- [I-D.ietf-ccamp-llcsm-yang]  
Lee, Y., Lee, K., Zheng, H., Dios, O., and D. Ceccarelli, "A YANG Data Model for L1 Connectivity Service Model (L1CSM)", draft-ietf-ccamp-llcsm-yang-13 (work in progress), November 2020.

- [I-D.ietf-teas-actn-pm-telemetry-autonomics]  
Lee, Y., Dhody, D., Karunanithi, S., Vilata, R., King, D.,  
and D. Ceccarelli, "YANG models for VN/TE Performance  
Monitoring Telemetry and Scaling Intent Autonomics",  
draft-ietf-teas-actn-pm-telemetry-autonomics-04 (work in  
progress), November 2020.
- [I-D.ietf-teas-actn-vn-yang]  
Lee, Y., Dhody, D., Ceccarelli, D., Bryskin, I., and B.  
Yoon, "A YANG Data Model for VN Operation", draft-ietf-  
teas-actn-vn-yang-10 (work in progress), November 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+)  
Service", draft-ietf-teas-enhanced-vpn-06 (work in  
progress), July 2020.
- [I-D.ietf-teas-ietf-network-slice-definition]  
Rokui, R., Homma, S., Makhiyani, K., Contreras, L., and J.  
Tantsura, "Definition of IETF Network Slices", draft-ietf-  
teas-ietf-network-slice-definition-00 (work in progress),  
January 2021.
- [I-D.ietf-teas-ietf-network-slice-framework]  
Gray, E. and J. Drake, "Framework for IETF Network  
Slices", draft-ietf-teas-ietf-network-slice-framework-00  
(work in progress), March 2021.
- [I-D.ietf-teas-rfc3272bis]  
Farrel, A., "Overview and Principles of Internet Traffic  
Engineering", draft-ietf-teas-rfc3272bis-10 (work in  
progress), December 2020.
- [I-D.ietf-teas-te-service-mapping-yang]  
Lee, Y., Dhody, D., Fioccola, G., WU, Q., Ceccarelli, D.,  
and J. Tantsura, "Traffic Engineering (TE) and Service  
Mapping Yang Model", draft-ietf-teas-te-service-mapping-  
yang-05 (work in progress), November 2020.
- [I-D.ietf-teas-yang-te]  
Saad, T., Gandhi, R., Liu, X., Beeram, V., and I. Bryskin,  
"A YANG Data Model for Traffic Engineering Tunnels, Label  
Switched Paths and Interfaces", draft-ietf-teas-yang-te-25  
(work in progress), July 2020.

- [I-D.wd-teas-transport-slice-yang]  
Bo, W., Dhody, D., Han, L., and R. Rokui, "A Yang Data Model for Transport Slice NBI", draft-wd-teas-transport-slice-yang-02 (work in progress), July 2020.
- [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>>.
- [RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <<https://www.rfc-editor.org/info/rfc7665>>.
- [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>>.
- [RFC8299] Wu, Q., Ed., Litkowski, S., Tomotaki, L., and K. Ogaki, "YANG Data Model for L3VPN Service Delivery", RFC 8299, DOI 10.17487/RFC8299, January 2018, <<https://www.rfc-editor.org/info/rfc8299>>.
- [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>>.
- [RFC8466] Wen, B., Fioccola, G., Ed., Xie, C., and L. Jalil, "A YANG Data Model for Layer 2 Virtual Private Network (L2VPN) Service Delivery", RFC 8466, DOI 10.17487/RFC8466, October 2018, <<https://www.rfc-editor.org/info/rfc8466>>.
- [RFC8639] Voit, E., Clemm, A., Gonzalez Prieto, A., Nilsen-Nygaard, E., and A. Tripathy, "Subscription to YANG Notifications", RFC 8639, DOI 10.17487/RFC8639, September 2019, <<https://www.rfc-editor.org/info/rfc8639>>.

- [RFC8640] Voit, E., Clemm, A., Gonzalez Prieto, A., Nilsen-Nygaard, E., and A. Tripathy, "Dynamic Subscription to YANG Events and Datastores over NETCONF", RFC 8640, DOI 10.17487/RFC8640, September 2019, <<https://www.rfc-editor.org/info/rfc8640>>.
- [RFC8641] Clemm, A. and E. Voit, "Subscription to YANG Notifications for Datastore Updates", RFC 8641, DOI 10.17487/RFC8641, September 2019, <<https://www.rfc-editor.org/info/rfc8641>>.
- [RFC8650] Voit, E., Rahman, R., Nilsen-Nygaard, E., Clemm, A., and A. Bierman, "Dynamic Subscription to YANG Events and Datastores over RESTCONF", RFC 8650, DOI 10.17487/RFC8650, November 2019, <<https://www.rfc-editor.org/info/rfc8650>>.

## Authors' Addresses

Daniel King  
Old Dog Consulting  
  
Email: [daniel@olddog.co.uk](mailto:daniel@olddog.co.uk)

John Drake  
Juniper Networks  
  
Email: [jdrake@juniper.net](mailto:jdrake@juniper.net)

Haomian Zheng  
Huawei Technologies  
  
Email: [zhenghaomian@huawei.com](mailto:zhenghaomian@huawei.com)

Adrian Farrel  
Old Dog Consulting  
  
Email: [adrian@olddog.co.uk](mailto:adrian@olddog.co.uk)



None  
Internet-Draft  
Intended status: Informational  
Expires: April 21, 2018

K. Makhijani, ed  
J. Qin  
R. Ravindran  
Huawei Technologies  
L. Geng  
China Mobile  
L. Qiang  
S. Peng  
Huawei Technologies  
X. de Foy  
A. Rahman  
InterDigital Inc.  
A. Galis  
University College London  
G. Fioccola  
Telecom Italia  
October 18, 2017

Network Slicing Use Cases: Network Customization and Differentiated  
Services  
draft-netslices-usecases-02

Abstract

Network Slicing is meant to enable creating (end-to-end) partitioned network infrastructure that may include the user equipment, access/core transport networks, edge and central data center resources to provide differentiated connectivity behaviors to fulfill the requirements of distinct services, applications and customers. In this context, connectivity is not restricted to differentiated forwarding capabilities but it covers also advanced service functions that will be invoked when transferring data within a given domain.

The purpose of this document is to focus on use cases that benefit from the use of network slicing.

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 April 21, 2018.

#### Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

#### Table of Contents

1. Introduction . . . . .	3
1.1. Requirements Language . . . . .	3
1.2. Terminology . . . . .	3
2. Scope . . . . .	4
3. A Generalized Network Slice as a Service . . . . .	6
3.1. Resource Centric Service Concept . . . . .	6
3.2. Strict Resource Demand . . . . .	7
3.3. Network Customization . . . . .	7
3.4. NSaaS of Different Granularity . . . . .	7
3.5. Service customization across multi-provider multi-domains as NSaaS . . . . .	8
4. Network Slicing in 3GPP Mobile Network . . . . .	10
4.1. Network Slices in 3GPP Systems . . . . .	10
4.2. Creating, Managing and Operating 3GPP Network Slices . .	11
5. Role of Virtualization in Network slicing . . . . .	12
5.1. Virtualized Customer Premise Equipment . . . . .	12
5.2. Enhanced Broadband . . . . .	14
6. Services with Resource Assurance . . . . .	16
6.1. Massive Machine to Machine Communication . . . . .	16
6.2. Ultra-reliable Low Latency Communication . . . . .	18
6.3. Critical Communications . . . . .	20
7. Network Infrastructure for new technologies . . . . .	23
7.1. ICN as a Network Slice . . . . .	23
7.2. New Verticals - ICN based service delivery . . . . .	24

7.2.1. Required Characteristics . . . . .	25
8. Overall Use Case Analysis . . . . .	26
8.1. Requirements Reference . . . . .	26
8.2. Mapping Common characteristics to Requirements . . . . .	26
9. Conclusion . . . . .	28
10. Security Considerations . . . . .	28
11. IANA Considerations . . . . .	29
12. Acknowledgements . . . . .	29
13. References . . . . .	29
13.1. Normative References . . . . .	29
13.2. Informative References . . . . .	30
Authors' Addresses . . . . .	31

## 1. Introduction

Network Slicing enables the creation of (end-to-end) partitioned network infrastructure that may include the user equipment, access/core transport networks, edge and central data center resources to provide differentiated connectivity behaviors to fulfill the requirements of distinct services, applications and customers. In this context, connectivity is not restricted to differentiated forwarding capabilities but it also spans service, management and control plane support offered to a slice instance.

### 1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

### 1.2. Terminology

Please refer to [I-D.geng-netslices-architecture] for related terminologies and definitions.

Additionally, the following terms are used:

- o V2X (Vehicle-to-everything): Is a communication of information from a vehicle to any other entity that may be another vehicle, road-side network element or application end point.
- o ITS (Intelligent Transportation Systems): Considered as an aspect of how using Internet of Things resource like road sensors can creates a smart transport network. The network offers services related to transport and traffic management systems through flow of information between road-side sensors, vehicles, smart devices and humans.

- o Over-the-top (OTT): A service, e.g., content delivery using a CDN or a social networking service, operated by a different service providers to which the users of the NSP service are attached to, and to whom it serves as a communication (or bit pipe) provider
- o Industry vertical: A collection of services or tools specific to an industry, trade or market sector. also, referred to as Service Verticals in this document.
- o TETRA: Terrestrial trunked radio is a digital trunked mobile radio standard to meet needs of public safety, transportation and utilities like organizations.
- o SLA: Service Level Agreement - A contract between a service provider and an end user that stipulates a specified level of service, support option, a guaranteed level of system performance as relates to downtime or up-time.

## 2. Scope

To maximize resource utilization and minimize infrastructure cost, services will need to operate over a shared network infrastructure, as against the traditional monolithic model operated either as dedicated network or as an overlay. Service operators can utilize or benefit from Network Slicing through multi-tenancy, enabling different customized network infrastructures for different group of services across different network domains and operating them independently.

In this document, multi-domain refers to combination of different kinds of connection-technology network domains. For example, it may be a RAN, DSL etc. in access; a fixed, wireless or mobile service provider network; as well as different technology domains, in transport networks such as carrier Ethernet, optical, MPLS, TE-tunnel etc. Often, a combination of technology domains is under the same administrator's control but may also belong to different administrative systems and may require cross-domain coordination.

The document covers generalized as well as resource guaranteed service scenarios that can benefit by applying Network Slicing principles as below in Figure 1

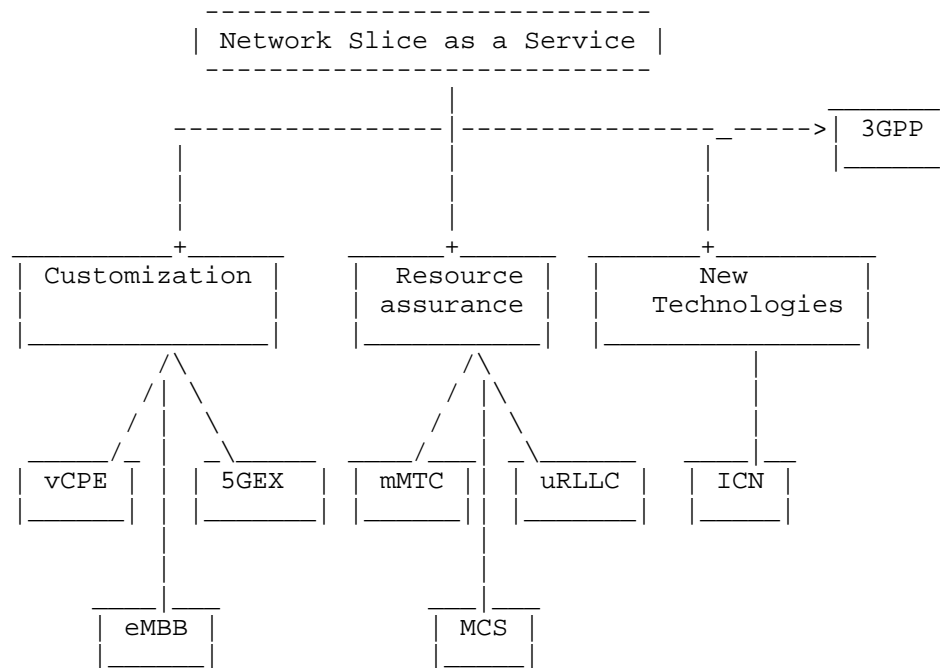


Figure 1: Use case organization in the document

The remaining document is organized as below:

- o In Section 3, Network Slice as a Service(NSaaS) delivery model is described.
- o Section 3.5, is a scenario for multi-domain network slice coordination.
- o In Section 4, 3GPP architecture for 5G is discussed as a use case so that those requirements may be taken into consideration during slicing activities in IETF.
- o Other use cases are discussed from 2 perspectives
  - a Existing scenarios: Several already deployed use cases that would further benefit operators when deployed through Network Slice paradigm are discussed in Section 5.
  - b Differentiated service scenarios: that must absolutely meet strict resource requirements, as if they use a dedicated

infrastructure. The example use cases are categorized in Section 6.

- o Section 7, has an example use case of cases where new technologies can be verified or deployed using network slicing concept.
- o In Section 8, the use case requirements are summarized which are inputs to the [I-D.qiang-netslices-gap-analysis].

### 3. A Generalized Network Slice as a Service

Network slicing instances share a common infrastructure, which provide flexible design of a logical network with specific network functions customized to support differentiated performance requirements of vertical industry through logical or physical system isolation and certain OAM tools.

Traditionally, vertical industries run their services in a shared network environment upon which infrastructure owners and service providers offer standalone network capabilities including connections, storage and etc. Network slicing paradigm enables supporting the requirements of a network slicing tenant to be met individually. Hence it is anticipated that this type of new business model where network slice instances are leased to industry verticals as a service (i.e. Network Slicing as a Service, NSaaS) may become a norm in the near future.

#### 3.1. Resource Centric Service Concept

Network services specify a set of resource requirements to offer desired Quality of Experience (QoE) to its consumers, using features offered by the control and forwarding planes. Traditional service guarantees are associated with resource attributes such as throughput, packet loss, latency, network bandwidth/burst or other bit rates and security. In addition, redundancy and reliability are provided by the infrastructure to improve overall QoE. More recently, concepts such as edge computing allow opportunistic placement of services to meet stringent requirements of low latency and/or high bandwidth applications.

Clearly the description of service delivery is more diverse now than before and demands higher degree of resource engineering and agility. The motivation behind Network slicing paradigm is to enable new service deployments without having to build new network infrastructures or causing disruptions to already deployed services in the network. In this regard, there are two primary characteristics NS should satisfy, a) Strict demand for network resource, b) Network Customization.

### 3.2. Strict Resource Demand

Several services are sensitive to response times and/or amount of bandwidth, e.g. real time interactive multimedia, high bandwidth video feed or remote access to an enterprise network. Failure to meet these criteria lead to service degradation. Moreover, new industry verticals are evolving due to technological advancements in sensors, IoT, robotics and multi-media, along with new type of network interactions (both human-human or human-machine). These impose even stricter resource and connectivity requirements. The challenge lies in utilizing common network infrastructure and judiciously allocating available infrastructure resources.

### 3.3. Network Customization

To a network slice tenant, the ability to customize services dynamically is important. Customization gives control to the operator of a slice to create, provision and change network resources to suit their service demands. Customization enables decomposition of resources from an underlying network infrastructure and logically aggregate them as part of a slice. These customizations also include placement and logical connection of the network functions based on the service requirements.

### 3.4. NSaaS of Different Granularity

In order to meet various requirements from the network slice tenant, NSaaS should be provided with different granularities. Some typical examples of granularities that a provider may offer are as follows.

- o Network Domain - Network slice instances of different networks i.e. access (wireless, fixed) network, transport network and core network.
- o Access technologies - Network slice instances of different generations of cellular and fixed network technologies, i.e. 4G, WiFi, Passive Optical Network (PON) and DSL.
- o SLA requirements - Network slice instances of different SLA requirements, i.e. low-latency network, legacy best-effort network and network with guaranteed-bandwidth.
- o Vertical applications - Network slice instances of different industry verticals. i.e. manufacturing site, V2X, industrial IoT and smart city.

- o OTT services - Network slice instances of different applications provided by OTT, i.e. messaging, payment, video streaming and gaming.
- o Cross domain services - Network slice instances of different services across multi-provider domains such as L2, L3 VPN services.

During the realization of network slice instance, it is also very important that sub-instance of a more general one can be provided with a finer granularity. In practice, it is up to the provider to decide the granularity to lease the network slice instances.

The customization of different granularities of a network slice introduce many challenges, especially in terms of network management and orchestration. As a network slice provider (provider of end-to-end slice service), it is essential to have a comprehensive understanding of the network capability. This requires that network connectivity and resources can be exposed to the network slice tenants (as the differentiated services). Accordingly, network slice provider is able to orchestrate specific instances based on these exposed capabilities.

### 3.5. Service customization across multi-provider multi-domains as NSaaS

L2 and L3 connectivity services can be deployed in a multi-provider multi-domain scenario and, in the SDN era, this implies the decoupling of network resources for different service provider and domain orchestrators. The allocation of network resources within the domain of each service provider, involved by the end-to-end service, can be defined as a network slice.

Within a single domain, provider is aware of the entire topology and its own resource availability and has complete control over those resources. However, in a multi domain scenario, the overall knowledge of the resources and topologies cannot be made across providers. Therefore, the exchange of information across these providers have to be enabled, as shown in Figure 2, inspired by [I-D.bernardos-nfvrg-multidomain] and [I-D.ietf-opsawg-service-model-explained].



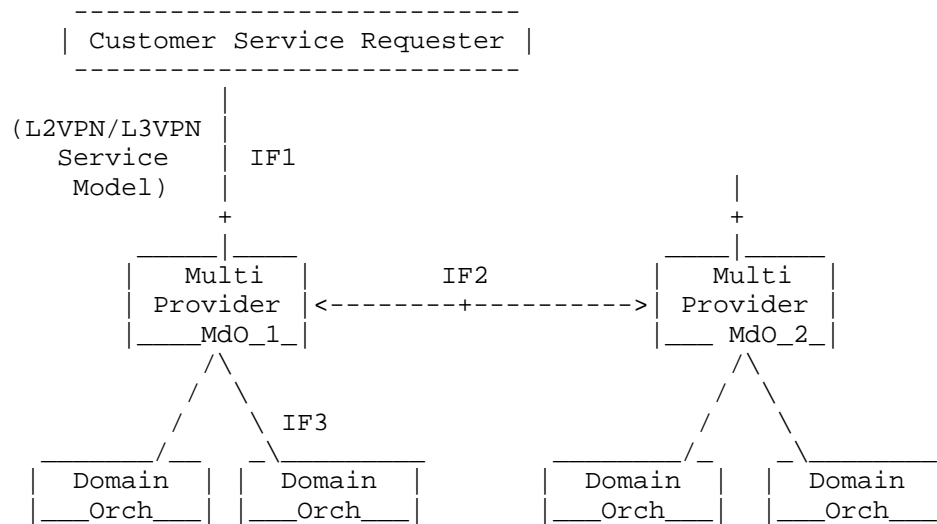


Figure 2: Multi-domain, multi-provider connectivity services

The Figure 2 shows a multi provider MdO (MP-MdO) exposing an interface 1 (IF1) to the tenant, interface 2 (IF2) to other multi provider MdO (multi domain orchestrator) and an interface 3 (IF3) to individual domain orchestrators. IF1 is exposed to the tenant who could request his specific services and/or slices to be deployed. IF2 is between the orchestrators and is a key interface to enable multi-provider operation. IF3 focuses on abstracting the technology or vendor dependent implementation details to support orchestration in each network domain (see [!@I-D.bernardos-nfvrg-multidomain] for details). The coordination alternatives between MP-MdOs are: \*

- \* Bilateral Cascading: providers can have long-lasting business agreements only with their direct neighbors.
- \* Full Mesh between MP-MdOs: Providers can have long-lasting business agreement with any provider (neighboring or remote).

This reference architecture is the main focus of the 5GEx European Project.

Among applications, L2VPN and L3VPN wholesale end-to-end services in a multi-provider and multi-domain scenario needs the following characteristics for network slice management

- o An automatic activation test and verification functionality (by customer or orchestrator).

- o Interface to modify parameters of L2VPN or L3VPN service such as bandwidth or path redundancy.

Looking at Figure 2, the customer needs a new L2 end-to-end service between CPEs across two domains (MP-Md01 and MP-Md02). As MP-Md01 receives the service request, it is deployed as a network slice. In this regards MP-Md01 has prior knowledge of topology and resource across domains in some form, it then splits the service request into a slice across each of the involved domain. Once service is set up: MP-Md01 allocates resources for the slice on SP1 domain while MP-Md02 allocates on SP2 domain respectively.

[I-D.ietf-l2sm-l2vpn-service-model] and [RFC8049] can describe IF1 For L2 and L3 end-to-end services respectively. The ability to map such services as network slice will be considerable opportunity for dynamic cross-domain operations.

#### 4. Network Slicing in 3GPP Mobile Network

Network Slicing is a core capability of the currently under development 3GPP 5G phase 1 mobile system, as it makes it possible for different service verticals, such as IoT and broadband applications, to be deployed over a common shared infrastructure. More details can be found in [TS\_3GPP.23.501], [TS\_3GPP.23.502], [TR\_3GPP.38.801], [TR\_3GPP.33.899] and [TS\_3GPP.28.500].

3GPP is currently defining its own solution for network slicing. An IETF effort in this field may, however, still be complementary in the long run as IETF focuses on the IP infrastructure and protocols which are generally out of scope of 3GPP. Challenges relevant to the IETF include isolation between network slices, supporting sharing network functions between several slices, building slices recursively from smaller slice subnets, implementing slicing across different domains for roaming, etc.

##### 4.1. Network Slices in 3GPP Systems

In 3GPP systems a network slice is a complete logical network which provides telecommunication services and network capabilities. Distinct Radio Access Network (RAN) slices and core network slices interwork to provide mobile connectivity. A device may access multiple NS simultaneously through a single RAN. 3GPP defines slice IDs (NSSAI) composed of a Slice Service Type (SST) and a Slice Differentiator (SD). SST refers to an expected network behavior in terms of features and services (e.g. specialized for broadband or massive IoT), while SD helps distinguishing among several NS instances.

Figure 3 describes the general layout of Network Slicing in mobile networks. A core network slice includes, a Session Management Function (SMF), which manages PDU sessions, and a User Plane Function (UPF). Some functions such as the Access and Mobility management Function (AMF) are common and shared between multiple RAN and core network slices.

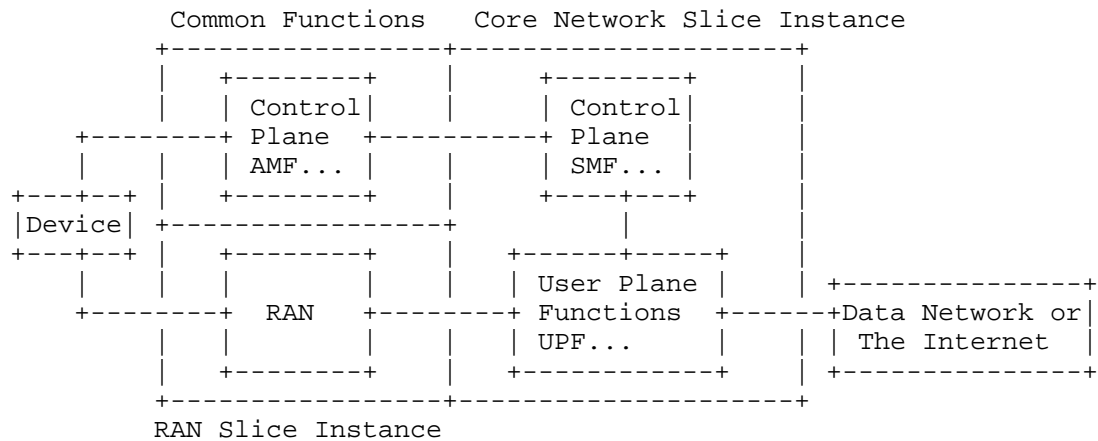


Figure 3: 3GPP Network Slices

#### 4.2. Creating, Managing and Operating 3GPP Network Slices

To create a network slice instance, mobile network operators define "Network Slice Subnets" into OSS/BSS management system. NS subnets are NS components including NFs and reserved network resources. OSS/BSS communicates with the orchestrator, which, through the rest of the NFV-MANO system, configures compute and network elements to create, compose and activate slices.

Mobile network operators can modify the configuration of a RAN or core network slice, while it is in use. To support this, the operator needs to measure QoS/SLA data for hosted network services, and associate results with the relevant network slice. Example of operations include increase or decrease network capacity or compute capacity of NFs; update the configuration of NFs; add, replace or remove a NFs or a Network Slice Subnet.

Slice selection occurs in 2 phases: first, common functions (including AMF) and available network slices are pre-selected when the device registers with the network. Later on, the network dynamically selects network slices when a device initiates

communication, based on a slice ID associated with the application (on the device) that requests a new flow.

## 5. Role of Virtualization in Network slicing

Virtualization is a key enabler of network slices; Many network services can be easily deployed using components of NFV framework like network functions, hardware decoupling and resource placement [1]. When deployed as a network slice, the resources associated with virtualized network services are managed uniformly by network slice provider. One such use case is described below.

### 5.1. Virtualized Customer Premise Equipment

A CPE is an equipment that connects the customer premises to the provider's network. A CPE may either be a layer-2 or a layer-3 device (the routing gateway) performing different network functions depending on the access technology (DSL modem, PON modem, etc.). Any services provided such as Internet access, IPTV, VoIP, etc. or network functions for example, local NAT, local DHCP, IGMP proxy-routing, PPP sessions, routing, etc. are also part of CPE. The installation of different on-premise devices, entails a high cost for service providers in terms of both initial installation and operational support, since they are typically responsible for the end-to-end service.

Traditional CPE deployments are service provider network functions installed on customer site to provide above mentioned functionalities along with remote site connectivity. Communication Service provider (CSP) is responsible for management and administration of connections and state with proper policy, bandwidth, security and QoS requirements.

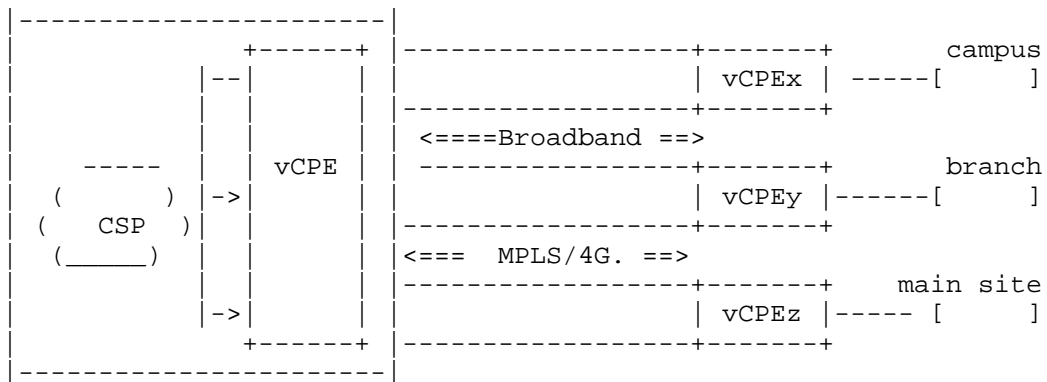


Figure 4: Virtualized CPE with distributed architecture

Figure 4 shows a virtualized architecture in which many functions are moved to CSP's cloud simplifying CPE on premises tremendously. Additional details of deployment architecture models are captured in [I-D.pularikkal-virtual-cpe] where full dissemination of data path and control plane functions is described. The figure shows vCPE<sub>x</sub>, vCPE<sub>y</sub>, vCPE<sub>z</sub> are virtualized CPEs on multiple sites of a specific customer, there may be set of different network functions in each x, y and z CPE. The vCPE instance in CSP cloud is integrated to each site performing service chains of network functions and resource allocations specific for ingress and egress path of each site.

A vCPE is a well-known concept[VCPEBBF] which when combined with WAN technologies provides end to end visibility and reachability to remote sites. However, there is no standard approach to connectivity or management of various CPE functions. Using network slicing, a greater level of agility can be achieved, with each customer dynamically managing its own network with the assistance of network slicing framework.

The benefit of self-managing a vCPE network slice is the capability to move network functions on premise or to the cloud. An obvious use case will be customer initiated gradual migration of network functions from a site to CSP cloud.

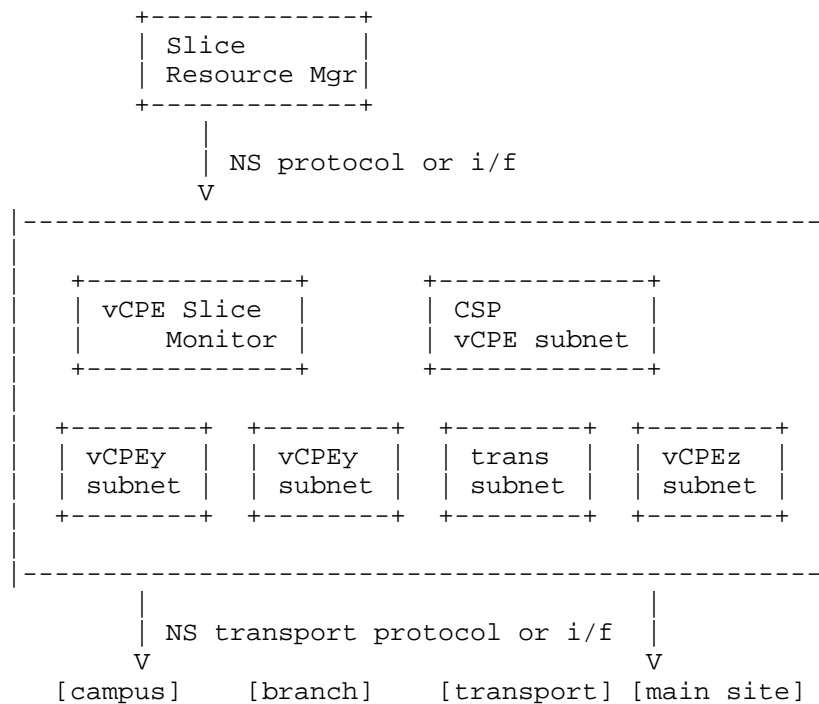


Figure 5: vCPE as a Network Slice

In Figure 5, a slice for vCPE is shown. Using slice subnet approach, each vCPE site instance may be considered as an abstracted subnet, along with the WAN transport as another subnet. The network functions are chained in a distributed fashion between site vCPEs and CSP vCPE subnet. A monitoring function interfaces with CSP's global slice manager for resource management. A south-bound interface through network slice transport protocol, realizes these functions on the infrastructure.

## 5.2. Enhanced Broadband

Today, video consumes the largest amount of bandwidth over the Internet. As the higher resolution formats enter mainstream, even more bandwidth will be needed to stream 4K/8K/360 degree formats. For example, connected Virtual Reality(VR)/Augmented Reality(AR) is the future use case of eMBB services. Notably, media processing for AR/VR will require in-network processing functions and high latencies between components could lead to downgrade of user experience. Therefore, an AR/VR stream requires a special infrastructure that differs from best-effort network.

A purpose-built network slice for eMBB streaming shall ensure to minimize processing overheads, it may be done by placement of network functions closer to subscribers. Resource scaling for eMBB should be dynamic because bandwidth is expensive and such vertical service operators may not want to pay for unutilized bandwidth. Therefore, slices should be able to monitor, negotiate and adjust the scale for both bandwidth and service functions. Latency guarantees vary from general services, therefore, as a first step, monitoring for quality of service is needed and more advanced operation would involve recovery and reparation of paths.

A typical eMBB slice Figure 6 from a network operator is a performance oriented service customization. An eMBB service slice template will allow a tenant to request or specify

- (1) CDN components (as service functions)
  - \* Regional network locations of CDN, encoders etc.
  - \* Location of acquired content.
  - \* Describes transport constraints for its own distribution network comprising of connectivity between content acquisition and Fan-out points.
- (2) An interface to subscriber database perhaps as a network function, from multiple access network types (cellular, fixed).
- (3) Live performance monitoring and resource negotiation loop.
- (4) A well-coordinated network slice protocol that enables resource allocation across different network domains.

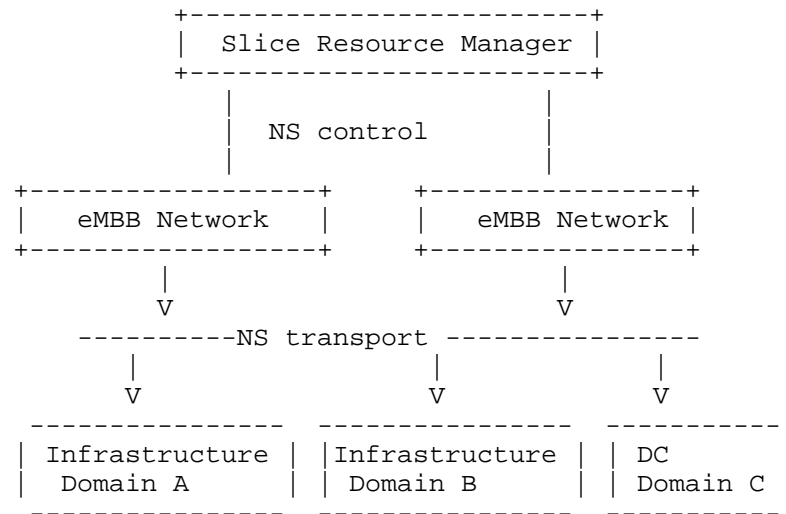


Figure 6: Transport provider network operator view.

## 6. Services with Resource Assurance

### 6.1. Massive Machine to Machine Communication

Sensor networks are widely deployed in industries such as agriculture, environmental monitoring and manufacturing. The general workflow of wireless sensor network is provided in Figure 7.



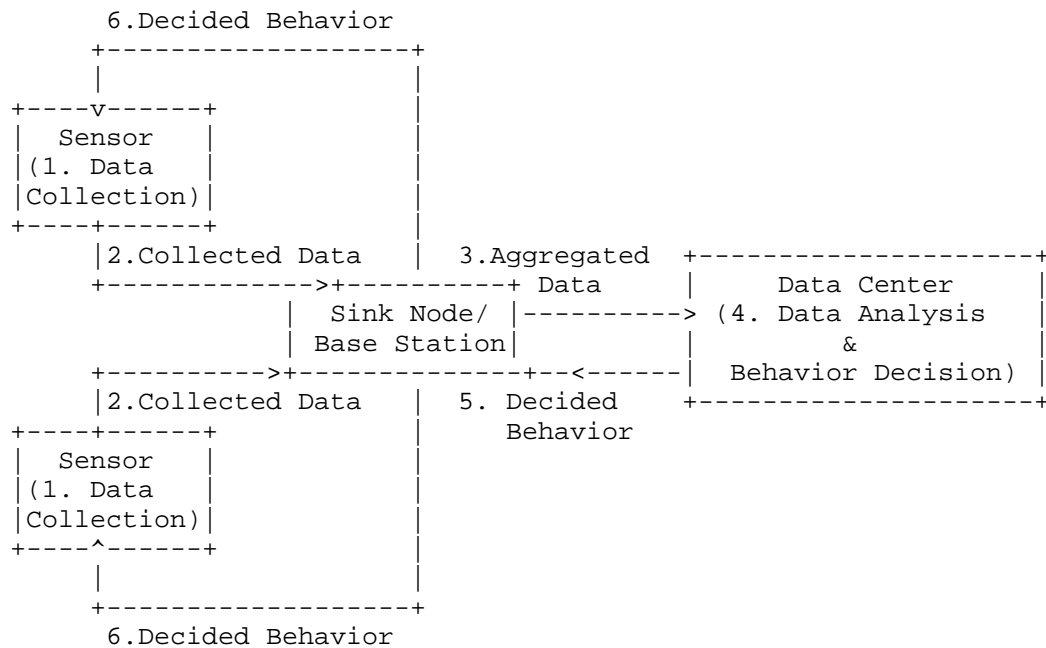


Figure 7: Workflow of wireless sensor network

Figure 7 shows, control of sensor data & behavior at scale, requiring wide area coverage and power constrained communication. A few new types of scenarios that require unique infrastructure are:

- o Smart city networks: an integration of several public infrastructures together through M2M communications. For example Automatic metering (for gas, energy, water, etc.), environment monitoring (for pollution, temperature, humidity, etc.), traffic signal control etc.
- o E-health communications that remote monitor the physical conditions (e.g., heart rate, pulse, blood pressure etc.), and accordingly take necessary measures remotely. E-health communication network must be secure, reliable and fast but small-size of data exchange.

mMTC Type Slices involves potentially a large number of small and power-constrained devices, therefore, resource allocation at scale is of particular importance in mMTC type slices. Furthermore, different kind of IoT devices may exhibit delay sensitivity in industry operations etc. The mMTC type slices should be conscious of requirements of scale, variable data pattern, and energy efficient communications.

## 6.2. Ultra-reliable Low Latency Communication

In uRLLC scenarios, data loss is not acceptable. Both data and control planes may require significant enhancements to transmission or information distribution protocols. [TR\_3GPP\_38.913] specifies access network user plane latency as 1ms and reliability factor of 99.999% for transmission of a packet of size 32 bytes. The slices of this type must be ensured that shared infrastructure absolutely does not cause any adverse effects.

In the following sections three new uRLLC scenarios are described.

- (1) Industrial operation: Operations in remote sites usually need combined support of cellular and transport network. Operational accuracy is characterized by
  - \* Requires high-quality communication links between the control site.
  - \* Low latency and low jitter in communication path
  - \* Closed control loop (Sensor -Controller - Actuator) as shown in Figure 8, a typical control cycle time where network is involved should be below 10ms [Tactile-Internet].

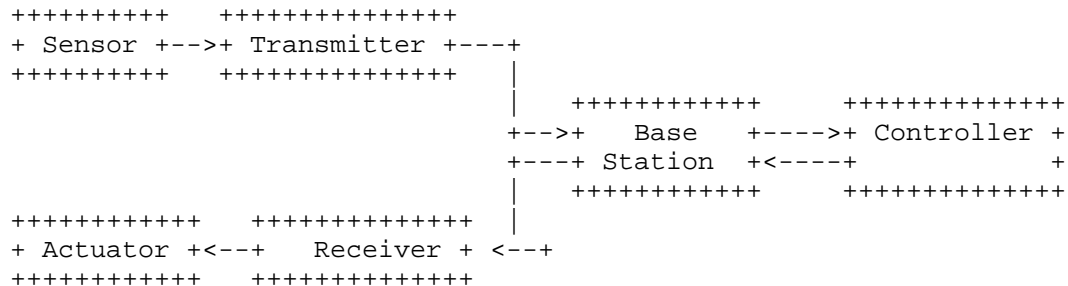


Figure 8: Industrial closed control loop

- (2) Remote surgery enables surgeons to perform critical specialized medical procedures remotely, providing accurate control and haptic feedback.

A uRLLC network slice only accepts service specific traffic and must not receive any other type of traffic to avoid negative impact on the service operation. Capabilities required by uRLLC service provider include

- o Locations of the access nodes for terminals (devices, vehicles) to the transport network and locations of the controller to construct its own network topology within the network slice. In high mobility scenario such as automotive verticals, the dynamic topology adjustments are required without loss of data.
- o Each service vertical has different performance requirements in terms of latency, reliability and data rate etc., therefore, the uRLLC network slice should allow customization for these parameters.
- o A uRLLC service provider should be able to registers self with access rights to resource monitoring and negotiation loop.

A network slice provider offers a uRLLC Slice with the following considerations

- o Should support/provide specific data and control planes protocols with significant enhancements for deterministic latency and reliability (e.g. DetNet[I-D.dt-detnet-dp-sol] in data plane).
- o Allow uRLLC service operator to access user admission and authentication to its network slice in advance.
- o The network coverage for a uRLLC service provisioning may be limited to a confined area, either indoor or outdoor, network operator needs to be able to coordinate resource allocation across different access types and network domains.

A high-level Figure 9, shows a uRLLC slice provider and service view of the network. The monitoring of resources is done in the context of performance. A performance degradation would require resource adjustment. As shown in Figure 9, in one possible sliced model will have its own customizer that uses internal performance observing logic with in its slice by coordinating with different subnets/ domains using southbound NS transport protocol and transfers this information to operator via a northbound NS protocol for resource adjustment.

It is implied that domains maybe different access technologies and need for a common performance metric propagation and resource allocation is important for a uRLLC slice to function properly.

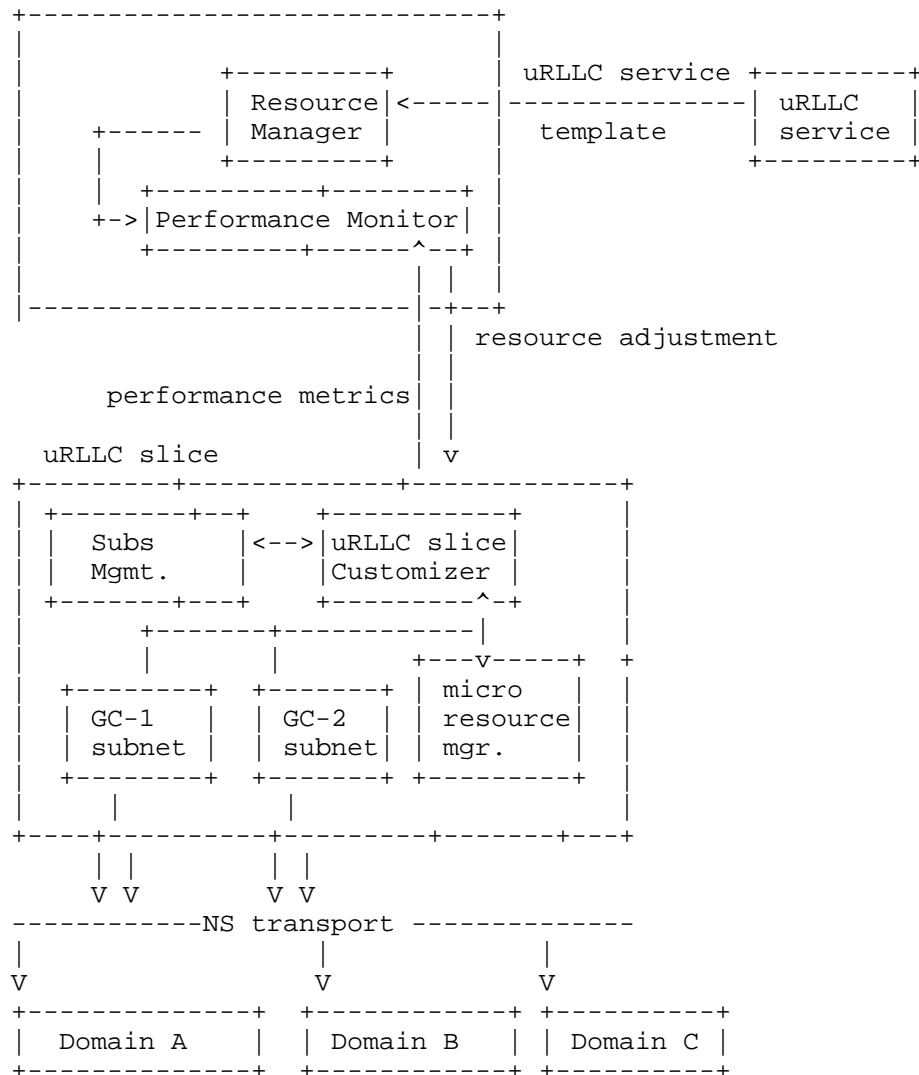


Figure 9: Reference for uRLLC Network Slice.

### 6.3. Critical Communications

Critical communications are associated with emergency situations. Often referred to as mission critical, the communication has to be reliable and non-disruptive. Different scenarios of critical communications relate to public safety responders (e.g. fire fighters, paramedics), military, utility or commercial applications,

mainly using reliable voice or short data messaging over wireless communication systems.

Next-generation public safety communications are planned to be built with enhanced broadband voice, data and video communications services beyond narrowband LMR with broadband LTE networks for high speed data (ref 22.179 and FirstNet).

3GPP defined on-network critical communication can be established both via (a) over the network infrastructure to manage the call, (b) off-network, where the terminals communicate directly to each other. In the network slicing context, over the network, involves transport networks for an always available, reliable, and zero packet loss quality of traffic support to meet critical services requirements.

Maintaining a separate broadband infrastructure for critical communications incurs a heavy deployment cost. Especially, as the coverage of this separate network has to be extended to large-scale nationwide geographies and remain interoperable is too expensive. As new communication technologies emerge, public safety systems will have to bear the state of the art adoption cost. A separate infrastructure lacks flexibility to add new value-added services or to take advantage of available commercial services.

While shared infrastructure, brings out challenges of these kind:

- (1) Reliable support: Of basic mission critical services: Such as loss of information in voice communication is not acceptable in emergency services, if common infrastructure is to be used, it must assure no loss of information.
- (2) Zero congestion: It is not acceptable for critical calls to be delayed at call setup times or be subjected to any other congestion scenarios.

Having the Mission Critical Service (MCS) as a network slice benefit from the following:

- o Insertion and authorization of subscribers in a group communication: In a critical infrastructure, the subscriber authentication may be done earlier at the entry point automatically through slice selection functional entity.
- o Pre-allocated QoS Class Identifiers (QCIs): Generally, QCIs are requested on per session basis which could slow down overall call control setup and is undesirable for emergency services. When operating in a slice, these resources maybe reserved ahead of time in a coarse-grained manner instead of per session.

MCS network slices are relatively straight forward as it only concerns with guaranteed bit rate (GBR) on per media basis and management of groups. From transport they should be able to request transport services based on GBR for reliable communication. A reference network slice in Figure 10 below, shows a mission critical (MC) organization providing service agreement through a network slice template with resource specification. The MCS slice sets up different subnetworks of different subscriber groups and manages its membership. These subnets are realized into the infrastructure across different domains through a network slice transport mechanism. The MCS must be capable of active resource monitoring to prevent congestions to ever occur as well as request additional transport resources in case of emergency event occurrence.

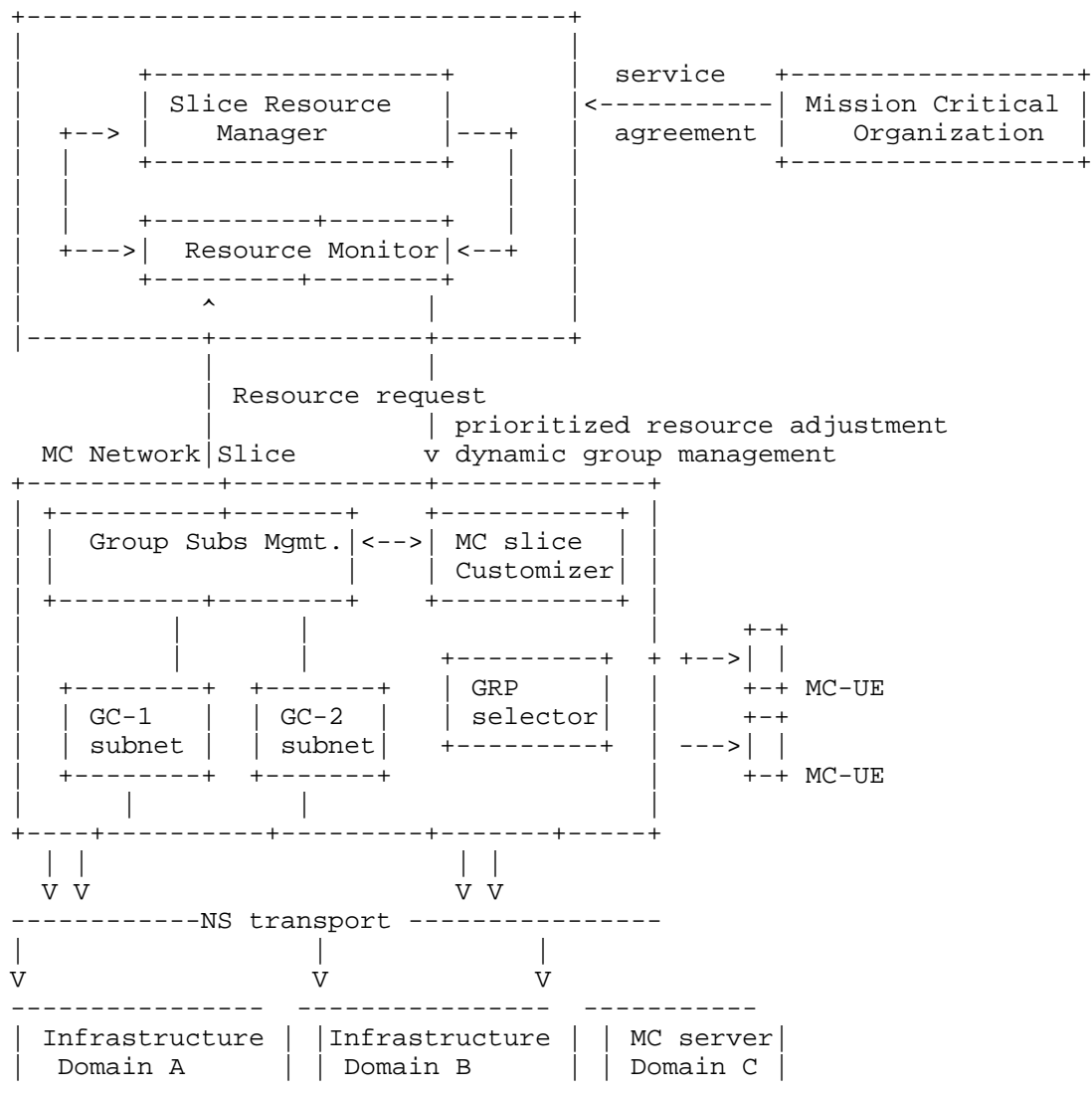


Figure 10: Reference for Mission Critical Network Slice.

## 7. Network Infrastructure for new technologies

### 7.1. ICN as a Network Slice

ICN as in Information-Centric Networking is a culmination of multiple future Internet research efforts in various parts of the world, now being pursued under IRTF's research task group called [ICNRG].

Information-Centric Networking (ICN) addresses Internet's network architectural design gaps based on evolving applications requirements and end user behavior that is significantly different from what IP was designed for - which was optimized for host-to-host communication paradigm. ICN is a non-IP paradigm based on name-based routing and offers many desirable networking features to applications such as naming, security, caching, mobility, multicasting and computing in a manner different from traditional host-centric communication model. ICN's name-based abstraction to application minimizes bootstrap configuration from the network, making it suitable to several communication modalities such as multi-point-to-multi-point, AR/VR, D2D and Ad hoc communication.

## 7.2. New Verticals - ICN based service delivery

Services over ICN slices can take advantage of its features such as:

- (1) In ICN, applications, services and content are addressed using names, hence end host resolution services like DNS can be avoided, this achieves name resolution to edge content or services without incurring additional RTT delays.
- (2) Service flows will be offered mobility and multicasting support, as the networking is session-less and optimized towards efficient movement of named data or networking named services and host level communication.
- (3) Services can be deployed at the very edges with ease as ICN routers are compute friendly, this is because states in the forwarding table can be that of either content or service resources.
- (4) Further saving bandwidth in the upstream link through opportunistic caching is an inherent feature of ICN, this also leads to energy efficient networking.

When offered as a programmable and customizable logical network slice, ICN based services can be offered as a network slice in parallel with traditional IP based services. ICN can be realized as a slice [\_5GICN\_] based on the choice of data plane resource offered by the operators in different domains of the network such as the access, core network or main data centers. While the same resources can be used to support services over IP, proper resource isolation shall allow it to co-exist with ICN slices as well. ICN slices can be offered over a network slicing framework built upon a programmable pool of software and/or hardware based data plane resources.



### 7.2.1. Required Characteristics

In ICN, applications use Interest/Data or Get/Put abstractions over named resources resolved by ICN's routing plane. An ICN slice shall be a programmable ICN-domain, in which content learning and distribution will be done using existing or new ICN aware distributed routing logic or through centralized application controllers. As a result, it should be possible to deploy software or hardware based network functions such as ICN routers and content producers and distributors that serve and speak ICN protocols, or enabled through service gateways at the edges of the network. Just as multiple service instances can be part of a slice, an ICN slices can multiplex heterogeneous services; on the other hand an ICN slice can be as granular as a single service instance too. The latter approach has implications with respect to consumer privacy, access control of name data objects, and granularity of mobility handling [\_5GICN\_].

A basic ICN slice can be manifested as a resource isolated logical network while sharing resources with other connectivity or IP based service slices. An ICN slice relies on programmability and virtualization framework to manage the service slices, to allow maximum flexibility through ICN aware logically centralized control plane for ICN service and slice management.

- o Through a network slice template -ICN service providing entity could specify specific locations (edge of network domains) to deploy ICN-routers or other ICN-NFs (ICN aware network functions). Its service definition varies with the type of service.
- o Application driven connectivity between ICN network elements in all segments and create an ICN based virtual topology.
- o Mechanisms to deliver ICN user traffic over the infrastructure such as overlay or, ICN NFs can be tightly integrated with the RAN such as the eNodeB or implicitly using traffic classification function at the edge and tunneled to ICN User Plane Function (UPF).
- o In addition, bandwidth and other network resources may be requested from the underlay depending on its capability of providing deterministic or statistically guarantees.

How multiple services will be deployed within an ICN aware slice may or may not be exposed to the network operator, depending on if the ICN slices are natively managed by it or a by other service providers.

## 8. Overall Use Case Analysis

The discussion in above use cases can be summarized as following in terms of the requirements for network slicing framework.

### 8.1. Requirements Reference

The following functional requirements are derived from discussions in above sections. They are described in details in [I-D.qiang-netslices-gap-analysis] document.

The differentiated services described in this document demonstrate several common functionalities. Therefore, a homogeneous approach towards deployment and management is absolutely necessary.

### 8.2. Mapping Common characteristics to Requirements

- (1) Resource Reservation: Compute and network resources are reserved as part of initial creation and subsequently during the maintenance of a slice. For example, a service may initially reserve resources for its own control plane, and then later it may reserve user plane flows for applications on demand. Reference use cases: Differentiated services discussed in section "Services with Resource Assurance". A network slice aware infrastructure shall be able to support mechanisms for elastic scaling (up/down) of resources and their non-disruptive provisioning.
- (2) Resource Assurance: A network slice aware infrastructure allows operators to allocate part of the network resources to meet stringent resource characteristics. Scenarios in both Section 5 and Section 6 require on demand and dynamic adjustments. It may not be possible to achieve this using centralize or API approach with finer granularity of resources participating in constrained path computation.
- (3) Multi-dimensional service vertical: Network slicing supports dynamic multi-services, multi-tenancy and the means for backing vertical market players.
- (4) Multi-domain coordination: Multi-domain refers to different technology related network domains. For example, it may be RAN, DSL etc., mobile core network, ISP or different domains in transport networks such as carrier Ethernet, MPLS, TE-tunnel etc. Often, they are under same administrator's control but may require coordination across different administrations. Furthermore, capabilities of each domain must be known in order to validate if a slice can be created or not. All scenarios

mentioned require multi-domain coordination to connect and administer different subnets.

- (5) Operational Isolation: A network slice represents logical group of network resources, functions and corresponding configurations separating its behavior and hence operation from the underlying physical network. Each network slice may have its own operator that sees this slice as a complete network (i.e. with router instances, policies, programmability, placement of virtual network functions according to traffic patterns etc.) and can manage as its own network.
- (6) Transparency: Network slicing does not change the functionality of a scenario; It only facilitates creation of an isolated, an independently run infrastructure for that use case over a common network. Transparency promotes inter-operability and a common resource specification enables it.
- (7) Reliability: It is an important resource attribute in the type of service verticals described above. Many services verticals cannot deliver functionality unless the network is reliable (See remote industry operation, remote surgery and other uRLLC applications). In this regard, monitoring probes are needed of each network slice and resources associated with it.

Requirements Illustrated above	Aggregated Requirements
1) Resource reservation 6) Transparency 4) Multi-access knowledge 3) Multi-dimensional service vertical	Req 1. Network Slicing Specification
4) Multi-Domain coordination 2) Resource Assurance	Req 2. Network Slicing Cross-Domain Coordination
5) Operational/performance Isolation	Req 3. Network Slicing Performance Guarantee and Isolation
7) Reliability	Req 4. Network Slicing OAM

Figure 11: Mapping Common Characteristics to Requirements

NSaaS is a key for network operators to deploy network slices. Having standard means to realize these use cases, enables (a)

different usecases to be uniformly understood by a network slice provider, and (b) similar use cases to be understood in a similar fashion by different network slice providers. Both these cases should allow common mechanisms to map and allocate network slices over the network infrastructure.

Due to the availability of diverse technologies in control and data planes; the first step should be a top-down means to realize a slice with a common technology independent information model. It may describe a resource-centric slice with connectivity, storage, and compute resources, network functions, and operational requirements, that further get mapped to infrastructure resources and capabilities for run-time operations and monitoring. This model may be used by an orchestrator onboarding function for creating instances of network slice services and distributing to network infrastructure providers.

## 9. Conclusion

A service should typically need a network slice for one of those reasons:

- (1) The service cannot provide optimal experience on a best-effort network.
- (2) It is inefficient and expensive to build a separate infrastructure.

The separation from a generalized network, should allow new services to use newer or different protocols in network, transport and management layer/plane for that service (as in the case of ICN, mMTC, uRLL). The goal of Network slices is to offer enriched service verticals with very different network capability and performance demands but also simplify from the traditional service delivery models.

There is need for a uniform framework for end to end network slicing specifications that spans across multiple technology domains and can drive extensions in those technology-areas for support of Network slices.

## 10. Security Considerations

The security considerations apply to each kind of slice. In addition general security considerations of underlying infrastructure whether isolated communication with in a slice apply for links using wireless technologies.

## 11. IANA Considerations

There are no IANA actions requested at this time.

## 12. Acknowledgements

Note, the 5GEX L2VPN and L3VPN usecase is an independent contribution by authors and is not endorsed by 5GEX. Many thanks to the following reviewers for providing details for several use cases and for helping with the review of the document.

Stewart Bryant (stewart.bryant@gmail.com), Hannu Flinck (hannu.flinck@nokia-bell-labs.com), Med Boucadair (mohamed.boucadair@orange.com), Dong Jie (dong.jie@huawei.com).

## 13. References

### 13.1. Normative References

- [I-D.bernardos-nfvrg-multidomain]  
Bernardos, C., Contreras, L., Vaishnavi, I., and R. Szabo, "Multi-domain Network Virtualization", draft-bernardos-nfvrg-multidomain-03 (work in progress), September 2017.
- [I-D.dt-detnet-dp-sol]  
Korhonen, J., Andersson, L., Jiang, Y., Finn, N., Varga, B., Farkas, J., Bernardos, C., Mizrahi, T., and L. Berger, "DetNet Data Plane Encapsulation", draft-dt-detnet-dp-sol-02 (work in progress), September 2017.
- [I-D.geng-netslices-architecture]  
67, 4., Dong, J., Bryant, S., kiran.makhijani@huawei.com, k., Galis, A., Foy, X., and S. Kuklinski, "Network Slicing Architecture", draft-geng-netslices-architecture-02 (work in progress), July 2017.
- [I-D.ietf-l2sm-l2vpn-service-model]  
Wen, B., Fioccola, G., Xie, C., and L. Jalil, "A YANG Data Model for L2VPN Service Delivery", draft-ietf-l2sm-l2vpn-service-model-03 (work in progress), September 2017.
- [I-D.ietf-opsawg-service-model-explained]  
Wu, Q., LIU, W., and A. Farrel, "Service Models Explained", draft-ietf-opsawg-service-model-explained-05 (work in progress), October 2017.

- [I-D.pularikkal-virtual-cpe]  
Pularikkal, B., Fu, Q., Hui, D., Sundaram, G., and S. Gundavelli, "Virtual CPE Deployment Considerations", draft-pularikkal-virtual-cpe-02 (work in progress), February 2017.
- [I-D.qiang-netslices-gap-analysis]  
Qiang, L., Martinez-Julia, P., 67, 4., Dong, J., kiran.makhijani@huawei.com, k., Galis, A., Hares, S., and S. Slawomir, "Gap Analysis for Transport Network Slicing", draft-qiang-netslices-gap-analysis-01 (work in progress), July 2017.
- [RFC8049] Litkowski, S., Tomotaki, L., and K. Ogaki, "YANG Data Model for L3VPN Service Delivery", RFC 8049, DOI 10.17487/RFC8049, February 2017, <<https://www.rfc-editor.org/info/rfc8049>>.

### 13.2. Informative References

- [\_5GICN\_] IEEE Communication, "Delivering ICN Services in 5G using Network Slicing. 'Asit Chakraborti, Syed Obaid Amin, Aytac Azgin, Ravi Ravindran, G.Q.Wang'", May 2017, <<https://arxiv.org/abs/1610.01182>>.
- [ICNRG] IRTF, "ICN Routing Group", November 2016, <<https://irtf.org/icnrg>>.
- [Tactile-Internet]  
ITU-T, "Technology Watch Report, The Tactile Internet", August 2014, <<https://www.itu.int/oth/T2301000023/en>>.
- [TR\_3GPP.33.899]  
3GPP, "Study on the security aspects of the next generation system", 3GPP TR 33.899 0.6.0, November 2016, <<http://www.3gpp.org/ftp/Specs/html-info/33899.htm>>.
- [TR\_3GPP.38.801]  
3GPP, "Study on new radio access technology Radio access architecture and interfaces", 3GPP TR 38.801 1.0.0, March 2017, <<http://www.3gpp.org/ftp/Specs/html-info/38801.htm>>.
- [TR\_3GPP\_38.913]  
3GPP, "Study on scenarios and requirements for next generation access technologies", 3GPP TR 38.913 14.2.0, March 2017, <[http://www.3gpp.org/ftp/Specs/archive/38\\_series/38.913](http://www.3gpp.org/ftp/Specs/archive/38_series/38.913)>.

- [TS\_3GPP.23.501]  
3GPP, "System Architecture for the 5G System", 3GPP  
TS 23.501 0.2.0, February 2017,  
<<http://www.3gpp.org/ftp/Specs/html-info/23501.htm>>.
- [TS\_3GPP.23.502]  
3GPP, "Procedures for the 5G System", 3GPP TS 23.502  
0.2.0, February 2017,  
<<http://www.3gpp.org/ftp/Specs/html-info/23502.htm>>.
- [TS\_3GPP.28.500]  
3GPP, "Telecommunication management; Management concept,  
architecture and requirements for mobile networks that  
include virtualized network functions", 3GPP TS 28.500  
1.3.0, 11 2016,  
<<http://www.3gpp.org/ftp/Specs/html-info/28500.htm>>.
- [VCPEBBF] Broadband Forum, "TR-345 Broadband Network Gateway and  
Network Function Virtualization", Dec 2016,  
<[https://www.broadband-forum.org/technical/download/  
TR-345.pdf](https://www.broadband-forum.org/technical/download/TR-345.pdf)>.

## Authors' Addresses

Kiran Makhijani  
Huawei Technologies  
2890 Central Expressway  
Santa Clara CA 95050  
USA

Email: [kiran.makhijani@huawei.com](mailto:kiran.makhijani@huawei.com)

Jun Qin  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095

Email: [qinjun4@huawei.com](mailto:qinjun4@huawei.com)

Ravi Ravindran  
Huawei Technologies  
2890 Central Expressway  
Santa Clara CA 95050  
USA

Email: [ravi.ravindran@huawei.com](mailto:ravi.ravindran@huawei.com)

Liang Geng  
China Mobile  
Beijing 100095  
China

Email: gengliang@chinamobile.com

Li Qiang  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: qiangli3@huawei.com

Shuping Peng  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: pengshuping@huawei.com

Xavier de Foy  
InterDigital Inc.  
1000 Sherbrooke West  
Montreal  
Canada

Email: Xavier.Defoy@InterDigital.com

Akbar Rahman  
InterDigital Inc.  
1000 Sherbrooke West  
Montreal  
Canada

Email: Akbar.Rahman@InterDigital.com



Alex Galis  
University College London  
London  
U.K.

Email: a.galis@ucl.ac.uk

Giuseppe Fioccola  
Telecom Italia  
Italy

Email: giuseppe.fioccola@telecomitalia.it

none  
Internet-Draft  
Intended status: Informational  
Expires: January 4, 2018

L. Qiang, Ed.  
Huawei  
P. Martinez-Julia  
NICT  
L. Geng  
China Mobile  
J. Dong  
K. Makhijani  
Huawei  
A. Galis  
University College London  
S. Hares  
Hickory Hill Consulting  
S. Kuklinski  
Orange  
July 3, 2017

Gap Analysis for Transport Network Slicing  
draft-qiang-netslices-gap-analysis-01

Abstract

This document presents network slicing differentiation from the non-partition network or from simply partition of connectivity resources. It lists 7 standardization gaps related to 4 key requirements for network slicing in transport network. It also presents an analysis of existing related work and other potential solutions on network slicing.

This gap analysis document aims to provide a basis for future works in transport network slicing.

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 <http://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 January 4, 2018.

#### Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

#### Table of Contents

1. Introduction . . . . .	3
2. Terminology and Abbreviation . . . . .	4
3. Overall Requirements in Network Slicing . . . . .	4
4. Network Slicing Specification . . . . .	7
4.1. Description . . . . .	7
4.2. Related Work in IETF . . . . .	8
4.2.1. YANG Data Models . . . . .	8
4.2.2. Building NSS from Protocol Independent Traffic Engineering Models . . . . .	9
5. Network Slicing Cross-Domain Coordination . . . . .	10
5.1. Description . . . . .	10
5.2. Related Work in IETF . . . . .	11
5.2.1. Autonomic Networking Integrated Model and Approach (ANIMA) . . . . .	11
5.2.2. Connectivity Provisioning Negotiation Protocol (CPNP) . . . . .	12
5.2.3. Abstraction and Control of Traffic Engineered Networks (ACTN) . . . . .	13
5.3. Other Potential Solutions . . . . .	14
6. Network Slicing Performance Guarantee and Isolation . . . . .	14
6.1. Description . . . . .	14
6.2. Related Work in IETF . . . . .	15
6.2.1. Virtual Private Networks . . . . .	15
6.2.2. NVO3 . . . . .	15
6.2.3. RSVP-TE . . . . .	15
6.2.4. Segment Routing . . . . .	16
6.2.5. Deterministic Networking . . . . .	16
6.2.6. Flexible Ethernet . . . . .	17
7. Network Slicing OAM with Customized Granularity . . . . .	17
7.1. Description . . . . .	17

7.2. Related Work in IETF . . . . .	18
7.2.1. Overview of OAM tools . . . . .	18
7.2.2. Overlay OAM . . . . .	19
7.2.3. Service Function Chaining . . . . .	19
7.2.4. Slice Identification . . . . .	19
8. Summary . . . . .	20
9. Security Considerations . . . . .	21
10. IANA Considerations . . . . .	21
11. Acknowledgements . . . . .	22
12. References . . . . .	22
12.1. Normative References . . . . .	22
12.2. Informative References . . . . .	22
Authors' Addresses . . . . .	26

## 1. Introduction

Network slicing is an approach to enable flexible isolation of network resources and functions for dedicated services, providing a certain level of customization and quality guarantee. It establishes customized dedicated network upon a common infrastructure for vertical industries with flexible design of functions, different performance requirements, system isolation and OAM tools.

Several SDOs have investigated network slicing. To list a few: NGMN initiated a study of network slicing in the context of 5G from the mobile network point of view [NGMN-2016]. Around the same time ITU-T IMT 2020 and ITU-T SG13 studied network softwarization that also included network slicing concept. ITU-T has issued a number of recommendations, such as: Gap Analysis [IMT2020-2015], Network Softwarization [IMT2020-2016], Terms & Definitions [IMT2020-2016bis]. Open Network Foundation (ONF) has developed a recommendation on applying SDN architecture to Network Slicing [ONF-2016]. Finally, 3GPP standards development for 5G includes network slicing in radio access and core networks. 3GPP issued TS 23.501 [TS23-501] about the system architecture for 5G in 2017. BBF started the project SD-406 focusing on the end-to-end architecture enhancement and requirements gathering for transport networks. Although these SDOs have done a lot of work, potential requirements especially in the transport network and end-to-end enabling need to be investigated in order to elicit and identify the technical gaps in IETF for transport network slicing.

In order to establish a network slice that meets various customer's demands, an infrastructure owner needs to understand how these demands map with the available network resources and accessible capabilities. This also requires end-to-end coverage and inter-domain coordination. Meanwhile, the slice provider provides customized OAM to the tenants under provisioning. Slicing OAM

approach is a fundamental capability to guarantee stable, effective and reliable services for the vertical industries. It is also expected to be capable of operations with customized granularity levels that provides robust management flexibilities.

This document presents the identified key requirements and investigates potential technical gaps accordingly. To assist understanding of this document, Section 2 outlines the terminology. Section 3 introduces overall requirements of network slicing. Sections 4~7 illustrates resource specification, end-to-end consideration, performance guarantee and OAM concerns respectively. Section 8 summarizes the identified gaps.

## 2. Terminology and Abbreviation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

All of the network slicing related words used in this document are to interpreted as described in [NS-Framework]

## 3. Overall Requirements in Network Slicing

This section introduces 4 key requirements of network slicing derived from [NS-UseCase] as shown in Table 1. These 4 requirements are organized according to a general network slice working process as shown in Figure 1:

- 1: describe network slicing resource/functions and capture requirements (Req. 1)
- 2: network slicing cross-domain coordination (Req. 2)
- 3: construct a performance guaranteed and isolated end-to-end network slice (Req. 3)
- 4: provide necessary Operation & Maintenance & Administration (OAM) (Req. 4)

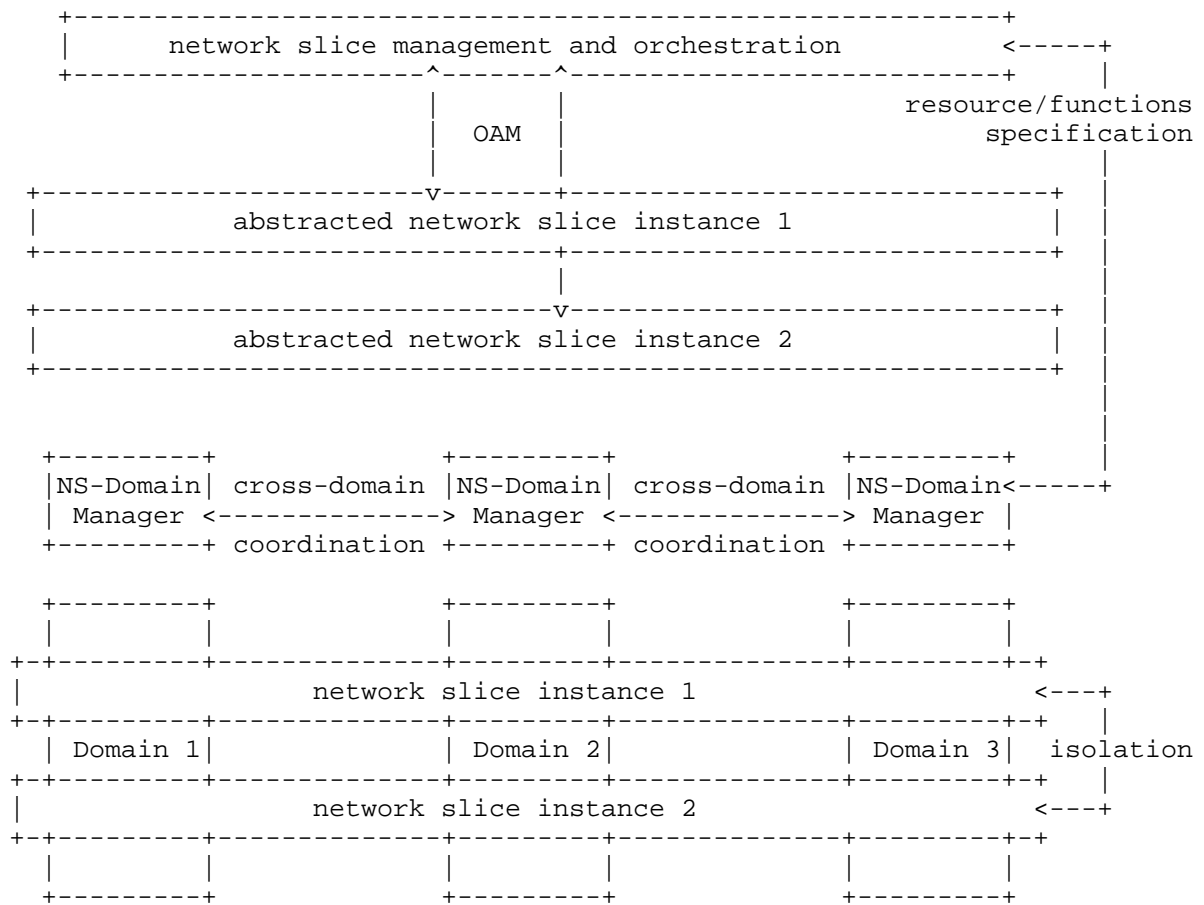


Figure 1: Illustration of Key Requirements

Table 1: Requirement Association

Requirements Illustrated in NS UseCase	Extracted KEY Requirements
1) Resource Reservation 2) Abstraction 3) Multi-Access Knowledge 4) Multi-Dimensional Service Vertical 5) Agile Resource Adjustment	Req 1. Network Slicing Specification
6) Multi-Domain Coordination 7) Resource Assurance	Req 2. Network Slicing Cross-Domain Coordination
8) Performance/Operation Isolation	Req 3. Network Slicing Performance Guarantee and Isolation
9) Independent Slice Management Plane Reliability	Req 4. Network Slicing OAM

Table 1: Requirements Association

- o Req 1. Network Slicing Specification (NSS) - The management systems of both network slice providers and operators need to know what and how much resources/network functions they have, so that they can accurately and abstractedly describe the available resources/network functions to tenants or peers. The objective of NSS is to deliver the network slicing requests without incurring any over-utilization of resources. In order to cooperate and provide consistent network slicing service, the way that resources/network functions are described should be homogeneous and compatible among all of the involved technology-specific domains, provides, and slicing platforms.
- o Req 2. Network Slicing Cross-Domain Coordination (NS-CDC) - From terminal to server (or other terminal), an end-to-end network slice will involve different infrastructural domains (e.g., AN, TN, CN, etc. ) that may be owned by different providers/operators. Each infrastructural domain may be further divided into different administrative domains. That is an end-to-end slice is a logical entity composed by multiple separated components, and the cross-domain coordination is a way to integrate these components together.
- o Req 3. Network Slicing Performance Guarantee and Isolation (NS-PGI) - In order to enable the safe, secure, privacy-preservation service for multi-tenancy on a common physical network, the

isolation among network slices in each of the Data/Control/Management/Service planes are needed. Furthermore, network slices that provide differentiated services usually require different resources. The resources allocated to a network slice must be able to guarantee the service performance requirement.

- o Req 4. Network Slicing OAM (NS-OAM) - On one end of the spectrum we have those operators that will require a finalized service that they will simply commercialize. On the other end we have those operators that may want to fine-tune all the low-level aspects of the network resources that form their system or service. Moreover, in the middle there is plenty of room for variations. Therefore, the underlying network layers must offer different levels of granularity for the management of their resources, that the upper layer operators can choose according to their needs and objectives.

#### 4. Network Slicing Specification

##### 4.1. Description

Network Slicing Specification (NSS) is meant to describe the network slicing resources and capture requirements from tenants or peer networks to characterize the service expected to be delivered by a network. These requirements include (non-exhaustive): reachability scope (e.g., limited scope, Internet-wide), direction, bandwidth requirements, performance metrics (e.g., one-way delay [RFC2679], loss [RFC2680], or one-way delay variation [RFC3393]), protection and high-availability guidelines (e.g., uRLLC service restoration in less than 50 ms, 100 ms, or 1 second), traffic isolation constraints, and flow identification. NSS is used by a network provider to decide whether existing network slice instances can be reused or (some of them) even combined, or if another network slice instance is needed for a given service.

Technology-specific actions are then derived from the technology-agnostic requirements depicted in an NSS. Such actions include configuration tasks and operational procedures. A standard definition of NSS is needed to facilitate the dynamic/ automated negotiation procedure of NSS parameters, but also to homogenize the processing of service requirements.

To explain by an example, a network slice may cross multiple domains:

- o A cloud deployed, NFV enabled, chain of network functions in a virtualized 5G core.



- o A segment routing [I-D.ietf-spring-segment-routing] based IGP network transport/aggregation or slice-specific application functions.
- o A PCE [RFC4655] monitored TE-tunnel with ingress and egress points.
- o Optical, carrier Ethernet or cellular networks.

The network slice is a combination of the above technologies. It creates a compelling need for a common resource specification interface across these domains.

#### 4.2. Related Work in IETF

##### 4.2.1. YANG Data Models

As rightfully discussed in [I-D.wu-opsawg-service-model-explained], the IETF has already published several YANG data models that are used to model monolithic functions as well as very few services (e.g., L2SM, L3SM, EVPN). These models may be used in the context of network slicing if corresponding technologies are required for a given network slice, but none of them can be used to model an NSRD.

[RFC7297] describes the Connectivity Provisioning Profile (CPP) and proposes a CPP template to capture connectivity requirements to be met within a service delivery context. Such a generic CPP template is meant to

- o facilitate the automation of the service negotiation and activation procedures, thus accelerating service provisioning;
- o set (traffic) objectives of Traffic Engineering (TE) functions and service management functions;
- o improve service and network management systems with 'decision-making' capabilities based upon negotiated/offered CPPs.

[RFC7297] may be considered as a candidate specification for NSRD. Releasing a RFC7297-bis to take into account specific requirements from network slicing is needed. Since [RFC7297] may not be implemented by all providers, the [SLA-Exchange] may be adopted to implement indirect SLA negotiation and SLA events report. [SLA-Exchange] provides an in-band method to exchange the SLA parameters, and then by the receiving devices to translate SLA in technical specific provisioning languages. However, there still does not exist any standard protocol to translate SLA agreements into technical clauses and configurations.

#### 4.2.2. Building NSS from Protocol Independent Traffic Engineering Models

The NSS requirement for reachability, direction, bandwidth requirements, performance metrics, traffic isolation constraints, and flow identification can be built utilizing protocol which can perform operations (read, write, notification, actions (aka rpcs)) on a yang service layer that supports these traffic engineer and resource definition at the service layers. The network slicing service data model can extend existing work in the TEAS and I2RS working group for protocol-independent topology models. These models support configuration or the dynamic datastores defined in [NMDA] which will be abbreviated as NMDA in this section. This section provides the detail on how the NSS can be built from these models and the RESTCONF protocol.

##### 4.2.2.1. Basic Topology Model

The basic topology model is defined in [I2RS-Yang] to include a service layer. This topology model is protocol independent and can be utilized as a configuration data model or a dynamic datastores model. The configuration data model must abide by the configuration persistence and referential requirements. The dynamic datastores do not need to abide by the same requirements as the configuration datastore. I2RS is defining a dynamic datastores reference model for a data store which ephemeral. The network slices may want to use configuration, ephemeral datastores, or define a third type of dynamic datastores. The I2RS WG provides a place to collaborate this work on the dynamic datastores.

##### 4.2.2.2. TEAS Model Utilization of Basic Topology Model

The TEAS topology model [TE-Yang] provides a general description of a Traffic engineering model that provides:

- o abstract topologies with TE constraints (bandwidth, delay metrics, links to lower layers, some traffic isolation constraints, and some link identifiers);
- o templates for links or resources;
- o functionality to read, write, notification, and rpcs.

Options that need to be consider are:

Augmenting TEAS - The TEAS models provide substantial traffic engineering. It was envisioned in the early topology model that a service resource model would be part of the service layer. This

work was delayed until the maturation of the service requirements from L2VPN, L3VPN, and EVPN plus the maturation of resource requirements from 5G. Network slicing provides a good application use case for this work.

Why not Augment TEAS - The TEAS models are TE specific, lack of the abstraction for Layer 3+ resources.

Dynamic models to combine TEAS models for network-slicing - The network slicing controller operating across domains may wish to create a multiple-domain data model based on the service layer data models exposed by different providers. These service models would not need to be configured, but only learned as providers exchange data with one another. The rules for combining these models could be defined as part of the dynamic datastore for network-slicing.

Protocol within a domain - The RESTCONF and NETCONF protocol can support read, write, notification and actions (rpcs) within a domain.

Protocol across domains: The RESTCONF protocol currently supports Configuration protocols and 90% of the dynamic datastores. The RESTCONF protocol is being enhanced to support the push of telemetry messages. The RESTCONF protocol could be used to exchange a specific Yang network-slicing service-layer topology (TE and Resources) and for the I2NSF security capabilities between domains.

If a multicast of telemetry data is required between domains, then the push model for telemetry information or the IPFIX protocol may be utilized.

## 5. Network Slicing Cross-Domain Coordination

### 5.1. Description

The network slicing cross-domain coordination (NS-CDC) requirement includes the following aspects:

- o Network slice resource/functions coordination: for example, a tenant requests for a network slice with at most 10 ms latency from terminal to server. Different infrastructure/administrative domains should coordinate and negotiate to reach an agreement such as RAN provides at most 2 ms service, TN domain I provides at most 4ms service, TN domain II provides at most 2 ms service and CN provides at most 2 ms service;

- o Configuration information coordination: for example, for a given TN domain, the configuration information such as VLAN ID, remote IP address, physical port ID, etc. need to be coordinated with other TN domains;
- o Other coordination: for example, RAN (or other access network) needs to notify TN about the information of new attachment point when user moves.

From terminal to server, an end-to-end network slice will involve different infrastructure domains (e.g., RAN, TN and CN). An infrastructure domain may be further divided into multiple domains due to geographic isolation, administrative isolation and other reasons. There are two ways to enable an end-to-end network slice: based on a common platform or based on cross-domain coordination.

If all of the involved domains belong to the same operator or the same operator union, the common platform solution may be work. In this case, all of the domain controllers only need to communicate with the common platform, and follow the coordination management of this common platform. Whilst the most common case is that the domains belong to different owners/operators/administrators, making it difficult to realize such a common platform. Consequently, the cross-domain coordination will be essential throughout the whole lifecycle of an end-to-end network slice.

## 5.2. Related Work in IETF

There are some related works studies the inter-operation/coordination between different entities. Coordination of different components of a slice requires automation. It can be achieved either by

1. Coordination protocols such as ANIMA, CPNP
2. Or through abstraction and corresponding interfaces as in ACTN.

This subsection will briefly review these related work to provide a basis for the gap analysis.

### 5.2.1. Autonomic Networking Integrated Model and Approach (ANIMA)

Autonomic Networking Integrated Model and Approach (ANIMA) WG provides a series of tools for distributed and automatic management, which includes: Generic Autonomic Signaling Protocol (GRASP), Autonomic Networking Infrastructure (ANI), etc.

GRASP [ANIMA-GRASP] is a protocol for the negotiation between ASAs (Autonomic Service Agent). In GRASP, ASAs could be considered as

"APPs" installed on a device. Different ASAs fulfill different management tasks such as parameter configuration, service delivery, etc. Based on GRASP, the same purpose ASAs that installed on different devices are able to inter-operate and negotiate with each other. Network slicing could make use of GRASP for the coordination among devices in the underlying infrastructure layer, as well as the negotiation among different domain managers. However, the security issue incurred by cross-domain usage should be fixed in GRASP.

ANI [ANI] is a technical packet consisting of BootStrap (for authentication, domain certification distribution, etc.), ACP (a separate control plane), and GRASP (for control message coordination). ANI could be used to construct the management tunnel among devices in underlying infrastructure layer within a single domain. While the network slicing and cross-domain oriented extensions are necessary.

#### 5.2.2. Connectivity Provisioning Negotiation Protocol (CPNP)

[I-D.boucadair-connectivity-provisioning-protocol] defines the Connectivity Provisioning Negotiation Protocol (CPNP) that is meant to dynamically exchange and negotiate connectivity provisioning parameters, and other service-specific parameters, between a Customer and a Provider. CPNP is a tool that introduces automation in service negotiation and activation procedures, thus fostering the overall service provisioning process.

CPNP runs between a Customer and a Provider carrying service orders from the Customer and respective responses from the Provider to the end of reaching a connectivity service provisioning agreement. As the services offered by the Provider are well-described, by means of the CPP template, the negotiation process is essentially a value-settlement process, where an agreement is pursued on the values of the commonly understood information items (service parameters) included in the service description template.

The protocol is transparent to the content that it carries and to the negotiation logic, at Customer and Provider sides, that manipulates the content.

The protocol aims at facilitating the execution of the negotiation logic by providing the required generic communication primitives.

CPNP can be used in the context of network slicing to request for network resources together with a set of requirements that need to be satisfied by the Provider. Such requirements are not restricted to basic IP forwarding capabilities, but may also include a characterization of a set of service functions that may be invoked.

### 5.2.3. Abstraction and Control of Traffic Engineered Networks (ACTN)

ACTN [TEAS-ACTN] is an information model proposed by TEAS WG, which enables the multi-domain coordination in Traffic Engineering (TE) network. In order to enable network slicing in transport networks, portion of transport domain will need to be engineered. In particular, building a TE entity and stitching service for this entity is within the scope of ACTN. As an end-to-end network slicing solution, ACTN is able to provide cross-domain coordination. In ACTN, each physical transport network domain is under the control of a Physical Network Controller (PNC) as shown in Figure 2. A Multi-Domain Service Coordinator (MDSC) controls multiple PNCs. Although the MDSCs may form a hierarchical structure, a hierarchical MDSC can still be regarded as a logical common platform. As Section 5.1 discussed, such a common platform solution has a strict presumption that all domains are assumed to follow a common coordination management.

While ACTN does carry out network slicing-related work, some proposed concepts are similar the concepts of today's network slicing: in particular, the virtual network (VN) is similar to a slice instance. ACTN enables VN based on LSP technique, different LSP tunnels correspond to different VNs. However, ACTN focuses on resource abstraction and management on Layer 2 and Layer 1. For transport network slicing, resources abstraction and management on Layer 3+ (e.g., IP routing table, etc.) may also be necessary but have not been addressed by ACTN.

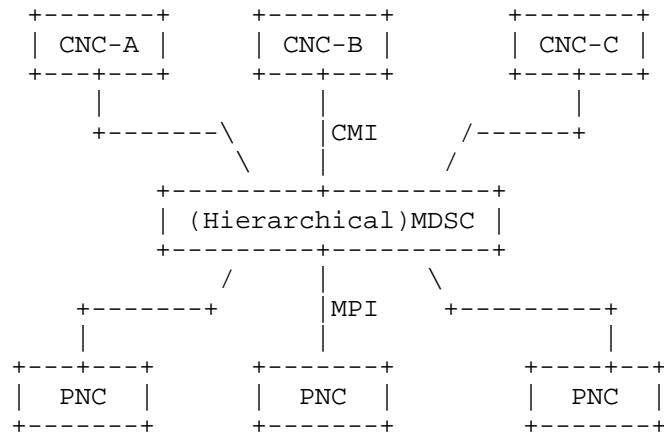


Figure 2: A Three-tier ACTN Control Hierarchy

### 5.3. Other Potential Solutions

5G Exchange (5GEx) [FGEx] is a 5G-PPP project which aims to enable cross-domain orchestration of services over multiple administrations or over multi-domain single administration networks. The main infrastructure considered in 5GEx is the NFV/SDN compatible software defined infrastructure, which limits the scope of network slicing to SDN based architecture.

## 6. Network Slicing Performance Guarantee and Isolation

### 6.1. Description

Network slicing is expected to enable the deployment of various services with diverse requirements, independently on a common physical network. Each network slice is characterized by particular service requirements, which usually are expressed using in the form of several key performance indicators (KPIs) such as bandwidth, latency, jitter, packet loss, etc., and different degrees of isolation. Isolation requirements include performance isolation, which means performance guarantee are maintained regardless of activity in other slices, as well as secure isolation (e.g., including privacy), and management (or OAM) isolation. Additionally, performance isolation in network slicing has to maintain while scaling up or down computing capabilities of a slice (i.e., for elastic scaling). Moreover, since IoT is also a use case for NS, and since some IoT applications are sensitive to data plane or bits on wire overheads, data path encapsulation in the form of labels, VLANs, VxLANs should be optional, or minimized for those cases.

As we will discuss in the detailed sections below, each of these technologies can address some but not all performance and isolation requirements:

- o RSVP-TE, Segment Routing (SR), DETNET, FlexE are mostly related to performance guarantee and performance isolation requirements
- o Virtual Private Networks (VPN), NV03 are mostly related to security and management isolation requirements

A Network Slicing solution, to support performance guarantee and isolation requirements, will therefore need to merge in some way characteristics from these two families of technologies, through the combination of (possibly enhanced) existing technologies and/or specifically developed ones. We can also consider the possibility that multiple such technology stacks may be deployed in different domains, and rely on cross-domain coordination, as described in Section 5, to form a single abstracted network slice.

## 6.2. Related Work in IETF

### 6.2.1. Virtual Private Networks

VPN technologies such as L3VPN [RFC4364], L2VPN [RFC4664], EVPN [RFC7432], etc. have been widely deployed to provide different virtual networks on common service provider networks. Although VPNs can provide logically separated routing/bridging domains between different VPN customers, essentially it is an overlay network technology with little control of the network resources, so it is challenging for VPN to meet the performance and isolation requirement of some emerging application scenarios such as industrial verticals. VPNs essentially are private networks of enterprises by connecting remote sites. The following two issues illustrate limitations of VPNs for network slicing:

- o An end-to-end VPN tunnel competes with other traffic in the network and end-to-end network resource policies cannot be guaranteed.
- o The reachability and resource reservation protocols are not tightly integrated and often solutions require centralized PCE-P like methods.

### 6.2.2. NVO3

[NVO3-WG] defines several network encapsulations which support the network virtualization and multi-tenancy in the data center networks. Similar to the VPN technologies of service provider networks, NVO3 is also an overlay network technology, which relies on the performance characteristics provided by the IP-based underlay networks. Thus NVO3 may not meet by itself the performance and isolation requirements of network slicing.

### 6.2.3. RSVP-TE

RSVP-TE [RFC3209] is the signaling protocol to establish end-to-end traffic-engineered Label Switched Paths (LSPs). It can reserve the required link bandwidth along an end-to-end path for specific network flows, which is suitable for services with particular requirement on traffic bandwidth. RSVP-TE LSPs can be used as the underlay tunnels of the VPN service connections. However, the requirement of some emerging services is not only about traffic bandwidth, but also has quite strict requirement on latency, jitter, etc. Such requirements can hardly be met with existing RSVP-TE.



#### 6.2.4. Segment Routing

[I-D.ietf-spring-segment-routing] provides the ability to specify a traffic-engineered path by the source node of data packets. It can provide traffic-engineering features comparable to RSVP-TE with better scalability, by eliminating the per-path state in the transit network nodes. It is therefore a candidate method of creating an NSI, mapping a packet into an NSI and specifying the passage of the packet through the resources dedicated to the NSI. Further study will be required to determine if/how SR as designed today can be used as a core technology for building an NSI. With respect to performance guarantee and isolation, some further investigation may be needed to understand whether SR can provide the same or better performance characteristics as RSVP-TE. In addition, it is not clear whether SR-based LSPs can provide the guaranteed latency and jitter performance required by network slicing.

#### 6.2.5. Deterministic Networking

[DETNET-WG] is working on the deterministic data paths over layer 2 and layer 3 network segments. Such deterministic paths can provide identified flows with extremely low packet loss rates, low packet delay variation (jitter) and assured maximum end-to-end delivery latency. This is accomplished by dedicating network resources such as link bandwidth and buffer space to DetNet flows and/or classes of DetNet flows. DetNet also aims to provide high reliability by replicating packets along multiple paths. It is a characteristic of DetNet that it is concerned solely with worst-case values for the end-to-end latency.

The primary target of DetNet is real-time systems and as such average, mean, or typical latency values are not protected, because they do not affect the ability of a real-time system to perform their tasks. This contrasts with a normal priority-based queuing scheme which will give better average latency to a data flow than DetNet, but, on the other side, the worst-case latency can be essentially unbounded. As such DetNet seems to be a useful technique that may be applied to either a complete NSI, or to part of the traffic within an NSI to address the emerging low latency requirement for real time application.

DetNet can therefore address some of the requirements of NS. It was however not designed with network slicing in mind, which means a mapping between an NSI and a DetNet service may need to be defined.

#### 6.2.6. Flexible Ethernet

[FLEXE-1.0] was initially defined by Optical Internetworking Forum (OIF) as an interface technology which allows the complete decoupling of the Media Access Control layer (MAC) data rates and the standard-based Ethernet Physical layer (PHY) rates. The channelization capability of FlexE can be used to partition a FlexE interface into several independent sub-interfaces, which can be considered as a useful component for the slicing of network interfaces. Currently there is ongoing work in IETF to define the control plane framework for FlexE [FlexE-FWK], which aims to identify the routing and signaling extensions needed for establishing FlexE-based end-to-end LSPs in IP/MPLS networks.

### 7. Network Slicing OAM with Customized Granularity

#### 7.1. Description

In accordance with [RFC6291], OAM is used to denote the following:

- o Operations: refer to activities that are undertaken to keep the network and the services it deliver up and running. It includes monitoring the underlying resources and identifying problems.
- o Administration: refer to activities to keep track of resources within the network and how they are used.
- o Maintenance: refer to activities to facilitate repairs and upgrades. Maintenance also involves corrective and preventive measures to make the managed network run more effectively, e.g., adjusting configuration and parameters.

As per [RFC6291], network slicing provisioning operations are not considered as part of OAM. Provisioning operations are discussed in other sections.

Maintaining automatically-provisioned slices within a network raises the following requirements:

- o Ability to run OAM activities on a provider's customized granularity level. In other words, ability to run OAM activities at any level of granularity that a service provider see fit. In particular:
  - \* Per slice OAM: An operator must be able to execute OAM tasks on a per slice basis.

- \* Per domain OAM: These tasks can cover the "whole" slice within a domain or a portion of that slice (for troubleshooting purposes, for example).
  - \* Per service OAM: When a given slice is shared among multiple services/customers, an operator must be able to execute (per-slice) OAM tasks for a particular service or customer.
  - \* For example, OAM tasks can consist in tracing resources that are bound to a given slice, tracing resources that are invoked when forwarding a given flow bound to a given network slice, assessing whether flow isolation characteristics are in conformance with the NS Resource Specification, or assessing the compliance of the allocated slice resource against flow/customer requirements.
  - \* An operator must be able to enable differentiated failure detect and repair features for a specific/subset of network slices. For example, a given slice may require fast detect and repair mechanisms (e.g., as a function of the nature of the traffic (pattern) forwarded through the NS), while others may not be engineered with such means.
- o Ability to automatically discover the underlying service functions and the slices they are involved in or they belong to.
  - o Ability to dynamically discover the set of network slicing that are enabled within a network. Such dynamic discovery capability facilitates the detection of any mismatch between the view maintained by the control plane and the actual network configuration. When mismatches are detected, corrective actions must be undertaken accordingly.
  - o Ability to efficiently OAM on shared resources. If multiple network slices share some resources, the same kind of OAM operations from different network slices should be performed only once for efficiency. For example, several network slices share a link. We only need to execute once status query, and directly return the queried result to other status query requests.

## 7.2. Related Work in IETF

### 7.2.1. Overview of OAM tools

The reader may refer to [RFC7276] for an overview about available OAM tools. These technology-specific tools can be reused in the context of network slicing. Providers that deploy network slicing capabilities should be able to select whatever OAM technology-

specific feature that would be address their needs. No gap that would legitimate specific requirements has been identified so far.

#### 7.2.2. Overlay OAM

[I-D.ooamdt-rtgwg-ooam-header] specifies a generic OAM header that can be used if overlay technologies are enabled. Obviously, this effort can be reused in the context of network slicing when overlay techniques are in use. Nevertheless, For slice designs that do not assume an overlay technology, OAM packets must be able to fly over the appropriate slice and for a given service/customer. This is possible by reusing some existing tools if and only if no specific fields are required (e.g., carry a slice identifier as Req. 5 stated).

#### 7.2.3. Service Function Chaining

SFC WG [SFCWG] is chartered to describe data plane service encapsulation, control and manageability aspects of service functions. Extensions that will be specified by the SFC WG will be reused in the context of network slicing. Nevertheless, The current charter of the WG does not imply work on the automated discovery of SF instances and their capabilities, nor the automatic discovery of control elements. An additional specification effort is therefore required in this area.

#### 7.2.4. Slice Identification

A network slice data plane, may or may not follow traditional data plane tagging/labeling. However, each network element (router/switch) still has to classify an incoming packet and associated with the slice instance for proper treatment. Network slice instance identification is essential for network element to make local decisions on forwarding policies, QoS mechanism and etc. The performance requirements of a network slice instance can therefore been met by making the correct decision. Meanwhile, it is also important for OAM so that configuration and provisioning can be delicately performed to particular network slice instances by their identifications.

For flow identification, many existing technologies provide mature solutions. These approaches might be able to be re-used in network slicing by adding an additional layer of mapping to a network slice instance ID. The network slice instance ID further maps to a group of performance requirements and OAM profiles, based on which the network elements within the slice can make local decisions. However, per flow level identification could have adverse impact on the scale of the forwarding entries in the routers.

With traditional IP/MPLS VPNs, the set of Route Targets configured for the VPN can be used as some sort of identifier of the VPN in the control plane, and in the data plane, the VPN service labels can be used to identify the data packets belonging to a particular VPN. NVO3 uses the Virtual Network Identifiers (VNIs) in the header of data packets to identify different overlay network tenants. However, It is not clear if the existing identifiers can meet the requirements of network slicing in terms of making local decisions on forwarding policy, QoS and OAM mechanisms, etc.

## 8. Summary

The following table is a summary of the identified gaps based on previous analysis in this document.

Requirements	Gaps
Req 1. Network Slicing Specification (NSS)	1) A detailed specification of NSS 2) A companion YANG data model for NSS
Req 2. Network Slicing Cross-Domain Coordination (NS-CDC)	3) A companion data model for NS-CDC
Req 3. Network Slicing Performance Guarantee and Isolation (NS-PGI)	4) Slicing specific extension on existing technologies
Req 4. Network Slicing OAM (NS-OAM)	5) Mechanisms for dynamic discovery of service function instances and their capabilities. Mechanisms for dynamic discovery of instantiated network slices 6) non-overlay OAM solution 7) Mechanisms for customized granularity OAM

Table 2: Summary of Gaps

## 9. Security Considerations

This document analyzes the standardization work on network slicing in different WGs. As no solution proposed in this document, no security concern raised.

## 10. IANA Considerations

There is no IANA action required by this document.

## 11. Acknowledgements

The authors wish to thank Hannu Flinck, Akbar Rahman, Ravi Ravindran, Xavier de Foy, Young Lee and Igor Bryskin for their detailed and constructive reviews. Many thanks to Mohamed Boucadair, Christian Jacquenet and Stewart Bryant for their valuable contributions and comments.

## 12. References

### 12.1. Normative References

[NS-Framework]  
"NS Framework", <<https://datatracker.ietf.org/doc/draft-geng-netslices-architecture/>>.

[NS-UseCase]  
"NS Use Case", <<https://datatracker.ietf.org/doc/draft-netslices-usecases/>>.

### 12.2. Informative References

[ANI]  
"A Reference Model for Autonomic Networking",  
<[https://datatracker.ietf.org/doc/draft-ietf-anima-reference-model/?include\\_text=1](https://datatracker.ietf.org/doc/draft-ietf-anima-reference-model/?include_text=1)>.

[ANIMA-GRASP]  
"A Generic Autonomic Signaling Protocol (GRASP)",  
<<https://datatracker.ietf.org/doc/draft-ietf-anima-grasp/>>.

[DETNET-WG]  
"Deterministic Networking",  
<<https://datatracker.ietf.org/wg/detnet/about/>>.

[FGEx]  
"5G Exchange (5GEx) - Multi-domain Orchestration for Software Defined Infrastructures",  
<[https://www.researchgate.net/publication/296486303\\_5G\\_Exchange\\_5GEx\\_-\\_Multi-domain\\_Orchestration\\_for\\_Software\\_Defined\\_Infrastructures](https://www.researchgate.net/publication/296486303_5G_Exchange_5GEx_-_Multi-domain_Orchestration_for_Software_Defined_Infrastructures)>.

[FLEXE-1.0]  
"Flexible Ethernet 1.0", <<http://www.oiforum.com/wp-content/uploads/OIF-FLEXE-01.0.pdf>>.

[FlexE-FWK]  
"FlexE-FWK", <<https://datatracker.ietf.org/doc/draft-izh-ccamp-flex-e-fwk/>>.

- [I-D.boucadair-connectivity-provisioning-protocol]  
Boucadair, M., Jacquenet, C., Zhang, D., and P. Georgatsos, "Connectivity Provisioning Negotiation Protocol (CPNP)", draft-boucadair-connectivity-provisioning-protocol-14 (work in progress), May 2017.
- [I-D.ietf-spring-segment-routing]  
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-12 (work in progress), June 2017.
- [I-D.ooamdt-rtgwg-ooam-header]  
Mirsky, G., Kumar, N., Kumar, D., Chen, M., Yizhou, L., and D. Dolson, "OAM Header for use in Overlay Networks", draft-ooamdt-rtgwg-ooam-header-03 (work in progress), March 2017.
- [I-D.wu-opsawg-service-model-explained]  
Wu, Q., LIU, W., and A. Farrel, "Service Models Explained", draft-wu-opsawg-service-model-explained-06 (work in progress), May 2017.
- [I2RS-Yang]  
"A Data Model for Network Topologies",  
<<https://datatracker.ietf.org/doc/draft-ietf-i2rs-yang-network-topo/>>.
- [IMT2020-2015]  
"Report on Gap Analysis", <<http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx>>.
- [IMT2020-2016]  
"Draft Technical Report Application of network softwarization to IMT-2020 (O-041)",  
<<http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx>>.
- [IMT2020-2016bis]  
"Draft Terms and definitions for IMT-2020 in ITU-T (O-040)", <<http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx>>.
- [NGMN-2016]  
"Description of Network Slicing Concept",  
<[https://www.ngmn.org/uploads/media/160113\\_Network\\_Slicing\\_v1\\_0.pdf](https://www.ngmn.org/uploads/media/160113_Network_Slicing_v1_0.pdf)>.



- [NMDA] "Network Management Datastore Architecture",  
<<https://datatracker.ietf.org/doc/draft-ietf-netmod-revised-datastores/>>.
- [NVO3-WG] "Network Virtualization Overlays".
- [ONF-2016] TS, "Applying SDN Architecture to 5G Slicing",  
<[https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/Applying\\_SDN\\_Architecture\\_to\\_5G\\_Slicing\\_TR-526.pdf](https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/Applying_SDN_Architecture_to_5G_Slicing_TR-526.pdf) >.
- [RFC2679] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, DOI 10.17487/RFC2679, September 1999, <<http://www.rfc-editor.org/info/rfc2679>>.
- [RFC2680] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Packet Loss Metric for IPPM", RFC 2680, DOI 10.17487/RFC2680, September 1999, <<http://www.rfc-editor.org/info/rfc2680>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<http://www.rfc-editor.org/info/rfc3209>>.
- [RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, DOI 10.17487/RFC3393, November 2002, <<http://www.rfc-editor.org/info/rfc3393>>.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <<http://www.rfc-editor.org/info/rfc4364>>.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <<http://www.rfc-editor.org/info/rfc4655>>.
- [RFC4664] Andersson, L., Ed. and E. Rosen, Ed., "Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, DOI 10.17487/RFC4664, September 2006, <<http://www.rfc-editor.org/info/rfc4664>>.

- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <<http://www.rfc-editor.org/info/rfc5440>>.
- [RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu, D., and S. Mansfield, "Guidelines for the Use of the "OAM" Acronym in the IETF", BCP 161, RFC 6291, DOI 10.17487/RFC6291, June 2011, <<http://www.rfc-editor.org/info/rfc6291>>.
- [RFC7276] Mizrahi, T., Sprecher, N., Bellagamba, E., and Y. Weingarten, "An Overview of Operations, Administration, and Maintenance (OAM) Tools", RFC 7276, DOI 10.17487/RFC7276, June 2014, <<http://www.rfc-editor.org/info/rfc7276>>.
- [RFC7297] Boucadair, M., Jacquenet, C., and N. Wang, "IP Connectivity Provisioning Profile (CPP)", RFC 7297, DOI 10.17487/RFC7297, July 2014, <<http://www.rfc-editor.org/info/rfc7297>>.
- [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <<http://www.rfc-editor.org/info/rfc7432>>.
- [SFCWG] "\Service Function Chaining (sfc)", <<https://datatracker.ietf.org/wg/sfc/about/>>.
- [SLA-Exchange] "Inter-domain SLA Exchange Attribute", <<https://datatracker.ietf.org/doc/draft-ietf-idr-sla-exchange/>>.
- [TE-Yang] "YANG Data Model for TE Topologies", <<https://datatracker.ietf.org/doc/draft-ietf-teas-yang-te-topo/>>.
- [TEAS-ACTN] "Information Model for Abstraction and Control of TE Networks (ACTN)", <<https://datatracker.ietf.org/doc/html/draft-ietf-teas-actn-info-model>>.
- [TS23-501] "System Architecture for the 5G System", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>>.

Authors' Addresses

Li Qiang (editor)  
Huawei

Email: qiangli3@huawei.com

Pedro Martinez-Julia  
NICT

Email: pedro@nict.go.jp

Liang Geng  
China Mobile

Email: gengliang@chinamobile.com

Jie Dong  
Huawei

Email: jie.dong@huawei.com

Kiran Makhijani  
Huawei

Email: Kiran.Makhijani@huawei.com

Alex Galis  
University College London

Email: a.galis@ucl.ac.uk

Susan Hares  
Hickory Hill Consulting

Email: shares@ndzh.com

Slawomir  
Orange

Email: slawomir.kuklinski@orange.com